Integration of Artificial Intelligence in E-Procurement of the Hospitality Industry: A Case Study in the UAE

دمج الذكاء الاصطناعي في الشراء الإلكتروني لصناعة الضيافة: دراسة حالة في دولة الإمارات العربية المتحدة

by

ELEZABETH MATHEW

A thesis submitted in fulfilment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE

at

The British University in Dubai

May 2020

Integration of Artificial Intelligence in E-Procurement of the Hospitality Industry: A case study in the UAE

by

ELEZABETH MATHEW

A thesis submitted to the Faculty of Engineering and Informatics

in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE

The British University in Dubai, UAE

at

The British University in Dubai

May, 2020

Thesis Supervisor

Prof. Sherief Abdullah

Approved for award:

____________________________	____________________________
Name Designation	Name Designation
____________________________	____________________________
Name Designation	Name Designation

Date:________________

DECLARATION

I warrant that the content of this research is the direct result of my own work and that any use made in it of published or unpublished copyright material falls within the limits permitted by international copyright conventions.

I understand that a copy of my research will be deposited in the University Library for permanent retention.

I hereby agree that the material mentioned above for which I am the author and copyright holder may be copied and distributed by The British University in Dubai for the purposes of research, private study or education and that The British University in Dubai may recover from purchasers the costs incurred in such copying and distribution, where appropriate.

I understand that The British University in Dubai may make a digital copy available in the institutional repository.

I understand that I may apply to the University to retain the right to withhold or to restrict access to my thesis for a period which shall not normally exceed four calendar years from the congregation at which the degree is conferred, the length of the period to be specified in the application, together with the precise reasons for making that application.

____Elezabeth___

Signature of the student

COPYRIGHT AND INFORMATION TO USERS

Abstract

The hospitality industry is growing at an increasingly fast pace across the world which results in accumulating a large amount of data, including employee details, property details, purchase details, vendor details, and so on. The industry is yet to fully benefit from these big data by applying Machine Learning (ML) and Artificial Intelligence (AI). The data has not been investigated to the extent that such analysis can support decision-making or revenue/budget forecasting. The data analytics maturity model is used as the conceptual model for evaluating both data analytics and data governance in this research. In this paper, the author has explored the data and produced some useful visual reports, which are beneficial for top management, as the results provide additional information about the inventoried data by applying ML. Demand forecasting is done using deep learning techniques. Long short-term memory (LSTM) is used to find the demand forecasting of spend and quantity using time lags.

The research proposes an extended framework for integrating AI within the e-procurement of the hospitality industry. The AI integrated technologies will enable stakeholders of the industry to be interoperable with all the providers and sub-providers to obtain information easily and efficiently to identify the best solution for their requirements. The proposed framework of integrating AI in the conceptual framework could be used by medium to large enterprises for interoperability, interconnectivity and to take optimum decisions.

This paper has uses six ML methods to check the accuracy scoring of the predicted duration of purchase. The duration is predicted using feature variables, including recent purchases, frequency of purchases, spend per purchase, days between the last three purchases, and mean and standard deviation of the difference between purchase days. Logistic Regression, XGBoost, and Naïve Bayes models have proven to be useful for this kind of study where three different scenarios are drawn. Other major results of the research include an answer to what to buy when to buy and how much to buy using demand forecasting for the e-procurement in the hospitality industry. The novel LSTM time series algorithm proved to work best for demand forecasting. Various descriptive, diagnostic, predictive, and prescriptive analysis is done on the e-procurement data. The deep learning model developed can perform thousands of routine and, repetitive tasks within a fairly short period compared to what it would take for a human being without any compromise on the quality of work.

Finally, an interview with a subject matter expert is done to evaluate the result and confirm the importance of the study. A survey is also conducted with people involved in the procurement process as part of triangulation. The survey revealed 92% of participants agreed that having an integrated e-procurement framework is very important for the hospitality industry. The integration of AI and ML in e-procurement will revolutionise the hospitality industry.

Keywords – artificial intelligence, data analytics maturity model, hospitality, systematic literature review, big data, predictive analytics, e-procurement, machine learning, conceptual framework, demand forecasting

نبذة

تنمو صناعة الضيافة بوتيرة أسرع في جميع أنحاء العالم مما يؤدي إلى تراكم كمية هائلة من البيانات من حيث تفاصيل الموظف وتفاصيل العقار وتفاصيل الشراء وتفاصيل المورد، إلخ. يمكن للصناعة الاستفادة بشكل كبير من هذا الكم من البيانات من خلال تطبيق التعلم الآلي والذكاء الاصطناعي خاصة فيما يخص اتخاذ القرارت والتنبؤ بالميزانيات. يُستخدم نموذج نضج تحليلات البيانات كنموذج مفاهيمي لتقييم كل من تحليلات البيانات وإدارة البيانات. في هذه الورقة ، حاول المؤلف استكشاف البيانات وإنتاج بعض التقارير المرئية المفيدة التي تفيد الإدارة العليا التي تقدم معلومات إضافية حول البيانات المخزنة من خلال تطبيق التعلم الآلي. يتم التنبؤ بالطلب باستخدام تقنيات التعلم العميق. كما يتم استخدام LSTM للعثور على توقعات الطلب على الإنفاق والكمية باستخدام الفواصل الزمنية.

يقترح هذا البحث إطار عمل جديد لدمج الذكاء الاصطناعي في المشتريات الإلكترونية لصناعة الضيافة. ستتيح تقنيات الذكاء الاصطناعي المتكاملة لأصحاب المصلحة في الصناعة أن يكونوا قابلين للتشغيل المتبادل مع جميع مقدمي الخدمات والموردين الفرعيين للحصول على المعلومات بسهولة وكفاءة للحصول على أفضل حل لمتطلباتهم. يمكن للمؤسسات المتوسطة والكبيرة استخدام الإطار المقترح من خلال دمج الذكاء الاصطناعي في الإطار المفاهيمي من أجل التشغيل البيني والترابط واتخاذ القرارات المثلى.

استخدمت الورقة 6 طرق للتعلم الآلي للتحقق من دقة تسجيل المدة المتوقعة للشراء. يتم توقع المدة باستخدام متغيرات الميزات مثل الشراء الأخير ، وتكرار الشراء ، والإنفاق على الشراء ، والأيام بين آخر ثلاث عمليات شراء ، والانحراف المعياري للفرق بين أيام الشراء. أثبت نموذج الانحدار اللوجستي و XGBoost و Naive Bayes أنه مفيد لهذا النوع من الدراسة حيث يتم رسم ثلاثة سيناريوهات مختلفة. تتضمن النتائج الرئيسية الأخرى للبحث إجابة على ما تشتريه ، ومتى تشتري ، وكم تشتري باستخدام التنبؤ بالطلب للمشتريات الإلكترونية في صناعة الضيافة. أثبتت خوارزمية السلسلة الزمنية LSTM الجديدة أنها تعمل بشكل أفضل للتنبؤ بالطلب. يتم إجراء العديد من التحليلات الوصفية والتشخيصية والتنبؤية والوصفية على بيانات المشتريات الإلكترونية. إن نموذج التعلم العميق الذي تم تطويره قادر على أداء آلاف المهام الروتينية المتكررة في غضون فترة زمنية أقصر إلى حد ما مقارنة بما يتطلبه الإنسان دون أي مساومة على جودة العمل.

أخيرًا ، يتم إجراء مقابلة مع ذوي الخبرة في الموضوع لتقييم النتيجة وتأكيد أهمية الدراسة. تم إجراء المسح مع الأشخاص المشاركين في عملية الشراء كجزء من التثليث. كشف الاستطلاع أن 92٪ من المشاركين وافقوا على أن وجود إطار عمل متكامل للشراء الإلكتروني مهم جدًا لصناعة الضيافة. إن دمج الذكاء الاصطناعي و التعلم الآلي في الشراء الإلكتروني سيحدث ثورة في صناعة الضيافة.

Dedication

This dissertation is dedicated to my father, (late) Mr. Mathew Ninan M who was my constant support and confidence all through my life. I wouldn't have completed this without the never-ending blessing of my beloved father from wherever he is now.

Acknowledgments

Thank God for providing me with enough courage and patience to finish my research. I would like to express my gratitude to my husband, Babu Varghese, who has been the pillar and cheerleader to complete my research.

I would like to express my deepest gratitude to my supervisor Prof. Sherief Abdullah, Professor, The British University in Dubai, UAE, for his patience, professional guidance, and motivation. His works, rigorous scholarship, knowledge, and wisdom fascinated me intensely.

My special thanks to Prof. Khaled Shalaan and Prof. Sufian Forawi for their continuous encouragement. Also, all other respected professors who have been my constant inspiration. I would like to extend my sincere appreciation to the rest of the staff at the British University for their support and understanding.

This thesis is for my dearest husband, I wouldn't have made it without his continuous and tireless support. It is also for my loving children, Daniel and Anina, for their support, patience, and understanding. I would like to thank my entire family including my mother and brothers, for their love, support, encouragement, and trust are everlasting.

Table of Contents

Chapter 1: INTRODUCTION ... 1
1.1 Problem Statement ... 4
1.2 My Contributions ... 7
1.3 Significance of the Study ... 9
1.4 Theoretical Framework ... 13
1.5 Research Aims and Objectives ... 15
1.6 Research Questions ... 16
1.7 Organisation of the Thesis ... 17
Chapter 2: LITERATURE REVIEW ... 19
2.1 Integration of AI ... 19
2.2 E-procurement in the Hospitality Industry ... 20
2.2.1 Benefits & Barriers of E-procurement ... 22
2.2.2 Quality of Data in E-procurement of Hospitality Industry ... 24
2.2.3 Data Analytics in the Hospitality Industry ... 25
2.2.4 Predictive Analytics in the Hospitality Industry ... 29
2.2.5 Deep Learning and Machine Learning Techniques in E-procurement of the Hospitality Industry ... 33
2.2.6 Optimisation in e-procurement ... 35
2.2.7 IoT in Hospitality ... 35
2.2.8 Agent-Based Technology (ABT) in the hospitality industry ... 39
2.2.9 Ecosystem in Hospitality ... 41
2.3 Gap in the literature ... 43
2.4 Chapter Summary ... 45
Chapter 3: A SYSTEMATIC REVIEW ... 46
3.1 Structured Review ... 46
3.1.1 Statistics of the trend in publishing ... 47
3.1.2 Major areas of research ... 48
3.1.3 Content analysis in the selected publications ... 49
3.1.4 Comparing various studies in the field of AI /e-procurement/supply chain /hospitality .. 50
3.2 Machine Learning and Deep Learning used for Prediction and Forecasting ... 64
3.2.1 A conceptual framework for applying machine learning and deep learning ... 68
3.2.2 Classification algorithm selection ... 69

3.2.3 Cross-Validation.. 71
3.2.4 Fine Tuning the hyperparameters with GridSearch.. 72
3.2.5 LSTM & RNN.. 72
3.3 Current Level of Adoption 73
3.4 Major Gaps in the Literature 75
3.5 Chapter Summary.. 77
Chapter 4: RESEARCH METHODOLOGY 78
4.1 An Exploratory Case Study.. 78
4.2 Research Paradigm... 78
4.3 Challenges Faced During Research.. 79
4.4 Situating the Research 80
4.5 Exploring the Setting.. 80
4.6 Data Collection... 81
4.7 Proposed Method for Designing the Framework 82
4.7.1 Use Case: Initiate Request and Submit Request for e-procurement................................ 85
4.7.2 Components of AI Integrated Framework... 88
4.8 Proposed Method for Forecasting Problems 89
4.9 Validity of the Results... 92
4.10 Ethical Considerations... 92
4.11 Chapter Summary.. 92
Chapter 5: DATA ANALYSIS 93
5.1 Data Analytics... 93
5.2 Data Selection, Slicing, and Filtering.. 94
5.3 Data Clustering.. 94
5.4 Forecast Error Measures.. 94
5.4.1 Root Mean Square Error (RMSE) 95
5.4.2 Adjusted R Squared.. 95
5.5 Baseline Prediction... 96
5.5.1 Linear regression 96
5.5.2 Persistence forecast 96
5.6 Model Selection 96
5.7 Chapter Summary.. 97
Chapter 6: DESCRIPTIVE ANALYTICS 99

6.1 Supplier Spend Summary 99
6.2 Spend summary per property 100
6.3 Spend report by purchase type 101
6.4 Spend report by category 102
6.5 Spend by product name 103
6.6 Discussions 104
6.7 Chapter Summary 105
Chapter 7: DIAGNOSTIC ANALYTICS 106
7.1 Pie chart 106
7.2 Time Series Graph 107
7.3 Stacked Bar Graph 110
7.4 Line Graph 112
7.5 Discussions 113
7.6 Chapter Summary 115
Chapter 8: PREDICTIVE & PRESCRIPTIVE ANALYTICS 116
8.1 Duration of purchase 116
8.2 Demand Forecasting 126
8.3 Cognitive Analytics 132
8.4 Discussions 133
8.5 Chapter Summary 134
Chapter 9: NEW PROPOSED CONCEPTUAL FRAMEWORK 135
9.1 New proposed conceptual framework for the hospitality industry 135
9.2 Conceptual Model for E-procurement in the Hospitality Industry 138
9.3 Discussions on the Extended Framework 140
9.4 Summary of the Interview 142
9.5 Summary of Survey 147
9.6 Statistical Analysis of the Survey 153
9.7 Validate the qualitative study 155
9.8 Further Discussions 157
9.9 Limitations of Deploying the Proposed Framework 159
9.10 Chapter Summary 160
Chapter 10: LIMITATIONS AND RECOMMENDATIONS 162
10.1 Limitations 162

10.2 Recommendation for Implementing the AI Integrated Framework 163
10.3 Validity, Reliability, and Generalisability.. 165
10.4 Chapter Summary.. 166
Chapter 11: CONCLUSION AND FUTURE WORKS 167
11.1 Conclusions 167
11.2 Impact of the Contributions.. 169
11.3 Further Works 171
11.4 Chapter Summary.. 172
REFERENCES.. 173
APPENDICES... 185
I. Descriptive Analytics.. 185
A. Buyer Summary.. 185
B. Spend Summary 185
C. Spend report (yearly).. 185
D. Spend by Purchase type.. 185
II. Diagnostic Analytics 186
A. Pie Chart... 186
B. Time Series... 186
C. Trend Analysis 186
D. Stacked Bar Chart... 186
E. Line graph 187
III. Machine Learning Algorithm for Finding Purchase Duration.................................... 187
A. Import data 187
B. Import Libraries.. 187
C. Date conversion to datetime data type 188
D. Cut off entire data into two [9 month and 3 months] for finding purchase duration........... 188
E. Finding latest purchase date and earliest purchase date from the cutoff point.................... 188
F. Finding purchase duration... 189
G. Adding Initial feature 189
H. Plot Features... 189
I. Elbow method for making cluster 190
J. Finding overall score.. 191
K. Scatter plot between each features. 191

L. Adding more features 192
M. Adding label 192
N. Create Heat map 193
O. Split into test and train data 193
P. Baseline Model 194
Q. Machine Learning Model Selection 195
R. Finding Accuracy 195
S. Fine Tuning the model 195
Scenario 1: Duration of purchase for the frequent buyer 196
Scenario 2: Duration of purchase for the most spend product 198
Scenario 3: Duration of most purchase commodity 200
IV. Machine Learning Algorithm for Demand forecasting 202
A. Data Wrangling 202
B. Aggregate spend by month 202
C. Data Transformation to make stationary data 203
D. Baseline Prediction 204
E. Adding time lags 205
F. Finding optimum Adjusted R^2 205
G. Train-Test-Split 206
H. Scaling the data 206
I. Building LSTM Model 206
J. Fit the LSTM model 206
K. Use the model for prediction 207
L. Inverse Transform 207
M. Plot Actual and Predicted data for Spend forecast 208
N. Plot Actual and Predicted data for Quantity forecast 209
V. Survey Consent Form 210
VI. Survey Design 210
A. Demographic details 211
B. Importance of the integrated system 212
C. Knowledge and recommendation 212

List of Illustrations

Fig.1: Main concepts in e-procurement adapted from Podlogar, 2007

Fig.2: Five stages of analytics continuum adapted from Krol and Zdonek, 2020

Fig.3: Data Analytics Maturity Model adapted from Król & Zdonek, 2020

Fig.4: Initial framework for predictive analytics adapter from International Research Organisation's Symposium, 2019

Fig.5: Initial AI framework adapted from International Research Organisation Symposium, , 2019

Fig.6: Stakeholders of the hospitality sector adapted from Buhalis and Leung, 2018

Fig.7: Smart hospitality framework adapted from Buhalis and Leung, 2018

Fig.8: Stages of the selection process of literature for review

Fig.9: Conceptual processes involved in forecasting problems

Fig 10: Use case: initiate request and submit request adapted from International Research Organisation (2019)

Fig.11: Method proposed for forecasting

Fig.12: Supplier spend summary of X Hotel

Fig.13: Spend summary of each property in the UAE

Fig.14: Spend by purchase type report

Fig.15: Spend by category report

Fig.16: Spend by product name visual report

Fig.17: Pie chart on spending by category

Fig.18: Time series graph on spend vs occupancy for the year 2018 and 2019

Fig.19: List of UAE national holidays in 2018 and 2019

Fig.20: Occupancy trend in 2019 in each property

Fig.21: Stacked bar graph of top 25 category of the year

Fig.22: Price distribution of two top most products across various properties

Fig.23: Data visualization of feature variables

Fig.24: Scatter plot between each feature variables

Fig.25: Heatmap to check collinearity between feature variables and label

Fig.26: Accuracy of prediction from various models in different scenario

Fig 27: Box and whisker plot to perceive the spread of accuracy score

Fig.28: The increasing trend for spending

Fig.29: Baseline graph for spend forecasting

Fig.30: Stationary time series graph

Fig.31: Spend forecast for last 6 months

Fig.32: Quantity forecasted for last 6 months

Fig.33: Conceptual framework of AI integrated hospitality system

Fig.34: The proposed conceptual model for e-procurement in the hospitality industry

Fig.35: Response from participants for question 4

Fig.36: Response from participants for question 5.

Fig.37: Response from participants for question 6

Fig.38: Response from participants for question 7

Fig.39: Response from participants for question 8

Fig.40: Response from participants for question 9

Fig.41: Response from participants for question 10

Fig 42: Descriptive statistics on survey responses

Fig 43: Frequency distribution of organizational size

Fig 44: Correlation between responses

List of Charts

Chart 1: Statistics of literature selected

Chart 2: Major topics of research in the selected articles

Chart 3: Keywords in the publication

Chart.4: Stages of Adoption (International Research Organisation- Artificial Intelligence Trends: Decision Augmentation,2020)

Chart.5: Decision Augmentation barriers by International Research Organisation (published 2020)

List of Tables

Table 1: Recommended IoT location from Buhalis and Foerste, 2015

Table 2: Integrated software agents and their description (Bokolo Anthony Jnr, 2018

Table 3: Insight into works of literature in the area of AI in procurement/supply chain /hospitality

Table 4: A synopsis of machine learning used publications

Table 5: Actors: Users, Buyers, Suppliers, Applications, and other stakeholders

Table 6: Actual vs Predicted values for spend and quantity

Table 7: Mean and SD of the spread of accuracy

Table 8: Actual vs Predicted values for spend and quantity

List of Abbreviations

AI	-	Artificial Intelligence
ML	-	Machine Learning
DL	-	Deep Learning
SCM	-	Supply Chain Management
BDBA	-	Big Data Business Analytics
LSCM	-	Logistic and Supply Chain Management
SCA	-	Supply Chain Analytics
PA	-	Predictive Analytics
LR	-	Logistic Regression
ABT	-	Agent Based Technology
LSTM	-	Long Short-Term Memory
XGBoost	-	eXtreme Gradient Boost
NB	-	Naïve Bayes
SVM	-	Support Vector Machine
RMSE	-	Root Mean Square Error
MSE	-	Mean Square Error

Publications

1. Intelligent Transport Systems and Its Challenges
   2020 | book-chapter

   DOI: 10.1007/978-3-030-31129-2_61 *[unrelated]*

2. Vision 2050 of the UAE in Intelligent Mobility

   2018 Fifth HCT Information Technology Trends (ITT)

   2019-02-25 | conference-paper

   DOI: 10.1109/ctit.2018.8649542ISBN: 9781538671474 *[unrelated]*

3. Big Data Analytics in E-procurement of a Chain Hotel

   EIDWT 2019: Advances in Internet, Data and Web Technologies

   2019-02-06 | book-chapter

   DOI: 10.1007/978-3-030-12839-5_27

   Part of ISBN: 9783658251635

   Part of ISBN: 9783658251642

   Part of ISSN: 2197-6708

   Part of ISSN: 2197-6716 ***[related]***

   {included parts in Chapter 1,2,4,5,6,7}

4. Computer Information Sciences Views on Community Services and Computer literacy

   2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)

   2018-11-01 | conference-paper

   DOI: 10.1109/iccceee.2018.8515809ISBN: 9781538641231 *[unrelated]*

ORCID iD
https://orcid.org/0000-0002-8490-7843
Researcher ID: M-8229-20

Chapter 1: INTRODUCTION

The notion of hospitality precedes that of tourism (King, 1995; Higgins, 2009), which came into use in the 1700s[1]. The Oxford Learners Dictionary 2020 defines 'hospitality' as "friendly and generous behaviour towards guests", while the Chambers English Dictionary 2020 (13th edition) defines it as "the friendly welcome and entertainment of guests or strangers, which usually includes offering them food and drink". In contrast, 'tourism' means people traveling for leisure and includes activities such as sightseeing and camping (Wikipedia). The total contribution of the hospitality sector to Global GDP in 2019 was US$9.258 trillion and it is forecast to rise by 4.9 percent per annum, as per UAE government tourism statistics (2019).

The hospitality business purchases products from its suppliers, such as food and drink, in bulk (Hassanien, Dale & Clarke, 2011). The process of obtaining supplies is referred to as procurement. Organisations firmly believe that applying information technology to their procurement processes can produce vital advantages in their operations (Sanchez-Rodriguez, Martinez-Lorente & Hemsworth, 2020). E-procurement is a competitive factor in all types of businesses in which information technology and purchasing resources like purchase orders, supplier details, and catalogues are involved to exchange information and to make purchasing decisions (Sanchez-Rodriguez, Martinez-Lorente & Hemsworth, 2020).

E-procurement provides opportunities to access purchasing networks for suppliers and buyers, expands the selection of products, and makes information more easily obtainable (E. Mathew,2019). E-procurement is one process involved in supply chain management.

[1] (2020). Available at: https://study.com/academy/lesson/hospitality-industry-history-origin.html

“Supply chain management (SCM) is the process and activity of sourcing the raw materials or components an enterprise needs to create a product or service and deliver that product or service to customers” (Perkins & Wailgum, 2020). SCM has six components, namely: planning, sourcing, making, delivering, returning, and enabling[2]. Therefore, the process involves manufacturers, distributors, resellers, and suppliers, according to Boris Evelson, Vice President and Principal Analyst at Forrester Research (2018). This means if anybody wants to apply Artificial Intelligence (AI) in SCM, they need to collect data from all of these, which makes the process tedious and time-consuming. “The supply chain is the entire process, while procurement is a part of it” (Tom, 2020). This concept is depicted in Figure1.

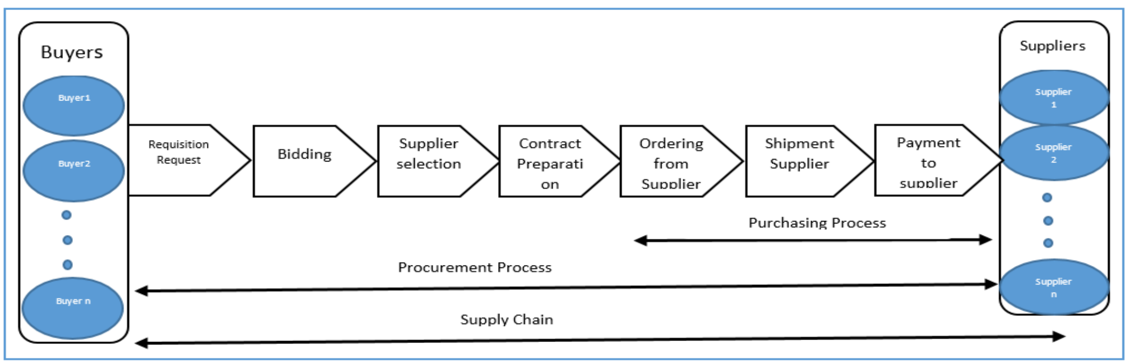


Fig. 1 Main concepts in e-procurement adapted from Podlogar, 2007

For instance, an epidemic outbreak, such as a novel coronavirus (COVID-19), has a massive impact on SCM. When the epidemic is at its peak it will compel companies to initiate some kind of regulation or shut assembling and manufacturing plants, at least temporarily. Moreover, the imports and exports from the affected countries will decrease drastically[3]. An

[2] (2017) https://www.cio.com/article/2439493/what-is-supply-chain-management-scm-mastering-logistics-end-to-end.html
[3] (2020) https://hbr.org/2020/03/what-are-companies-legal-obligations-around-coronavirus

Agent-Based Technology (ABT) for risk analysis can predict the risk in procuring commodities from the affected region and changes in exchange rates. These predictions will help the procurement managers to take proactive steps in planning the demand and supply based on the aftermath of economical fluctuations and travel controls.

In this paper, e-procurement has been used interchangeably for the supply chain to obtain a broader view. This research uses e-procurement processes from a chain hotel in the UAE. E-procurement has an important place in the hospitality industry, which contributes over $9.258 trillion to Global Domestic Product (GDP) per annum. In international hotel chains, the importance of e-procurement is much greater as it enables strategic growth and supports hotels to segregate the final products and services from competitors (Daghfous & Belkhodja, 2019).

Integration of AI is peaking, especially in the UAE, as the visionaries of the country proclaimed that "Artificial Intelligence becomes an important aspect for building the future, and we focus our efforts on building the future of the UAE based on a forward-looking vision that adopts global trends and developments, and supports brightest young minds and encourage innovation and creativity"[4]. As quoted, the UAE is striving to integrate AI in almost all fields. It is worth mentioning that Dubai is named as the leading city in international visitor spending worldwide by Statista - the No.1 business data platform - and the occupancy rate has increased in past years in Dubai. Hence, the integration of AI in the hospitality industry is essential.

[4] (2018) https://u.ae/en/about-the-uae/the-uae-government/government-of-future/innovation-in-the-uae

1.1 Problem Statement

There are several problems in the e-procurement of the hospitality industry. McKinsey analytics who works together with clients to build analytics-driven organisations and achieve better performance through data published an article in 2018 following a survey on 'Business functions where AI has been adopted'. As per the study, business functions in which AI has been adopted by industry do not even mention the hospitality industry. Only 18% of travel, transport, and logistics industry respondents specified adoption of AI and, even in other industries, the adoption of AI in the supply chain is comparatively less. This study proposes an AI integrated e-procurement system in the hospitality sector by using the data analytics maturity model to eradicate several issues faced in the procurement processes. The enumerated problems with sub-problems are given below in separate paragraphs.

Product taxonomy in e-procurement is a major issue because there is no standard taxonomy setup per product ["A taxonomy in procurement is the categorical hierarchy of spend and sourcing groups, from general to specific "(Wiki)]. Therefore, individual buyers or suppliers are free to choose any category for a product. For example, the product 'Fresh Chicken' could be categorised as a 'Coffee' item and the system will accept it. The user can key in any category and the system will not reject it even if it is wrong. This impacts the analytics and requires tremendous effort in constructing an item master. ["An item master is a record that lists key information about an inventory item" (Wiki)].

Currently, the buyers can buy from listed suppliers or from local suppliers who are part of a vast administration structure. For example, the suppliers need to provide a lot of documentation to confirm their compliance with the qualification requirements. Complete automation of this process will increase availability and transparency. Moreover, there exists a significant amount of duplication of work taken by either a supplier or a buyer. Furthermore, there exist duplicate suppliers and products in the system.

The prices for commodities bought from the same vendor at different properties showed variations at a given time. A single system that integrates across all regions with a standard rate for each commodity from individual vendors was missing, where vendors can maintain their negotiated fixed price for that period in the system.

There is always a world of difference between what is on the documentation and the actual product. The issue is that supplier compliance is not recorded anywhere in real-time because such non-compliance from the supplier will not reach top management for review during the supplier selection process.

Another problem is budget development for procurement. The budget for the given financial year should be developed and approved beforehand. Most of the time, buyers over- or under-budget due to the bullwhip effect ["The bullwhip effect is a distribution channel phenomenon in which forecasts yield supply chain inefficiencies" (Wiki)]. There exists no mechanism to predict the budget accurately.

Furthermore, their contingency plans are not fool-proof against cyberattacks which are quite common nowadays. As per the study conducted by Issabayeva, Yesseniyazova, and Grega (2019), during three quarters of 2018, there were more than 14,000 cyberattacks while the number of registered attacks in 2019 exceeded 16,000. "Lloyd's study describes cyberattacks as one of the costly risks. According to their calculations – it could cost the global economy more than £120 billion (Guardian 2017)" (Issabayeva, Yesseniyazova & Grega, 2019).

The X Hotels [*a pseudonym to maintain anonymity*] have a database server to collect big data for purchasing requests. This system is limited in its reporting facility for managers' extract data from the Oracle database into an Excel sheet. In their opinion, Oracle is limited to data visualisation and reporting together. In other words, their current system cannot display the report as per the requirement of managers. Therefore, the managers extract data from the database to compile a monthly report of expense records. Foremost businesspeople are seriously thinking about incorporating the data analytics maturity model, as it offers limitless possibilities of the commercial revolution and functioning efficiency. Headlines on Amazon and Walmart using data analytics has captured the attention of both practitioners and researchers (Ittmann, 2015). SCM industries are flooded with data to such a degree that a recent report noted that "businesses collect more data than they know what to do with" (Fosso Wamba et al. 2018). This is particularly true for the hospitality industry as they use collected data only for biweekly or monthly reports, which are used for comparative study. Although most companies have high expectations for data analytics the actual use is restricted and many multinationals struggle to reveal its business worth.

To bring awareness to practitioners of e-procurement in the hospitality sector, this paper is mostly focused on 1) analysing previously published papers on AI connected with e-procurement and hospitality, 2) using the data analytics maturity model to apply Machine Learning (ML) techniques on historical data of a chain hotel to optimise e-procurement data, 3) investigating how the current hospitality framework can integrate AI in its functionality using data analytics maturity model and, finally 4) identifying the perception of a subject matter expert(s) on the results obtained. Furthermore, the study explores the hotel's purchasing function to understand the purchasing process and identify the limitations of using e-procurement data.

Nowadays, the passionate race in an industrial setting means that tourism and hospitality businesses have to toil rigidly to sustain and progress their competitiveness. As the industry includes addressing the human interest to travel and to embrace the warmth of new experiences, understanding customer interests and being able to harness both, can promote guest loyalty by proactively adding value to their expectations using cognitive technology. Hence, digital or electronic technology supports the organisation in managing information dynamically and influences business competitiveness by assisting decision-makers to make appropriate investments and decisions.

1.2 My Contributions

This author's major contributions include the integration of AI and applying ML to optimise e-procurement in the hospitality industry by developing an extended framework as a roadmap for all stakeholders of procurement processes. Having such a system incorporated following the data analytics maturity model will help to solve the enumerated problems.

My first contribution is an extended framework integrated with AI in the e-procurement of the hospitality industry. Item master and product taxonomy should be defined clearly and accurately as part of data models. Real-time supplier compliance will be recorded, which will support the calculation of supplier rating and supplier selection, while negotiations will be easier. The completely integrated intelligence system will be interconnected and interoperable, as well as able to sense, accumulate, explore, and infer data dynamically. Standardised communication protocols agreed among stakeholders across the ecosystem, will increase interoperability, interconnectivity, and thus, trustworthiness. Thereby, the key stakeholders or decision-makers can make use of the AI integrated decision support system and use management software for a variety of cases to improve their business and strategic planning.

The framework follows the data analytics maturity model. The data analysis based on a descriptive, diagnostic, predictive, prescriptive, and cognitive level is my other contributions. This will help the managers to obtain varied reports with numeric and graphical representations of data. Descriptive and diagnostic numerical and visual reports are my second and third contributions.

An ML algorithm for demand forecasting using LSTM time-series predictions is my fourth contribution. ML is applied to historical data collected over 5 years for demand forecasting of spend and quantity. The ML algorithm ensemble with deep learning gives an accurate prediction which will help the top management in budget forecasting.

An ML algorithm for finding purchase duration is my fifth contribution. Purchase duration is found using a supervised ML technique which will give top management an idea of what to buy when to buy and how much to buy. This result will help the stakeholders be ready for better cash flow.

A systematic literature review on the relevant topic is my next (sixth) contribution. A systematic literature review is conducted on literature related to AI, e-procurement, and hospitality. This systematic study details various results showing the trend of publication, most used keywords, and major topics researched.

An interview with a subject matter expert and a survey with people working in the procurement process are further contributions (seventh) of this study. A detailed interview with a subject matter expert is conducted to evaluate the results obtained in this research. The same is validated by surveying people involved in procurement processes. The survey conducted with procurement professionals will be my eighth contribution.

Overall, this thesis makes significant contributions. The conceptual framework proposed in this thesis is beneficial for future research in AI linking hospitality and e-procurement. The recommendations will be helpful for all stakeholders to maintain a long-term business outlook by having automated decision making, interoperability and interconnectivity, and ML to optimise procurement in the supply chain of the hospitality sector, and implementing suitable actions. Moreover, the framework and its components are generalisable to the hospitality industry.

1.3 Significance of the Study

Most of the previous research that studied IT in hospitality, focused on the adoption of IT (Alsaadi & Tubaishat, 2020; Helo, Gunasekaran & Rymaszewska, 2017; Korte, 2013) while my work focuses on the integration of AI in particular and applying ML in e-procurement within the hospitality industry. The truth is the hospitality industry has grown extensively in most of the world, however, there is limited research published in the scientific world on hospitality linked with procurement and AI (Gomezelj, 2016). However, the travel and tourism industry has initialised many kinds of business and technology research, which will impact revenue directly (Mariani, Baggio, Fuchs & Höepken, 2018). The stakeholders involved in the procurement process of the hospitality industry are numerous (Boris Evelson, 2018), and obtaining approval from all of them to use their data is a real challenge (McKinsey, 2018). There is no doubt that there is a high demand for state-of-the-art techniques and technology wherever these are found to be apt and useful (Korte, 2013). Adapting such a state-of-the-art infrastructure is crucial in this era when technology is growing exponentially (Treiblmaier, 2018).

Moreover, such advancement is especially required for competitive operations in the hospitality sector, as well as to strategise the marketing and distribution of these sectors on a global scale. Over the last few decades, research and development (R&D) in the area of AI have drastically advanced (Bogetić, Antić & Lekić, 2017). As a result, in many fields, R&D departments have been integrated for the same purpose, especially as a large amount of data is being accumulated in all industries (Mariani, Baggio, Fuchs & Höepken, 2018). The hospitality industry has only recently integrated digital technologies in a systematic manner (Edghiem & Mouzughi, 2018). Few early studies were anecdotal and did not make any contribution to the industry, nor to academic research, as they focused on the individual operation or the locality only (Brandon-Jones & Kauppi, 2018).

Another major reason for late adoption is that adoption requires a large amount of time and money, so the industry, especially small and medium enterprises, showed reluctance (Lamba

& Singh, 2017). Furthermore, senior management often finds it hard to understand the concept of AI in the hospitality industry as they cannot relate it to any business benefits or profit (Mariani, Baggio, Fuchs & Höepken, 2018). As per an International Research Organisation (2018), another reason for such changes in any industry is the lack of talent or skilled people in the organisation.

As discussed, one of the major problems of this topic is limited academic studies exploring the hospitality industry linked with AI in procurement processes and at the same time not having skilled personnel to implement any state-of-the-art technology. This study is vital for hospitality as it provides the stakeholders of the procurement process of the hospitality industry with a roadmap on how to integrate AI and ML in e-procurement. The study shows that knowledge gathered online or offline should be used further for data analytics and predictions and thus to bring greater benefit to organisations efficiently and effectively. Content analysis on recently published articles and visual representations of data is conducted intensively. Demand forecasting to optimise e-procurement data by applying deep learning will ease the decision-making for procurement stakeholders. The author proposes a new conceptual framework for the integration of AI and applying ML in the hospitality industry and particularly in e-procurement.

The spread of IT generates opportunities and pressures in the hospitality industry that are aggressive and influencing IT implementation. To enhance customer satisfaction, online data retrieval, online booking, feedback, and so on, have been adopted across the hospitality industry. Integration of AI is considered the latest trend and a lot of opportunities for data analytics have been forecast. Even though integration has substantially expanded information availability and transparency, intelligent analytics has been trialled to a minimal level in e-procurement, especially in hospitality. This paper tries to highlight the e-procurement operation in the running of a chain of hotels. Furthermore, it investigates the use of procured data and most importantly to optimise e-procurement data with ML.

The required information will be collected through qualitative and quantitative methods. Intense content analysis is conducted to identify the integration level of AI and ML. Furthermore, document analysis is undertaken on the proposed new framework with AI to review its integration with the current architecture within the industry. Data is collected for a chain of hotels in the UAE over the past five years for descriptive, diagnostic, predictive, and prescriptive analysis. Demand forecasting is done by applying ML with deep learning. Finally, an interview and survey are conducted with a subject matter expert to understand the perception of the results derived in the study. Furthermore, the researcher reviews the feasibility for future works that can be derived from this study with a synopsis into cognitive analytics.

The study is valuable for e-procurement within the hospitality industry, as it will enhance strategic decision-making capabilities, risk assessment, cash-flow management, and thus, reduce costs and increasing revenue. Additionally, with the latest spread on developing refined ML-based techniques and, in particular, with the deep learning algorithms, higher accuracy and powerful results obtained can be an eye-opener to stakeholders within the hospitality industry. These results will help them to foresee many other issues to be solved in e-procurement across the industry. Though the novel centralised framework that is integrated with AI will take time for complete implementation, it will solve problems on product taxonomy, duplication of supplier, product and work, programme roll-out compliance, procurement budget forecasting, demand forecasting, and many more.

1.4 Theoretical Framework

The research can be divided into two main sections. One for creating the framework that integrates AI in the e-procurement system and the other for big data analytics. In this paper, the data analytics maturity model is used as the conceptual model for developing both studies. The data analytics maturity model is described under five main classifications: descriptive, diagnostic, predictive, prescriptive, and cognitive analytics. There are several data analytics maturity models developed by various organisations using a variety of terminology. It is interesting to note that all of these summarise similar concepts. The author in her research has followed the model defined by Krol and Zdonek (2020), specifically influenced by International Research Organisation's maturity model. Traditional analysis methods include self-assessment, qualitative interviews, and quantitative studies. Data analytics that are fully developed can be defined as the progression of an organisation in integrating, managing, and leveraging all significant internal and external data sources into key decision points. The ecosystem thus formed should facilitate insight and action. Hence, data analytics maturity is not just having technology in place, rather it should have the technology, data management, analytics, governance, and organisational components. Moreover, it may take a long time for proper implementation. Analytics maturity refers to how intensely and successfully the organisation uses tools, people, processes, and strategies to accomplish and scrutinise data to inform business decisions. Maturity models are used to direct this conversion process.

The maturity model can be assessed to check the organisation's current maturity level. The assessment is usually conducted using various lengths and widths. These dimensions may include an assessment of technical infrastructure, assessment of organisational issues, and an assessment of human resources. Technical infrastructure would include all the equipment, software, and data collection method. Organisation issues include analytics culture, the degree of support and democratisation of analytics, and the level of acceptance towards an analytics culture in the entire enterprise. Human resources consist of the staff's analytical competencies. When assessing an organisation, it can fall under five categories depending on their level of data analytics maturity level. The five stages of data analytics are depicted in Figure 2. An organisation that is unaware of the benefits of data analytics or that has a negative attitude towards adopting data analytics is referred to as analytically lagging or learning; this is also referred to as the initial stage of an organisation. The next stage is the infected stage where the executives are interested in the idea of data analytics and slowly start to develop models. The third stage is the acceleration stage where the organisation uses data analytics for competitive advantages. The fourth stage is known as the momentum or impulse stage where the data analytics is developed and implemented throughout the organisation providing motivation and inspiration. The fifth stage, known as being ahead of data analytics, is when the company regularly gains benefits of its enterprise-wide analytics capability and focuses on continuous analytics reviews.

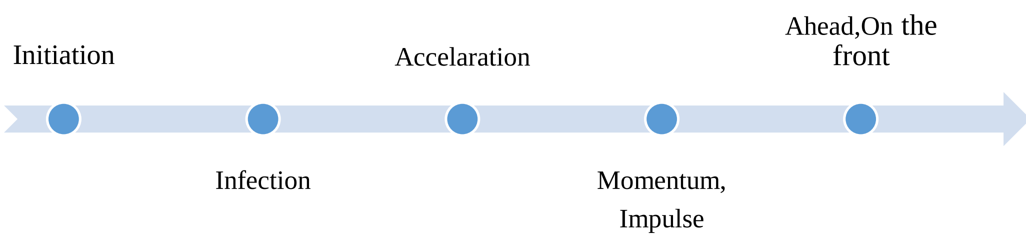


Fig. 2. Five stages of analytics continuum adapted from Krol and Zdonek, 2020

The model can be used by any organisation without regarding their current stage of analytics continuum as the model that is proposed is a roadmap to be ahead of data analytics. Data analytics indicates strategies and directions of development and is a source of innovation that would eventually result in revenue enablement.

1.5 Research Aims and Objectives

The research papers evaluated suggest that for the stakeholders to take advantage of the procurement big data being acquired, more effective and efficient studies need to be conducted. The stakeholders may not have enough background to conduct a realistic study on assessing the effectiveness of the data and the system being used. Thus, this paper intends to provide a proper systematic analysis to suggest efficient ways to enhance the data management system in the hospitality industry, especially in e-procurement by using ML methods and AI. In the cited articles, several authors note the importance of having the latest state-of-the-art technology in general, but the gap is felt in the hospitality supply chain. This paper aims to identify how the latest technological advancement can be used to enhance SCM in the hospitality industry to make better decisions. Another major aim is to compute descriptive, diagnostic, predictive, and prescriptive analysis of historical data of e-procurement within the hospitality industry.

The study is conducted across the 20 UAE-based properties of a globally well-known hotel chain. The spend related details were available, which will help in understanding the cash flow requirements of each property at different times. There was, however, no access to the consumption data, that would have otherwise assisted with a detailed cost-benefit analysis.

The name of the hotel is referred to as ‘The X Hotels’ for this study; the actual name and other non-disclosure details are kept confidential. The X Hotels has more than 20 properties in the UAE ranging in various luxurious statuses from 3 to 7-star rating. Each of the local properties has between 100 to 750 guest rooms.

The objectives of this research are listed below:

1. Create a clear and precise framework on how to integrate a coherent added value AI system to the selected industry.
2. Apply descriptive, diagnostic, predictive and prescriptive analysis in e-procurement data
3. Suggest a noble technology on ML analytics that can use more automatic prediction using AI:
    a. What is applied demand for goods?
    b. ML to optimise procurement
4. Find the perception of seasoned professionals in the selected industry

1.6 Research Questions

To meet the purpose of the study, the author has considered the following as the research questions:

RQ 1) How can the state-of-the-art technologies be part of the contemporary e-procurement system within the hospitality industry?

1. What is the adoption level of AI in the hospitality industry?
2. How can the industry improve performance by integrating AI?

RQ 2) How can the data analytics maturity model be adapted in e-procurement data of the hospitality industry?

a. What visualisation techniques are most effective in exploring e-procurement data in the hospitality industry?

b. How can Machine Learning (ML) analytics optimise procurement efficiency in the hospitality industry using Deep Learning?

RQ 3) What is the perception of subject matter expert(s) on the proposed integrated system with AI and ML in e-procurement of the hospitality industry?

The required information will be collected through qualitative and quantitative methods. As part of the qualitative method, intense document analysis and reviews are completed plus an interview/survey with an expert(s) in the relevant area on the results obtained. Thus, the triangulation is undertaken to validate the model developed. Quantitative analysis will involve various descriptive, diagnostic, predictive, and prescriptive analytics, using the data collected over five years from the hospitality industry, using Python deep learning methods.

1.7 Organisation of the Thesis

The remainder of the thesis is organised as follows. Chapter 2 reviews current research works on related topics. The systematic review is separated from the literature review and kept in Chapter 3. A detailed narration of the research methodology is included in Chapter 4. Chapter 5 provides details of data analysis. The rest of the chapters from 6 to 9 represents the results and discussions of the study. Chapter 6,7 and 8 show results of data analytics including, descriptive, diagnostics, predictive and prescriptive analytics, respectively. Chapter 9 discusses the proposed extended network that integrates AI and ML in the e-procurement of the hospitality industry. The results from the survey and interview of subject matter expert(s)'s perception of the findings in this research are presented in the same chapter. Chapter 10 contains the limitations and recommendations of the research. Chapter 11 provides conclusions and discusses possible future works.

Chapter 2: LITERATURE REVIEW

The review demonstrates that several technology-based innovations could be used in the hospitality sector as a whole, which could benefit from improved decision support. The state-of-the-art technology and techniques being discussed in this paper are big data analysis, AI, ML, and ABT. The chapter also highlights the importance of having quality data for analysis.

2.1 Integration of AI

AI has been formally defined as "technologies [that] aim to reproduce or surpass abilities (in computational systems) that would require 'intelligence' if humans were to perform them. These include: learning and adaptation; sensory understanding and interaction; reasoning and planning; optimization of procedures and parameters; autonomy; and creativity"[5]. AI methodologies established in the earlier decade predominantly used algorithms have drastically developed the proficiency of AI to identify complex patterns, optimize for particular results, and create automated decisions. To conduct all these needs massive amount of relevant and quality data, a strong algorithm or a machine learning method, a narrow domain and a concrete objective. The consolidation can result in intense enhancements in reliability, efficiency and productivity and thus becomes outcome driven (WEF, 2019).

Operationally, big data allows the researchers to work with the entire population under analysis as per Gerald et al. (2016). Furthermore, the studies make it more reliable and powerful to address novel research questions that will give innovative frameworks to the enhancement of knowledge, ultimately to top management decisions as mentioned by Gerald

[5] Definition from the Engineering and Physical Science Research Council, a UK government research funding body

et al. (2020). For that, it is important to have an excellent business intelligence system that handles the big data to extract the specific knowledge required for managerial decisions (Verhoef et al., 2016). Although the competition is significant within the hospitality industry, the question is still what kind of knowledge can be evaluated and managed? Should practitioners look for direct or indirect effects? Do practitioners need AI in place? The paper focuses on finding the importance of integrating AI systems in e-procurement within the hospitality industry worldwide.

In the industry, AI thoughts need to be elaborated to clinch inter-organizational matters concerning the frameworks and courses within organisational networks (Edghiem & Mouzughi, 2018). Big data analytics within the data science, in general, are overtaking many technological headlines and it is an important concept in which AI has a broad scope and cannot be overlooked, especially in the field of procurement in the hospitality industry as it is the major cost invested (Ittmann, 2015). Most available data have been formed in the past few years while the term "Big Data" has been around since 2005, introduced by O'Reilly Media after creating the term web 2.0. Nonetheless, the practice of data and knowledge has been around much longer (big-data-history). Although the adoption of a system is evidenced in the research, to the best of this author's knowledge, very few papers have been published connecting big data and AI in e-procurement within the hospitality industry. There exists a need to understand the research gap between AI, e-procurement, and hospitality.

2.2 E-procurement in the Hospitality Industry

Kothari, Hu, and Roehl (2018) conducted an exploratory study on adopting e-Procurement technology in a chain hotel and the results are shown herewith. The hospitality industry is becoming more technology-based in the guest-facing areas while there is significant scope for improvement within the back of the house. Thereby, it was identified to be very important to understand e-procurement practices and their related processes. In this study, the author attempted to find the challenges encountered by top management while adapting and implementing e-procurement. The major finding was that the company lacked standardisation in purchasing processes across various properties and, also, not all suppliers showed interest to be part of the centralised e-procurement initiative. Kothari, Hu, and Roehl's (2018) study revealed ambiguities in the company's audit system. A centralised purchasing system would be a solution to this ambiguity. This system will control the audit system from a corporate level and at the same time, simplify day to day communication between accounting and operations staff in different properties. Furthermore, "adopting an e-Procurement system would enable the hotel company to be more efficiently and accurately know how much they are spending corporate-wide in the various purchasing product area, allowing them to use the leverage of their buying power to reduce costs", as narrated by Kothari, Hu, and Roehl (2018). The hotel industry then can negotiate with the vendors for better prices and deals for bulk purchases.

The property under study being small had many variations in both what supplies procured and how it is purchased. Even the cost in various properties was different even if it was bought from the same vendor. Thus, it is obvious that there was a communication gap between the corporate office and operations at the hotel properties. Simultaneously, this was resulting in increased operational costs at the property level. The data transferred to the corporate office was inconsistent which made the office difficult to track the progress of the given property. The standardisation lacked a common unit of measurement and categorisation also.

Furthermore, in this study "these indicators argue for a standard centralized purchasing system that allows for a possibility of audit control at the corporate level and facilitation of regular communications between accounting and operations personnel at the corporate office and the various properties" (Kothari, Hu, and Roehl, 2018).

2.2.1 Benefits & Barriers of E-procurement

The research aim is to define how e-procurement enhancement at this stage for the hospitality industry can be done. Several journal articles have been analysed to perform further system evaluation and it was identified as necessary to assess current systems capabilities.

As per Katru Kauppi, hypothetical studies are significant which implies the importance of e-procurement factors in vigorous e-procurement technology reception. Peleg et al. (2002) and Yu, Yevu, and Nani (2020) discovered a few advantages of integration being vital to the organisation, help the client in the online hunt. Simultaneously, Attaran and Roche (2001) stated a few benefits as it enables paperless processes, integration of supply chain, and eases operational management. Yu, Yevu, and Nani (2020) in their study state the benefits of e-procurement as control of vendor relationship, ease and accurate order fulfillment, enhanced efficiency in the usefulness of the purchasing order. Further research done by Croom and Brandon-Jones in 2007 suggests few more benefits including cost reduction, better access to supply chain, minimising unnecessary purchases, diminishing request order process cost, and so on.

The perceived barriers to e-procurement integration found in several articles are explained herewith. Liao et al. (2003), Croom and Brandon-Jones (2007) states the major obstacles are incorrect floor prices, the items sent as replacements may not be of the same quality, the procedure to choose the contracts may be inappropriate, breach of information, unkind management, supply propagation and lack of top management backing. Toktaş-Palut and Ülengin (2015) state organisational dysfunctionalities in practice, cost/benefit concerns, incompatibility, and inadequate IT infrastructure and inadequate business process are major barriers in managing intelligence related to e-procurement. Whereas Anagappa and Eric (2016) mentioned other barriers like lack of knowledge and skills, resistance to change, the time needed for the change, and lack of system integration are identified.

The critical success factor in data management in terms of e-procurement would be a previous experience that includes social, tacit, and explicit knowledge. This will apply common sense knowledge with technical knowledge to make the e-procurement system a success factor. Some factors that affect the successful implementation of e-procurement include high investment cost of IT, infrastructure, and software, customers are hesitant to accept change as they are happy with the current system.

The advantages of the perceived organizational performance of procurement can be listed from various authors that seemed similar are listed here. The study conducted by Shukla. et al. (2015) observed the following changes: large procurement costs have decreased, can have customary strategies, less time for request handling, less administration cost, and stock expense. Whereas Brandon et al. (2018) found out the feasible measure, step-by-step transformation, promotion incentives, and government support are few another factor that enhances the performance on the adoption of e-procurement.

2.2.2 Quality of Data in E-procurement of Hospitality Industry

The usage of data and benefit depends on the quality of data collected (Pramdana, 2019). The weaknesses in data can end up in direct or indirect damages to decision-making in an organisation (Hazen et al., 2016). Nevertheless, SCM also foresees problems in having poor data being analysed, which leads to later bad impacts on the organization. Moreover, a recent survey depicts that the managers claim that poor data in the pool is one of the major obstacles to adapt business strategies by data analytics (Hazen et al., 2016). Therefore, SCM needs quality data control for fostering the results gained by data analytics.

Accuracy, timeliness, consistency, and completeness are the four dimensions that come under intrinsic data quality. Relevancy, value-added, quantity, believability, accessibility, and reputation of the data are dimensions being considered in contextual data quality. Contextual data quality refers to the information being derived from the data more than the data quality. Hence the intrinsic quality of data is more relevant when considering data analytics as a means to make strategic decisions (Hazen et al., 2016). The major things to consider before any data analytics are whether the data is free of errors, is it up to date, are all data presented in the same format, and are there any important missing data. These abovementioned measurements should be evaluated to get a good impact from the analysed data that will help in making strategic decisions for the top management. This is taken care of as the first task in data analytics.

The relevance of data and the significance must be made aware to all resources involved in data entry so that these data used by analysts can derive better effective decision-making tools (Hazen et al., 2014).

2.2.3 Data Analytics in the Hospitality Industry

The research that is conducted in big data business analytics (BDBA) is always linked with either logistic supply chain management (LSCM) or supply chain analytics (SCA). In this paper, analytics is discussed under five main classifications: descriptive, diagnostic, predictive, prescriptive, and cognitive analytics. Wang et al (2016), the three categories are descriptive, predictive, and prescriptive. Greasley (2019) suggests adding a diagnostic analysis into the analytics model. However, in a recent study by Król and Zdonek (2020) suggests adding cognitive analytics as to the fifth data analytics while explaining the analytics maturity model. A combined data analytics maturity model adapted from International Research Organisation (2018) and Król and Zdonek (2020) are given below in Figure 3.

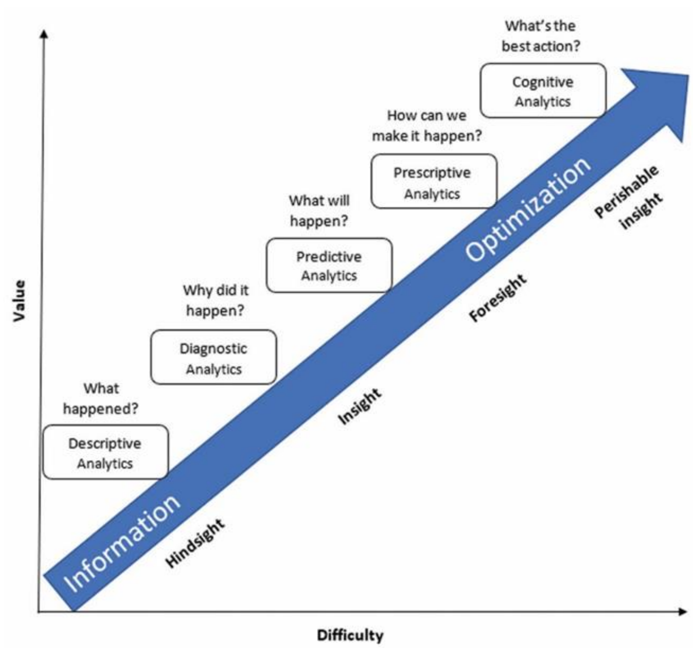


Fig. 3: Data Analytics Maturity Model adapted from Król & Zdonek, 2020

a. Descriptive analytics is conducted on existing data or processes to identify problems and opportunities. In most cases, descriptive analysis is undertaken in most organisations as part of report generation which itself is part of online analytical processing (OLAP). It explains "what happened"?
b. Diagnostic analytics is another traditional analytics in which decisions taken by assured delays. The delay is due to the necessity to gather and analyse data and then interpret them. Diagnostic analytics supports the finding of consistencies and measurable relations between variables via historical data analysis. It explains "why did it happen"?
c. Predictive analytics is mainly used to forecast and predict using carefully worked-out algorithms and programming to determine illustrative patterns inside the data. Various techniques and programs can be used to do this, which include web/data mining using the Python data analytics tool. It explains "what will happen in the future"?
d. Prescriptive analytics is used for high-level decision-making and finding alternatives to meet the strategic goals, which are described by high dimensions and density to enhance business performances. It explains "what action to be taken"?
e. Cognitive analytics are based on real-time analytics. Data is collected, organised, analysed, and interpreted mainly to find regularities and patterns. These models are kept in the data stream which affects the collaboration between guests and the organisation. That is the way of communication with the guests and the reception of a brand which involves real-time monitoring

> of the situation and a guest's behaviour patterns and finally selecting a behavioural pattern that is optimal. This is also named "perishable insight".

Prescriptive and predictive analytics plays a significant role in fostering the level of importance an organisation attributes to making effective decisions as per Demirkan and Delen (2013). In this paper, the researcher focuses on integrating BDBA and SCA to manage uncertainties in the organisation.

BDBA itself has two dimensions: big data (BD) and business analytics (BA).

> "BD refers to high volume, high-velocity, and high variety sets of dynamic data that exceed the processing capabilities of traditional data management approaches" (Russom, 2011; Chen and Zhang, 2014; Wang et al, 2016).

> "BA is the study of the skills, technologies, and practices used to evaluate organization-wide strategies and operations continued to obtain insights and guide business planning of an organization through evidence-based data, statistical and operations analysis, predictive modelling, forecasting, and optimization techniques" (Russom, 2020; Chen et al., 2018; Wang et al., 2016).

The difference in foreign exchange commodity fluctuations, which in turn affect commodity pricing, is a big challenge in the supply chain. The solution is to have real-time pricing to ensure the best commercial terms and to protect the supply chain. "With globalization, country and region-specific economic development impact global trade flows and the global economy more than ever" (Blackburn et al, 2015). Changes occur daily, which is challenging as the social, economic, and political worlds are closely interlinked and cover more multicultural environments than ever before (Blackburn et al, 2015).

The important dimensions of this current economic environment are revealed in the term VUCA: volatility, uncertainty, complexity, and ambiguity (Blackburn et al, 2015). Volatility refers to the rapidity and degree of change in information and surroundings. Uncertainty is the opposite of certainty. Complexity is linked with several aspects and their interactions in cause and effect systems concerning time. Ambiguity articulates the unclearness, lack of transparency, and opacity in the meaning and interpretation of events and ambient conditions. VUCA presents severe challenges for different forecasting methods (Blackburn et al, 2015). The author will try to devise appropriate data analytics which is an indispensable tool for SC decision-making.

Demand uncertainty is the overriding cause of uncertainty (Blackburn et al, 2015). Furthermore, the economy holds growing challenges for demand forecasting. The increased rate of product innovation leads to a shorter product life cycle and the volatility of customer preferences leads to a reduction in the use of historical data, which is the basis of forecasting data. One will need close interaction with multiple sources of real-time data to generate a beneficial forecasting process that will impact performance positively (Blackburn et al, 2015; Acar and Gardner, 2012). The study intends to explain demand uncertainty with a predictive methodology. The techniques that mainly reside within predictive analytics are i) classical statistics, ii) knowledge discovery from databases (KDD) (data mining), and iii) ML. A synthesis between these three techniques bears the immense potential for predictive data analytics (Perner, 2018). Perner's study investigated how supply chain analytics methodology based on forecasting with covariates can benefit from the inclusion of internal and external data into a statistical model based on the predictor variable. To predict the demand, the variables considered are public holidays, the relationship between customer and supplier, industrial value chain, industry-specific indicators, changes in the regulatory environment, future enhancement in the company and, leading indicators as suggested by R. Blackburn (2015).

2.2.4 Predictive Analytics in the Hospitality Industry

Various situations, such as optimisation of operations, revenue, cost, and competitiveness, win significant support from big data. Business opportunities and revenue can be forecast by prediction models. Previously, internal big data from past years are used for decision support and forecasting on pricing, rate rules, distribution channel management, and inventory optimisation. However, recently very few organisations have started to use a neural network to analyse the given input with the expected output to obtain a better multi-attribute decision or prediction as to the result. Contextual information can be used to calculate the best price from vendors to gain long-term profit for those parties involved (Bendoly, 2013). That is why the organisation needs to combine both internal big data and contextual data to generate an efficient result.

AI not only fosters human-computer interaction but also accelerates machine-to-machine communication and at the same time it combines multiple-source data automatically. For example, the system can tabulate not just revenue from the room but also income from different divisions too. Predictive analysis is used in various services including customer services, strategic planning, and forecasting. Voice recognition can be used to interact with a robot concierge, for example. AI technology integrated with forecasting algorithms can be used to identify the demands and requirements in SCM. The outcomes are better-publicising plans, financial management, and manpower adjustments (Claveria et al., 2015; Huang, 2014). Furthermore, a predictive analysis could be used to identify suspicious behaviour among employees (Collins et al. 2019), and apply information management for decision-making (Stalidis et al.,2015). The paper also focuses on how AI technology can be integrated with predictive analysis that could be used to identify the demands and requirements in e-procurement.

A. The initial framework for Predictive Analytics

To develop a clear roadmap to the findings, an International Research Organisation has suggested a few steps to follow and an initial framework is given in Figure 4 on getting started with perspective analytics. Moreover, another framework for developing a business intelligence system is also provided.

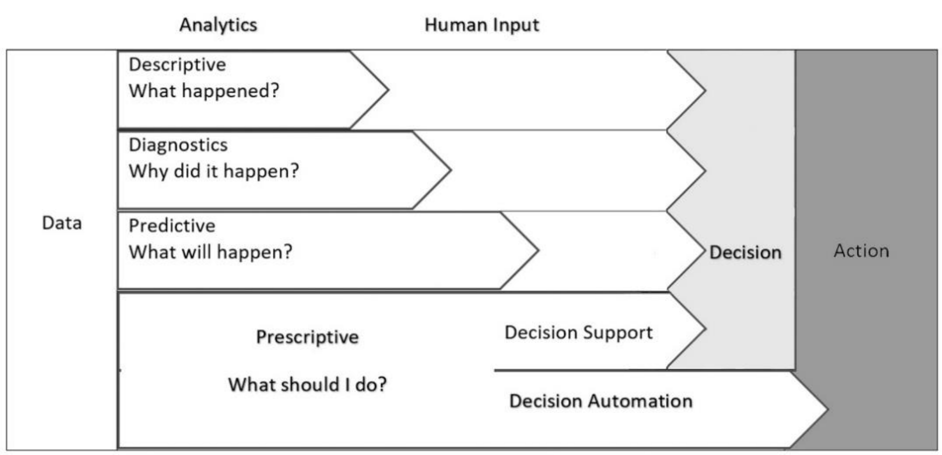


Fig. 4: Initial Framework for Predictive Analytics Adapted from International Research Organisation Symposium, 2019

Big data can undergo much analytics as shown in figure 4 (Elliot & Andrews, 2019). Descriptive analytics will give an idea of what has already happened in the organisation. Diagnostics analysis will provide a rationale for what has happened. Predictive analytics not only uses a lot of statistical tools but also a deep neural network algorithm. The prescriptive analysis provides a decision support system that gives suggestions and recommendations on a particular prediction. If the situation demands automation of the process it is also possible by decision automation systems. Most of the above analytics need human input for action except in automated decisions. Data science usually has multiple participants in various roles. The result is to know how to manage the data effectively as an employee with a tactical background. A business expert, who has a broad knowledge of data extraction, is important to have on the team (Pauleen & Wang, 2017).

B. The initial framework for Integrating AI

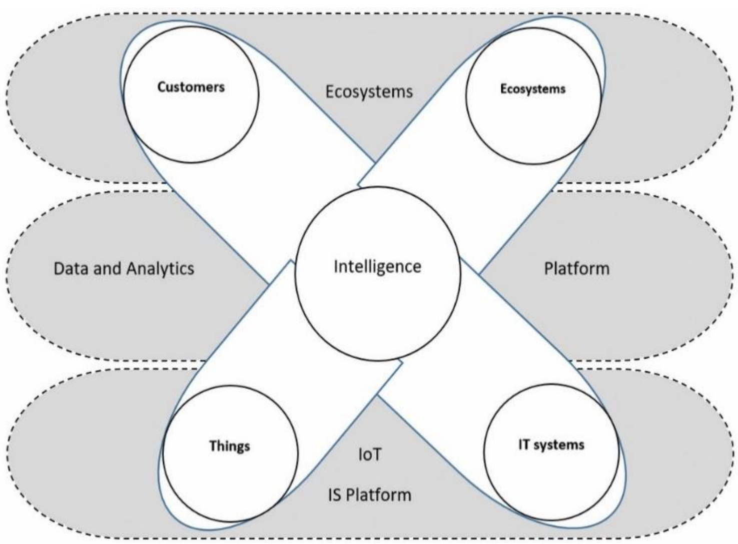


Fig. 5. Initial Framework for Integrating AI Adapted from International Research Organisation Symposium,2019

The framework shown in Figure 5 adapted from International Research Organisation Symposium (Dubai, 2019) with the title “Digital Platforms: The role data and analytics play in their success”. Intelligence is effectively data and analytics. There is a central role in the way an organisation explains how data and analytics evolve, starting data and analytics programme initiatives, or even making a data and analytics strategy successful. This involves connecting a lot of components.

Traceability, awareness, and consistency are very important in making a decision, so it is critical to have a skilled workforce in place to manage and analyse the data. The ecosystem clearly can be defined as the stakeholders of the organisation who will interact with the business intelligence systems and so is a critical component of the platform. The language in which we capture our data captured is very relevant to depict the particular organisation’s scenario. Demystifying one angle makes it relevant to have more at one time to increase the relevance for the given situation. Data Analytics is a great opportunity to often see the operating model in a format that is innovative and tailored to the nature of data that is specific to the organisation. Data analytics is the best way to connect all different components in the organisation, such as IT platforms, IoT, customers, and stakeholders. The decision should be taken as per the context of a given organisation. Decisions can again be used to generate further decisions, for example, tactical decisions can be used for operational decisions. For this kind of decision, real-time data or near real-time data should be used, for example, what the best next action is. Evolution of the idea of continuous next, where humans and computers work together, is important to optimise and resolve opportunity leveraging analytics by effectively applying augmented intelligence. From this given platform the research seeks to develop a framework that gives solutions for digital and analytical decisions effectively for the selected sector.

2.2.5 Deep Learning and Machine Learning Techniques in E-procurement of the Hospitality Industry

Several techniques exist as per Hadian et al. (2019) in AI as suggested solutions to problems in SCM, such as genetic algorithm (GA), neural network (NN), rough set theory (RST), and grey system theory (GST). These tools are accepted worldwide as search techniques to recognise the rough way out for multifaceted optimisation obstacles. GA is considered a heuristic technique that cannot provide the best solution all the time, but GA is included in all decision-making techniques. Several authors suggest the GA method for supplier selection in SCM. The technique in which there are input and output connections is used in NN where each link is assigned with a weight. These weights help in the predictive analysis to reach a final decision. This method is used mainly to predict supplier bid prices (Lee & Ouyang, 2017; J Chai et al.,2013).

GST is used to provide interval values, which is not a precise value but rather a point estimation. This technique is mainly used in supplier selection from a different perceptive, while RST is used to find operational connections within inaccurate or noisy data. There are several minor AI techniques too, which are used in SCM for various purposes, including a decision tree, case-based reasoning, association rule, and ant colony algorithm.

Cyberspace offers a limitless commercial setting and a robust contentious market. The hospitality industry is increasingly innovative in finding ways to distinguish and give importance to properties among an enormous number of opponents (Chai, Liu & Ngai 2013).

They can take advantage of cyber-physical interoperability and interconnectivity to gain access to their customers with their preferences and conduct market analysis to better inform decision-making processes (Buhalis & Leung,2018). While some hoteliers have adopted technology in ambiance and intelligence as the central point of their progress some others are still using the inherited technology. The 4th industrial revolution is disturbing the smooth running of the hospitality sector, especially the usage of the Internet of Things and the Internet for Everything. Currently, data procured is considered the most important asset in the hospitality industry. There are a lot of possibilities with the data gathered if it is processed efficiently and effectively.

"Big data collected from both internal and external services enable hospitality practitioners to make use of historical databases to forecast and predict business trends such as occupancy, rates and yield, labour costs and investment decisions" (Zhang et al., 2015; Buhalis & Leung, 2018). However, only limited research has been done on the data gathered during procurement, despite the numerous possibilities for this. Moreover, the available data does not follow any standardised format, so it is a challenge to retrieve and process it to make a reliable sense of this large data. The hospitality industry involves a large number of stakeholders in the form of employees, suppliers, managers, dealers, customers, guests, and so on. The data collected can be helpful to all of these stakeholders only if they can access and analyse it. The management relies on the historical and contextual data for prediction and forecasting of future trends in pricing to attract customers.

With state-of-the-art technology, it is possible to generate dynamic data, and thus to establish a decision support system that can maintain business operations to exploit the value for all participants and intelligence. Such a system will ignite a healthy competition between those who are involved and to attain the values of each stakeholder. The hospitality industry is considered a highly connected and interoperable industry to make use of all these collaborative systems. Marketing and distribution have improved drastically in recent years due to the smart systems available.

2.2.6 Optimisation in e-procurement

One of the most important factors of the accomplishment of a business is effective SCM. Managing revenue and inventory in a multi-national chain hotel structure is a tedious task as it is mostly multifaceted to predict the demand of the majority of the commodities (Ampazis, 2015). Globally, travel and tourism have evolved massively due to social, political, and technological advancement (Song, Qiu & Park, 2019). As a result, cost and benefits may rise too, due to an unusual demand for resources (Song, Qiu & Park, 2019). Hence accurate forecasts are vital for each stakeholder where they try to exploit the growths in the market demand and balance local ecological and supply chain capacities (Hemmington, 2007). The optimisation of the supply chains is vital for any organisation that involves buying/selling since these procedures may openly affect customer service, inventory and cost, and reaction to the ever-changing situations. Therefore, decision-making SCM should think through basic uncertain events while collaborating on the goals and objectives of the various processes involved. (Ampazis, 2015).

2.2.7 IoT in Hospitality

A surrounding monitoring system using the IoT technology would add value in carrying out automatic activities. “IoT functions into three different layers including smart systems (data acquisition); connectivity (data transmission); and analytics (actuate other IoT objects)" (Chuah et al., 2016). A smart hotel should be enabled with intelligence that can enhance interaction amongst everything existing in the environment, trade, and basic needs of humankind and thus provide values to its stakeholders (Nguyen & Simkin, 2017). Other features of an intelligent hotel would be to keep it environmentally friendly, proper usage of space, excellent daily operations, depends on natural energy, highly cautious on health and safety of everyone around it, stakeholders' outlooks are taken care of, adapt ad-hoc requests, and so on (Ghaffarianhoseini et al., 2017). Simultaneously, smart hotels also find historical data to be useful in identifying customers' preferences to make the ambiance more attractive, to go green, and to have effective waste management systems (Ghaffarianhoseini et al., 2017).

The suggestion is to have this kind of smart and intelligent system as part of the hospitality industry. Sensors should be used to examine the activities inside and outside the premises of a building. Sensors can be a good support to have a seamless flow of information within the network. Having the cyber-physical systems in place will provide human-computer interaction and interaction with machines, which are interoperable and inter-communicable by remotes (Hersent, Boswarthick & Elloumi, 2011; Samie, Bauer & Henkel, 2019). The humungous amount of sensory information, calculations, and processing are all essential for IoT. Since all of these are possible with IoT, human intervention to manage such an enormous amount of data will be decreased drastically (Alsaadi & Tubaishat, 2015). Ever since the World Wide Web started disseminating information in the form of reviews and stars from customer experience, travel and tourism have directly impacted on the hospitality sector

(Noone & McGuire, 2013). Therefore, the hospitality sector is becoming more and more competitive enabling the scrutinization of situations now and then and make strategic changes proactively to maximise the benefit out of these data attained (Buhalis & Foerste, 2015).

The four major macro-environmental factors are namely, political, economic, social, and technological (PEST), which are factors that leaders must proactively adapt to based on the strategic plans and key performance indicators. PEST has a direct or indirect influence in the hospitality industry. That is any change in PEST can affect the smooth running of the hotel and so hoteliers should act proactively by reviewing and revising the situation when it dynamically conflicts (Alvarez & Campo, 2014; Cheng et al., 2016). Economic instability also affects the decision of the customer to stay in a particular hotel (Tang et al., 2016) by choosing a high-quality ambiance. Although several factors influence the decision of customers in choosing a place to stay, technology advancement is frontier choice as they can do everything in a few clicks and even edit their choices. "Technology can enhance and co-create hotel guest's stay experience" (Neuhofer, Buhalis & Ladkin, 2015).

Nowadays, online reputation is a big factor that influences everyone's decision as social network plays an important role in decision-making when it comes to entertainment and accommodation. Hence, technological changes should be adopted quickly to keep the customer satisfied and to maintain loyalty that would influence the guest to revisit the property. The customer's stay history could be retrieved and can be used for analysis and prediction on preferences. IoT and sensors would collect a vast amount of internal and external data that influences the overall satisfaction of the customer's stay (Jin et al., 2014). With instant service recovery responses from a hotel's end, guests' services can be more personalised and customisable which would eventually leave an unforgettable experience that the guests would cherish for a longer period. Virtual assistance could be used to improve their overall stay. Sensors can be used to examine the expiry date of food and beverage items and help the chef with making purchases and consumptions (Fan et al., 2014). The RFID tags (radio frequency identification) par stock levels can be done and examined to send out purchase orders directly to suppliers so that there is no problem of out of stock items and hence, reducing the time taken for inventory and manpower required to do this task. Details of IoT usage are given in the table below. At least eight locations are recommended to keep IoT devices which include four inside the hotel and four outside the hotel premises. The recommended IoT location is given in table 1.

Table 1: Recommended IoT location adopted from Buhalis and Foerste, 2015

IoT Location	Type of IoT/sensor	Function
Inside Hotel		
Guest Room	Movement sensor	Energy management system adjusted in-room environment and ambiance according to guest presence and their location inside guest room
	Voice sensor	Voice activation controlling in-room devices such as curtain, lightings, room temperature etc
	Temperature sensor	Measure room temperature ensure guest can stay with confortable environment
	Door lock	Mobile app can act as keyless card for door lock system
	Wearable sensor	Monitor guest health situation during their workout and provide
Restaurant & Lobby	Location sensor	Identify registered members presence and send push welcome message or events invitation
	Promotion beacon	
Hotel facilities	Availability beacon	Delivery availability notifications to hotel guest
Warehouse	Inventory tag	Detect item profile and location; Examine expiry date and par-stock level
Outside Hotel		
Building	Temperature sensor	Measure external temperature and make adjustment on energy management system
	Light sensor	Detect the sunlight and adjust the blinds and brightness of the lighting system
Roadside	Traffic sensor	Detect parking space and traffic situation
Social Network	Content sensor	Monitor social netwrok and UGC sites content related to hotel and alert manager to feedback promptly; customer's stay history can be extracted from internal big data for management review
PEST Data	Contextual data extractor	Extract PEST contextual data around the world and store in hospitality big data

2.2.8 Agent-Based Technology (ABT) in the hospitality industry

ABT has become crucial in AI (Anthony Jnr 2019). ABT consists of computational entities that sense and act intelligently to finish their pre-defined tasks (Rudenko & Borisov, 2006; Bokolo, 2018). The characteristics of agents as per Anthony Jnr, which are described by Yang et al (2015) are as follows: autonomy, co-operability, reactivity, being pro-active, and social ability (Woodridge & Jennings, 1995; Anthony Jnr, 2019). The agents collect information from internal and external environments. For example, agents will get a sales alert that quantities of a particular product are low and send a notification to the seller. A single agent collaborates with multiple agents that react in real-time to customer procurement orders (Plinere & Borisov 2011). Anthony Jnr has shown that the 10 different agents required in a procurement process are as follows: user interface agent, searching agent, monitoring agent, query retrieval agent, negotiating agent, repository agent, evaluating agent, acquisition agent, presentation agent and information agent. The description of each of the agents is shown in table 2.

The supply chain has to collaborate with several business processes that have a common aim. For the same reason of having various components and processes, the progression is an intricate process, as per Anthony Jnr. The breadth and width of these processes to interoperate and interconnect makes its implementation complex. To reduce diversity there should be a better way to synchronise and standardise the core business processes (Blanc et al., 2019). Recently, "an application [ABT] of AI is considered as the fastest growing area of the next generation" (Alsetoohy & Ayoun, 2018). Moreover, their study reveals that there is a positive relation between ABT and a hotel's food procurement practices and performance.

Table 2: Integrated software agents and their description (Bokolo Anthony Jnr, 2018)

Software Agents	Description
User Interface Agent	This agent provides access to buyers by serving as a connection between buyers and the procurement system. Hence, buyers can search for product information through the interface agent. Additionally, a buyer can view the current product status through this agent.
Searching Agent	This agent is responsible for a search query executed by the buyer. Following this, the searching agent accepts the query from the buyer through the interface agent. This agent searches and collects related product information from the procurement product knowledgebase.
Monitoring Agent	The agent mainly sends a notification to the seller and procurement manager when the buyer makes an order for a particular product added by the seller. Subsequently, this agent also monitors the procurement transaction prominence and detects possible errors.
Query Retrieval Agent	This agent retrieves procurement product information, such as a product category seller’s location on Google maps, product images, and other product information for the buyer.
Negotiating Agent	This agent requests product price information from the seller or procurement manager to the buyer to negotiate product procurement. The negotiation or order status is confirmed by the buyer if he/she is satisfied with the product after the buyer makes the procurement order.
Repository Agent	This agent enables the sellers to add products into the procurement product knowledgebase. The repository agent also collects and stores product orders confirmed by the seller.
Evaluating Agent	This agent enables the procurement manager to approve products newly added by the seller. This agent validates the seller’s products added to the system to be procured by buyers. This agent also verifies sellers and then sends the seller’s product information to the procurement product knowledgebase.
Acquisition Agent	This agent enables the procurement manager to search, add, update, and delete procurement products in the procurement product knowledgebase. With the support of this agent, the procurement manager can update product information, such as price, category, and sellers’ data.
Presentation Agent	This agent simplifies procurement operations in the system by sending a notification to the procurement manager on every procurement product sale in the system. Furthermore, this agent generates procurement reports to the procurement manager in regards to buyer procurement orders.
Information Agent	This agent provides information to the procurement manager by retrieving procurement-related information, such as sales information and product delivery information from the buyer and seller.

2.2.9 Ecosystem in Hospitality

Figure 6. adapted from Buhalia and Leung (2018), reveals the major inclusion of the sub-ecosystem in orange. Hospitality runs on all resources including manpower, money, and materials. Suppliers are one of the largest groups of stakeholders within the hospitality industry, as hotels cannot be sustained without food and beverage suppliers, heating, ventilation, and air-conditioning suppliers, technology vendors, maintenance and service providers, and many others. Moreover, these supply chains might also be linked to sub-contractors such as butchers, farms, wineries, transportation companies, warehouses, and so on (Zhang et al., 2015, Buhalis & Leung, 2018). Even though these subcontractors are not directly linked with hotels, but rather involved only as providers for services or goods, they also contribute significantly to hotel experiences.

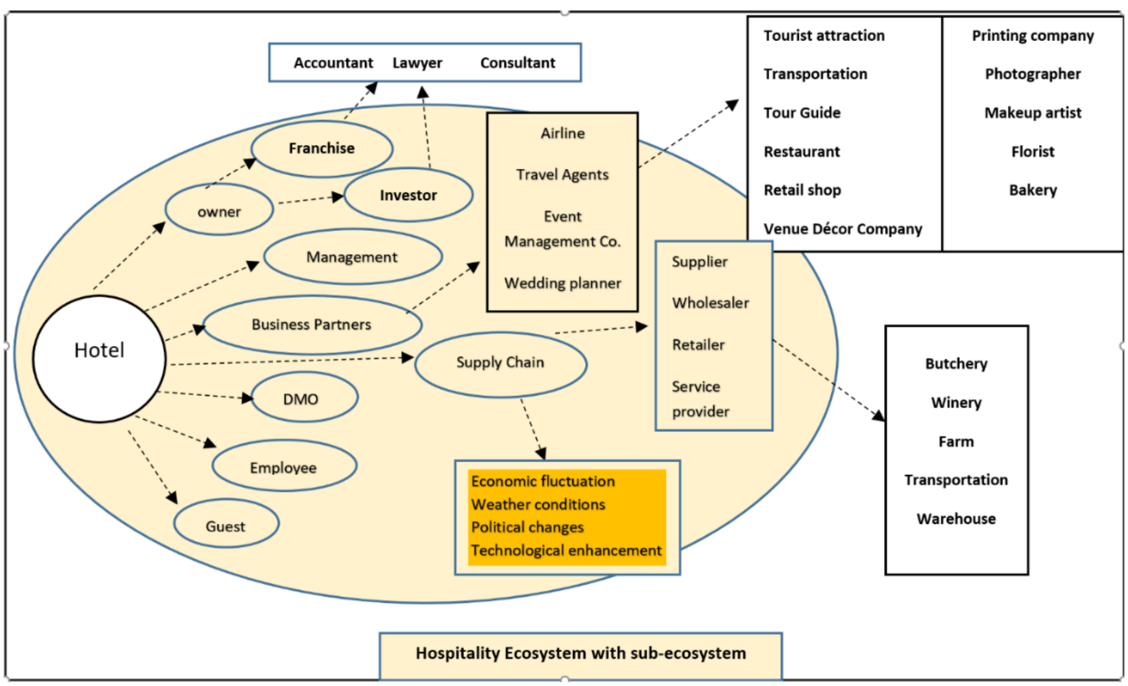


Fig. 6 Stakeholders of the hospitality sector adapted from Buhalis & Leung, 2018

Technology is a medium that dynamically improves supply chain efficiencies to enhance collaboration with these stakeholders to provide appropriate supplies within the budget and time limit. The smart technologies will enable them to be interoperable with all these providers and sub-providers to obtain information easily and efficiently to secure the best solution for their requirements. Through interconnectivity, the obstacles to working together are reduced by successfully supporting hotels to regularly assess the direct and indirect stakeholder's requests and approaches.

The suggested framework by Buhalis and Leung(2018) for a smart hospitality system is depicted in Figure 7. It contains three layers: the first layer is the network layer which inter-operates with several application systems of stakeholders to increase consistency and accuracy and to reduce redundancy. The second layer, being the cloud and data layer, helps in data aggregation from internal big data and contextual data and stores it for later usage. The third layer is the AI layer, which chooses the big data required for intelligence analysis and decision-making. Applications in hospitality may be adapted to situational demand and further decision support, as marketing or pricing decisions can then be broadcast by using beacons (low energy Bluetooth device). Situational demands can provide optimisation of hotel internal application systems and business intentions could be disseminated by beacons. Beacons can push information and location-based data points to pertinent stakeholders conferring to management decisions. Figure. 7. will be the base for developing a state of art novel framework for integrating AI in the e-procurement system of the hospitality industry.

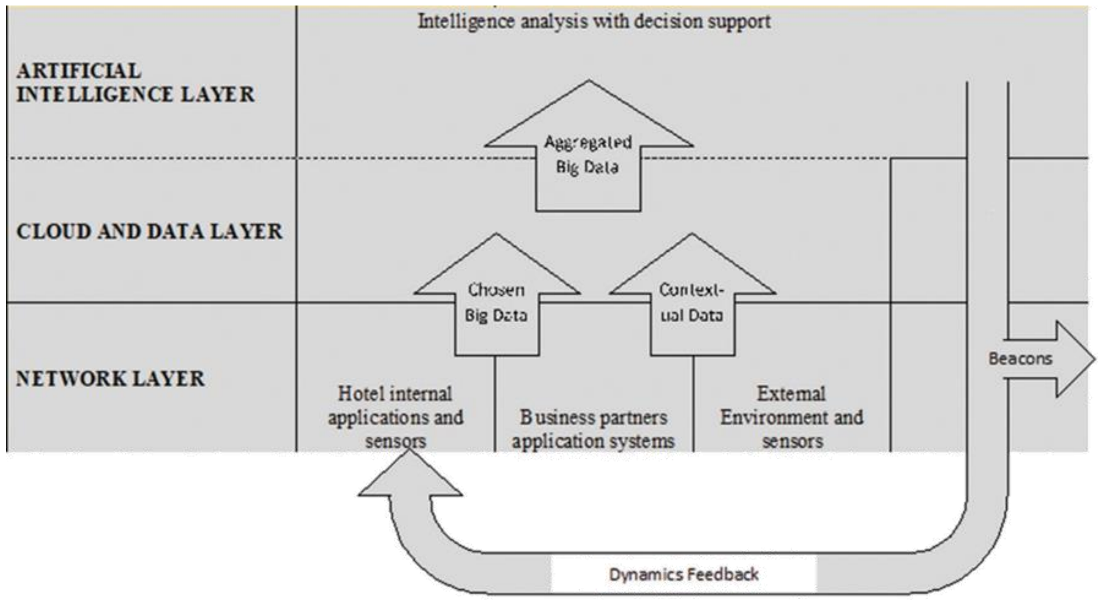


Fig. 7 Smart hospitality framework adapted from Buhalis and Leung, 2018

2.3 Gap in the literature

The framework by Buhalis and Leung does not cover the compliance with data analytics maturity model. Moreover, this framework is not deployable in any ecosystem as it does not speak of infrastructure compatibility or generalisability. The scope of the cloud layer and AI layer is not defined clearly. Overall, the framework needs some updates on state-of-the-art techniques. In this research, the researcher is leveling all these gaps. The data analytics, agent-based technology, and the entire framework proposed in this study are made in a way to revolutionize e-procurement in the hospitality industry. The existing framework is extended to have one more important layer for ABT with 14 agents for several jobs.

In my paper, a mixture of data analysis and data mining is undertaken to a great extent. Data analysis includes the analysis of data procured without regarding the size of data, for example, the effect of national holidays in occupancy rate; in contrast, data mining applies

ML and statistical models to reveal hidden patterns in VUCA data. To predict the demand using deep learning, the variables considered in my study are lags in spend or quantity for 12 iterations and month are used as other information like consumption, wastage of resources, events, etc. were not available which would have added more accuracy to the results. Purchase duration in days' calculation is another predictive analytics conducted in my study which uses recency, frequency, spend, occupancy rate, etc as feature variables. Apart from these analytics, my study also includes a database, data pre-processing, descriptive, and inferential statistics incorporating a representation of data in numerical data tables and visualization. LSTM (Long Short Term Memory), a version of RNN (Recurrent Neural network) is used for demand forecasting and train_test split with cross-validation is used as deep learning techniques in this research.

The initial framework for Predictive Analysis and AI integration by International Research Organisation is used as the backbone in my study. The methodology of proposing the extended AI integrated framework is based on these frameworks. The smart hospitality framework by Buhalis and Leung is the base framework for proposing the extended framework for integrating AI with ML. The ecosystem will include all stakeholders depicted in figure 6. The proposed extended framework will enhance interoperability, interconnectivity, and thus traceability, awareness, and consistency.

In my study, the barriers in the framework proposed by Buhalis and Leung are intra-organization, inter-organisation, technical and political issues such as lack of trust in the new system, exposure of information to competitors is mitigated by the introduction of IoT and ABT. Anonymity and confidentiality of stakeholders are upheld by introducing more agents apart from the ones suggested by Anthony Jnr. Although Buhalis and Leung talk about big data analytics, there is no clear roadmap on how to achieve the objective to make it convincing for stakeholders. In my study, several ML techniques are used to show the impact to optimise the e-procurement data, which is undoubtedly impressive for stakeholders, especially the study is unique in this area. There is no single proposal on having state-of-art technology combining data analytics maturity models, agents, or multi-agents and that integrate AI for complete efficient and smooth functionality. The data analytics, agent-based technology, and the entire framework proposed in my study are made in a way to reform e-procurement in the hospitality industry.

2.4 Chapter Summary

The literature review has revealed a framework to integrate AI in the hospitality industry which will be used in this study as the base to propose the novel conceptual framework for e-procurement in the hospitality industry. The fundamentals of descriptive, diagnostic, predictive, and prescriptive analytics are narrated. Moreover, an overview of predictive and forecasting analytics along with the advantages and steps involved in processing which is presented together with an insight into cognitive analytics. The major objective of the study is to conduct demand forecasting in several ways in the hospitality industry. The author has tried to highlight the advantages of having an integrated system, current trends in the industry, and applying ML in e-procurement within the hospitality industry along with gaps in the works of literature reviewed.

Chapter 3: A SYSTEMATIC REVIEW

3.1 Structured Review

At the time of scripting the paper (January 2020), 5,950 results appeared for the search with the keywords “AI in e-procurement” on the Google Scholar site. A further search adding the keyword “hospitality” gave 286 results. Further research for papers after 2015 resulted in 121 results. From the result, any article with at least two keywords combinations not in the body or if mentioned only once were removed. Forty-eight results had these keywords only in the reference page. Books, thesis papers, and non-English language papers were also removed, eight in total. Eleven articles did not allow access even with university credentials. Few papers were repeatedly listed as well, leaving 52 articles. “AI is a combination of technologies machine learning is one of the most prominent techniques utilized” (Reavie, 2020) with other technologies of data analysis and predictive analysis. Thus, from the results obtained, big data, ML, deep learning, and sub-areas of AI are included, and 52 articles were shortlisted. Figure 8 shows the stages of the selection process of the literature review.

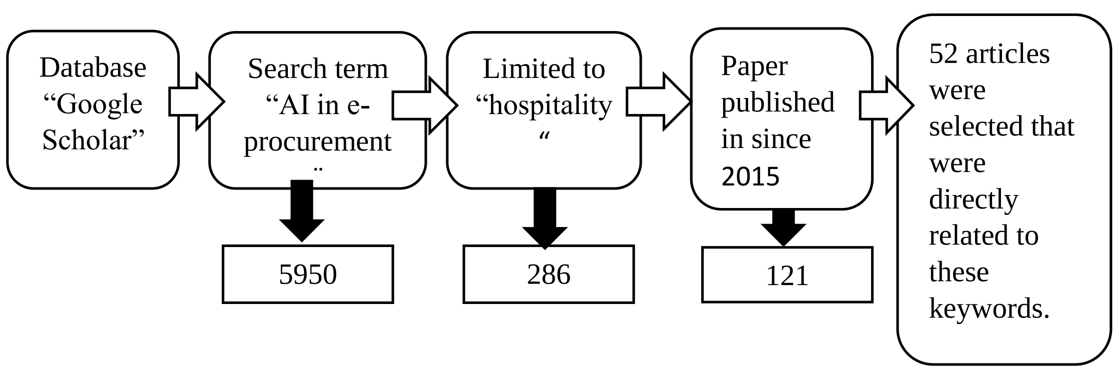


Fig. 8: Stages of the selection process of literature for review

AI is a technology or computer system used to operate in a way that reflects how the human brain works. The area of AI is an overarching field of computer science that includes extensive classifications comprising natural language processing, ML, deep learning, neural nets, content abstraction, decision-making, and more (Deer, 2020).

3.1.1 Statistics of the trend in publishing

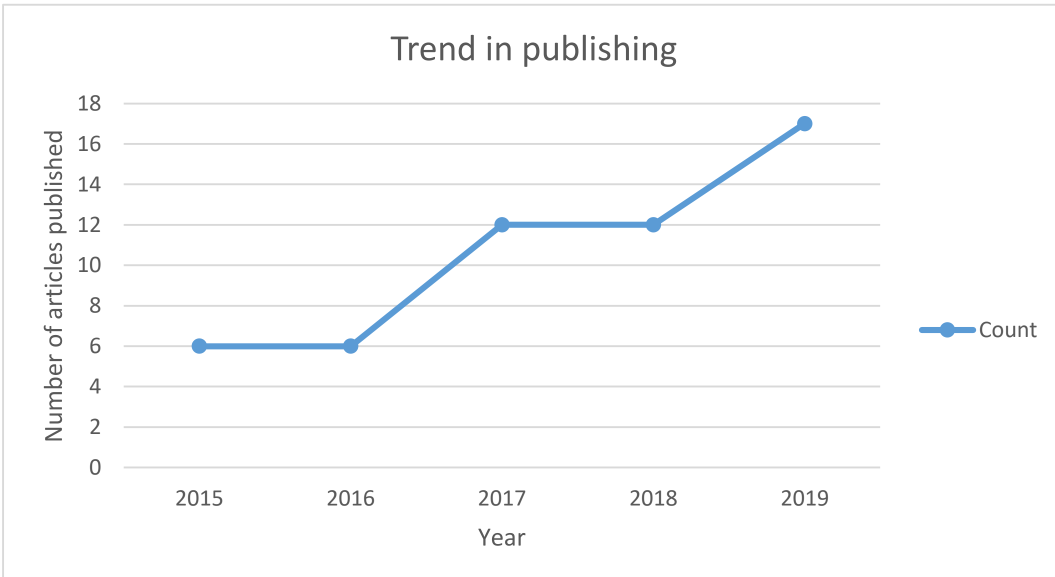


Chart 1: Statistics of literature selected

In this study, 52 articles were selected from the Google Scholar search results. The line graph above (Chart 1) shows the statistics of articles selected from the search which were relevant for the subject matter. It is interesting to notice that the number of articles published has increased by 20% in the last decade. This clearly shows more researchers find this topic relevant for studies and that it requires even more attention. This is a clear indication that in the coming years there will be more publications in the field of hospitality and e-procurement.

3.1.2 Major areas of research

A pie chart is displayed below (Chart 2) showing the major areas of research for the past five years in e-procurement and hospitality concerning AI.

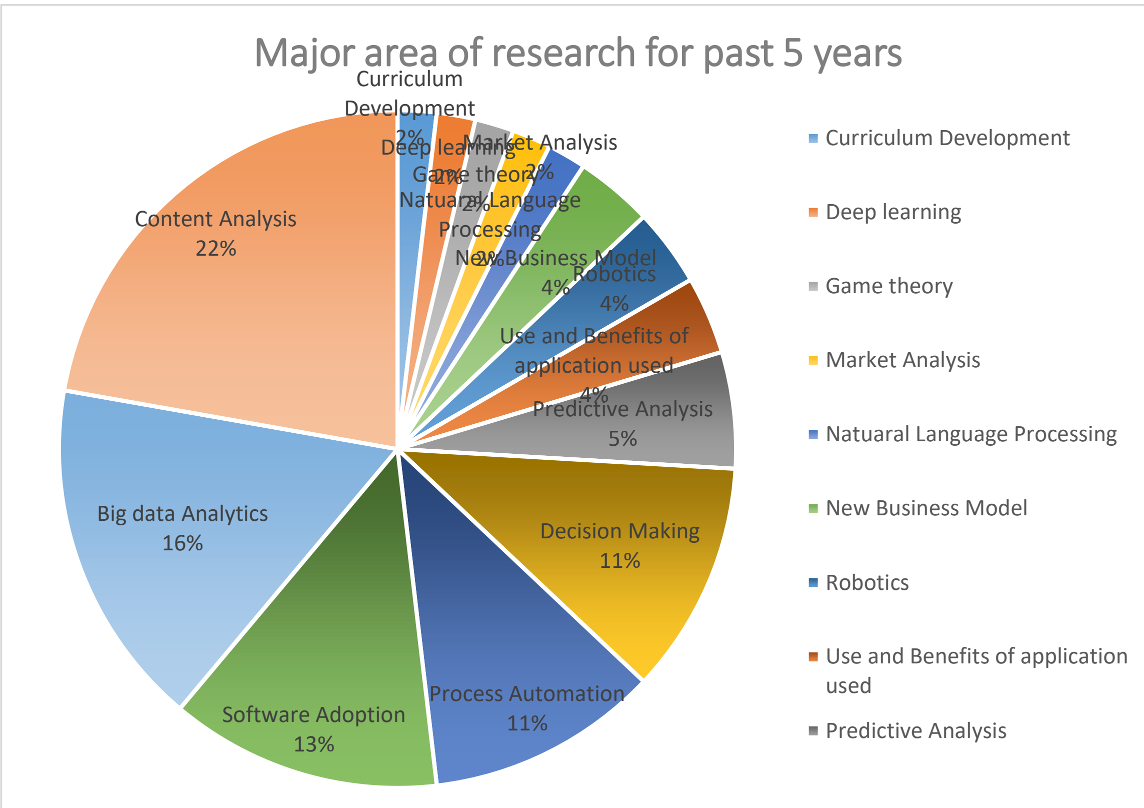


Chart 2: Major topics of research in the selected articles

The literature shortlisted contained almost 14 various topics out of which content analysis was the most highly researched at 22% overall. After that, the most common topic was big data analytics with 16% and software adoption with 13%. It is interesting to note that although various areas of AI are included in the research, the numbers are less compared to the ones listed above. There were papers on deep learning, ML, natural language processing process automation, predictive analysis, and robotics. The topic seems to be covered less compared to what occurred because the search was for e-procurement and hospitality in particular. Intense studies are on-going in other industries and departments.

3.1.3 Content analysis in the selected publications

The keywords from each paper are collated to display the most researched content in the past five years in the selected paper as displayed in Chart 3.

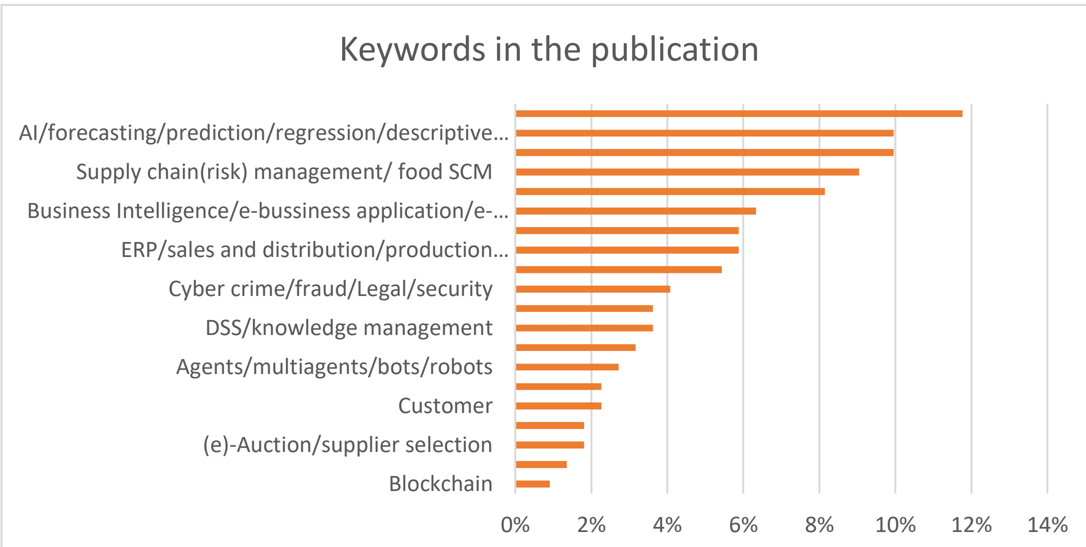


Chart 3: Keywords in the publication

In a content analysis of the papers produced more than 50 various keywords were evident. Then similar keywords were grouped to produce a cluster of topics. From there, the most repeated research work was observed on the topic related to AI. Moreover, keywords repeated at least three times are listed, while others were categorized as miscellaneous. Blockchain was the least used keyword, which does not belong to the miscellaneous group.

3.1.4 Comparing various studies in the field of AI /e-procurement/supply chain /hospitality

Here fifty-two different research papers are displayed in the table with the author name, the major focus of research, and the topic of research. The papers are varied and with little repetition. It was one of the limitations of this research to find appropriate papers. The search was broadened to travel and tourism as well. Papers that are not related to either supply chain or hospitality are not listed though they are used for reference. The comparison is displayed in Table 3.

Table 3: Insight into works of literature in the area of AI in procurement/supply chain /hospitality

#	Research Topic	Article Title	Authors, Year	Type of Paper (Conceptual /Emperical)	Source and Type of Data	Data Collection Method	Data Analysis Method	Data Reprting & Visualization	Major Results
1	Process Automation	A Developed Software Agent Knowledge-Assisted Procurement Management Tool For Retailing Enterprise A Feasibility Study	Anthony Jnr, 2019	Emperical	Structured	Self-Administered Surveysurvey From	Php And Mysql	Basic Tables Of Descriptive Statistics And Graphs, Uml, Use Case And Class Diagram.	Agent Based Model Support Supply Chain
2	Content Analysis	Systematic Literature Review On Electronic Reverse Auction: Issues And Research Discussion	Aital, Pawar & Behl, 2017	Emperical	Un-Structured	Literature Review 2005 To 2015	N/A	N/A	Benefits And Challenges Of E-Reverse Auction
3	Decision Making	Using Multiple Criteria Decision Making Approaches To Assess The Quality Of Web Sites.	Rekik Et Al., 2016	Emperical	Un-Structured	Article Published Between 2009-2015	Multiple Criteria Decision Making (Mcdm)	Detailed Tables	Quality Of Website Is Listed
4	Process Automation	Developing Design Principles For The Digitalisation Of Purchasing And Supply Management.	Srai & Lorentz, 2019	Conceptual	Un-Structured	Psm Literatures	Focusgroup-Type Full-Day Workshops	Graphs & Proposed Design Grids	Psm Digitalisation Grid With Application-Design Implications
5	Content Analysis	New Business Models In Supply Chains: A Bibliometric Study.	Delafenestre, 2019	Emperical	Un-Structured	292 Documents	Vos Viewer	Structured Maps	The Authors Present Insights And Deduce New Perspectives In The Potential Search For New Business Models.

6	Content Analysis	When Digital Government Matters For Tourism: A Stakeholder Analysis	Kalbaska Et Al., 2017	Conceptual	Un-Structured	Digital Government Stakeholder Analysis	N/A	N/A	Discussions On Digital Innovation To Advance The Tourism Sector
7	Content Analysis	E-Gov And Sustainability: A Literature Review	Pougel, Bonnel & Beier, 2020	Emperical	Structured	30 Articles Published Between 2012 Ans 2018	Manual	Statistical Analysis Reports	The Systematically Analysed And Presented Insights Into The Current State Of Research In The Context Of The Digitalisation-Governance & Sustainability.
8	Content Analysis	The Role Of Mobile Technology In Tourism: Patents, Articles, News, And Mobile Tour App Reviews	Kim & Kim, 2017	Conceptual	Un-Structured	Multiple Sources Are Used Such As Patents, Academic Articles, And News,	Netminer, T-Lab Program For Content Analysis And Leximancer Program	Clusters And Descriptive Tables	Suggest Future Research And Strategy Directions For Academia And Managers In Practice
9	Curriculum Development	Curriculum Development Of Scm Master'S Degree Program In Turkey	Özdemir, 2018	Emperical	Un-Structured	Curriculum Comparison	N/A	Statistical Analysis Reports	Highlight The Main Importance And Priority Of Scm In An Industry
10	Big Data Analytics	Big Data Analytics In E-Procurement Of A Chain Hotel	Mathew, 2019	Emperical	Structured	About 1 Million Data	Manual	Descriptive And Disgnostic Vsualization Reprot	Establishes There Is A Lot To Do With E-Procurement Data.
11	Use And Benefits Of Application Used	The Use And Benefits Of E-Technology Business Applications	Wokabi & Fatoki, 2019	Emperical	Structured	330 Businesses	Spss	Statistical Analysis Reports And Tables	Utilise E-Business Applications So That They Are Able To Create And Deliver Value Propositions In An Efficient And Effective Manner

12	Use And Benefits Of Application Used	Examining The Antecedents Of The Technology Acceptance Model Within E-Procurement	Brandon-Jones & Kauppi, 2018	Emperical	Structured	Survey Data Collected From 139 E-Procurement Users	Partial Least Squares	Summary Descriptive Tables	Results Confirm The Core Tam Relationships Within An E-Procurement Context
13	Content Analysis	Text Mining Applied To Literature On Sustainable Supply Chain (1996–2018)	Lis-Gutiérrez Et Al., 2020	Both	Structured	Literature On Sustainable Supply Chain From 1996 To 2018	Clustering	Descriptive, Analytical Visual Reports And Content Analysis	Provides Researchers In The Field With Elements And Analysis That Facilitate The Understanding Of Knowledge Production Dynamics In The Subject Matter.
14	Process Automation	Rethinking Supply Chains In The Age Of Digitalization	Hennelly Et Al., 2019	Conceptual	Un-Structured	11 Papers That Met Ppc Quality And Editorial Requirements.	N/A	N/A	Summaries All Paper Evaluated
15	Robotics	Intelligent Agent Technology: What Affects Its Adoption In Hotel Food Supply Chain Management?	Alsetoohy Et Al., 2019	Emperical	Un-Structured	Survey In A Hotel -88 Participants	Exploratory Factor Analysis, Simple Regression	Statistical Analysis Represented In Tables	Benefits And Challenges Of Iat
16	Deep Learning	How The Hotel Website Management Influence Hotel Supply Chain Management And Tourism Industry?	Roespinoedji Et Al., 2019	Emperical	Un-Structured	Survey With Hotel Employees About 300 Employees	Pls-Sem	Descriptive And Descriptive Tables	Establishes The Topic Of Interests And Work During The Period With Various Studies Conducted Across The World In Summarized

17	Robotics	Agent Based Fraud Detection And Reporting In E-Procurement	James K, 2018	Both	Un-Structured	Interviews,Observations,Prototype System,Questionnaire	Java And Oracle	Use Case Diagrams	The Study Demonstrated Abt Can Be Used To Detect And Stop Fraud/Corruption In Public Entities Thereby Deriving Maximum Value For Taxpayers Money
18	Game Theory	C-Negotiation Game: An Educational Game Model For Construction Procurement And Negotiation	Dzeng & Wang, 2017	Both	Structured	72 Participating Students	Php And Mysql	Game Model Archotecture	Developed A Web-Based Negotiation Game, Named C-Negotiation Game, For Enabling Students To Make Simulated Decisions In Construction Procurement And Negotiation Processes.
19	Process Automation	Ontology Based Multi Agent System For Improved Procurement Process: Application For The Handicraft Domain	Dhaouadi, Benmiled & Ghédira, 2014	Emperical	Structured	N/A	N/A	The Multi Agent System Interactions Diagrams	Designed And Developed A Multi Agent System (Mas) For The Supply Chain Automatization.
20	Predictive Analytics	Proactive Supply Chain Performance Management With Predictive Analytics	Stefanovic, 2	Conceptual	Un-Structured	Extensive Literature Review	Data Mining Predictive Analytics	Data Mining Tables And Forecasting Graphs	New Predictive Supply Chain Model

21	Big Data Analytics	Big Data And The Perceived Expectations Gap In Digital Authentication Processes	Calderon & Onita, 2020	Emperical	Un-Structured	15,463 Tweets	Ibm Watson Analytics	Sentiment Analysis	The Investigation Uncovered An Expectation Gap In The Perception Of The Efficiency And Effectiveness Of Different Authentication Methods.
22	Decision Making	Business Strategy Of Top Indian Company: L&T Infotech	Radhakrishnan, Aithal & L.M, 2018	Conceptual	Un-Structured	L & T Company	Manual	N/A	Analysed The Operational Strategy, Business Strategy, Financial Stability, Marketing Mix, Competitors, Training, And Recruitment Strategy Of The Company Briefly Using Theory A.
23	NLP	Characterizing Customer Experience Management In Business Markets	Witell Et Al., 2019	Conceptual	Un-Structured	Document Analysis	N/A	N/A	The Paper Draws Out The Theoretical Implications And Develops Managerial Implications For B2B Firms
24	Develop New Model	E-Tourism: Definition, Development And Conceptual Framework	Kazandzhieva & Santana, 2020	Conceptual	Un-Structured	Document Analysis	N/A	N/A	A Conceptual (Theoretical) Framework Of The E-Tourism System, Introducing Basic Groups (Subsystems)

25	Process Automation	Intelligent Agent Technology:The Relationships With Hotel Food Procurement Practices And Performance	Omar Alsetoohy & Baker Ayoun , 2017	Emperical	Structured	Survey Of Managers At Luxury Hotels In Florida	Hypothesis Testings	N/A	The Study Provided Academia With A Comprehensive Review Of The Prior Research On Iat Benefits In Food Supply Chain Management
26	Content Analysis	Shaping Industrial Relations In A Digitalizing Services Industry - Challenges And Opportunities For Social Partners	Holtgrewe Et Al., 2020	Conceptual	Un-Structured	Literature Review And Interview With Experts	N/A	N/A	Recommendations On Digitalisation Services To Belgium Strategic Decision Makers
27	Content Analysis	Supply Chain Management: Implementation Issues And Research Opportunities In Tourism Industry	Sutono, 2019	Emperical	Un-Structured	Survey With Hotel Employees In Indonesia	Statistical Study	Descriptive And Statistical Data Tables	The Most Important Impact Factors Are Supply Chain Marketing Planning Capabilities And Better Supply Chain Marketing Implementation.
28	Process Automation	A Utility-Driven Approach To Supplier Evaluation And Selection: Empirical Validation Of An Integrated Solution Framework	Ulutas Et Al., 2015	Emperical	Structured	Supplier Evaluation Selection Related Articles	Fuzzy Functional Assessments	Tables Of Validations	The Results Generated Using The Proposed Framework Is Compared With The Actual Historical Data Collected From The Company
29	Software Evaluation	Enterprise Resource Planning (Erp) Systems In The Egyptian Higher Education Institutions: Benefits, Challenges And Higher Education Institutions: Benefits, Challenges And Issues	Soliman & Karia, 2016	Conceptual	Un-Structured	Document Analysis	N/A	N/A	The Research Contributes To The Advance Of Concepts And Dimensions For Erp System From Heis' Standpoint

30	Market Analysis	Essence Of Digital Transformation—Manifestations At Large Financial Institutions From North America	Pramanik, Kirtania & Pani, 2019	Conceptual	Un-Structured	Annual Reports Four Large Banks In North America	Manual	Detailed Qualitative Analysis	An Emergent Structure For A Digital Transformation Maturity Model (Dtmm)
31	Software Evaluation	Forecasting Patronage Factors Of Islamic Credit Card As A New E-Commerce Banking Service	Jamshidi & Hussin, 2016	Conceptual	Structured	327 Bank Customers	Statistical Analysis	Tables Representing The Descriptive Analysis	The First Study That Proves The Applicability Of Tam For Explaining Adoption And Usage Of The Islamic Credit Card.
32	Software Evaluation	Legal Technologies In Action: The Future Of The Legal Market In Light Of Disruptive Innovations	Hongdao Et Al. 2019	Conceptual	Un-Structured	Qualitative Content Analysis	N/A	Charts And Tables	This Study Adds Theoretical And Practical Implications To The Research Discussing The Advent Of Legal Technologies.
33	New Business Model	New Business Models For Cultural And Creative Institutions	Nikiel, 2019	Conceptual	Un-Structured	Survey Technique, Using A Categorized And Standardized Questionnaire.	N/A	N/A	New Business Models For Cultural Institutions
34	Software Adoption	Nuance Of Government Procurement Ethics In India	S. Padhi, 2020	Conceptual	Un-Structured	Questionnaire Surveys	N/A	N/A	Models For Procurement Auction Of Construction Projects

35	Content Analysis	Study Of Various General-Purpose Technologies And Their Comparison Towards Developing Sustainable Society	Aithal & Aithal, 2020	Conceptual	Un-Structured	Identified, Analysed, And Compared Information Communication And Computation Technology (Icct), And Nanotechnolog y (Nt) In Various Industries	N/A	Summary Descriptive And Narrative Tables	The Study Led To Development Of The Concept Of 'Universal Technology' Model.
36	Software Adoption	What Drives Ict Adoption By Smes? Evidence From A Large-Scale Survey In Greece	Giotopoulos Et Al., 2017	Emperical	Structured	Largescale Survey On 3,500 Greek Smes	Statistical Analysis	Detail Tables	Innovation And R&D Activities And Collaborations, Well-Educated And Skilled Workers, Decentralized Decision-Making And Visionary Leadership Increase The Likelihood Of Adopting New Technologies In Smes.
37	Comparativ e Analysis	Co-Evolutionary Prospects In Tourism	Valeri & Fadlon, 2018	Conceptual	Un-Structured	Document Analysis	N/A	N/A	The Co-Evolutionary Process Implies The Identification Of A Governance Body Able To Exploit And Enhance The Systemic Resources Made Available By The Territory And To Inspire The Management Approach Of The Different Tourist Enterprises.

38	Big Data Analytics	Consumers' Attitudes Towards The Introduction Of Robots In Accommodation Establishments	Ivanov, Webster & Seyyedi, 2020	Emperical	Structured	Survey Of 393 Iranian Consumers	In-Depth Statistical Analysis	Analysis Results Displayed In Tables	No Demographic Variables Explored Seem To Play A Role In Shaping Attitudes Towards Service In Hotels By Robots.
39	Technology Adoption	Development And Validation Of A Formative Scale Of Technological Advancement In Hotels From The Guest Perspective	Ruiz-Molina Et Al., 2018	Emperical	Structured	197 Spanish Hotel Guests	Survery Analysis	Least Squaredregressio n	The Study Validates That Hotel Guests Mainly Associate With Highly Technified Establishments.
40	Content Analysis	Procurement Oversight Agencies And Growth Of Mses In Transnzoia County, In Kenya	Kissinger Amayi, 2020	Emperical	Structured	155 Respondents	Self Administered Quentionnare	Descriptive And Inferential Statistical Results	The Study Concluded That There Exists Strong Relationship Between Procurement Oversight Agencies And The Growth Of Mse In Trans Nzoia County , In Kenya
41	Big Data Analytics	Quantifying Potential Tourist Behavior In Choice Of Destination Using Google Trends	Padhi & Pati, 2017	Emperical	Un-Structured	63 Relevant And Semantically Related Keywords On “Kerala Tourism” Using Google Trends Data	Four Time-Series Constructs	Trend Analysis	The Analyses Are Expected To Guide Policy Makers In Understanding And Making Appropriate Decisions To Deploy Resources At Potential Tourist Destination Sites To Enhance The Potential Experience Of The Tourist.

42	Data Envelopment Analysis	An Integrated Data Envelopment Analysis Mathematical Programming Approach To Strategic Biodiesel Supply Chain Network Design Problem	Babazadeh Et Al., 2017	Both	Un-Structured	Document Analysis	Mathematical Programming Techniques And Data Envelopment Analysis	Statistical Graphs	Proposed A Method In Assisting The Policymakers To Take Suitable Strategic And Tactical Level Decisions Related To Bidiesel Supplu Chain Planning
43	Big Data Analytics And Decsion Making	Big Data And Supply Chain Decisions: The Impact Of Volume, Variety And Velocity Properties On The Bullwhip Effect	Hofmann, 2015	Conceptual	Un-Structured	Literatures On Big Data And Bullwhip Effect On Current And Past Released Papers	Research Frameworks And Models	Schematic Representation	Operationalize Big Data In The Control Engineering Analysis
44	Big Data And Predictive Analysis	Big Data And Predictive Analytics Applications In Supply Chain Management	Gunasekaran Et Al., 2016	Conceptual	Un-Structured	A Rigorous Review Process	N/A	N/A	N/A
45	Big Data Analytics	Customer Profitability Forecasting Using Big Data Analytics: A Case Study Of The Insurance Industry	Fang, Jiang & Song, 2016	Emperical	Structured	The Tenures Of The Health Insurance In 2007 In Taiwan	Dataanalytics Random Forest Regression	Vrious Regression Models	This Paper Proposed A New Customer Profitability Method For The Insurance Industry By Adding Liability Reserve.
46	Big Data Analytics	An Analytic Infrastructure For Harvesting Big Data To Enhance Supply Chain Performance	Zhan & Tan, 2020	Emperical	Un-Structured	Databases, Social Media, Mobile And Sensor Data	Data Analytics	Tables And Descriptive Resutls	The Study Propose An Integrated Infrastructure For Breaking Down The Information Silos, In Order To Enhance Supply Chain Performance.

47	Software Adoption	Barriers To Big Data Analytics In Manufacturing Supply Chains: A Case Study From Bangladesh	Moktadir Et Al., 2019	Emperical	Un-Structured	Data Analytics	Delphi Based Analytic Hierarchy Process	Sensitive Analysis , Tables And Charts.	Barriers And Potential Benefits Of Using Bda And To Make Policy Regarding Bda Adoption In Manufacturing Supply Chains.
48	Decision Making	A Data Mining-Based Framework For Supply Chain Risk Management	Er-Kara, Oktay Fırat & Ghadge, 2019	Emperical	Un-Structured	Datamining Based Scrm Model (Supply Chain Risk Management)	Flowchart And Models	Data Mining Models	Showcases How Datamining Supports In Making Structured Intelligent Management Decisions.
49	Process Automation	Heuristic Modeling For Sustainable Procurement And Logistics In A Supply Chain Using Big Data	Kaur & Singh, 2018	Emperical	Structured	42 Randomly Generated Data Instances Possessing Essential Characteristics Of Big Data.	Minlp (Mixed Integer Non Linear Program) And Milp (Mixed Integer Lin-Ear Program)	Flowachart And Models As Solutions	This Paper Proposes An Environmentally Sustainable Procurement And Logistics Model For A Supply Chain.
50	Big Data Analytics	Big Data Analytics And Application For Logistics And Supply Chain Management	Govindan Et Al., 2018	Conceptual	Un-Structured	Overview Of Highly Cited Paper	Various Descriptive Analysis	Charts And Tables	Summarises The Discussions On The Big Data Attributes, On Effective Practices For Implementation, And On Evaluation And Implementation Methods.

51	Decision Making	Investment Decision-Making And Coordination Of A Three-Stage Supply Chain Considering Data Company In The Big Data Era	Liu & Yi, 2018	Conceptual	Un-Structured	Three Stage Supply Chain Stakeholder	Various Model Analysis	Statistical Equations	Four Benefit Models About Big Data Information And Centralized & Decentralized Supply Chains
52	Content Analysis	Doctoral Dissertations In Logistics And Supply Chain Management: A Review Of Nordic Contributions From 2009 To 2014	Rajkumar Et Al., 2016	Emperical	Un-Structured	150 Dissertation Papers From 2009 To 2014	Manual	N/A	A Score For Measuring The Significance Of Article-Based Dissertations Is Proposed

The listed articles suggest having flexibility and decentralised decision-making as one of the success factors in the adoption of any technology (Er-Kara, Oktay Fırat & Ghadge, 2019; Ulutas et al., 2015; Moktadir et al., 2019). In this study, the proposed framework has the flexibility in decision-making concerning purposes and conditions used to adapt it to suit various contexts without considerably reducing its efficacy. The studies by Ruiz-Molina et al. (2018) and Ivanov, Webster and Seyyedi, (2020) suggest that adoption of agents does not depend on any demographic details but purely on people’s attitudes and guests usually associate with highly technically sound establishments. This study supports my framework as it is involved in the inclusion of the latest state-of-the-art technology and making stakeholders' life easier is in the plan. The listed authors have used structured or unstructured and both empirical or conceptual models in developing their studies and making a conclusion. This clearly shows that there is no preferred method without strictly inclining to one. But as per Creswell (2008), for a competitive result, it is always preferred to have a mixed approach when the problem and solution are defined quantitatively and qualitatively; this approach is followed in my research. It is important to select a set of strategies that fit the research type. In this study, most of the questions are on *how what,* and *why* so it is better to use case studies, experiments of histories (Creswell & Creswell 2008).

As per Stefanovic (2014), a business intelligence system that is capable of tracing, evaluating, modeling forecasting and delivering information that helps in decision-making must be maintained in an organisation supporting the supply chain to achieve the goal in stipulating accurate information to the stakeholders at the exact time. Mainly, the indicators are backward-looking without any prediction or forecasting. Stefanovic (2014) states that the efficiency of the supply chain can be promoted by having proactive decision-making through powerful data mining algorithms that compare, analyse, and forecast accurate predictive insights. Few of the researchers have contributed to the research world on the supply chain performance framework. While Lambert and Pohen introduced a framework for progressing metric, Boone et al.'s (2016), framework analyses supply chain performance accomplishments. Gunasekaran et al.'s (2016) framework was developed by considering the four major processes including planning, sourcing, making, and delivering. These frameworks were introduced for enabling managers to make better decisions, with an extensive outlook of both tangible and intangible assets. As per Stefanovic (2014), regularisation of supply chain performance is one of the essential criteria for real supply chain performance measurement. Data analysts estimate that from the large data aggregated in supply chain transactional software, 80% of data is irrelevant for decision-making

(Altin et al. 2018). Moreover, further analyses have to be done on the remaining 20% to make it useful. Stefanovic's, inventory management is an essential process in the supply chain. Incorporating predictive analysis in inventory management has advantages like cost reduction, higher customer satisfaction, optimal reordering policy, enhanced productivity, and, ultimately, increases cost-effectiveness. The study also states that "there have not been many research projects related to predictive performance management". The suggestions of Stefanovic and Gunasekharan et al. had a great influence on my research and tried my best to include each point in the proposed framework to be proactive in decision-making.

3.2 Machine Learning and Deep Learning used for Prediction and Forecasting

Data analytics is crucial especially in this era as in all fields are getting smarter day by day by the introduction of intelligent systems. The usability of big data is enormous and finding ways to make it informative is also very important. For making this big data consistent by data mining, slicing and cleansing are critical in the decision support system. Making these data available to optimise the operations, revenue, cost, and competitiveness is a major task for all organisations. A limited amount of this was undertaken in a traditional revenue management system which usually uses the past five years' data to forecast recommendations on pricing, rate, rules, distribution channel management, and inventory optimization (Guillet & Mohammed, 2015) but by using traditional methods. It is recommended that artificial neural networks, analytic network processes, and fuzzy goal programmes are used to mine and improve a multi-attribute decision-making prototype to meet the aims and objectives of the organisation. To enhance the decision-making and revenue management of any organisation it would be better to collaborate internal big data and external contextual data. This big data analytic reports would then be made into a smart dashboard for the users to make their reporting life easier.

ML specific search did not give much on the relevant topic. A sample of fifteen papers that were found in the Google Scholar search is included herewith in Table 4 with details showing the techniques used in each study.

Table 4: A synopsis of machine learning used publications

	ML algorithms used in prediction - a synopsis view				
	Area of research	Researchers	Machine Learning Models Used	Results	Software's Used
1	Stock Market Trend Prediction	Pahwa & Agarwal, 2019	1. Support Vector Machine 2. Random Forest 3. K-Nearest(KNN) 4. Naïve Bayes 5. Softmax	The random Forest algorithm performs the best for large datasets and Naïve Bayesian Classifier is the best for small datasets.	Not specified
2	Breast Cancer Prediction	Bharat, N & Reddy, 2018	1. Decision Tree 2. Algorithm (C4.5) 3. Seleukos (Clustering Algorithm) 4. Naïve Bayes 5. SVM 6. KNN	The Seleukos algorithm accelerated the learning curve for breast cancer and 4 percentage points better than C4.5 in the early stage.	Weka
3	Disease Prediction --Heart Disease --Breast Cancer --Diabetes	Kohli & Arora, 2019	1. Logistic Regression 2. Decision Trees 3. Random Forest 4. Support Vector Machines(SVM) 5. Adaptive Boosting	Logistic regression is good for Heart disease prediction. SVM is good for Diabetes prediction. AdaBoost for Breast cancer prediction	Python
4	Car Popularity Prediction	Mamgain, Kumar &	1. Logistic Regression 2. KNN	SVM was found to be more accurate	Python

		Manjari Nayak, 2018	3. Random Forest 4. SVM		
5	Class Result Prediction	Pushpa et al., 2018	1. SVM 2. Naïve Bayes 3. Random Forest 4. Gradient Boosting	Random Forest Classifier gives the most accurate result	Python
6	Urban Traffic Prediction	Lee & Min, 2017	1. Random Forest 2. Gradient Boosting 3. KNN 4. Multi-Layer Perception	Gradient Boosting Regression	Not specified
7	Prostate Cancer Survivability Prediction	Wen et al., 2017	1. KNN 2. Decision Tree 3. SVM 4. Neural Network	Neural Network has the best accuracy	Not specified
8	Black Friday Sales	Wu, Patil & Gunaseelan,2018	1.Linear Regression 2. MLK Classifier 3. Deep Learning Model using Keras 4. Decision Tre	deep learning with a neural network is far better than other methods	Python
9	Coconut Sugar Quality Assessment and prediction	Alonzo et al., 2018	1. Artificial Neural Network 2. Stochastic gradient descent 3. k-nearest neighbours 4. Support vector machine 5. decision tree 6. random Forest	SGD is superior in terms of accuracy but falls short to KNN and SVC in terms of running time.	Python
10	Travel Time Prediction	Goudarzi, 2019	1. Nearest Neighbourhood 2. Windowed Nearest Neighbour 3. Linear Regression 4. Artificial Neural Network	In the comparison, the best results were obtained by applying a shallow neural network with three hidden layers.	Google Maps API and C++

11	Predicting traffic and user mobility in telecommunication networks	Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z. & Zhang, H. ,2019	1. SVR 2. ARIMA 3. FFNN 4. LSTM	Deep Learning with Long Short-Term Memory for Time Series Prediction outperform other methods.	Not specified
12	Trajectory Prediction	Yeom, H., Kim, J. & Chung, C. 2020	1. Random Forest 2. Gradient Boosting 3. KNN 4. MLP 5. LSTM	LSTM Improves Accuracy of Reaching Trajectory Prediction From Magnetoencephalography Signals	Python
13	Forecasting Economics and Financial Time Series	Siami-Namini, S. & Namin, A. ,2018	1. ARIMA 2. LSTM	LSTM proves to predict better than ARIMA in time series forecasting	Python
14	Electricity price forecast	Lihong, D. & Qian, X. ,2020	1. Dickey-Fuller Test 2. LSTM	LSTM proved to be quite accurate at predicting fluctuations in electricity consumption.	Not specified
15	Multiple Seasonal Patterns	Bandara, K., Bergmeir, C. & Hewamalage, H., 2020	1. MSTL 2. Auto STR 3. TBATS 4. Prophet	LSTM-MSNet computationally more efficient than many univariate forecasting methods.	Python

The studies sampled show it is difficult to judge which techniques are more useful than others. After applying several ML techniques, choose the algorithm with the best accuracy. In other words, some ML is better than others based on the available data. Nevertheless, it is interesting to note that all sequential time series-based research suggests LSTM as one of the best techniques. The descriptive and diagnostic analysis is used for representing visual reports in my study. It is the primary source of information for the management team. ML and deep learning techniques are applied for predictive and prescriptive analytics. ML techniques include descriptive, diagnostic, and advanced predictive data analytics. Prescriptive analytics uses simulation and ML to suggest actions to be taken to achieve the desired results. Collecting data would be the first step in all data analyses. Then organise and analyse the data collected. The concept involved in this process is narrated below.

3.2.1 A conceptual framework for applying machine learning and deep learning

The descriptive and diagnostic analysis is used for representing visual reports. It is the primary source of information for the management team. ML and deep learning techniques are applied for predictive and prescriptive analytics. ML techniques include descriptive, diagnostic, and advanced predictive data analytics. Prescriptive analytics uses simulation and ML to suggest actions to be taken to achieve the desired results. Figure 9 shows the steps involved in applying ML and deep learning by several authors mentioned in the table above. Collecting data would be the first step in all data analysis. Then organise and analyse the data collected. The concept involved in this process is narrated below.

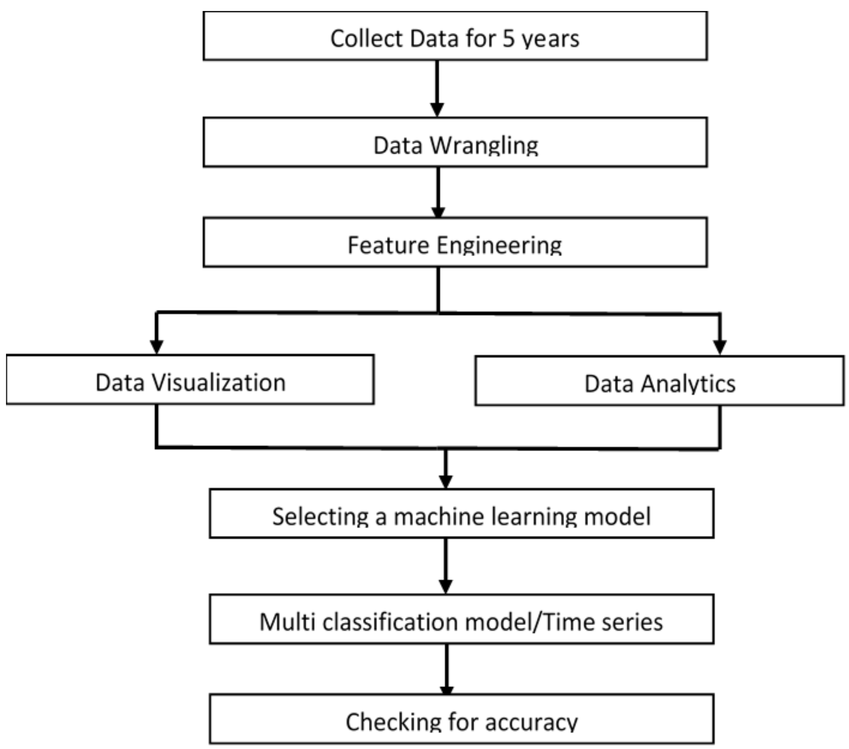


Fig. 9: Conceptual processes involved in forecasting problems

3.2.2 Classification algorithm selection

The literature review covered in this section on concepts for ML techniques used two books namely "Python for Machine Learning" by Raschka, Julian and Hearty (2016), and Python data science cookbook by Subramanian (2015). The section includes the relevant modules used in this research (only).

No one can say a specific classifier is the best, as it depends on many conditions. This might vary in the features or samples, the noise in the dataset, and so on. Moreover, it is important to see if the data is linearly separable or not. Therefore, one has to execute trial and error with various classifiers and choose the best that suits the given situation. In the end, the precision and accuracy depend on the data available.

The five main steps that are involved in training an ML algorithm can be summarised as follows:

1. Selection of features.

2. Choosing a performance metric.

3. Choosing a classifier and optimisation algorithm.

4. Evaluating the performance of the model.

5. Tuning the algorithm.

The *scikit-learn* API helps in the implementation of various classification algorithms and has got many convenient functions to pre-process data and to fine-tune and evaluate any models. The *train_test_split* function from scikit-learn's cross_validation module is used in the research with 20% test data. Furthermore, applying feature scaling using *StandardScaler* class from scikit-learn's pre-processing module is important to optimise performance. Instead of misclassification error, classification accuracy is found in this study. The dataset used is not continuous data and so not linearly separable. Thus, linear regression and logistic regression performs well on linearly separable classes. The most commonly used machine learning models and described below:

a) Logistic Regression (LR): This is a technique borrowed by ML from statistics. It is commonly used for binary classification problems. The name came from the function used in it, also called the sigmoid function that has an S shape and takes any value between [0,1].
b) Naïve Bayes Classification (NB): This is the most commonly used, most popular, and simple probabilistic classifier. This model is good for predicting a particular label with the given feature variables. One of the assumptions in NB is that each feature is independent of each other and in this research, it has worked better than other models as that is the case with the dataset considered.
c) Support Vector Machine (SVM): This is another most commonly used in supervised learning techniques both in classification and regression. In this model, each feature is considered as co-ordinates in n-dimensional space. Classification is done by finding the hyperplane by which the classes are differentiating. If the classes are linearly separable many hyper-planes can be drawn. The optimum hyperplane is found by the margin lengths.

d) Random Forest (RF): The decision tree model uses a classification model that works by splitting the data into a minimum of two sets. Decision tree's internal nodes indicate a test on the features, a branch depicts the result and leaves are decisions made after the succeeding process on training. Whereas RF grows several classification trees. The input vector is kept down in each tree to give a classification that is considered as a vote. RF is usually fast and does not overfit. RF is considered more accurate than NB and liked by many researchers.

e) Gradient Boosting (GB): Boosting develops a collection of tree-based models by training each of the trees in the collection of different labels and then combining the trees. This is the same idea used in GB to improve robustness over a single estimator. Thus, it combines several weak predictors to build a strong predictor.

f) k-Nearest Neighbours (KNN): The classification is predicted by a predefined number of training samples nearest in distance to a new point. This algorithm is simple and considered good even on instances of the irregular decision boundary.

g) Decision Trees (DTree): This is considered as a simple representation for classification where the data is continuously split according to a certain parameter. The outcome in this type of classifier is fit or unfit for categorical /discrete decision variables.

3.2.3 Cross-Validation

The various classification models are evaluated by cross-validation. This requires inputs and targets. For every fold, one unique partition is used as the test set, and the remaining (k-1) partitions were used for training. The accuracy achieved for all folds were averaged to evaluate the model that used a certain classification method. This method is iterated k times so that k models and performance estimates are obtained.

Normally, cross-validation is used for model tuning, which is finding the optimal hyperparameter values that produce an adequate generalization performance. Once this is done, the model is retained on the complete training set and obtain a final performance estimate using the independent test set.

3.2.4 Fine Tuning the hyperparameters with GridSearch

The parameters of the learning algorithm are optimised separately which are called tuning parameters or the hyperparameters of a model. GridSearch can additionally enhance the performance of a model by finding the optimal combination of hyperparameter values. In this research, the models are evaluated using model accuracy which is a useful metric to quantify the performance of a model in general. Using *param_grid* the parameters for tuning are specified. After this process, the score of the best performing model with the *best_score* attribute is obtained and the parameters that can be accessed through the *best_params* attribute are checked. Finally, the independent test dataset is used to estimate the performance of the best-selected model.

3.2.5 LSTM & RNN

RNN can be considered as feedforward neural networks with feedback loops or backpropagation through time. Before the neurons get momentarily disabled, it fires for a limited time and enables other neurons that fire at a later stage. It has an additional time variable that is not in multi-layer perception (MLP). This additional feature of the “memory line” will allow the model to not just use the current input but also the earlier input.

Hochreiter and Schmidhuber introduced LSTM in 1997. LSTM is a better version of RNN. RNN has an architecture that embeds the memory concept, yet they suffer from a certain problem called "the vanishing and exploding gradient problem", LSTM has a solution that avoids this problem. The notion in LSTM is to pile the last output in memory and use that as the input for the following step. The study conducted by Namin and Namin on forecasting economic and financial time series in 2018 depicts that the LSTM algorithms are superior and outperform traditional-based other time series algorithm as the error reduction rates returned are between 84-87% (Hua et al., 2019; Yeom, Kim & Chung, 2020). Moreover, the study also specified that “epoch” the number of training times has no effect on the performance of the trained forecast, and it shows just an unsystematic behaviour (Siami-Namini & Namin 2018). Bandara, Bergmeir, and Hewamalage (2020) states that the LSTM techniques are more efficient than many other univariate forecasting methods in their study on multiple seasonal cycles.

3.3 Current Level of Adoption

As per the recent study published by International Research Organisation (2020), decision augmentation is the most widely adopted AI technology, which uses predictive and prescriptive analytics. The study gave the following results as shown in Chart 4. The key findings of the study undertaken in 447 companies in North America are as follows. Forty-eight percent of the selected subjects reported either deploying or piloting an integrated system. Forty-three percent of the subjects surveyed mentioned cost as one of the topmost barriers for adoption. Hitherto, most companies focus on descriptive and diagnostic use cases and predictive to a limited extent. Hardly any companies fully explore prescriptive decision augmentation.

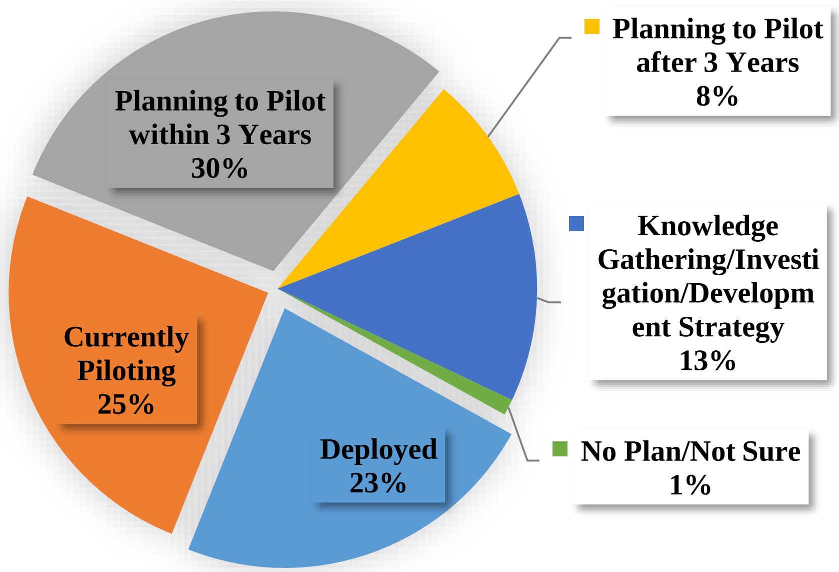


Chart.4: Stages of Adoption (International Research Organisation- Artificial Intelligence Trends: Decision Augmentation,2020)

In the same study of "Artificial Intelligence Trends: Decision Augmentation" (International Research Organisation, 2020), barriers to adoption were explored, which are depicted in Chart 5. Although adoption has a lot of advantages, most of the companies surveyed mentioned that acquiring skilled people was the greatest barrier. It is worth mentioning that the respondents do not consider data collection as a high barrier, or it was the lowest considered barrier. Other barriers to their rate of the barrier are given in the following chart.

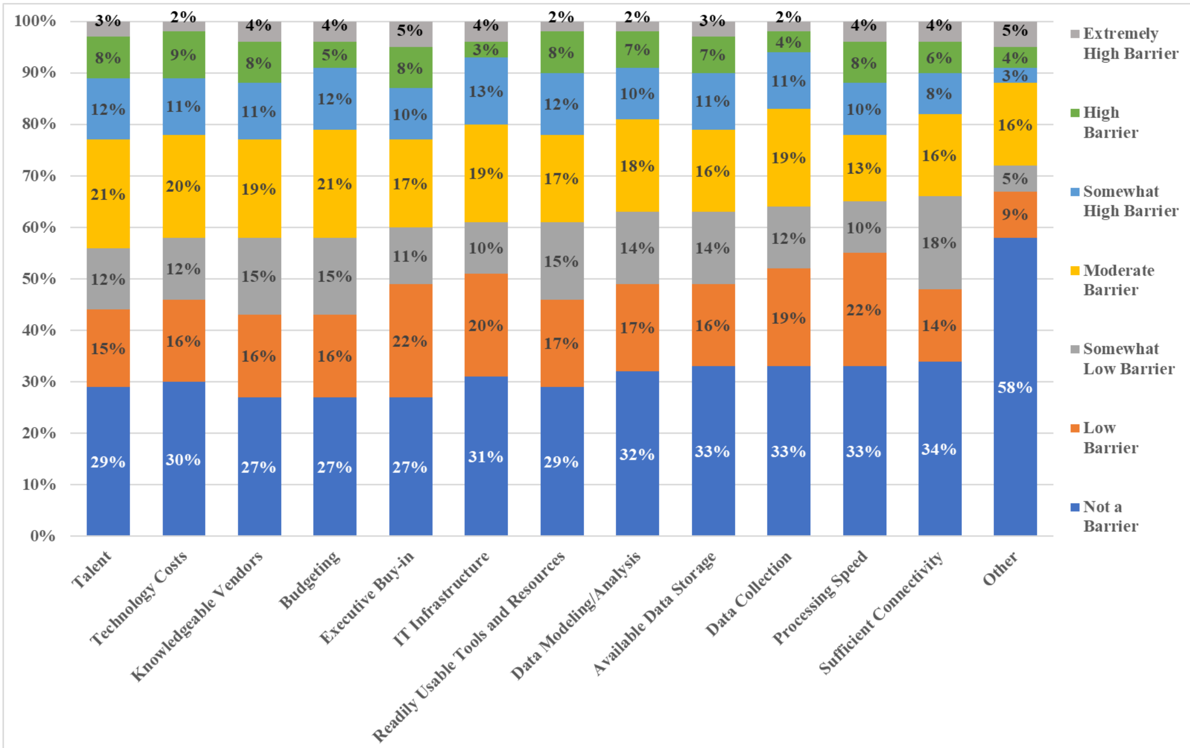


Chart.5: Decision Augmentation barriers adapted from International Research Organisation (2020)

3.4 Major Gaps in the Literature

Certainly, it was found that ML in the e-procurement of hospitality needs more attention and development in the field of research, especially in academia. Though hospitality is gaining attention, the studies are mainly to forecast hotel occupancy, visitor flow, travel package recommendations, tourism marketing knowledge, and so on (Grover & Kar, 2017). It is also noticeable that there are industrial publications, but these too are restricted to a few of the major streams like travel and tourism, and not particularly e-procurement in the hospitality industry. The pieces of literature reviewed either focus on framework or data analytics. There is not a single paper that commends the enhancement of the e-procurement systems while giving importance to data analytics too. This would be the major gap in the literature and my study is unique in this research area. There could be several reasons for this to happen like unclean data, difficulty in filtering valuable data, privacy, security, digital anonymity, internet access disparities, and so on (Mergel, Rethemeyer & Isett, 2016; Jarmin & O'Hara, 2016; Grover & Kar, 2017). The linkage between AI and e-procurement is yet to be established and hence a theoretical framework that helps to identify problems and thus AI integration in e-procurement of hospitality is still to be proven. Any worries from the members of the ecosystem can worsen their participation and so a standardised communication protocol must be established. Moreover, a reliable and secure cloud-based data warehouse should be maintained with compliance with privacy regulations, to gain trust and credibility from the key members.

An AI system aims to assist in decision-making, creating an easier and fruitful process that is vital to the eventual success of the multifaceted business systems. The system should help the organisation in taking tactical, strategical, and operational level decisions. In the swift and unpredictable business environment, it is very important to be proactive rather than reactive. “Predictive analytics is a natural complement to traditional software and processes” (Nenad, 2014). However, most of the systems in the supply chain display information about what has happened and not on what will happen. Predictive analytics can forecast what is going to happen and why it is so. The benefits of having an AI integrated agent and a predictive tool as are as follows:

a. E-procurement holds a large amount of data, pulling out extra value-added information from an existing data repository to predict prospect routines.
b. Standardisation of procurement data measures across the globe will enable a cleansed and consolidated data repository that can be used by agents to attain a common goal.

c. It will allow timely reactions towards ad-hoc situations and quick adaptations to changes because they are prepared due to proactive thinking.
d. Inter-organization, intra-organization, technical and political issues can lead to lack of trust and impede implementation of immense data analytics

This will enable a system that is collective and has relevant intelligence that provides wide-ranging, impulsive, and joint business ecosystems that cover the understanding of predictive analysis to advise corporate decisions throughout the supply chain in general and particularly in procurement.

3.5 Chapter Summary

A systematic review of articles published on topics with e-procurement and hospitality connected with AI is selected. The articles selected has undergone various statistical analysis and presented in the chapter. Similarly, a literature review for ML and deep learning is also conducted. The current level of adoption of AI and gaps in the literature is also presented in this chapter.

Chapter 4: RESEARCH METHODOLOGY

In this section, the methodology adopted in developing the framework and application of ML is narrated. With the overview of the methodology, sampling strategy, data collection and analysis procedures, verification procedures, and ethical considerations are narrated.

4.1 An Exploratory Case Study

"A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not evident; and in which multiple sources of evidence are used." (Creswell & Creswell, 2008). In my research, the hotel chain studied is referred to as *The X Hotels.* The X Hotels is a global brand having more than 2,500 properties across the world. In this study, only The X Hotels properties located in the UAE are considered. The name of the hotel company, its vendors, and the suppliers have been disguised throughout the study for the sake of confidentiality. The data collected from The X Hotels is extracted, cleaned, analysed for further reporting and prediction purposes which could be beneficial for the key stakeholders in this study. The data analytics is done using Python 3.6.

4.2 Research Paradigm

The research will be conducted with an exploratory mixed-method approach in the pragmatic research paradigm. In the research, the author is trying to find the nature of knowledge and limits of knowledge that could be derived from the data gathered. The paradigm uses various methods with qualitative and quantitative analysis.

The research can be considered as a case study in the chain of hotels, which will give an in-depth look at the test subject. The test subject here is the various properties (hotels) of the same chain of the hotel brand. Therefore, data is collected from various sources and compiled using the details to create a bigger conclusion. This could be then added to the scientific world for the dissemination of knowledge.

4.3 Challenges Faced During Research

The major challenge was not getting approval to conduct surveys/interviews with managers of The X Hotel. Getting approval to access extra information can always lead to an increase in the scope of study and also to produce a more valid conclusion. Evaluation of the current e-procurement system to obtain the perceptions directly from users would have been beneficial to find the pros and cons of various users in a broader sense.

A shortage of scientifically published papers focusing on the hospitality industry remained a challenge, especially papers on the e-procurement and hospitality sector linked to AI. The initial search with hospitality and e-procurement gave 121 papers as the search result and it took about two minutes to display the query output. Further scrutiny on the results exposed that the displayed papers were out of focus. A narrowed search for the same topic for the past five years' reviewing journal articles gave a result of 52 sources. Hence, the author had to expand the research areas to e-procurement, big data analytics, and the latest trending technology with specific terms like IoT and AI, which were separately used with the supply chain. The same experience was encountered when obtaining academic publications on ML connected with hospitality and e-procurement.

4.4 Situating the Research

This study can be classified into two major pieces of research. The first one is to propose a framework to integrate AI in the e-procurement of the hospitality industry and the second one is to optimise e-procurement data using ML algorithms.

The framework is proposed based on pure document analysis. The major studies by International Research Organisation, Stefanovic, Gunashekaran, Anthony Jnr, Buhalis, and Leung are used extensively to integrate AI into the hospitality e-procurement system and, to propose the novel extended framework for e-procurement of the hospitality industry by overcoming the barriers in their study.

The research is conducted on a chain of hotels that has more than 20 properties with a various star rating in the UAE itself and is still expanding. The hotel chain has almost 2500 plus properties worldwide with a presence across 100+ countries. The data would be collected for the past five years for analysis and optimisation purposes. The data was available for almost 20 properties. 947869-row data in 2018 dataset and 996930-row data in the year of 2019. The five-year dataset has precisely 4503218. The years 2018 and 2019 are mentioned separately because these two years have an extra column of occupancy rate that was obtained. The data for 2018 and 2019 are used in descriptive, diagnostic and to find purchase duration analysis. The five years' data is used in demand forecasting using the spend and quantity of a product.

4.5 Exploring the Setting

The X Hotels are managed hotels in this region which means it does not own property. There are more than 20 properties staffed by around 10,000 employees with different brands ranging from 3 stars to luxurious hotels in the UAE. The X Hotels already has standardised policies and procedures on implementation and also for procuring items in all departments. The X Hotels has a database server to collect the big data which is collected for purchasing requests. This system is limited in reporting facility, the Field Operations managers extract data from an Oracle database into an Excel sheet. In their opinion, Oracle is limited to data visualization and reporting together. Or in other words, their current system cannot display the report as per the requirement of managers. Therefore, the Field Operations Managers extract data from their database to make a monthly report, spending records, and so on.

4.6 Data Collection

The research tools include an interview, document analysis, data analytics using Python, and a survey. The qualitative study is done based on an interview with a subject matter expert to evaluate the results obtained in the research. A short interview was conducted with the same domain expert in the preliminary study. Triangulation is conducted later for the validation of the results by document analysis and a survey with professionals related to procurement processes in the hospitality industry. The structured interview can then be summarised into themes that rise eventually and are cross-checked with the survey results and documented data.

To find the inclusion level and advantages of AI, intense document analysis was undertaken. Documents analysed include textbooks, study materials, and magazines. International Research Organisation released materials and journal articles are used extensively. “International Research Organisation is the leading research and advisory company that provides senior leaders across the enterprise with the indispensable insights, advice, and tools they need to achieve their mission-critical priorities and build the organizations of tomorrow" (International Research Organisation.Inc, 2019).

Data analytics and predictions take a longer portion in the research since the data needed to be organised and made consistent with the units and formats. The handling of data and value depends on the quality of data collected (Hazen et al., 2014). The data analytics research approach allows the user to overcome the complications in dealing with individual properties as samples. As big data can be used to analyse the entire population – all properties under analysis (Gerald et al, 2016). Apart from simple analysis on descriptive and diagnostic statistics, advanced approaches using ML, AI are considered for advanced predictive analysis like SVM (Support vector machine algorithm), NN (Neural network) classifiers, NB (Naive Bayes), LSTM (Long Short-Term Memory), and so on.

4.7 Proposed Method for Designing the Framework

In this research, the novel extended framework adapts several concepts from the analysed documents. The initial framework for integrating AI and predictive analysis is given in Chapter 2 section 2.2.4. A use-case is used to draw the methodology to develop the final integrated AI framework with e-procurement of the hospitality industry. The methodology is adapted from International Research Organisation (2019) and Andreadou, Papaioannou, and Masera (2018).

A single use-case will help in identifying the process and errors if any. The objective listed in the use-case can establish the complexity and cost of the system. The needs of the user and system can be found out by such a case analysis. The language used is easily understandable to all stakeholders as the case analysis is purely narrative. Extending a use-case and finding exclusions can help the developers easy to understand and at the same time easy to find requirements and develop. Moreover, there exists no scope creep as the case is narrated.

Data and analytics are a great opportunity to see operating models often in a format that is innovative by the nature of putting data and analytics at the center. Data and analytics are a way to connect the missing parts and connect dots by expanding the scope and to see if there is a better way to do it. Many companies deal with a lot of decisions, a decision influenced by other types of a decision like a tactical decision may be influenced over time from the strategic decision, many operational decisions are involved as well. Depending on the complexity of the given environment the decisions can be contextual based on the particular requirement of any organization. The effects of such decisions have an implication on the time that can be dedicated, not necessarily instigated by need human intervention but may be machine-related. Therefore, to adapt this framework some real-time operational decisions must be considered in many cases, such as the best next action, specific service, and product, and so on. This exchange is very critical for the organisation to respond positively or negatively. The series of decisions is based on questions and answers and depends on customer expectations. Any interaction creates value based on the decision taken. While creating a digital scenario there are endless possibilities because this depends on how you want to act to make things work based on the given situation. This eventually evolves into continuous intelligence enhancement that enables humans and machines to work together to optimise and resolve opportunities leveraging analytics to take continuous intelligent decisions.

To achieve the abovementioned objectives, follow the three essential elements in use-case methodology such as the actor, the goal, and the system. The process includes:

1. Identify all system users and their roles.

2. Identify the goals of each user.

3. Identify the course taken to reach the goal

4. Identify all alternative paths and extend use cases.

5. Identify commonalities in a journey to create a common course use case.

6. Iterate the above processes for all stakeholders.

The actors involved in this use case have identified in such a way that would depict a real system in the most possible accurate way and information from International Research Organisation has also taken into account. The list of actors is shown in Table 5.

Table 5: Actors: Users, Buyers, Suppliers, Applications, and other stakeholders

Actor Name Selection List	**Actor Type**	**Actor Description**
Buyer Analytics	Application	An application that provides useful analytic information dashboard to Buyers
e-procurement System (Core System)	Application	Core application used by buyers and adopted suppliers
Decision Models	Application	automated decision models developed.
Buyer & Supplier Analytics	Application	Analytical dashboard accessible to both buyers and suppliers
Partner-facing APIs	Application	An application program interface for all stakeholders
Agents	Application	The various agents developed in the agent-based technology layer
Support Portal & Apps	Application	Applications used to interact with the support team
Suppliers/Service Providers	External Actor	External actor interacting with the system functions and components in the ecosystem
Outsourced Service/Product	External Actor	External actors interacting with the system functions and components in the ecosystem. Will be acting when the purchased item not available in the property or with suppliers.

Third Party/Contractor	External Actor	External actors interacting with the system functions and components in the ecosystem. Will be acting when the purchased item not available in the property or with suppliers.
Users (Buyer)	Role	The user in the X hotel, who initiated the purchase request.
Buyers (Purchasing)	Role	The buyer is usually the property from which the user made the request
Approvers	Role	Approvers are the line manager of the user or the person in charge to approve the requests in field operations
Own asset	Role	The property itself from the request is initiated
Enterprise-run Ecosystems	System	Interoperable systems with stakeholders
Multichannel Interacting & Commerce	System	it involves an app that shows promotions and sales on every channel that guests/customers can interact with the X hotel.
Connected Things	System	Data aspects of devices and sensors involved in routing any information
Endpoint Computing	System	A device or a node that is connected to LAN or WAN to the Information system platform
Back office and operational Systems	System	All programs used by the X hotel
Employee Collaboration & Workplace	System	chatting/messaging platform for employees of the hotel
IoT Analytics	System	An application that provides useful analytic information dashboard about IoT
Industry & Partner-run ecosystem	External Actor	It involves suppliers, distributors, customers, competitors, government agencies, and so on—involved in the delivery of a specific product or service through both competition and cooperation.

4.7.1 Use Case: Initiate Request and Submit Request for e-procurement

Figure 10. depicts a use case of initiate request process and submit a request in The X Hotel. Here the use-case is on an effective digital and analytics platform. In this example, a use-case is a Buyer (the individual hotel) putting forward a request in the e-procurement system which is the core system. The dots or rings represent entities of the use case. When there is a request, the dots must be connected to complete the request process, and this happens as a little journey. The engagement, adoption, and acquirement happen in this path which is connected to the given ecosystem.

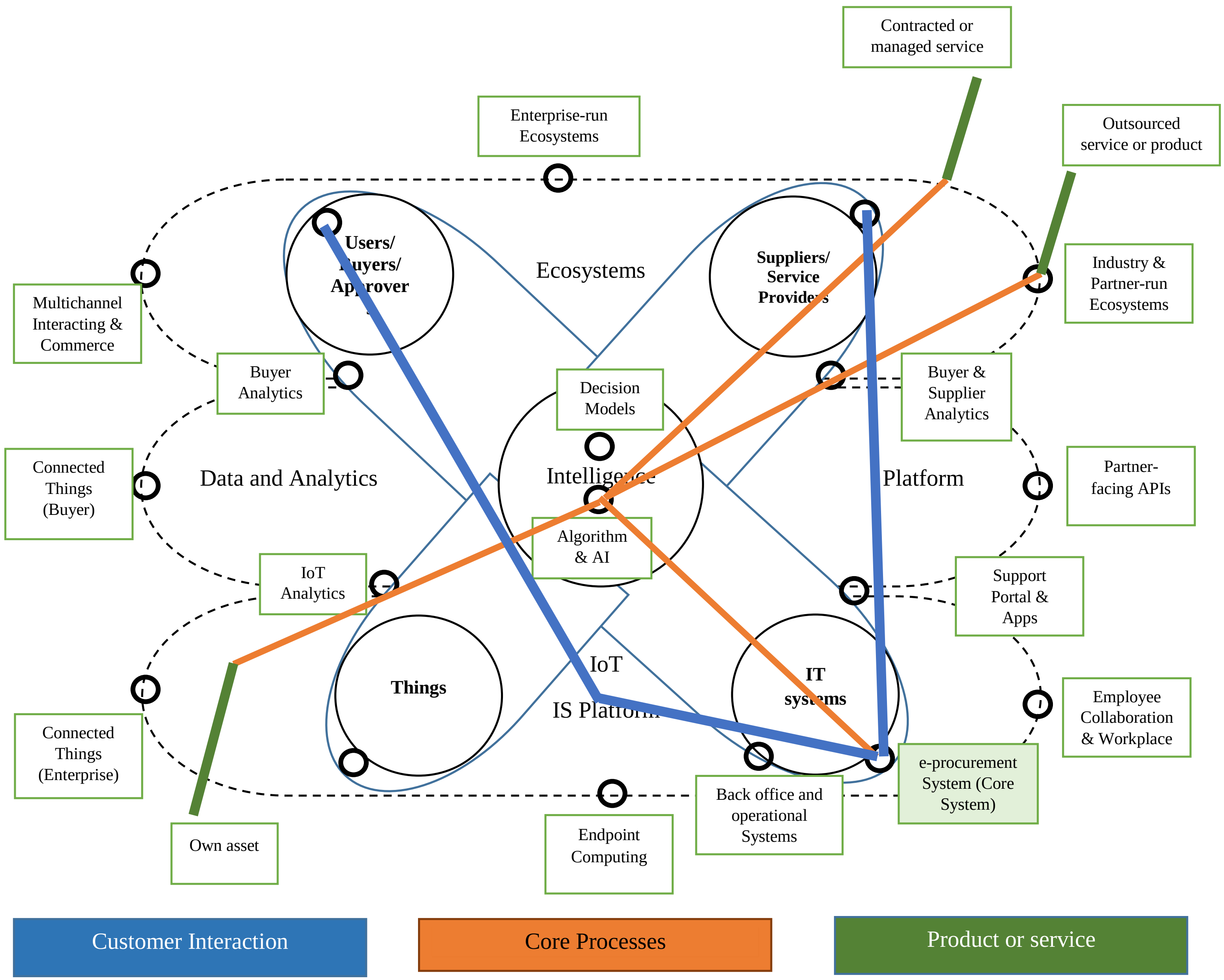


Fig 10: Use case: initiate request and submit request adapted from International Research Organisation (2019)

The ecosystem also needs additional elements like product and service from The X Hotels' assets or other potential leverage that exist like outsourced products/services or some other kind of platform that enables some connectivity. There should be a relevant connection between all entities involved in this process. The use-case shows a map of the critical journey by taking decisions first in which area should be focused on, then the specific capabilities must be combined. Therefore, the map of capability applied to a real scenario is to consider these engagements that are created with the model. The steps are not easy, but it is a multiplication of the opportunity to think in that way so that the process can be accelerated. To implement the next step, the next decision thinking should be considered.

Therefore, the sources that are to be looked at must be defined, as well as what is the data that can be used and so on. Traditionally, the infrastructure means the data sources itself. Digital thinking with data sources will be a digital twin- a simulation of environments, to apply analytics to an environment that is in data. Thus, a series of decisions already emerge from the data and it makes difficult to sense how to apply this in the journey. Contextualisation is the solution for this by making a specific situation for a specific ecosystem and finally, act accordingly. The triggers, preconditions, and post-conditions are also important in any use case. In this example, the trigger is the buyer indicating that they want to purchase a particular product. The precondition for the use case is the buyer selecting the specific product/service they eventually want in The X hotels. The post-conditions include the order being placed and the buyer receives confirmation with a tracking ID with an estimated delivery date.

4.7.2 Components of AI Integrated Framework

The list of analytics components that need to be in AI integrated framework adapter from International Research Organisation include:

1. Information Portal: The structured process's primary concern is to have standard outcomes. To resolve different types of problems you may work with flexibility and understanding what standard outcomes are. To execute, working with the processes that are known very well is a common practice. In this case, credible and consistent reports and dashboards are used to monitor the workflow.
2. Analytics workbench: All the components of analytics are leveraged and defined. This includes various elements that adapt the data analytics maturity model, i.e., descriptive, diagnostic, predictive, and prescriptive analytics. Thus, the different capabilities that can be implemented are obtainable. In this case, an agile and insightful matured data analytics model with self-service analytics and data preparation is explored.
3. Data Science Laboratory: This involves a maturity element as well as a machine element, automation will push in this area of autonomy and performance to go ahead of the maturity model by using ML and deep learning. An organisation needs to recognise that the capabilities and deliverables in analytic space are different for each organisation. In this case, advanced and comprehensive ML and DL are used to investigate the situation.
4. Artificial Intelligence Hub: There are different types of requirements in any organization. Leveraging many use cases to find all alternatives and going for the best would need much more capability than human interaction, especially when it is real-

time, and time depended. In this case, self-learning and autonomous personal digital assistants and agents are used to improve performance.

5. D & A Governance: One area that is not matured is Data and Analytics Governance. It means priority should be given to how best to work with data by knowing the data, the outcome, which policy to apply, and by considering the level of trust that needs to be considered. In this case, the least amount of data should be governed for maximum impact.
6. Data Management Infrastructure: Before even applying the infrastructure, the organisation should think about data anomalies like replication, consistency, missing semantic value of information. These anomalies cannot be fixed automatically since companies are organised in many different ways. Flexible infrastructure is needed to collect and build a connection with the AI framework. To accomplish a connection to any intelligence platform, first, a maturity assessment should be conducted as a pre-requisite then the investment to adapt a physical infrastructure can be defined. There are a few stages between collecting and connecting data through the infrastructure, namely: describe, organise, integrate, share, govern, and implement.

It is important to define the business outcomes, prioritise which data and analytics capabilities are required and to what degree. The data and analytics platform components should be right-sized. Finally, the focus should be given to workforce skills and organisational models not just on technology.

4.8 Proposed Method for Forecasting Problems

The methodology used for ML techniques is adapted from two books namely "Python for Machine Learning" by Raschka, Julian and Hearty (2016), and Python data science cookbook by Subramanian (2015). This section includes the relevant modules used in this research. The method proposed for forecasting is depicted in Figure 11. The proposed method is used for the construction of the purchase duration and demand forecast model for The X Hotels proceeds as follows:

- Exploration of the dataset: Dataset is opened up in the python environment and a data dictionary of the attributes are included as well. The sample consists of five years of data for demand forecasting problem which has almost 4.5 million records. The sample for purchase duration and other descriptive and predictive analysis raw data of one year is used which consists of about a million rows.
- Data Mining: Missing value concerns are not in the data as only mandatory fields from the system are used. The issues in the dataset would mostly be a man-made mistake in the data entry. This error will not be caught unless it is noticed during the data analysis. Making sure the CSV format maintained well was one of the issues that consumed time and the data set involved five years of data. The data for demand forecasting is grouped into months as monthly forecasting is undertaken. Thus, the record number reduces to 60 (5 x 12) from 4.5 million records. Similar weekly forecasting can be considered in a future study.
- Feature Selection: Feature selection is very important for any forecasting and predictive modelling. Features are selected carefully to avoid multi-collinearity and mitigate any redundant features that have a high correlation and thus, improving the overall performance. In this research, consumption, seasonal, room size, or other information were not available. The only available data was quantity, unit price, receiving date, property ID, product name, category name, purchase type. The occupancy rate of hotels

was available only for the last two years. So could not include this field in demand forecasting. Therefore, based on these variables all possible features are made. Adjusted R squared is calculated after adding each feature to find the optimum number of features.

- Model Fitting and Testing: After the feature selection, to find purchase duration seven classification algorithms including LR, NB, RF, SVC, DTree, XGB, KNN are used with selected features. Then a comparison between their prediction accuracy was done after the train/test split method. The test size for comparison was 0.2, which is 80% of the data set was for the training of the model and the remaining 20% of the dataset for testing. Fig 11. below summarises the steps for the proposed method.
- Forecasting: LSTM technique is used for demand forecasting that works well with the sequence of data representation w.r.t time. The predictions are conducted for six months and a time-series graph is plotted against actual data and predicted data for six months to visualise the accuracy in prediction.

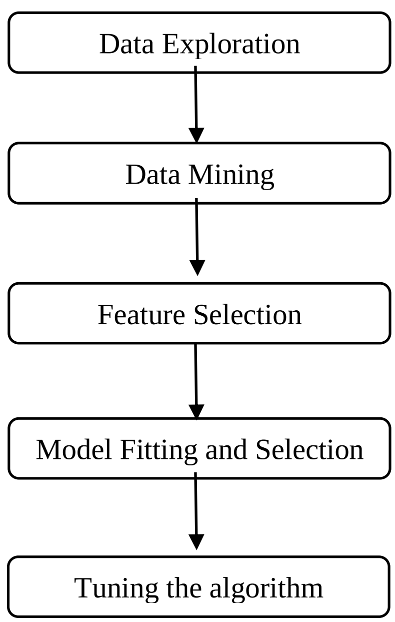


Fig. 11: Method proposed for forecasting

4.9 Validity of the Results

The methodology used for both developing the framework and for algorithm is taken from published works. The interview is conducted to obtain the perception of a domain expert on the proposed framework and forecasting. The reports would be evaluated manually by the e-procurement manager for correctness and relevance. Again, the same will be cross-checked using document analysis and finally, a survey with people in the procurement process of the hospitality industry to accomplish triangulation.

The prediction accuracy is checked by k-fold cross-validation and further analysis is done for the spread of accuracy with box and whisker plot. The accuracy of demand forecasting is done by plotting a time series graph to see the predicted and actual values in the same place.

4.10 Ethical Considerations

A declaration statement and consent letter were provided to the participant(s) in the research. Moreover, a non-disclosure agreement was made to get access and use a large amount of data in The X Hotels' database systems. The interviewee was given a letter of permission and assurance to keep the participant's name confidential. The company name, property names, and vendor names are hidden to maintain confidentiality.

4.11 Chapter Summary

The methodology chapter has clearly shown the steps involved in the model fitting and hypothesis used in proposing a new framework. The steps including data exploration and data mining are equally important in any ML research. The method of validating the results and ethical considerations are also depicted in this chapter. In Chapter 5, data analysis is presented to walk around the details of data exploration, the techniques used in data mining, and the detailed analytics on data to follow in later chapters.

Chapter 5: DATA ANALYSIS

5.1 Data Analytics

Intense document analysis is conducted on proposing a conceptual framework for the architecture of the hospitality industry. The collected data has five years of purchase details from 2014 to 2018 which is used for finding descriptive, diagnostic, predictive analysis, and prescriptive analytics. The same is used in demand forecasting. Two years of additional (2018, 2019) data with occupancy were used to identify the duration of a purchase and making time-series comparisons in which occupancy is relevant. The buyer names, supplier names and other confidential details are hidden/removed from the visuals.

In this research, data mining and data analytics are undertaken on the data collected from a chain hotel in the UAE. The data is extracted from The X Hotels e-procurement software in CSV format. Although the CSV contains a lot of unclean data, the author has tactically used the columns that are mandatory in their system, like Property ID, Received Date, Received Quantity, Item SKU (Stock Keeping Unit), Product Name, Category Name, Category ID, Unit Price, Invoice Amount and Occupancy Rate. The author has done this to make sure that there is no missing data in the used columns and make the result reliable. Common naming conventions are followed for initial adjustment of renaming the columns to consistent, understandable, and relevant attributes. The date column has been converted to the DateTime data type with dd/mm/yyyy format. Furthermore, the quantity column was recorded as an object by python which was converted to float before any further coding is applied. Since The X Hotels has hotels in other countries too, to avoid any inaccuracy and to use this model later at any given location, a query has been used to filter records from the UAE only. The results for each research question is shown in different chapters.

5.2 Data Selection, Slicing, and Filtering

To make the selection, slicing, and filtering of data easier pandas' data frames are used throughout the python program. Selection based on particular columns is used initially to initialise the work. Slicing is done with index and column names are done as well at various stages of the study. Slicing and grouping to dates are done immensely especially in time series analysis. Moreover, the same concerning columns are done in test and split data to rearrange the data frame into features and labels. Filtering is applied to specifically choose certain records, for example: limiting to a certain period by Receiving date and to a country (UAE only). No outliers were deleted, as The X Hotels wanted to know the influence of all properties with regards to spending or accruing.

5.3 Data Clustering

Real-life problems need non-linear models where outcomes are categorical. Grouping data points together based on similarities among them and differences from others is the concept behind data clustering. In this research, most features are discrete or ordinal, so clustering is done to be consistent in handling the data. To estimate the optimal number of clusters *k* for the given task, the elbow method is used in the research. Examples of clusters made are recency cluster, frequency cluster, spend cluster, and occupancy cluster. These clusters are used in finding purchase duration.

5.4 Forecast Error Measures

Forecasting accuracy can be impacted by various features such as season, weather conditions, economic factors, and random cases [like any pandemic] (Nandola, Koshal & Koshal, 1982). Any addition of these features to the research would surely add a lot more accuracy to the forecasting results, but due to the lack of data availability, these features have not been considered in this research. Thereby, in this research time lags for 12 months are used as a feature variable for demand forecasting of spend and quantity.

5.4.1 Root Mean Square Error (RMSE)

The **RMSE** is the square root of the variance of the residuals, is used as the error measure in finding the duration of purchase. It is an absolute error measure whereas adjusted R squared is a relative error measure. (Karen, 2020). It measures the unexplained variability by the regression. For the best fitting model, RMSE should be minimised. In this research, the linear regression method was used as the baseline model from which the RMSE score is taken as the error measure. The RMSE obtained is 1.57, the lower the better.

5.4.2 Adjusted R Squared

For demand forecasting, a linear regression model (OLS — Ordinary Least Squares) is fit and the adjusted R^2 is calculated. As an example, initially, only one feature variable was used to see how much it explains the variation in the column of the label. Then again, the number of feature variables is increased and adjusted R^2 is found until the adjusted R^2 became better with the available features. Adjusted R^2 is better than R^2 (coefficient of determination) because adjusted R^2 will not increase unless the feature is adding value to it. However, whereas in R^2 adding any feature will give an increased R^2 value (Hyndman & Koehler, 2006). Adjusted R^2 increases only if the variable is of use giving an idea of the optimal number of variables to use. Adjusted R^2 obtained in the study is 79% which was the highest after including all the features variables in the model against the label.

5.5 Baseline Prediction

For any forecasting problem, it is always recommended to have a baseline to compare with the complex model. A baseline is the result of a very basic model/solution. Establishing a baseline is essential in any time series forecasting problem. A baseline forecast should have three properties: simple, fast, and repeatable.

5.5.1 Linear regression

Linear regression is used for inferences and prediction and commonly considered as a good baseline to compare with. Linear regression is used as the baseline forecast for finding purchase duration. Finally, the performance measure to evaluate the forecasting is done using RMSE.

5.5.2 Persistence forecast

A common algorithm used in establishing a baseline performance is the persistence algorithm, especially for demand forecasting. First, develop a test harness that comprises of the dataset. Do resampling by train test split to have training set to predict on the test data. Apply performance measure to evaluate forecast, in this study MSE (Mean Squared Error) is used.

5.6 Model Selection

Model selection is a method for setting a blueprint to analyse data and then using it to measure new data. Selecting a proper model allows generating accurate results when making a prediction. First, the model needs to be trained by using a specific dataset. Then the data model must be tested against another dataset. In this study, *train_test_split* function from sklearn library is used for the first forecasting problem whereas manual splitting is used in the later forecasting problems. With this function, there is no need to split the dataset manually. Using the same dataset for both training and testing leaves room for miscalculations, thus increases the chances of inaccurate predictions. Several models are used to test the accuracy of prediction and then choose the model that gives maximum accuracy and minimum variation between the train and test data.

For the demand forecasting problem, the Long Short-Term Memory (LSTM) technique is used. First, the behavioural changes over time need to be removed by making it stationary data. Recurrent Neural Network (RNN) has the property of selectively remembering patterns for long durations of time. This method is said to be the best for trend projection or when the dataset represents long time series data.

The goal of any time series forecasting is to make accurate predictions. The fast and powerful methods that ML is relied upon for, such as using train-test splits and k-fold cross validation do not work in the case of time-series data. This is because they ignore the temporal components inherent in the problem. For this research, the Keras framework is used, which is a high-level neural network API written in Python.

5.7 Chapter Summary

The current chapter represented the particulars of data analysis explaining data slicing, filtering, clustering, error measures used, baseline predictions, and model selection.

The next few chapters cover results and discussions. The results for each research question are depicted in separate chapters. The chapters are arranged as follows: The results of the descriptive analysis are discussed in Chapter 6. Visual reports and data analysis using diagnostic analytics are presented in chapter 7. Chapter 8 explains and discusses predictive and prescriptive analytics conducted in the results with detailed visual representations and validations. The budget forecast and demand forecast is narrated in this chapter. Chapter 9 details an extended framework with complete narration on the integration of AI and ML in the e-procurement system of the hospitality industry, the major outcome. Chapter 9 also contains a summary of the interview and survey results from procurement professionals. Moreover, the discussions and validation of results are explained in the same chapter.

Chapter 6: DESCRIPTIVE ANALYTICS

Some of the reports listed are already generated at The X Hotels but for monthly or quarterly analysis. A few important and relevant descriptive data analytical reports are given below for yearly data analysis. The hotel uses Excel for such display after extracting data from the Oracle database. Here the author has used Python for all types of reporting.

6.1 Supplier Spend Summary

Products bought from each supplier and the total amount paid are collected. Figure 12 shows a report detailing the top ten products with suppliers of The X Hotels. These top suppliers would be the preferred suppliers as reliable and competitive-priced suppliers are vital to the success of any industry.

Supplier_Name	Category_Name	PurchaseAmount
[redacted]	Fruits, Fresh	7781391.92
	Vegetables	7133748.73
[redacted]	Shellfish, Fresh	3902390.92
[redacted]	Juice, Fresh Juice	3900616.10
[redacted]	Chicken, Frozen	3519235.36
[redacted]	Lamb, Portioned	3297473.83
[redacted]	Waters, Still	2755595.79
[redacted]	Beer draught	2531672.05
[redacted]	Beer draught	2417325.30
[redacted]	Beef, Portioned	2336307.59

Fig. 12: Supplier spend summary of the X Hotels

6.2 Spend summary per property

A summary of spend at each property based on overall spend can be extracted and easily understood, this is presented as a pie chart with percentages as displayed below in Figure 13. It is useful and gives a quick idea of how much each property in the UAE has spent each month.

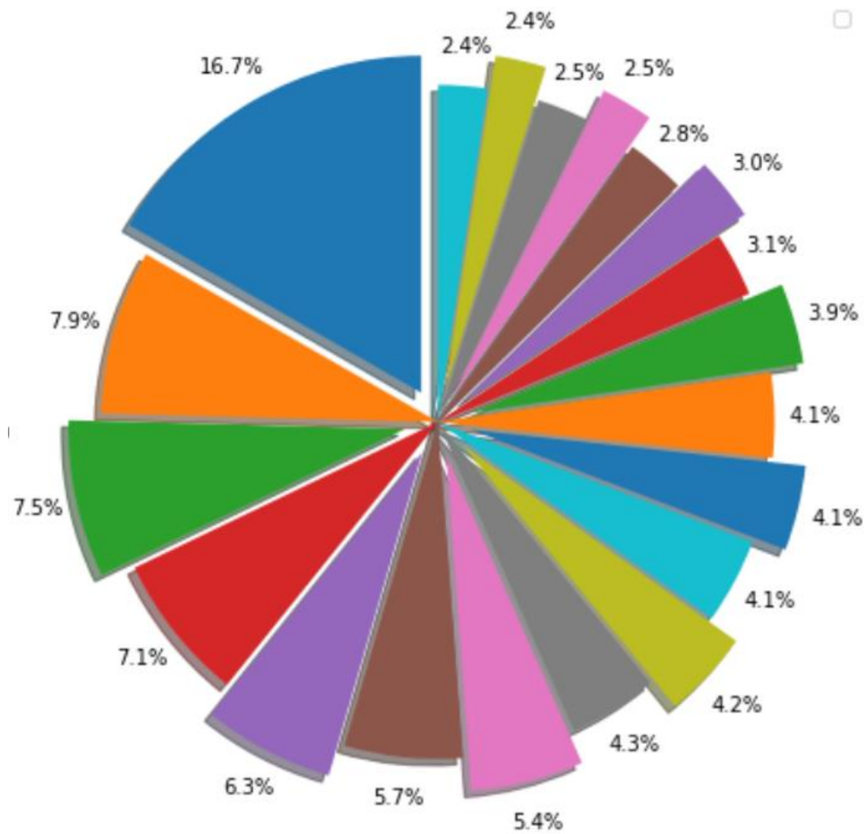


Fig. 13 Spend summary of each property in the UAE

6.3 Spend report by purchase type

The total purchased value is calculated for the entire year. Monthly expenses are extracted and can be used for a comparative study between months and even between years. Moreover, these data can still be used to find the variance in purchasing on a monthly and yearly basis. The top three purchase types and spend for the year are displayed below in Figure 14.

Purchasetype	Beverage	Food	General
Month			
01	[illegible]	[illegible]	[illegible]
02	[illegible]	[illegible]	[illegible]
03	[illegible]	[illegible]	[illegible]
04	[illegible]	[illegible]	[illegible]
05	[illegible]	[illegible]	[illegible]
06	[illegible]	[illegible]	[illegible]
07	[illegible]	[illegible]	[illegible]
08	[illegible]	[illegible]	[illegible]
09	[illegible]	[illegible]	[illegible]
10	[illegible]	[illegible]	[illegible]
11	[illegible]	[illegible]	[illegible]
12	[illegible]	[illegible]	[illegible]

Fig. 14: Spend by purchase type report

6.4 Spend report by category

The top 10 categories are displayed with spend for one year. The report is given below in Figure 15, only the top 10 purchase categories are displayed but it is not a difficult process to make changes to the report by product name or by purchase type, and so on.

Purchase type – The X Hotels has over 10 types of purchases including food, beverage, service, operating equipment, and so on.

Category name – Category gives a clear idea of the category of the purchase type. For example, the "Food" purchase type can have "fresh fruits", "frozen fruits, canned fruits" and so on, as a category.

Product name - Product name is a very specific category with all details, for example, "MINERAL WATER NATURAL SPARKLING, BRAND: PERRIER, ORIGIN: FRANCE, PKG: 20*330ML / GLASS, CASE"

Month	01	02	03	04	05	06	07	08	09	10	11	12	Total
Categoryname													
Fruits, Fresh	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Vegetables	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Advertising Services	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Beef, Processed, Raw	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Juice, Fresh Juice	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Beer draught	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Beef, Portioned	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Shellfish, Frozen	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Waters, Still	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Beer bottled	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Fig. 15: Spend by category report

6.5 Spend by product name

Any company needs to know where they are spending the most. Once the data is framed it is just a matter to know what exactly the top management wants to see and based on their request it would be extracted and either displayed with numbers or graphically. Figure 16 depicts the top 10 spends by product name.

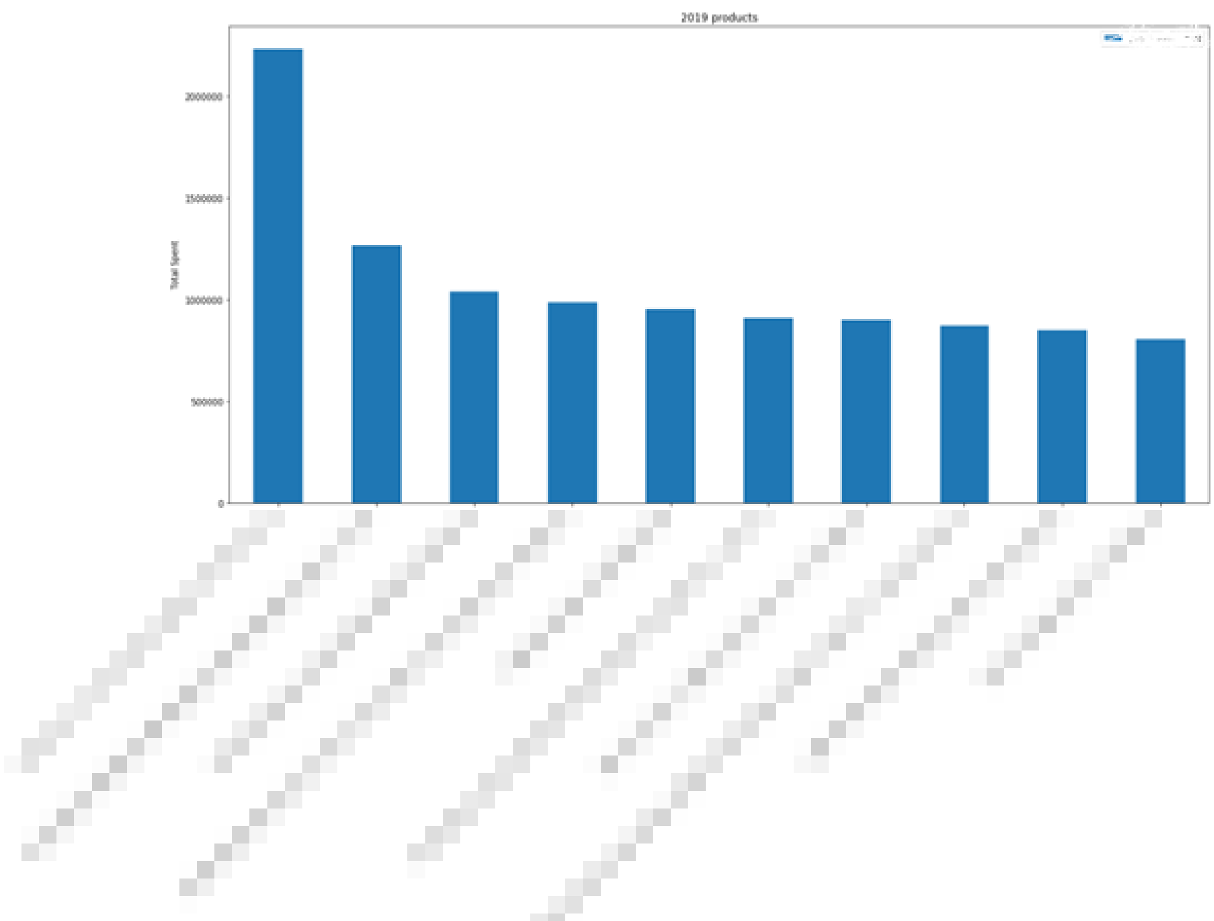


Fig. 16: Spend by product name visual report

6.6 Discussions

In the descriptive analysis, the author made an effort to display five reports. Descriptive analysis is quite a simple and easy representation of collected data and at the same time, it is pretty easy to understand.

1) Buyer Summary: This report is important to the top management as it shows the total spent on each buyer for the items bought for them with the subset of categories bought from them. It helps The X Hotels company with whom to negotiate and make a deal. Using analytical skills and experience the purchase manager can easily give the report on price, quality, speed of delivery, and other factors when a situation arises.
2) Spent summary: This report displayed in a pie chart gives management a clear picture of total money spent by each property in the UAE and what percentage that amount is among all properties. It helps the management to know which property spent the most during the month/quarter/year. Moreover, this report helps to do a comparative study with the budget given and money spent and forecast budget for the coming year.
3) Spend by purchase type: The spend for each category for the hotel chain for the year is displayed in the table. This report can be extended to include a comparative study of each category bought in the UA-based properties and compares it with the previous year same time. For example, 2019 quarterly spend on beverages is compared with the same quarter in the previous year. This would give management an idea if The X Hotels are getting more spend on a certain category or on what price The Hotels is procuring.
4) Top 10 spend by purchase type: This report describes the top 10 purchase type for a given year in the UAE-based properties. This helps the management to know who their close business partners are in procurement and who are the negotiators and deal makers.

The top categories would give the procurement manager a clear-cut idea of what items should be purchased during a period. Furthermore, to find alternate suppliers, if the category of items is not available with the preferred supplier.

5) Spend by product name: the product that is most frequently bought and spend most is important information when it comes to hospitality. The decision-makers are interested to know the suppliers from who these items are bought.

6.7 Chapter Summary

There are various possibilities for the descriptive representation of data. In this chapter, five data reports that are useful for top managers to know when assessing what has happened in past months/years are made some with numerical data and others using visual representation.

Chapter 7: DIAGNOSTIC ANALYTICS

Diagnostic analytics comprises mining data from present data sets to find trends and patterns. These trends and patterns are then used to further diagnose the outcome. Diagnostic analytics is traditional analytics that provides companies with the ability to understand the reason for a certain trend or pattern and to be able to drill down to the root cause.

The five diagnostic charts made are new to the company and are very beneficial to the top management for futuristic purposes.

7.1 Pie chart

This pie chart can be used to identify which category has been used most and the rate of spending as shown in Figure 17. Moreover, this could be expanded to show which month there is maximum expenditure and further trends can be analysed using the data obtained which is represented in a bar diagram and displayed later in the study.

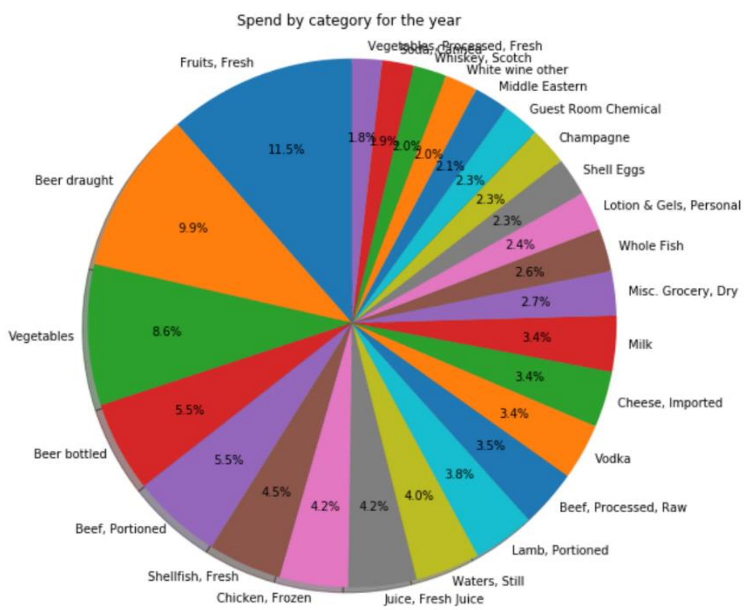


Fig. 17: Pie chart on spending by category

Although the pie chart is not easy to interpret if there are many categories. Pie charts are popular for top management as they are interested in seeing the percentage spend comparisons.

7.2 Time Series Graph

There are a lot of possibilities for a time series graph from this big data. Splitting the received date field into a date, month, and year can produce various graphs to time.

a) Occupancy vs spend: Total occupancy trend can be plotted to show trends in peak time and low time. A similar graph plotted for the next year can show the trend for occupancy and spend analysis could be done on what to expect for the coming years. A comparison is shown in Figure 18.

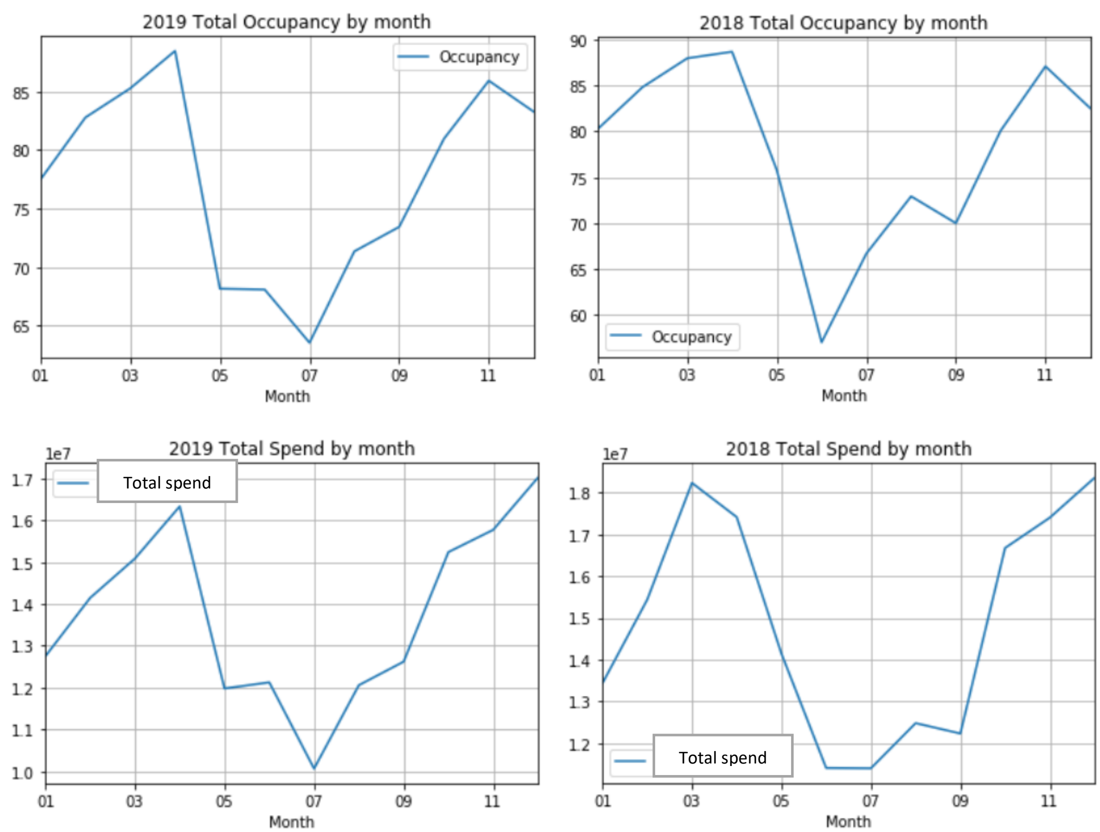


Fig. 18: Time series graph on spend vs occupancy for the year 2018 and 2019

In 2018, occupancy across the region was low in June and so was the spend whereas in 2019 the lowest occupancy was in July and so was the spend. Similarly, the highest spend and occupancy for the year 2018 was in March whereas the highest point in 2019 was in April.

A further comparison is done with national holidays to see if that is the peak time of occupancy. National UAE holidays taken from the official website is shown below in Figure 19.

List of holidays un United Arab Emirates in 2018 and 2019		
Holiday Name	**Date in 2019**	**Date in 2018**
New Year's Day	Jan-01	Jan-01
Eid Al Fitr	Jun-02	Jun-14
Eid Al Fitr	Jun-03	Jun-15
Eid Al Fitr	Jun-04	Jun-16
Eid Al Fitr	Jun-05	Jun-17
Eid Al Fitr	Jun-06	Jun-18
Afrah Day	Aug-10	Aug-20
Eid Al Adha	Aug-11	Aug-21
Eid Al Adha	Aug-12	Aug-22
Eid Al Adha	Aug-13	Aug-23
Hijri New Year's Day	Aug-31	Sep-13
Prophet Muhamad's Birthday	Nov-09	Nov-18
Commemoration Day	Dec-01	Nov-30
National Day	Dec-02	Dec-02
National Day Holiday	Dec-03	Dec-03

Fig. 19: List of UAE national holidays in 2018 and 2019

In 2018, there were no holidays in March but in June there was a popular holiday falling but that is the lowest time of occupancy. In 2019, there were no public holidays in April and July. The highest in both years was in March-April then a sudden increase in August and the highest was in December. The author depicts, this could be the vacation time for schools in the UAE and elsewhere in the world. So more than the national holidays, what matters for the highest spend and occupancy occurs in connection with schools and university spring and summer holidays. The only seasonal peak is in December when everyone has winter breaks and that is the most pleasant period to travel to the UAE.

Although the summer season is really hot and humid especially during the July-August period, the graph shows slight increments to denote the increase in occupancy and spend. As the autumn starts from September the temperature also starts to drop steadily. From September onwards the occupancy and spend increase until December. Thus, one can conclude, spend is directly related to occupancy in the hotel and both peaks in March-April and November-December period for X hotel.

b) The average occupancy of each property for the year 2019: This trend of average occupancy is useful when showing which property is performing well or to identify the preference of property by guests. It may be beneficial for top management to look into details of the performance of the top occupied properties to adapt the strategies to other low performing properties to improve their performance. It will also give an insight into whether guests prefer a particular property(s) because of the star attached to it to show the quality of services expected, or if it is related to rate differences in the various property. Moreover, it will help the top management to conclude if the occupancy is the reason for the increase or decrease in spend per property. Figure 20 depicts the average occupancy of each property for the year 2019.

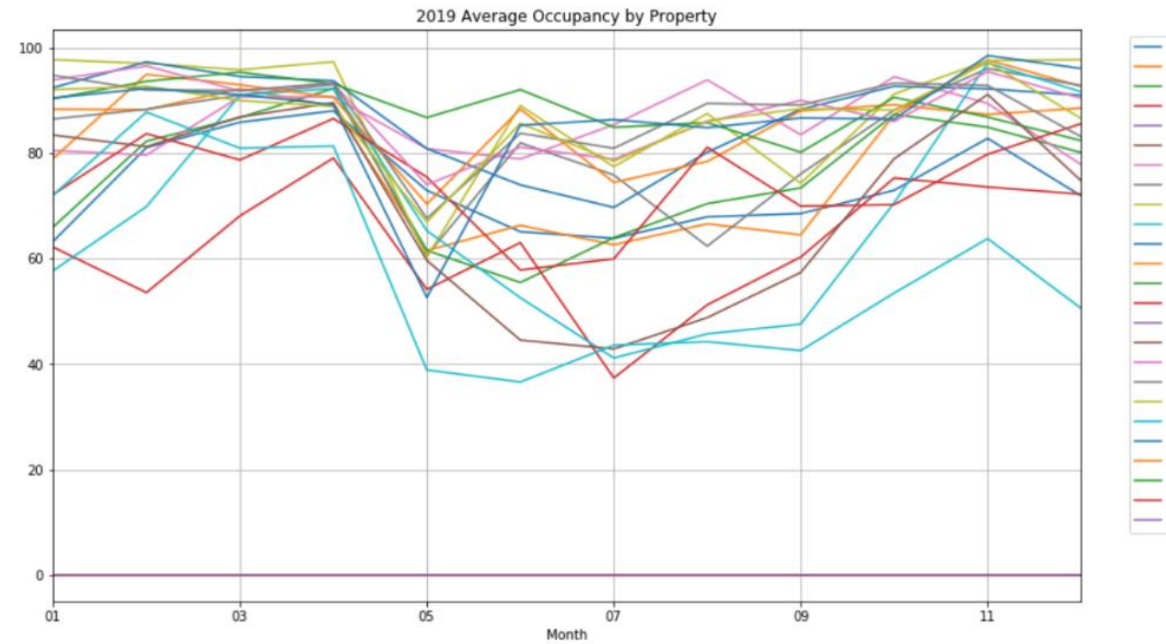


Fig. 20: Occupancy trend in 2019 in each property

7.3 Stacked Bar Graph

The stacked bar chart given in Figure 21 represents the top 25 categories for the year and their spend report. The analysis shows the top 25 categories for the year and their expenditure for each month for the fiscal year. This graph is excellent for studying the trend in spending per category. It can be used to visualize which category item has been used the most in the past year and in which month the most was spent for this category item. The code is shown in the appendix.

The item that has been purchased the most is calculated using total spend for each item. Then the items are grouped by category and month to find the total of each category. Later it is ordered by the total in descending order to select the top 25 from the list. In this finding, it is very clear that The X Hotels has spent the most on buying fruits and that minimum spend was in June, while April, and December show the maximum expense. Further detailing can also be done in the study to find the trend for each property and for each department purchases which would further be analysed for future forecasting. A stacked bar graph is produced which is easy to understand and visualized with various colours to represent 12 months of the year. The X Hotels has over 300 categories of items purchased per year. Various graphs could be plotted either with fewer or more categories based on the requirements from top management. Stacked bar charts are liked by top managers due to the following reasons: 1) easy to view, 2) outliers stand out, and 3) save time.

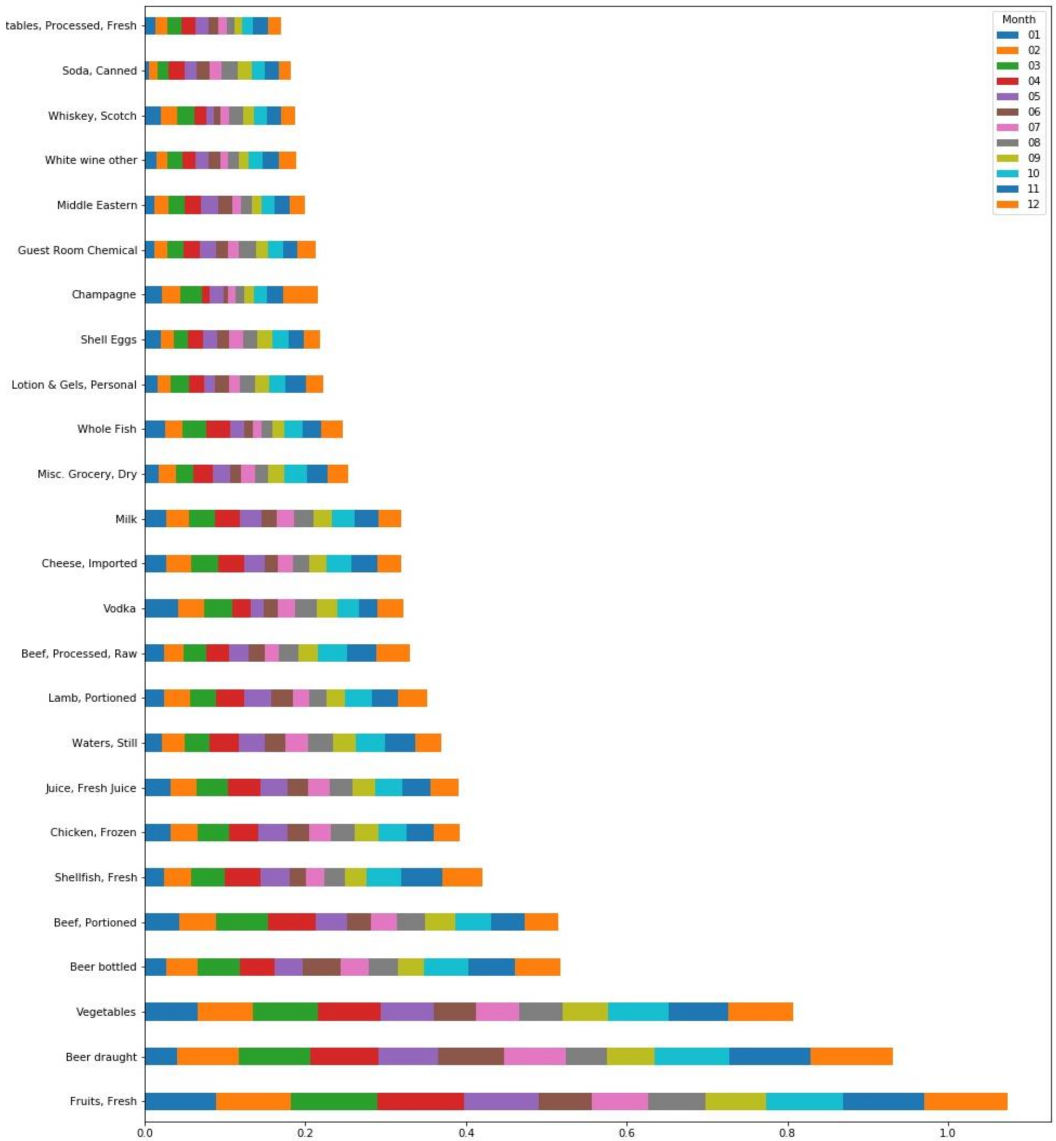


Fig. 21: Stacked bar graph of top 25 categories of the year

7.4 Line Graph

The price comparison of a commodity for various property is represented using the Line graph below where top purchased commodities are found. At the top of the list is mineral water which usually has a regular pattern for delivery and purchase across the country for business and personal use. Therefore, for the visualisation and comparison purpose, the second commodity in the list is taken which is ‘Chicken breast’. Figure 22 shows the price variation and the price distribution among different properties for the year 2018. The prices show fluctuation and vary slightly for various properties.

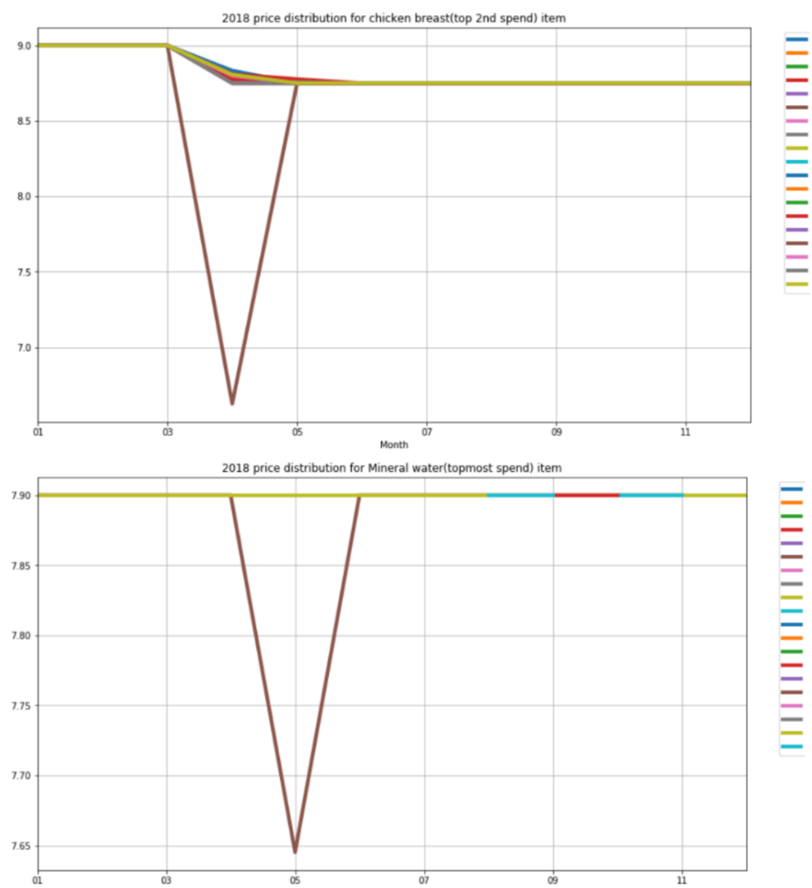


Fig. 22: Price distribution of two top-most products across various properties

A further analysis was done to review the reason for the drop in price for a certain product and property during a specific period. It was identified that the item was bought from the same supplier, using the same unit of measure for all the entries. The reason for variance was due to a supply of ‘free goods’ or ‘less price’ from the supplier, due to the property consuming an enormous amount of that particular commodity during that period. This shows the negotiation skills that should be adopted by other purchase managers as well. Moreover, these kinds of price differences can easily be seen in a line graph. This plotted graph takes the average for each month and each property. To double-check the graph, min, and max were also analysed. For any further understanding, the top management can do further investigation on this variance to identify any other reasoning.

The lowest price was found to be in April 2018. Further analysis showed that the occupancy and spend were highest during that period and so the properties secured a special deal. The brown lined property depicted in the line graph had the highest occupancy during that time across the region and thereby had the highest consumption. This gave the procurement stakeholders the extra edge to negotiate a special deal for one or more of those items. Top managers like line graphs especially as the data variables and trends are very clear and help them to easily identify trends and root causes for such variance (if any).

7.5 Discussions

The diagnostic analysis uses the current data and historical data to represent the trends to identify the core reason for such patterns. The benefit of each of the reports generated is explained below. All these reports are new to The X Hotels.

1) Pie chart: Category analysis is done using a pie chart. The report gives top management an idea on various categories bought with the total spent and percentage spent on other categories is displayed. The category analysis helps the management to procure supplies

more efficiently by giving time for well-organised planned procurement efforts with a better-dealt value and diagnostics, influencing the company's procurement power (Harrison 2013). A further study would be to find spends by subcategory.

2) Time series graph: The change of values of the total spent as time progresses for a quarter or a year is displayed. It makes it easy for the management to identify the trend or fluctuation in spending and use this for analysing the purpose. Moreover, a time series graph is used to compare if occupancy or national holidays influence spend.
3) The trend in occupancy: a time series line graph is plotted to represent the occupancy rate of various properties. This is beneficial for top management in quickly understanding the variations in occupancy rate. In turn, they can still make strategic plans to improve the occupancy by adapting plans from highly occupied property to other properties.
4) Stacked bar chart: Top 25 category for the year and its spend report is displayed in the stacked bar chart. Knowing how much is spent, on what, and how much is a significant baseline. These kinds of visual representations can be used to investigate further what value they have gained by these processes and decisions (Harrison, 2013). The chart depicts that "Fresh fruits" is the category where the company spends the most throughout the year and it also shows that Fresh fruits are bought every month. Detailed data analysis shows it is bought daily.
5) Line graph: price comparison of commodities at various properties is done using a line graph for a given year. The graph shows some fluctuations in the price at a particular period and for one property the variation is high. The graph is shown only for the two

topmost spend items for the year. The reason for fluctuation and assumptions are narrated in the result section.

7.6 Chapter Summary

One of the major advantages of diagnostic analytics is reducing the risk and prioritising workload. This is achieved, as it will allow catching suspicious trends before a loss occurs and thus achieve better partnership and control. In this chapter five diagnostic analyses are conducted, which are beneficial for The X Hotels' stakeholders.

Chapter 8: PREDICTIVE & PRESCRIPTIVE ANALYTICS

Predictive analytics comprises mining data from present data sets to find trends and patterns to predict future outcomes and trends. While it is not an absolute science, predictive analytics does provide companies with the ability to reliably forecast future trends and behaviours.

Prescriptive analysis is done to find the best course of action in the given situation (Wiki). Prescriptive analysis involves descriptive, diagnostic, and predictive analysis but is more focused on actionable comprehensions to maximise business value using advanced analytics.

8.1 Duration of purchase

The duration of purchase knowledge will assist top management, including the purchase manager in charge of forecasted cash flow, in their decision-making. Duration of purchase is done for three different scenarios-by spends, by frequency and by product. All three scenarios pass through similar steps apart from data wrangling. First, the outline of ML is given, and each difference will be explained in later sections. The steps involved in predicting the duration of purchase data pre-processing are as follows:

Step1: Data Pre-processing

a. Data: The columns used and their description are given below. All these fields are mandatory and so there is no missing value and data clean-up is minimal. The data are collected over 12 months of the year 2019. Ten attributes are used in the ML algorithm.

'Recv_Num' – Receiving Number [a unique number for purchasing]

'Item_Num' – Item Number

'ProductName' – Product Name

'Recv_Qty' – Received Quantity

'Recv_Date' – Received Date

'UnitPrice' – Unit Price

'PropertyID' – Property Id

'Country' – Country of the property

'OccupancyRate' – Rate of occupancy

'PropertyName' – Name of the property

b. Data Split: The first nine months are used to predict the customer's first purchase date in the next three months. This will identify if there was no purchase or if there was a purchase. Therefore, the cut-off date is taken as 30th September 2019 and the data is split. This split is used to find out the purchase duration between the last purchase date of the first nine months and the first purchase in the following three months.

c. **Target attribute** (duration of purchase in days): The last purchase in the first nine months and first purchase after that are calculated as the purchase duration, the label. The duration of purchase is found by taking the difference between the latest purchase date in the first nine months and the earliest purchase date in the last three months.

Step 2: Feature Engineering (Applied Machine Learning)

a. Select feature candidates

   i. Recency, Frequency, Revenue, and Occupancy of each property irrespective of a particular product is selected as the initial feature variables. Data visualisation of these features is shown below in Figure 23.

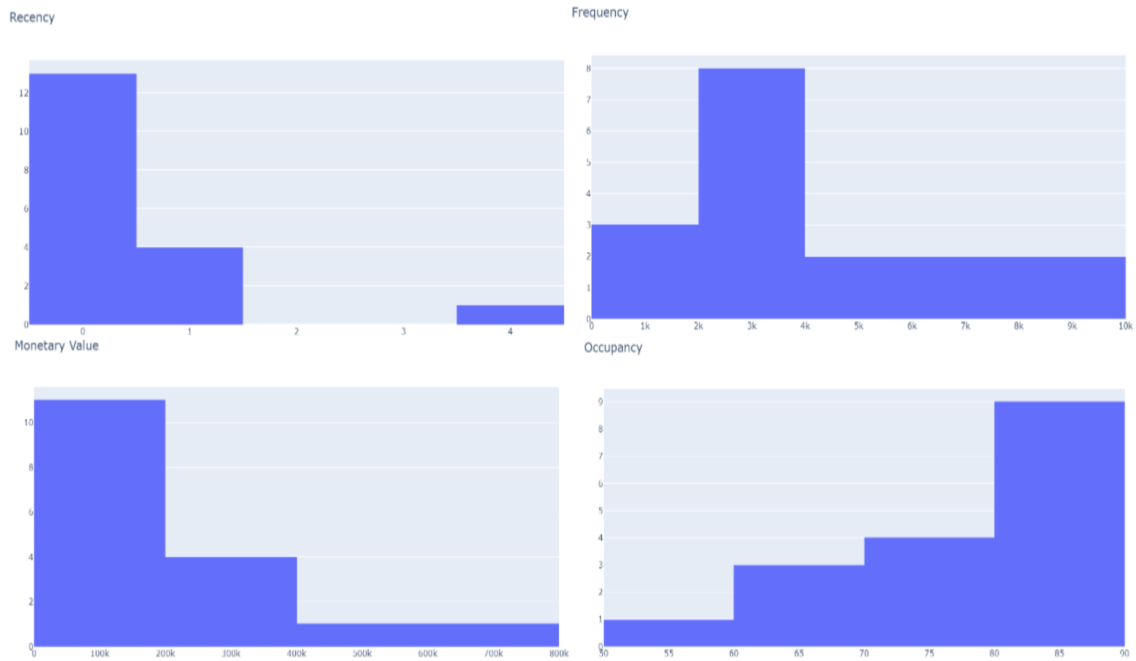


Fig 23: Data visualization of feature variables. Recency, Frequency, Spend, and Occupancy is plotted in the above graphs. Most properties are active buyers as the bar lies in the 0-1 for the Recency plot. The frequency of purchase in most of the properties is moderately high. Most properties spend between US$100k and US$200k in purchasing. The Occupancy rate is high for all properties. It is interesting to see the peak is at different positions for each feature.

The Recency graph shows that most properties are active buyers as data points for most properties reside in the 0-1 bar. The Frequency graph shows the number of times each property places an order, with most properties displaying moderately high purchase frequency. The Spend graph shows many properties belong to the low spending cluster. The Occupancy graph shows there are few properties that guests prefer over others. Unsupervised ML is applied to identify different groups (clusters) for each. Apply K-means clustering to assign a score for each feature. To know how many clusters are required the elbow method is used. The elbow method simply identifies the optimal cluster number for optimal inertia. An overall score is found by adding all these cluster values and then these are classified based on the overall score as grades.

Low Value: Property that is less active than others, not very frequent buyers, and generates very low zero, or maybe negative Spend and less occupancy.

Mid Value: In the middle of everything. Often using the platform (but not as much as the High Values), fairly frequent and occupied, and generates moderate Spend.

High Value: The group nobody wants to lose. High Spend, Frequency, Occupancy, and low Inactivity.

A Scatter plot of each pair of these features is visualized for better representation in Figure 24. These four features pairing up will produce six combinations as follows: Spend & Frequency, Spend & Recency, Spend & Occupancy, Frequency & Recency, Frequency & Occupancy, and Occupancy & Recency.

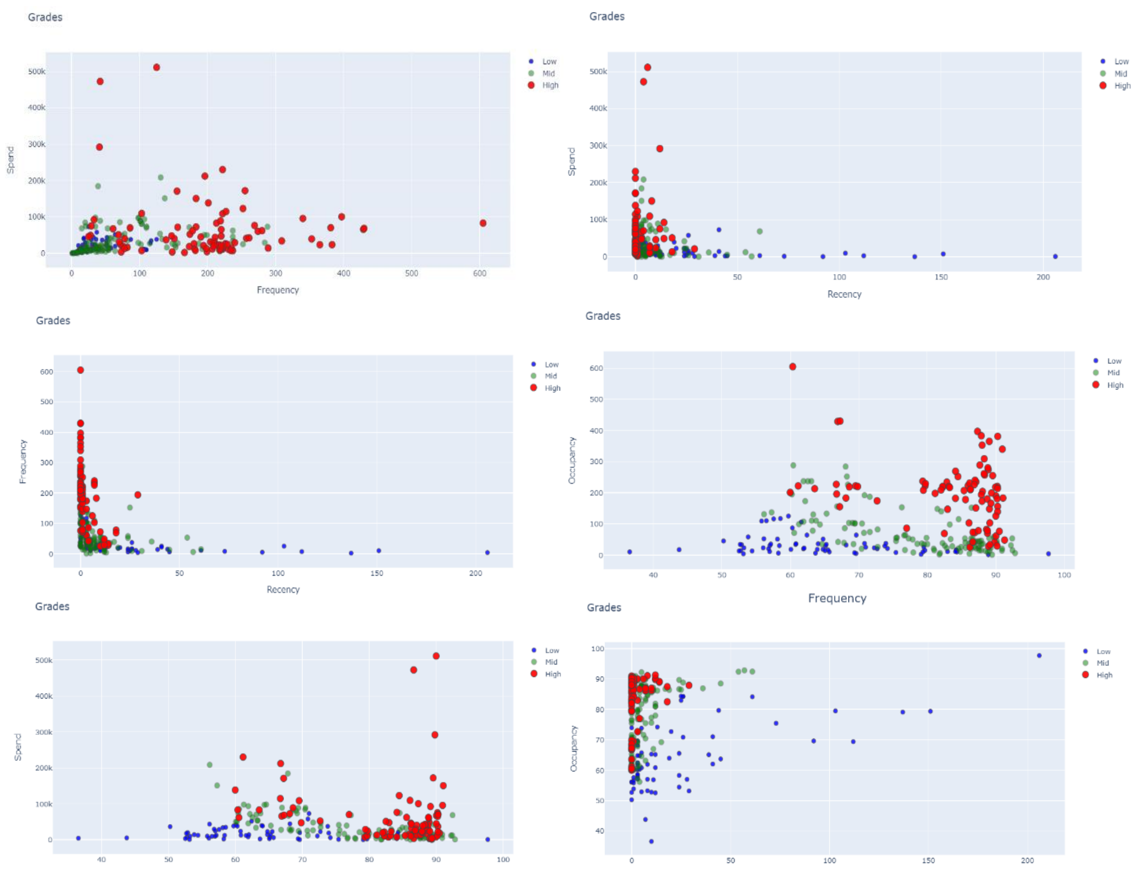


Fig 24: scatter plot between each feature variable. The dots in the scatter graphs above represent the overall score for each combination of feature variables. Red dots represent high scores, green dots represent medium scores and blue dots represent low scores. It is interesting to note that in some cases the score is lower than green or blue but still shows the red colour as each value influences other features.

ii. Days between the last three purchases: Adding more features to the model. First, a data frame with Property ID and Invoice Day was created. Later, the duplicates were removed as customers can complete multiple purchases in a day and the difference will become 0 for those instances.

iii. Add extra features by finding the mean and standard deviation of the differences between purchases in days: The calculation above is quite useful for buyers/property who have many purchases. However, the same cannot be said for the ones with 1–2 purchases. For instance, it is too early to tag a buyer as frequent who has only two purchases but back to back. Keep buyers who have more than three purchases.

iv. Categorical features are defined by applying dummies to create additional features. Categorical features like low, medium and high value are assigned to each property based on their overall score in feature candidates. Once the categorical features are transformed, the feature set is ready for building a classification model.

Step 3: Selecting a Machine Learning Model

Supervised learning using a classification technique approach is used for predicting the models, as per the following process:

a. Identify the classes in the label using percentiles to decide the boundaries: To generalize the results to any given data the percentile of the target dataset is found instead of specifying the actual number. It should make sense in terms of the first one and be easy to take action and communicate. In this research, the average purchase duration is 16 days with the minimum duration identified as one day and the maximum duration as 182 days. The classes are:

1. Class name: 2 ->Buyers with purchase duration less than the 25th percentile (0-3 days)

2. Class name: 1 ->Buyers with purchase duration greater than the 25th percentile but less than the 75th percentile (4-16 days)

3. Class name: 0 ->Buyers with purchase duration greater than the 75th percentile (above 16)

b. Find the correlation between the features and label: To find the collinearity between the features, it is important to know the correlation between each variable. Therefore, the last step is to observe the correlation between features and labels. The correlation matrix is one of the cleanest ways to show this wherein the data is visualized as a correlation matrix heatmap.

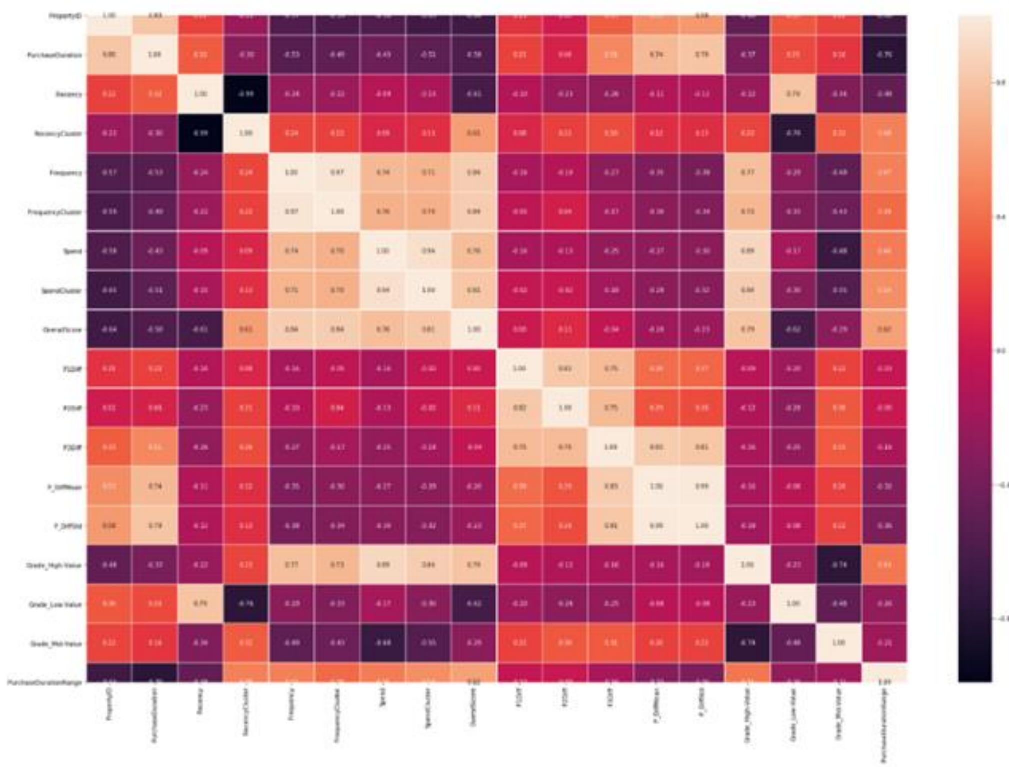


Fig 25: Heatmap to check collinearity between feature variables and the label. It can be observed that Pdiffmean and Pdiffstd have a high correlation with purchase duration giving us an indication of multicollinearity. Later, the model was tested with and without these features to see the effect of these variables in finding accuracy.

Pdiffmean and Pdiffstd have a value of 0.74 for correlation showing they have a strong relationship with the label. Even if these two features show a high correlation with the label, the variables were not removed from making the model as there is a limited difference in comparison with the prediction.

c. Test-Train Split: In this case, the author wants a model with the highest accuracy for various scenarios, as explained below. Therefore, as part of it, first, the data are split into train and test sets to measure the accuracy of different models. The train and test data accuracy demonstrates that most models have acceptable variations in the prediction. This reveals that the feature selection and label prediction is acceptable. The data set is split into 80:20 ratios where 80% of data is the test data. Cross-validation with 10 folds is conducted in the 80% train data and the best results for each model for various scenarios are given in Figure 26.

d. Select the ML for the best accuracy by cross-validation: The various classification models are evaluated by cross-validation. This requires inputs and targets. The inputs are Recency, Frequency, Spend, Occupancy, days between three purchases, mean and standard deviation of purchase days' differences, and overall score values. The target variable is the purchase duration range in 0,1 and 2 classifications. For every fold, one unique partition is used as the test set, and the remaining (k-1) partitions were used for training. The accuracy achieved for all folds were averaged to evaluate the model that used a certain classification method. This method is iterated k times so that k models and performance estimates are obtained. In this research, the models used are LR, SVC, NB, RF, DTree, XGB, and KNN.

Normally, k-fold cross-validation is used for model tuning, which involves finding the optimal hyperparameter values that produce an adequate generalisation performance. Once this is completed, the model is retained on the complete training set and a final performance estimate is obtained using the independent test set.

Scenario 1: Duration of purchase for the frequent buyer

In data pre-processing, the most frequent buyer is found using the ‘count’ function on the received quantity for the top 25 products. This is then merged with the data frame with the original data and the analysis is undertaken.

Scenario 2: Duration of purchase for the product accounting for the highest spend

In data pre-processing, the highest spend products are identified found by using sum function for total spending group by month. These products are then merged with the original data for the rest of the analysis.

Scenario 3: Duration of the most purchased commodity

In data processing, the highest spend item for the year is identified. Then the original data is filtered to display purchases of this commodity only. Then the rest of the analysis is conducted to find the duration of purchase.

		Test data accuracy		
Types	**Machine Learning Technique**	**Scenario 1** **Most Frequent Buyer**	**Scenario 2** **Most Spend Product**	**Scenario 3** **Most Purchased Product**
1	Logistic Regression (LR)	83%	74%	83%
2	Naïve Bayes (NB)	83%	70%	67%
3	Random Forest (RF)	83%	70%	33%
4	Support Vector Classification (SVC)	83%	53%	67%
5	Decision Tree (Dtree)	83%	72%	33%
6	XGBoost Tree (XGB)	83%	74%	67%
7	K-Nearest Neighbors (KNN)	83%	49%	67%

Fig 26: Accuracy of target variable from various models in a different scenario. This shows that the feature selection and label prediction is acceptable. Notice that NB and XGB have the best performing model with different scenarios. Scenario 3 has the most accuracy for both NB and XGB with 83%.

From the result, Logistic Regression and XGBoost are the best performing model of about 83%, 74 %, and 83% accuracy in different scenarios. The applied fundamental concept of ML which is cross-validation to get the result. Cross-validation is a way of measuring the stability of the

ML model across different datasets because it provides the score of the model by selecting different test sets. If the deviation is low, it means the model is stable. In this case, the deviation between scores is acceptable except for Dtree and RF. Scenario 1 is not convincing, as all models show the same accuracy. The buyer, in this case, is an individual hotel in the chain, and when grouped by the buyer the total number of rows is just the number of properties the chain has in the UAE, producing around 20 records for further analysis. This, fewer rows could be the reason for the unconvincing results for scenario 1. Scenarios 2 and 3 looks impressive with varied results, in which LR and XGB outperform other models.

Further Accuracy Check: Confirm the best performance model by further analysis for the given three scenarios. Further analysis with a box plot to find the spread of accuracy within the cross-validation would give additional information on the consistency of accuracy score. Box plot with mean and standard deviation for each accuracy score within 10 iterations are found and results are given in Table 7.

Table 7: Mean and SD of spread of accuracy

		Scenario 1		Scenario 2		Scenario 3	
Types	**ML Techniques**	**Mean**	**SD**	**Mean**	**SD2**	**Mean3**	**SD4**
1	LR	0.75	0.4	0.78	0.1	0.45	0.47
2	NB	0.75	0.4	0.71	0.11	0.45	0.47
3	RF	0.8	0.33	0.74	0.12	0.5	0.5
4	SVC	0.8	0.33	0.53	0.15	0.6	0.49
5	Dtree	0.85	0.32	0.77	0.13	0.45	0.47
6	XGB	0.8	0.33	0.76	0.14	0.45	0.47
7	KNN	0.8	0.33	0.49	0.16	0.6	0.49

The results are evident from the box and whisker plot. In scenario 1, DTree has the best performance with maximum average and least SD showing the consistent result whereas from the best accuracy result all the algorithms had the same result making it difficult to decide the best performing model. In scenario 2, Logistic regression out-perform all other models with the maximum mean of 0.78 and the least SD of 0.1 being the most consistent model with best accuracy score whereas the best accuracy value depicts both LR and XGB are best. In scenario 3, SVC and KNN are showing the best results but the best results show LR is the best. This gives the idea that scenario 3 needs further scrutiny or it gives non-comparable results. The box and whisker plot for all three scenarios are depicted in figure 27.

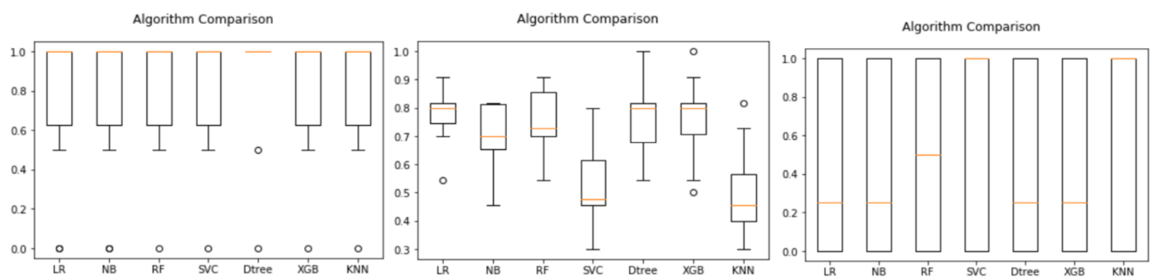


Fig. 27: Scenario 1: Dtree has the least SD showing the spread of accuracy is most consistent**.**

Scenario 2: LR is the most consistent model with the least SD. XGB is also equally good.

Scenario 3: None of them show any better result and scenario 3 could be annul

Step 4: Multi-Classification Model

The parameters of the learning algorithm are optimized separately which are called tuning parameters or the hyperparameters of a model. GridSearch can additionally enhance the performance of a model by finding the optimal combination of hyperparameter values. In this research, the models are evaluated using model accuracy which is a useful metric to quantify the performance of a model in general. Using *param_grid* the parameters for tuning are specified. After this process, the score of the best performing model with the *best_score* attribute is obtained and check the parameter that can be accessed through the *best_params* attribute. Finally, the independent test dataset which is 20% from the train_test split is used to estimate the performance of the best-selected model.

We tried to improve XGBoost performance by finding the optimum value for its chosen hyper-parameter. The algorithm says the best values are 3 and 1 for max_depth and min_child_weight respectively. In this research, tuning the algorithm for XGBoost has shown an improvement in the test dataset to accuracy to 93%.

The complete codes are available in the Appendix.

8.2 Demand Forecasting

Demand forecasting gives an estimate for the number of commodities and services that each property will purchase in the foreseeable future, which will optimize top management decision-making (Acar & Gardner 2012). The data considered are sequentially arranged in time. This dataset describes the monthly number of sales of all commodities over five years. First, a baseline of performance is developed for a forecast problem. Then data are organised, improved, and used to estimate an LSTM RNN for time series forecasting. The dataset has five columns showing the purchase dates, purchasing property, purchased items, quantity, and spend. The task is to forecast monthly total spend and total quantity. Data are aggregated at the monthly level and sum-up the spend and quantity column.

Step 1: As part of baseline checking a simple line plot is drawn, as displayed in Figure 28 to see if there is an increasing or decreasing trend.

Fig 28: An increasing trend for spending and ordering. The increasing trend shows that there are behavioral changes over time and the graph is not stationary. As part of data transformation, the data need to be converted to stationary before converting to supervised learning with feature sets for the LSTM model.

It clearly shows an increasing trend. Then the data is divided into training and test data. The experimental test setup data will do a model on training data and predict for test data. A good baseline forecast for a time series with a linear increasing trend is a persistence forecast where the prior time is used to predict the current time. A rolling forecast scenario is made by shifting the training spend data once. An error score based on Root Mean Square Error (RMSE) is calculated based on the model developed to summarise the accuracy. In this case, the error is more than 9221.876 over the test dataset. Finally, a plot is made to show the training dataset and the diverging predictions from the expected values from the test dataset.

From the persistence model predictions plot shown in Figure 29, it is clear that the model is one-step behind reality. There are a rising trend and month-to-month noise in the spend figures, which highlights the limitations of the persistence technique.

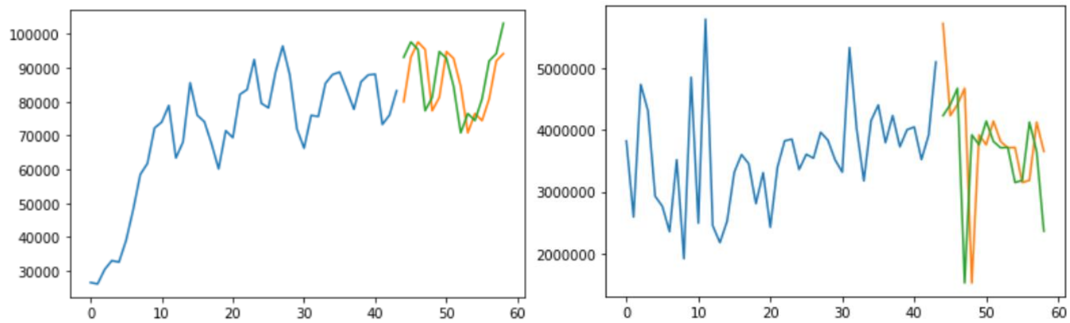


Fig. 29: Baseline graph for spend and quantity forecast. Notice that the predicted model (orange curve) is 1 step behind the actual values (green curve). This shows that the persistence technique has limitations in prediction as it did not avoid the noises in the values.

Step 2: Data Preparation for LSTM

The Long Short-Term Memory recurrent neural network has the potential of educating an extensive series of the dataset. The data preparation is done as follows. First, the data is made as a supervised learning model for ML, then make sure it is stationary, and finally, the dataset should be on a particular scale.

Supervised ML needs input and output variables and use a model to acquire the mapping function from the input to the output. The objective is to approximate the real underlying mapping even when the new dataset is used, it can do predictions. For time-series observations this is implemented by using prior time steps as input variables and the current or next time step as the output variable. This method is called the sliding window method or lag method. In this case, the multistep sliding window forecasting is done. By applying lags 12 times there exist 12 lag long input sequences to predict the output.

To transform the structure of the data into stationery is by differencing. To find the difference (diff variable) just subtract the prior data with current data to remove the trend and to show the differences in the dataset. Time series is stationary if they do not have a trend or seasonal effects. The stationary time-series graph is plotted as shown in Figure 30.

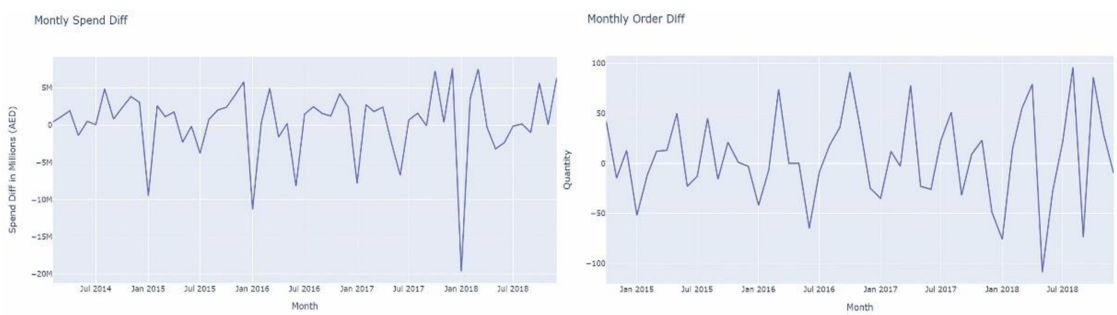


Fig. 30: Stationary time-series graph. Notice that there is no increasing or decreasing trend showing there are no behavioural changes over time. A stationary graph is a flat looking series, without trend, constant variance over time, and no periodic fluctuations**.**

Before using it for modelling it is important to check if it is useful for prediction for that, in this case, adjusted R^2 is found. Adjusted R^2 if greater than 0.5 is moderately good and more than 0.7 can be considered very good bonding.

The adjusted R^2 value which explains how much variation of diff variable is explained is notable as the score is 79%, which is better than the persistence model that achieved an RMSE of 9221.876 over the test dataset. Feature variables are ready to build a model after scaling data. Before scaling, the data should be split into train and test sets. As the test set, the last six months' Spend is selected.

LSTMs presume data to be within the rule of the activation function used by the network. The preferred range for a time series data is -1 to 1. In this case, MinMaxScaler is used for this transformation.

The Long Short-Term Memory network (LSTM) is a type of Recurrent Neural Network (RNN). To compile the RNN, loss function and optimisation algorithm must be given. MSE is used as a loss function as it is similar to RMSE. The code block used [given in Appendix I] also prints how the model improves itself and reduces the error in each epoch. Now the model is ready for prediction.

Step 3: Spend Forecasting

Results of prediction look similar, but it does not tell much because these are scaled data that shows the difference. To see the actual, spend prediction first, the inverse transformation for scaling is done. Second, the data frame is built to show the dates and predictions. Transformed predictions are showing the difference. Calculated predicted spend should also be shown in the same data frame as given in Figure 31. Table 8 shows the real value comparison of actual versus predicted for spend and quantity forecasting.

Table 8: Actual vs Predicted values for spend and quantity

Date	Actual Spend	Predicted Spend	Actual Quantity	Predicted Quantity
01/07/2018	22626740	18705284	317	306
01/08/2018	22781770	23740806	413	368
01/09/2018	21789480	25379430	339	427
01/10/2018	27425110	22904167	425	408
01/11/2018	27493800	32998242	454	450
01/12/2018	33823440	33063405	444	415

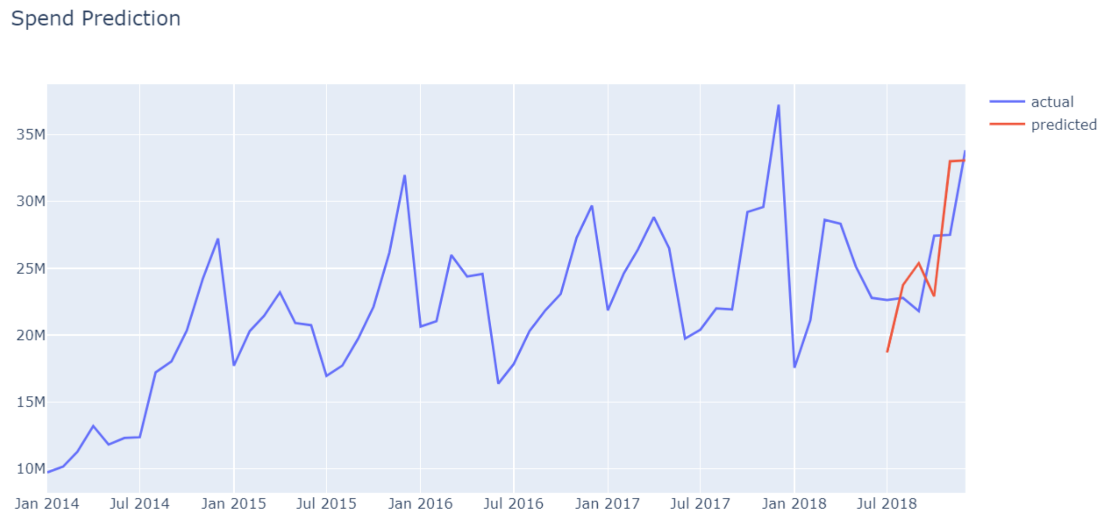


Fig. 31: Spend forecast for the last 6 months. From the plot, it can be observed that the actual spending went up while our model also predicted that the spend will go up. This clearly shows how powerful LSTMs are for analysing time series and sequential data.

This could be considered a good prediction as an increase is shown before it happened, which makes the top management ready to manage cash flow easier.

Step 4: Quantity Forecasting

A similar study is done using quantity ordered for a particular product from The X Hotels. The initial steps in data transformation included converting the data to stationary, convert time series to a supervised model for having a feature set for the LSTM model, and scaling the data. As in the earlier study of spend forecasting, lag 1 to lag 12 is assigned values by using shift command. Calculated predicted quantity is shown in the same data frame with the actual quantity ordered as given in Figure 32.

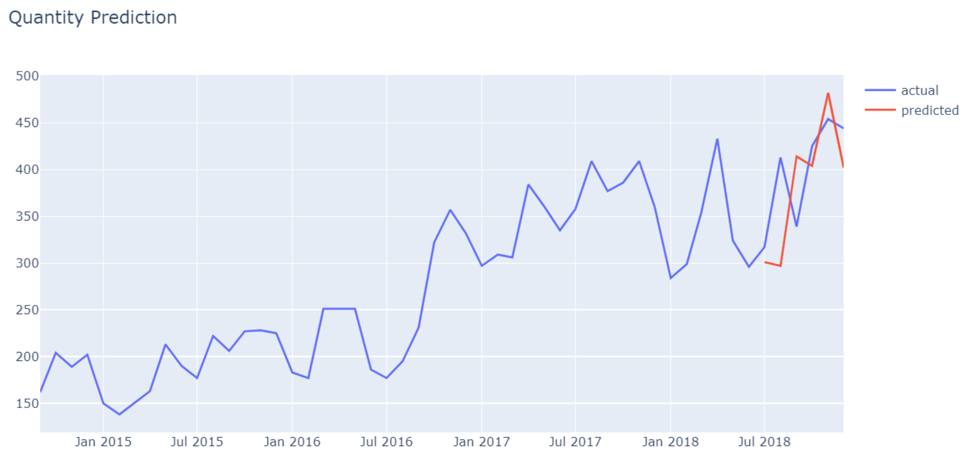


Fig. 32: Quantity forecasted for the last 6 months. From the plot, it can be observed that the actual quantity bought went up while the model also predicted that the quantity bought will go up. This clearly shows how powerful LSTMs are for analysing any variable in time series and sequential data.

8.3 Cognitive Analytics

Few conclusions in cognitive analytics, based on the results from prescriptive analysis follow in this section. The analysis shows that the adoption of such analytics can revolutionise the way the hotels operate currently and provide the stakeholders with enhanced real-time information to make learned decisions, showing them the alternate decision paths with their anticipated outcome. As an example: The Executive Chef, while walking into the hotel to understand the real-time updates on the expected changes in all related variables that can influence his order pattern and virtually guide him with a status on the stock levels that he needs to maintain and the critical stocks that he needs to confirm orders as on date. The study completed in this research would add value to augmented human interaction. Based on his confirmation, the system triggers an order with appropriate authorisation in real-time and a notification is sent to the preferred supplier with the confirmed order. The supplier was already alerted by the system on an anticipated order that was predicted using demand forecasting, and they would have retained their inventory par stock accordingly for delivery in compliance with the contractual terms.

8.4 Discussions

The predictive analysis illustrates the future trends or patterns and prescriptive analysis is useful to understand what actions can be taken to handle such occurrences. These results help the top management to take strategic and better decisions to enhance their cash flow.

1) Duration of purchase: The proposed ML algorithm is useful for top management to take proactive decisions and be prepared for any surprises. Knowing what to buy and when to buy it before the requirement will always help massively in purchasing negotiations and cash flow. The method of following supervised learning and identifying an accuracy score is a steppingstone to start using ML in the e-procurement process of the hospitality industry. Forecasting what to buy and when to buy it is very important in granting powerful insights and foresight. Furthermore, it guides top management in strategic spending to acquire products and services most effectively. It also enhances the cash flow management within the individual hotels. One improvement that can be done for this model is to add holidays, breaks, and other seasonal effects. These variables can be simply added as a new feature to improve the performance of the proposed algorithm.
2) Demand forecasting: In this research, the forecast for spend and quantity is done for e-procurement in the hospitality industry for which a novel LSTM algorithm with eight layers and 500 epochs are used. Six months' prediction is conducted with the least error

of about 0.0137. This information is important to all properties in anticipating developments and being proactive. Furthermore, if there was a noticeable fluctuation in the budget allotted and money spent, further investigation would assist the management in obtaining more information on predicting next year’s budget. Additionally, it guides the top management to take strategic decisions to spend to acquire products and services most effectively. At the same time, it enhances cash flow management within the different properties across the hotel chain. With the latest spread on evolving refined ML-based techniques and, in particular, deep learning algorithms using LSTM, higher accuracy, and powerful results can be obtained for sequential time series data.

8.5 Chapter Summary

Forecasting what to buy, when to buy, how much to buy, and what is best to buy are very important to give powerful insights and foresight. Furthermore, it guides the top management to take strategic efforts to spend to acquire products and services most effectively. It also enhances the cash flow management within the properties. One improvement that can be done for this model is to add holidays, breaks, and other seasonal effects. They can be simply added as a new feature. Moreover, an analysis of consumption data will also be beneficial for the X Hotels to make better predictions on quantity and spend. The author has attempted to put some insight into cognitive analytics as well. With the latest spread on developing refined ML-based techniques and in particular deep learning algorithms, higher accuracy and powerful results can be obtained. ML together with deep learning has good scope in e-procurement within the hospitality industry as the data has not undergone any exploratory data analytics study.

Chapter 9: NEW PROPOSED CONCEPTUAL FRAMEWORK

9.1 New proposed conceptual framework for the hospitality industry

Figure 33 shows the proposed integrated AI framework for the hospitality system that consists of four layers. The initial framework is shown in grey colour by Buhalis and Leung (2018) and the extended components are in brown. The first layer is the network layer which consists of all stakeholders' data including internal, business partners' and external data from sensors as well. Selected data and contextual data will make big data in the second layer of cloud computing. The second layers include all information systems and IoT which are also used as a medium for storage and retrieval of data. The third layer is an AI layer that performs data analytics with little intervention from humans that would augment the decision-making support for top management. The final and the fourth layer is the ultimate governance solution layer where human intervention is null or minimal. The third and fourth layer are separated by very thin dotted lines to depict these layers are interoperated. In this layer, there is process automation with decision support systems plus intelligent agent technology which includes robotics as well. The top layer shows that continuous enhancement of the intelligence is required and done via dynamic feedback from the ecosystem elements. A completely integrated intelligence system should be interconnected and interoperable as well as able to sense, accumulate, explore, and infer data dynamically.

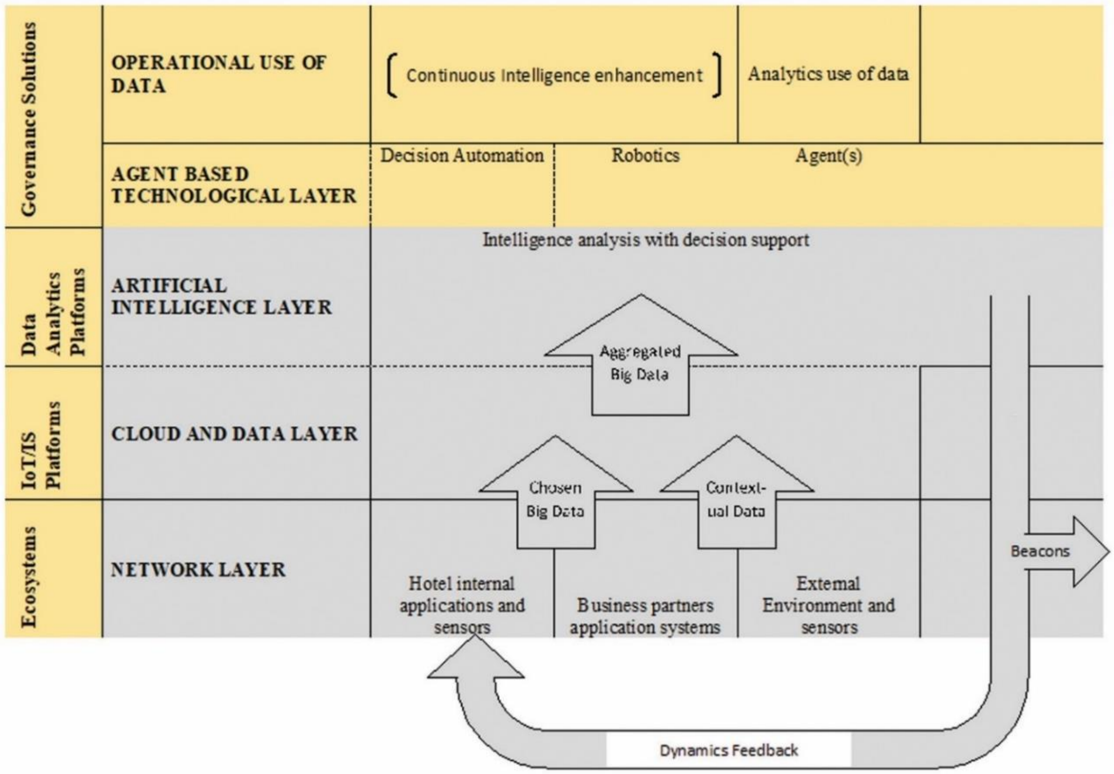


Fig. 33: Conceptual framework of AI integrated hospitality system

The internet brings endless opportunities in the business environment. The hype in the hospitality industry all over the world forces hoteliers to be proactive and innovative in ways to accommodate digital natives. Intelligence through interoperability and interconnectivity with stakeholders, including customers, suppliers, vendors, market conditions, and others makes decision-making more dynamic and efficient. However, the ecosystem consisting of all these stakeholders with information and financial systems and IT experts makes it challenging to interconnect as there is no standardisation yet among these members. The ecosystem of the hospitality industry is depicted in Figure 6. in section 2.2.9. The hospitality industry must work closely with all stakeholders to make sure all customers' needs are met efficiently and promptly. All stakeholders are customers to other curators at least once in the ecosystem. Although hotels have all departments in full swing, this study focuses on the supply chain, therefore the description is limited to the supply chain. However, in any case, the conceptual framework suggested could be used as a whole solution for integrating different areas of work and to other industry domains such as supply chain logistics, demand supply optimizations in e-commerce etc.

With the integration of AI, the hospitality supply chain can also be revolutionised. It is common sense to understand that no hotel can work without resources and purchases. In any kind of hotel, whether it be a small, medium or large enterprise, there is a large number of employees in the supply chain, including food and beverage, guest room suppliers, HVAC (heating, ventilating/ventilation, and air conditioning) vendors, technology suppliers, maintenance and service providers, and others. Moreover, they will, in turn, be connected with other sub-contractors like butchers, farms, wineries, transportation companies, maintenance services, and warehouse, which form further sub-ecosystems (Waller & Fawcett, 2013). However, these sub-contractors do not have a direct connection with customers or guests. Even then, they are equally important to sustaining the ecosystem. Integration with the latest technology will enable the collaboration with stakeholders who have the commodities required within the best time and price limits. Being interconnected and interoperable makes the system procure valuable information in real-time, quickly, and efficiently to arrive at the optimum solution for the situation.

Technological evolution has recently made several models that integrate programming to forecast and predict future business trends based on historical business trends. However, still, there is a long way to go to achieve the maximum benefit from technology integration. The new model displayed interacts with external data for example any recent changes in political, economic, social, technological, and any other external shocks, such as an epidemic, that will directly affect hospitality demand and supply. This means this model not only handles the internal big data but also external big data that can support decision-making on rates and increase profits. Hospitality can use sensors and beacons to communicate with stakeholders and enhance the relationship. Automating hotel operations itself has a wide scope, such as a robot for housekeeping, robot concierge, a robot receptionist, and so on, that integrates with internal and external international real-time information to serve the customer best. Sensors in the kitchen, for example, can identify the expiry date and also obtain information on guests and help the chef to be proactive in planning the consumption sequence. RFID tags can be used to check par stock levels and purchase orders can be sent out directly to suppliers to avoid being out of stock and thus can reduce human resource consumption.

9.2 Conceptual Model for E-procurement in the Hospitality Industry

Figure 34 is the proposed conceptual model for e-procurement in the hospitality industry derived from the proposed extended framework for the hospitality industry. The study by Anthony Jnr (2019), suggests having 10 agents to enhance e-procurement processes but in my study, four more agents are added and coloured in orange in the figure. These agents are included to address few of the problems stated in Chapter 1 in section 1.1. The agents are included to monitor price variations by catalogue agent, budget forecasting by optimisation agent, duplication of supplier/buyers by risk assessment agent and cyberattacks by authorisation/security agent. The data analytics platform will offer data analytics maturity models using descriptive, diagnostic, predictive, and prescriptive analytics by incorporating agents in the agent-based technological layer.

The first agent is the authorisation and security agent to keep the data encrypted to maintain confidentiality. This agent limits the person from accessing only the documents that they are authorised for. This agent is in charge of data protection and cybersecurity. The second agent introduced is a catalogue agent that is supposed to maintain the system catalogues. Whenever there is a commodity added/edited/deleted by a particular supplier this agent is in charge of notifying the Price regulator. Once the change has been authorised by the Price regulator, the catalogue is updated, which is accessible to the procurement stakeholders. Furthermore, this allows the supplier to be notified of whether their changes have been accepted.

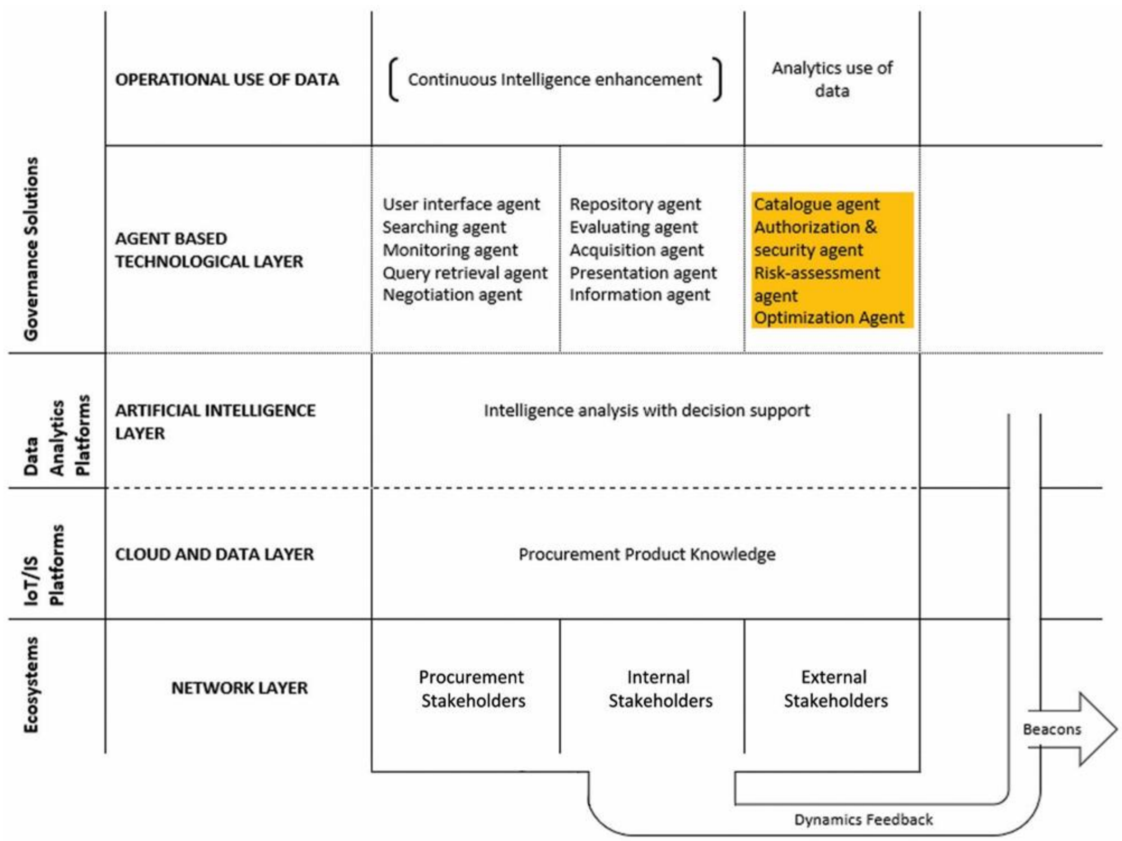


Fig. 34. The proposed conceptual model for e-procurement in the hospitality industry

The third agent introduced is the risk assessment agent, which will monitor the risks, such as currency fluctuation rates/pandemic alerts, and augment all related stakeholders to arrange for the mismatch in requirements/resources. The fourth agent added is the optimisation agent which will continue to monitor decisions taken to optimise the requirements and pre-empt alternative solutions to the given scenario. The optimisation agent has a continuous improvement feature to its responses in any given scenario based on its previous learnings from real-time responses or outcomes. The risk assessment agent and optimisation agent work with ML techniques to analyse risk and optimisation. These predictions will help the procurement managers to take proactive steps in planning the demand and supply based on the aftermath of economical fluctuation and travel controls. The demand forecasting algorithm and purchase duration algorithm is embedded to these agents.

9.3 Discussions on the Extended Framework

The large amount of data collected in the hospitality and especially in e-procurement is crucial in decision-making and necessitates in various situations concerning optimization of operations, revenue, cost, and competitiveness. For example, forecasting supports to estimate business opportunities and income and, efficiently tackle the convolutions (Buhalis & Leung, 2018). In Chapter 8, demand forecasting is already discussed using deep learning for spend and quantity of products using five years of historical data. These data can be helpful for a supplier to maintain optimal inventory and to efficiently manage cash flow. Based on the demand it will help buyers to decide if negotiation is possible and use the agent dedicated to it. It can also be furthered to other forecasting options, with additional data available for recommendations of pricing, rates rules, distribution channel management (Guillet &

Mohammed, 2015). Situational analyses of the state of demand and supply should involve all stakeholders vigorously to find optimal levels to maximise the profit. An immediate revenue-focused response can exploit revenue performance (Jang, Chen & Miao, 2019). Therefore, using AI in a framework where internal and contextual data are combined can have significant effects on revenue and cost management because the accumulation and merging are easier. For example, risk assessment agents can check the occupancy fluctuations daily and provide this as an input to the optimisation agent, which will give a forecast for the next few days or months on the expected inventory or stock level that should be maintained. Such an AI integrated framework will increase the accuracy of demand and supply at any hotel. The final result will be better market strategy planning, financial management, manpower adjustments, and optimization. Thus, the proposed AI integrated system is outcome-driven and focused on usability.

Furthermore, the use-case explained in Chapter 4 under section 4.7 can be used to extend further to a solution architecture that clearly shows the related dataflow and inter data dependencies. In section 4.7 the author has given detailed use-case methodology with an example and the components required to integrate AI in an industry. Moreover, the importance of having an expert who knows data extraction and data analytics is specified in section 2.2.4. Additionally, the initial framework to integrate AI and predictive analytics is also specified in section 2.2.4 A and section 2.2.4 B.

9.4 Summary of the Interview

An interview was conducted with the subject matter expert in e-procurement at The X Hotels in the UAE to explore the importance of integrating AI and ML in the hospitality industry. The interviewee has worked at both property level and corporate level. Moreover, he has worked within the hospitality industry for more than 20 years in various roles and at several properties. There were seven main questions which were structured. Specific names within various examples given by the interviewee are omitted as a part of the non-disclosure agreement signed with The X Hotels to maintain confidentiality. There was no point in spending time obtaining approval from each supplier and properties in the study as The X Hotels deals with at least 300 entities and, most importantly the author was denied the approval to get connected to the stakeholders of the X hotel.

The questions with summarised answers are as follows:

a) What are the pros and cons of the present purchasing methods at a set of properties owned and managed by The X Hotels Company?

There were a lot of pros with the current e-procurement system that exists across the properties, but the focus was on the cons to understand what can be done to improve them and why a further analysis would be required. Hence, the summary includes only the cons. The X Hotels has a database server to collect the big data which is collected for purchasing requests. This system is limited in reporting facility managers to extract data from the Oracle database into an Excel sheet. In their opinion, Oracle is limited to data visualisation and reporting together. The prices for commodities bought from the same vendor at different properties showed variations at times and was a constant challenge. A single system that integrates across all regions with a standard rate for each commodity from individual vendors was missing, where vendors can maintain their negotiated fixed price for that period in the system. These prices will then require authorisation from purchasing stakeholders to be visible to all buyers. This would entail the need to define preferred global or regional suppliers maintaining a standard catalogue, eliminating the need for the further negotiation for the given period. This also meant that The X Hotels group would then need to maintain assessment on compliance across properties to review the spend against these preferred global or regional suppliers. Thus, program rollouts across properties or regions will require monitoring the performance of each vendor for their supplied commodities.

Moreover, supplier rating is an area that needed definite improvement by reviewing seller relations and also to facilitate betterment in vendor fees and genuine discount plans based on purchase history. Simultaneously, a new process will also need to be enabled for the purchase stakeholders to ensure orders are released within the agreed supplier order lead time ["lead time is the latency between the initiation and completion of a process" (Wiki)] that could be monitored in real-time. Performance of lead-time is currently not being monitored at all and is an area that can help to better rate Supplier performance. Meeting the agreed lead time will also enable properties to understand the par stock requirement ["Par Stock was defined as a level of inventory items that a buyer believes must be on hand to maintain a continuous supply of each item from one delivery date to the next" (Kothari et al., 2007)]. Furthermore, a consolidated catalogue structure across the preferred list of suppliers and the ability to track real-time performance of suppliers would help maintain consistency and transparency towards achieving the strategic sourcing objectives across all properties of The X Hotels.

b) Why do you think the integration of AI and ML is important?

As of today, purchasing decisions are more reactive and purely judgmental and thereby this can lead to an urgency in purchasing decisions. Urgency always comes with a cost. This can result in many costs, such as lack of time to negotiate prices or payment or delivery terms with a supplier, additional administrative efforts to authorise orders; this could be an additional order on top of the standard order for that supplier and day, which might lead to an increased count of orders that causes more hassle for the Supplier / Receiver and an increased amount of invoices to be processed at the property. These additional incremental costs have a direct impact on the bottom line. The system should be able to predict the duration of the next purchase and augment the stakeholders with the list of products they need to order with the best alternatives and avoid last-minute emergency hassles.

c) How can purchasing duration be helpful for The X Hotel?

Buyers are currently relying purely on historical information and best judgment to initiate their orders. Often, this results in increased spend or shortage in stock that leads to emergencies. Knowing the duration of purchase will enrich the management to be proactive and make better decisions and thus better negotiations.

d) How can demand forecasting enhance The X Hotel's planning and strategy?

Forecasting accuracy is a part of The X Hotel's KPI. This will surely benefit the properties to achieve KPI and allow negotiation with the preferred suppliers. More accurate predictions will help build a better partnership with key suppliers. This will also help The X Hotels to perform a more accurate cash flow analysis. Demand forecasting will help the department plan their resource availability. It will also help The X Hotels maintain optimal inventory. Furthermore, analysing more accurate supplier lead times and incorporating that information into demand forecasting can improve the process of strategic sourcing.

e) Would these results not be feasible to improve upon good judgment?

It is usually not easy to get good judgment. Moreover, good judgment may not be as accurate as quantitative analysis. When it comes to accuracy quantitative analysis is preferred over qualitative analysis. Furthermore, sometimes it is more expensive than demand forecasting using data analytics. The top management will be naturally confident in the judgment by a subject matter expert, but not in all cases it is reliable. Thereby, predictive and prescriptive analytics will surely supplement and frame a stronger foundation for effective decision-making.

f) What recommendations can be made from this forecast model?

Adding a compliance agent that would monitor supplier performance by maintaining the compliance criteria against each product. This agent would interact between the catalogue agent and evaluation agent to ensure compliance and supplier rating based on real-time performance. It would be beneficial if it can also find supplier switching costs, as an additional feature in rating the supplier. This information will provide a lot of other possibilities in making better strategic sourcing decisions [Switching costs are costs that a consumer incurs from switching brands, products, services, or suppliers].

As of today, an executive spends a lot of their quality time understanding what, how soon and how much do they need to order. It would be a dream come true for many if the system could predict this for them and the foundation for such orders originate from more accurate demand forecasting. Now that the working model is ready, the possibilities to explore and use data are humongous. Expectantly, the duration of any particular product or even product quantities to order can be easily found by just making minor changes in the given model.

g) Do you think these results in this research can be generalised?

Having worked with more than five e-procurement applications that are widely used across the hospitality industry, and experiences gained in various roles in the international chain, the domain expert assures that the e-procurement systems have more or less the same datasets. In this research, the attributes used are a must in all hospitality e-procurement software. Those hotels that have a similar dataset or procurement system will need to be aligned to the used data structure to apply this ML model with few changes. Otherwise, customisation would be required either in their dataset or in the naming convention of the codes to make it up and running. However, most of the hotels have all these mandatory columns in their procurement system as they form the basic information required by any hospitality procurement system, and hence, the proposed framework can be generalised to all hospitality organisations.

As for the AI integrated conceptual framework, the implementation and execution would be multifaceted. The framework seems to be quite applicable across the hospitality industry as it uses state-of-the-art technology to optimise and provide cognitive analytics for human-computer interaction. The data analytics maturity model can be used to evaluate the current stage of any hotel and adapt accordingly. This framework will take the industry to the next level of the digital revolution where real-time information simulates immediate alternate decision paths for ideal results. In spite of the fact that the framework is adaptable, it would be expensive for SME.

9.5 Summary of Survey

A mixed survey was undertaken with 10 questions. The survey was constructed using the SurveyMonkey website and the link was forwarded to people working in the hospitality industry who are directly or indirectly related to the procurement process. The survey included scaler metrics for the participants to choose from stars depending on the importance of added features in the procurement process. The base of the AI integrated framework and 10 agents are already published in the scientific world. Hence, the survey included questions related to the additional features added. A consent form was given to all participants guaranteeing confidentiality and assuring that the data will be analysed collectively. The consent form and the survey can be found in Appendix V.

A total of 75 participants responded to the survey. The survey had three major sections with a total of 10 questions. The three sections in the survey are:

I. Demographic details

II. Importance of added features in the integrated system

III. Knowledge and recommendation

Four questions were demographic questions asking for the respondents’ email, the industry they work for, company size, and department in which they work. Out of 75, two refused to provide their email address and, due to confidentiality agreement, the identity is not revealed. Further details of the survey areas explored below.

All participants work for the hospitality industry. All participants are directly or indirectly involved in procurement processes except one. These details are important as the proposed integrated system is intended for the procurement system in the hospitality industry. The opinions of participants from this industry can be considered more valid than any other industry.

The company size varied for each participant. Half of the respondents work in large enterprises, which have more than 600 employees. The company size matters in adoption and perception towards a new system as it may not be cost-effective to invest in enormous changes for small medium enterprises. Participant responses to the company size question are shown in Figure 35. Majority of the respondents work in large enterprises.

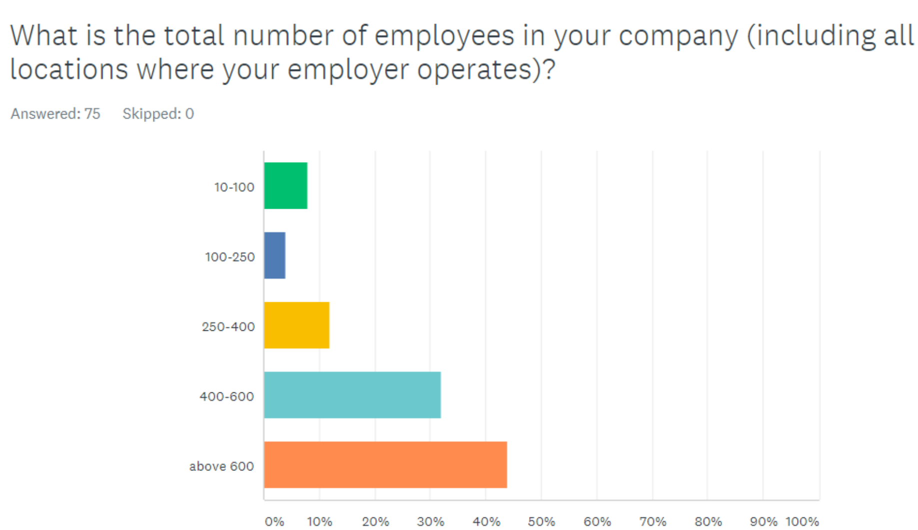


Fig. 35: Response from participants for question 4

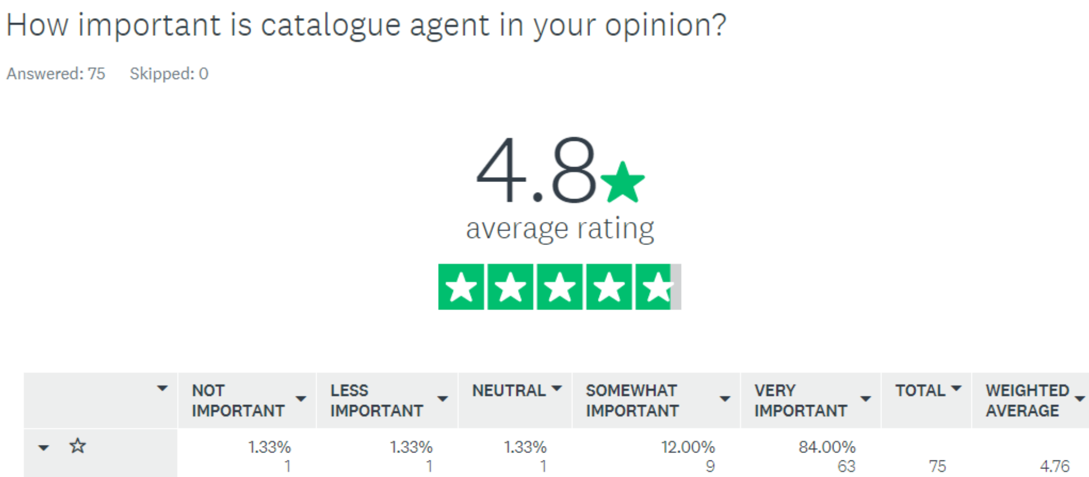
Section two had four questions, each one evaluating the importance of additionally added agents in the AI integrated system. In the survey, each agent with its definition was given to the participants. Moreover, the researcher's telephone number and email address were also provided so that respondents could seek clarification if necessary. These questions collected the participants' inputs on the importance and a textbox to write their comments were also provided. Each of the obtained results is shown below.

Authorisation and security agents received an average of 5 ratings out of 5 in the survey. Ninety-seven percent of the participants said the feature is very important, as depicted in Figure 36. This shows how important it is to have such an agent. The participants' comments include that this agent will have a financial impact and it is important to have high security and ensure privacy for users in the procurement process.

	NOT IMPORTANT	LESS IMPORTANT	NEUTRAL	SOMEWHAT IMPORTANT	VERY IMPORTANT	TOTAL	WEIGHTED AVERAGE
☆	0.00% 0	0.00% 0	0.00% 0	2.67% 2	97.33% 73	75	4.97

Fig. 36: Response from participants for question 5

The catalogue agent received an average of 4.8 ratings out of 5. Eighty-six percent of the participants said this feature is very important, as shown in Figure 37. Participants find it useful and important. The participants' comments include that the catalogue agent will ensure that they are ordering the correct product from the right supplier and notifying the concerned parties, which saves a lot of time.

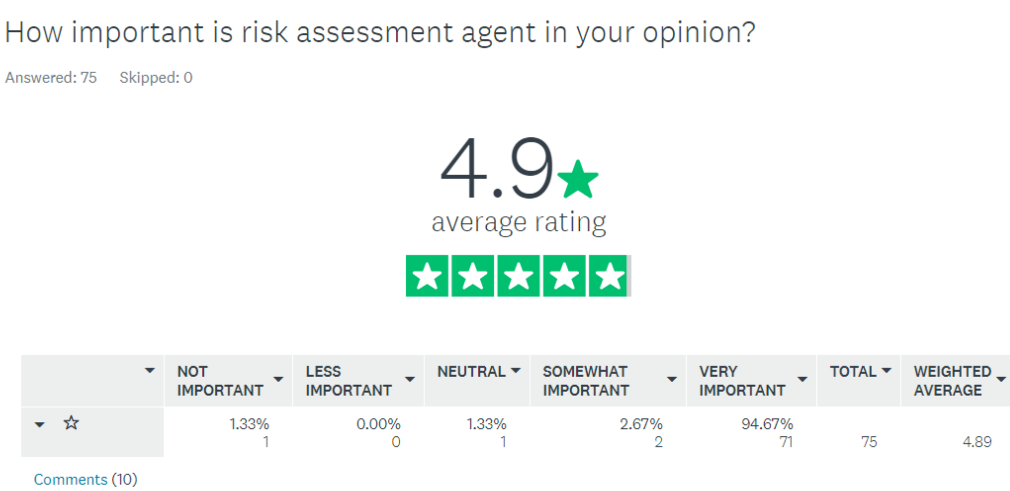

How important is catalogue agent in your opinion?

Answered: 75 Skipped: 0

	NOT IMPORTANT	LESS IMPORTANT	NEUTRAL	SOMEWHAT IMPORTANT	VERY IMPORTANT	TOTAL	WEIGHTED AVERAGE
☆	1.33% 1	1.33% 1	1.33% 1	12.00% 9	84.00% 63	75	4.76

Fig. 37: Response from participants for question 6

The risk assessment agent received an average of 4.9 ratings out of 5 from the survey. Ninety-five percent of the participants said this feature is very important, as displayed in Figure 38. The participants' comments include monitoring forex changes and ensuring the company's interests are protected; moreover, risk management has always been an issue. Therefore, it is always beneficial to have an agent to monitor these elements.


How important is risk assessment agent in your opinion?

Answered: 75 Skipped: 0

	NOT IMPORTANT	LESS IMPORTANT	NEUTRAL	SOMEWHAT IMPORTANT	VERY IMPORTANT	TOTAL	WEIGHTED AVERAGE
☆	1.33% 1	0.00% 0	1.33% 1	2.67% 2	94.67% 71	75	4.89

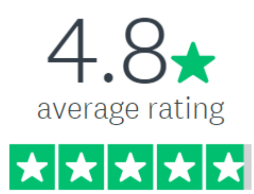
Comments (10)

Fig. 38: Response from participants for question 7

The optimisation agent received an average of 4.8 ratings out of 5 from participants. Ninety percent of participants said this feature is very important, as depicted in Figure 39. The participants' comments include that this element is a great addition that helps in making optimised decisions and is also beneficial as it may reduce the time taken for one workflow.

How important is optimization agent in your opinion?

Answered: 75 Skipped: 0



	NOT IMPORTANT	LESS IMPORTANT	NEUTRAL	SOMEWHAT IMPORTANT	VERY IMPORTANT	TOTAL	WEIGHTED AVERAGE
☆	1.33% 1	0.00% 0	2.67% 2	6.67% 5	89.33% 67	75	4.83

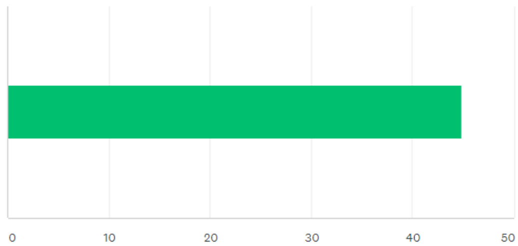
Fig. 39: Response from participants for question 8

The next two questions were to presented to acquire an understanding of the participants' expertise in AI and if they would recommend the proposed integrated system to their top management. These questions are important because the researcher is keen to know the opinion that they have provided is based on their AI knowledge or if it is purely experience-based. It is also important to know if they are willing to adopt such a system which can become a motivational factor for the system development team.

The average expertise of participants was 45%, as shown in Figure 40. This implies the participants' expertise is less than average. It is interesting to notice that even if their AI expertise level is not high, they all found these agents to be beneficial for their day-to-day work.

Concerning technical work in artificial intelligence, how would you describe your own expertise?

Answered: 75   Skipped: 0



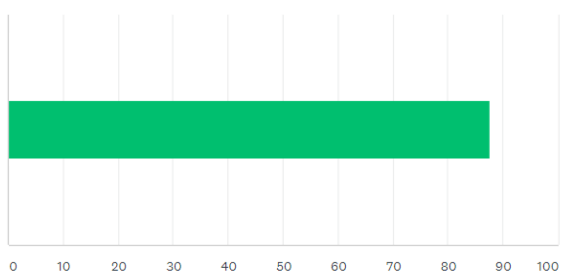
ANSWER CHOICES	AVERAGE NUMBER	TOTAL NUMBER	RESPONSES
Responses	45	3,362	75

Fig. 40: Response from participants for question 9

Eighty-eight percent of the participants confirmed that they will definitely recommend this AI integrated system to their company, as shown in Figure 41. It is interesting to note that although the participants are not experts in AI, they still would recommend the integrated system to their top management.

How likely are you to recommend such an integrated system with the agents to your company?

Answered: 75   Skipped: 0



ANSWER CHOICES	AVERAGE NUMBER	TOTAL NUMBER	RESPONSES
Responses	88	6,577	75

Fig. 41: Response from participants for question 10

The survey results show that the added agents are beneficial and add value to the business process as per the participant's opinion. The extreme responses maybe because none of the participants are experts in this field. This again shows that there is a need for skilled personnel in the selected organisation.

9.6 Statistical Analysis of the Survey

A further statistical analysis is conducted on the survey results using SPSS. For the analysis a value of 5 is given for 'very important', 4 for 'somewhat important', 3 for 'neutral', 2 for 'less important' and 1 for 'not important' responses.

a) Means and Standard deviation for each responses

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Q5	75	4.0	5.0	4.973	.1622
Q6	75	1.0	5.0	4.760	.6746
Q7	75	1.0	5.0	4.893	.5346
Q8	75	1.0	5.0	4.827	.6012
Valid N (listwise)	75				

Fig 42: Descriptive statistics on survey responses

The values clearly show that for questions 5 to 8 the means are very high and close to 5 denoting the added agents are very important for most participants. The descriptive statistics are displayed in figure 42.

b) Demographic distribution

Since surveymonkey allows only 10 questions, the author has considered using only very relevant questions.

- All respondents are from hospitality industry.

- All respondents are directly or indirectly related to procurement processes
- Variations in the size of organization is given below in figure 43

Q4

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	10-100	6	8.0	8.0	8.0
	100-250	3	4.0	4.0	12.0
	250-400	9	12.0	12.0	24.0
	400-600	24	32.0	32.0	56.0
	above 600	33	44.0	44.0	100.0
	Total	75	100.0	100.0	

Fig 43: Frequency distribution of organizational size

Forty-four percent of the participants worked in large enterprises. Majority of participants worked in medium to large enterprises. Here, as per the researcher any number above 400 is considered medium to large enterprise.

c) Correlation between responses

The 4 main questions are independent of each other and the author just wanted to explore any relation between these variables.

Correlations

		Q5	Q6	Q7	Q8
Q5	Pearson Correlation	1	.188	.123	.229*
	Sig. (2-tailed)		.107	.295	.048
	N	75	75	75	75
Q6	Pearson Correlation	.188	1	-.034	.329**
	Sig. (2-tailed)	.107		.769	.004
	N	75	75	75	75
Q7	Pearson Correlation	.123	-.034	1	.741**
	Sig. (2-tailed)	.295	.769		.000
	N	75	75	75	75
Q8	Pearson Correlation	.229*	.329**	.741**	1
	Sig. (2-tailed)	.048	.004	.000	
	N	75	75	75	75

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Fig 44: Correlation between responses

There is no much correlation between the responses except among question 7 and 8. The correlation matrix is shown in figure 44. Correlation between risk assessment agent and optimization agent is strong positive and that means participants have similar opinions about these agents.

Overall, an average of 92% of participants emphasised the importance of having the new agents and the standard deviation of the participants' views on agents is 4.3%. This clearly shows that the variation among the perception of people working in a procurement-related job is small and it is more consistent.

9.7 Validate the qualitative study

The qualitative research is a powerful method of study for the implicit and the explicit. The study depends on the personal perception and experiences gathered by the seasoned expert. In this study, the qualitative analysis is validated using the triangulation method. Member checking is also completed by transcribing interview answers for the interviewee to review to ensure that the original intent and meaning have been captured. The primary study in the research is evaluated through an interview, which is further validated by a survey. Furthermore, these findings are then evaluated using document analysis. The reference for each finding is given below.

The interview conducted with the subject matter professional is considered valid and reliable as he is a seasoned expert in the field and also strives to advance e-procurement. The rewritten interview response was given for the interviewee to review. Even then a further survey with team members was conducted to validate if their opinions match each other for triangulation purposes. Furthermore, these results are then cross-checked with document analysis. The summary of the interview is given below.

As stated by Bendoly (2013), contextual information can be used to calculate the best price from vendors to gain long-term profit for those parties involved and is achieved by integrating ML into the e-procurement framework. The same comment is specified by the interviewee. Buhalis and Leung (2018) and Raschka, Julian, and Hearty (2016) specified the importance of having a piece of accurate information rather than an estimate. Therefore, having an integrated system with AI and ML will ensure the provision of more accurate information as confirmed by the subject matter expert.

Russom (2011), Chen et al. (2012), and Wang et al. (2016) have explained the importance of business analytics importance of having skills, technologies, and practices to obtain insight and guide business planning through predictive and advanced analytics for forecasting and optimisation techniques. Blackburn et al. (2015) have specified the challenges in demand forecasting the importance of demand forecasting in the supply chain. The studies by Blackburn et al. (2015) and Perner (2008) confirm the importance of using historical data for predicting and forecasting purposes. Buhalia and Leung (2018) have proposed a framework that the author has further enhanced to include state-of-the-art technologies and techniques. The interviewee confirmed the same statement in his interview when he answered question 4 and 5. Nenad Stefanovic (2019) and Roeshartono Roespinoedji et al. explain how effective predictive and advanced analytics is in optimising decision-making in the supply chain system. The interviewee agreed with this in his response to question 6.

9.8 Further Discussions

The use of the data analytics maturity model will streamline and automate procurement processes with other operations of the hospitality industry. Besides, being ahead of the maturity model will replace humans with machines and result in employment reduction. The use of data analytics and integrating AI in e-procurement will create new jobs. Hence, the evolving analytics culture that involves deep learning will relieve more employees from tedious, routine, and repetitive activities, that in turn permits them to address new creative challenges.

The prices for commodities bought from the same vendor at different properties showed variations at times and was a constant challenge. The only way to make sure all properties have a standard rate for each commodity from a single vendor is to have a single system that integrates across all regions. All properties will maintain a similar amount of information regarding the purchases and vendors. Any addition to the information will affect all properties at the same time. Thus, all properties of The X Hotels can maintain consistent data that is transparent and easy to handle by having an integrated centralised system. There are additional incremental costs due to urgency that has a direct impact on the bottom line. The system should be able to predict the duration of the next purchase and augment the stakeholders with the list of products they need to order with the best alternatives and avoid last-minute emergency hassles.

More accurate predictions derived from demand forecasting will help to build a better partnership with key suppliers. This will help The X Hotels to perform more accurate cash flow analysis. Demand forecasting will help the department plan their resource availability.

It will also help The X Hotels to maintain optimal inventory. Furthermore, analysing more accurate supplier lead times and incorporating that information into demand forecasting can improve the process of strategic sourcing. Top management will be naturally confident in the judgment of a subject matter expert, but it is not reliable in all cases. Thereby, predictive and prescriptive analytics will surely supplement and frame a stronger foundation for effective decision-making.

AI enables better communication between a computer and human, as well as between machines, which enhances interoperability. AI will computerise the accumulation and consolidation of data from different origins. IoT is used for different functions, such as data acquisition, data transmission, and analytics, or actuates other IoT objects. No in-depth study on IoT is conducted in this research but the reality is that AI cannot stand alone without IoT support for optimum performance. While transporting foodstuffs from the farm to the warehouse or distribution centres or the customer, the items can be tagged using barcodes, radio frequency identification (RFID) tags, or near field communication (NFC) tags. The type of foodstuff, the farm of origin, and any other data that are mandated by local, regional, or global regulations and end-user requirements, can be identified with the data associated with each tag. As the tagged items move through the different components of the food supply chain, the related data points can be automatically verified, ensuring that a complete audit trail of handling is created and maintained. Temperature, humidity, location, power supply, and gas levels of the vehicles can all be continuously monitored using the IoT devices and those data points can be automatically compared to verify that the items are within the compliance limits and have been transported safely. Supplier smart contracts can be updated with real-time data and this information can automatically provide better insight into supplier compliance, performance or rating.

Hotels with a similar dataset or procurement system will need to be aligned to the used data structure, to apply this ML model with few changes. Generalisation of AI integrated framework and ML implementation are possible; what matters is that the ecosystem is ready for the change. Hence, the proposed integrated system is extremely useful as it is outcome-driven.

9.9 Limitations of Deploying the Proposed Framework

At present, the majority of the organizations are embracing e-procurement as it has turned into a status line in the market to adjust to the most recent technological advancement. There are a few frameworks accessible for it however one must be watchful in picking the correct programming for their association. One approach is to redistribute the improvement of e-procurement with the goal that the organisation would get a customised system that serves everything required. Few challenges in adoption of AI in e-procurement include:

a) The supplier whom the X company purchase may not have adopted the same system.
b) The most frequent supplier may not have the same smart systems adopted.
c) Supplier rating criteria is done using practical know-how and common sense.
d) Data entry may become a very tedious and time-consuming task.
e) The spend forecasting uses home currency whereas for quantity forecasting the UOM (unit of measure) of the purchased item may need to convert to the base unit before using it in forecasting.

To obtain the correct data for this intelligent system would be challenging. To make the implicit information to express data would be the task requiring the greatest effort. Just in case The X Hotel's integrated providers are not utilizing the integrated framework, the sub-procedure of choosing providers would turn out to be pointless. There ought to be progressively solid and enormous effort required to make the acquiring subtleties significant and helpful for the related stakeholder to achieve the required results. The framework should demonstrate its productivity at a time when, generally speaking, the best direction with the stakeholders is not going to be in line with the framework. The hurdle of the framework is, for the most part, difficulties in initial acceptance of the intelligent e-procurement system by the stakeholders.

9.10 Chapter Summary

This chapter has proposed a new framework with its advantages as well. The proposed AI Integrated framework empowers internal and external applications and collaborations of a large amount of data exchange from the cloud. For effectiveness, the recommendation is to add the IoT mentioned in Chapter 2 to the recommended locations. Standardization of communication protocols agreed among stakeholders across the Ecosystem, will increase interoperability, interconnectivity, and thus, trustworthiness. Thereby, the key stakeholders or decision-makers can make use of the decision support system and use management software for a variety of cases to improve their business and strategic planning. Furthermore, in this chapter qualitative analysis is done in the research is discussed. The results obtained are validated using triangulation. Member checking is also done to check the accuracy in converting the interview. The survey results show the stakeholders find the model useful and timesaving.

Chapter 10: LIMITATIONS AND RECOMMENDATIONS

10.1 Limitations

The major limitation is generalization as the research limits analysis to one organization, even if large. Although the research was conducted in the chain hotel in the UAE, the scope can be widened to other regions for the same chain without any changes in the proposal. The proposed framework and optimisation of e-procurement data are generalisable to the hospitality industry only. The requirements, functionality, and the data itself would be different for other industries and hence, there is no enough evidence to confirm the generalizability to other industries.

Another limitation of my study is not getting approval to conduct a survey(s)/interview in the organisation to evaluate the current e-procurement system. The consent was to access the receiving detailed report (only) in the e-procurement system and hence the predictors chosen are limited to the available data. To get a complete overview of the actual performance of the current e-procurement system is essential to propose a framework that completely overcomes the issues. The adoption level of AI couldn’t be evaluated at the X hotels.

The non-disclosure form agreed was only for accessing e-procurement data and did not get approval to seek any other data from the organisation. The next limitation is the limited access to data within the organization. To conduct thorough data analytics few more data values were essential. For example, information on events, occupancy rates for five years, consumption rates, number of rooms details, cost per room, and so on will add value to the study. This only shows that there is a lot more ML application development possible in the e-procurement of the hospitality industry. A wide selection of data parameters is vital for the data analysis process, especially for possible predictions.

10.2 Recommendation for Implementing the AI Integrated Framework

An organization can deploy the AI integrated framework depending on which stage the organization is in the maturity level. An organisation should do self-assessment to understand its level in the maturity model. Several organizations are either in the infected stage or in the acceleration stage. An organisation shall follow the steps given below unless the organization is ahead of the maturity level. The first four steps of evaluation steps are adapted from Grossman (2018). The framework introduced in this research which is based on data analytics maturity models can be evaluated by following the steps given:

1. Analytics models: The models should be first developed with the input that includes data and appropriate business requirements. A deadline should be defined for data collection to ensure that the roll-out happens within the stipulated project timeline. Likewise, make data collection stakeholders take ownership of their accomplishments to ensure that they input accurate and more relevant data. The data collection phase could be monotonous and exhausting, but the stakeholders have to ensure that there is enough personnel available to complete the task in line with the agreed timeline.
2. Analytic Infrastructure: the organization needs to make sure it has the required software components, software applications, and platform for managing data, processing data, producing analytical models and using it to generate business value.
3. Analytic operations: An organisation has to then make sure that the results of analytic models are integrating into an organisation's product, services, and operations to take

advantage of decision-making to take relevant actions that will decrease cost and improve operations.

4. Analytic governance: It is a repeatable process until analytic governance is achieved among all stakeholders in the ecosystem. A steady and thorough methodology is required until a definitive goal is accomplished. There might be additional intelligent devices that are integrated within the network to retrieve real-time information, the accuracy of integration of such communications among systems or individuals, is important to deliver learning and integrity into the data governance. Hence the system would contribute to the strategic advantage of all stakeholders and the results would show a positive impact across the entire ecosystem.
5. Analytic security: An organisation has to make sure that its implementation follows data security and privacy procedures to reduce risk and protect analytic assets.
6. Analytic strategy: An organisation should make sure that its strategy includes short and long-term requirements and opportunities. Furthermore, an assessment criterion or user acceptance testing ought to be agreed to understand system errors to track corrective action. The system should be subject to progressive improvement that should be achievable using online studies or on-site brainstorming meetings to generate new ideas until the framework ends up proficient and viable for the organisation.
7. Analytic training: There should be a training team to prepare the employees and to encourage them to utilise the implementation and use system. The training team ought to have functional knowledge and be responsible for applying change management

by finding what is turning out badly, why it is not proficient and effective, and so forth, and give important feedback for the enduring issues.

8. Analytics expectation: The most anticipated outcome by having an AI and ML integrated framework would be, that the system change would show immediate enhancement to the effective improvement in business processes and contribute immensely towards timesaving. Accordingly, the framework ought to demonstrate subjectively and quantitatively towards the system enhancement. The final system integration and implementation shall be collective and a collaborative decision between the company and its integrated suppliers to get the best of the system.

10.3 Validity, Reliability, and Generalisability

The predictions done in the findings accurately reflect the actual data. As such, the quantitative findings can be considered valid. At the same time, the qualitative study is validated by triangulation.

The results are consistent in the analytical procedures, including the qualitative and quantitative study. Hence, the study is reliable.

The transferability of the findings to other settings is confirmed by the subject matter expert and applicability, to any hospitality industry that wants to adopt data analytics maturity whereas it may not be transferable in other industries as the entire functionality and data set is different. Hence, the study is generalisable to the hospitality industry only. Having said that if anyone would want to get the roadmap in integrating AI into an industry, without using the examples given on data analytics, can definitely use the extended framework.

The economic factors included in the implementation is not considered in the study. The cost effectiveness would be for medium to large enterprises.

10.4 Chapter Summary

In this chapter, the researcher has given various stages of deployment for the AI integrated model and the predicting algorithm as recommendations and limitations. Furthermore, validity, reliability, and generalisability are established with the limitations of the study.

Chapter 11: CONCLUSION AND FUTURE WORKS

11.1 Conclusions

In this thesis, the researcher has proposed an extended framework by integrating AI and apply ML with deep learning techniques to optimise e-procurement in the hospitality industry. The author has used LSTM for building the ML model for demand forecasting. The LSTM model can be considered the best for time series prediction with various time lags. Furthermore, the author has used various ML classification models from sklearn kit for finding the accuracy of predicting the duration of purchase. Thus, two deep learning algorithm is developed for e-procurement data in the hospitality industry for prediction and forecasting purposes. The major findings in this research can be outlined as follows:

RQ 1) How can the state-of-the-art technologies be part of the contemporary e-procurement system within the hospitality industry?

E-procurement focused AI integrated framework is developed for the hospitality industry. This novel framework consists of all characteristics of an enterprise from external to internal factors. The framework that is developed in compliance with the data analytics maturity model meets the goals narrated below. The extended framework has an additional layer for agent-based technology and data & analytics governance is implemented.

The authorisation and security agent integrated into the framework provides security and compliance for e-procurement data assets. The risk assessment agent provides procurement strategies and opportunities for all stakeholders in the ecosystem. The optimisation agent enabled with the analytic model using ML and deep learning which can efficiently perform demand forecasting. The catalogue agent avoids duplication of information about the supplier, buyer, or their action. The framework operates an analytic governance structure by having a standardised and centralised system accessible by all stakeholders of the X hotel. The system manages and operates analytic infrastructure by enabling interoperability and interconnectivity among its components in the ecosystem. The integrated framework will have minimal to no man-made errors in data entry as all entries are mapped and will not require humans to key in any information for procurement.

RQ 2) How can the data analytics maturity model be adapted in e-procurement data of the hospitality industry?

A large number of scenarios were tested in different environments and each component that influences the e-procurement data was analysed by adopting the data analytics maturity model. The findings related to data analytics are given below. Other benefits of following data analytics maturity model is answered with other research questions.

Descriptive analytics provided five reports useful for top managers to know what has happened in the past months/years. Diagnostic analytics provided five visual reports that will enable the top management to take decisions in reducing the risk and prioritise workload. Predictive analytics is used to predict the purchase duration for a given commodity for a particular property by using ML techniques. K fold cross-validation is used to choose the best model with the best accuracy. Out of the seven models used for predicting purchase duration except for Dtree and LR, all other models found to be good models showed the model is good for predicting, especially the Naïve Bayes model, which had almost 83%

accuracy. Prescriptive analytics is done for demand forecasted on e-procurement data. Deep learning technique, LSTM is used for time series demand forecasting. LSTM proved to be a good algorithm for forecasting the spend and quantity of e-procurement data. Based on the results from the prescriptive analysis, insight into cognitive analytics is also provided.

RQ 3) What is the perception of subject matter expert(s) on the proposed integrated system with AI and ML in e-procurement of the hospitality industry?

The in-depth interview with subject matter expert confirmed the same with a detailed remark. As per the domain expert, the integrated system is outcome-driven with high usability. A survey was conducted with people in hospitality who are directly or indirectly related to procurement processes on the proposed framework to get their perceptions. The survey results show that staffs have less than average knowledge of AI. The expertise level of staff's averaged 45% only. Eighty-eight percent of participants said they will recommend the proposed framework to top management. Ninety-two percent of participants confirmed the added agents are very important for an e-procurement system and the variation was only 4.3%.

11.2 Impact of the Contributions

This research is a roadmap of how AI and ML can be integrated with the e-procurement system which can support decision-making and provides a blueprint to stakeholders on what needs to be monitored for such a change in the organisation. The amount of literature on the data analytics maturity model being evaluated in the e-procurement model of the hospitality industry is very scanty. This research focused on data from the hospitality industry would be a unique contribution to the scientific world.

The outcomes of implementing AI and ML integrated system may include but not limited to:

1) Exploit the given situation and improve business objectives by working together closely with the stakeholders.
2) Mitigate the happening of mishaps by prescriptive and cognitive analytics.
3) Make various reports or analyses based not just monthly or quarterly but also yearly.
4) Analyse the previous decisions made based on historical data and support in improving future decisions.
5) Budget forecast conducted using supervised machine learning can enhance planning and sourcing and hence optimize resources in procurement.
6) Demand forecasting can be used to ensure that all stakeholders are happy and getting the optimum results in their field of operation.
7) Demand forecasting using LSTM can perform a more accurate cash flow analysis.
8) The proposed integrated system can be implemented in any hotel(s) irrespective of size, but cost-effectiveness would be for medium to big enterprises.

Companies claim that integrated AI, especially for decision augmentation, has a high impact on better decision-making. Further, routine enhancement process to the proposed system should include:

1) Identifying more initiatives where this integrated system can explore more challenges and improve business objectives by working together closely with the stakeholders.

2) By simulating real-life situations, improve data usage in analytics, and continuously enhance system intelligence.
3) Where unexpected results are delivered by AI, evaluate them using human intervention, and monitor whether the predicting model is forecasted well.
4) Add more feature variables like the number of rooms in a property, seasonal value, consumption, holidays, wastage, etc. in the model to do various predictions with improved accuracy.

11.3 Further Works

Future works in the field of AI and ML integration in e-procurement of hospitality industry can be furthered by including the following technologies mentioned in International Research Organisation's Quantitative Analytics and Data Science team (2020) but not limited to:

a) Natural Language Processing: Currently the data is being captured into the system using data entry personnel. Enhancing the system to be interactive using NLP, can exponentially improve system efficiency and effectiveness. For example, the buyers currently search for catalogues or items by entering it within the application. NLP will make this efficient, by accepting voice commands from the Buyer to identify the requirement and respond accordingly.
b) Computer Vision: Currently this technology is rarely used across the hospitality industry. The usage of this technology could help make the processes simpler and faster. For example, a buyer currently searches for items by typing the name of the product in the system. With combined NLP and Computer Vision technology, the

system would acquire the details of the item to be ordered automatically as soon as the buyer dictates the name or shows the item to the application.

11.4 Chapter Summary

The author has constructed five reports in descriptive analysis and diagnostic analysis data visualisation. Furthermore, the author has constructed two very useful predictive analytics model that helps in prescriptive analysis of big data from The X Hotels. Moreover, the author has tried to put some light on cognitive analytics as well. The rich data collected at The X Hotels shows the flourishing or well-established business for them in the region of UAE. The X Hotels has thousands of properties across the world and there is a vast scope to study on the data being accumulated. Likewise, the proposed new conceptual framework is to be established and yet to be validated for its productivity and optimisation in real-time hospitality operations.

ML together with deep learning has good scope in e-procurement of the hospitality industry as the data has not undergone any exploratory data analytics study. LSTM has given good forecasting results and can be modelled to forecast occupancy, wastage of food, consumption rates, and so on. provided the organisation is willing to share the data with skilled people for ML. Thus, the integration of AI and ML in e-procurement will revolutionise the hospitality industry.

REFERENCES

Acar, Y. & Gardner, E. (2012). Forecasting method selection in a global supply chain. *International Journal of Forecasting*, vol. 28 (4), pp. 842-848.

Aital, P., Pawar, P. & Behl, A. (2017). Systematic literature review on electronic reverse auction: issues and research discussion. *International Journal of Procurement Management*, vol. 10 (3), p. 290.

Aithal, P. & Aithal, S. (2020). "Study of Various General-Purpose Technologies and Their Comparison Towards Developing Sustainable Society". *Papers.ssrn.com* [online]. [Accessed 16 March 2020]. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3238762

Alam, T. & Benaida, M. (2020). Blockchain, Fog and IoT Integrated Framework: Review, Architecture and Evaluation. *Internet of Things*, vol. 62 (2), pp. 1-13.

Alonzo, L., Chioson, F., Co, H., Bugtai, N. & Baldovino, R. (2018). A Machine Learning Approach for Coconut Sugar Quality Assessment and Prediction. *Engineering Research and Development for Technology (ERDT) of the Department of Science and Technology (DOST)*.

Alsaadi, E. & Tubaishat, A. (2020). "Internet of Things : Features , Challenges , and Vulnerabilities Authors". *Journal of advanced computer science and technology*

Alsetoohy, O. & Ayoun, B. (2018). Intelligent agent technology:The relationships with hotel food procurement practices and performance. *Journal of Hospitality and Tourism Technology*, vol. 9 (1), pp. 109-124.

Alsetoohy, O., Ayoun, B., Arous, S., Megahed, F. & Nabil, G. (2019). Intelligent agent technology: what affects its adoption in hotel food supply chain management?. *Journal of Hospitality and Tourism Technology*, vol. 10 (3), pp. 286-310.

Altin, M., Koseoglu, M., Yu, X. & Riasi, A. (2018). Performance measurement and management research in the hospitality and tourism industry. *International Journal of Contemporary Hospitality Management*, vol. 30 (2), pp. 1172-1189. [Accessed 27 January 2020].

Alvarez, M. & Campo, S. (2014). The influence of political conflicts on country image and intention to visit: A study of Israel's image. Tourism Management, vol. 40, pp. 70-78.

Ampazis, N. (2015). Forecasting Demand in Supply Chain Using Machine Learning Algorithms. International Journal of Artificial Life Research, vol. 5 (1), pp. 56-73.

(2020). Study.com [online]. [Accessed 15 May 2020]. Available at: https://study.com/academy/lesson/hospitality-industry-history-origin.html

Anthony Jnr, B. (2019). A developed software agent-knowledge-assisted procurement management tool for retailing enterprise. *VINE Journal of Information and Knowledge Management Systems*, vol. 49 (1), pp. 54-75.

"Artificial Intelligence Trends: Decision Augmentation". (2020). [Accessed 5 May 2020]. Available at: https://www.International Research Organisation.com/en/documents/3979508/artificial-intelligence-trends-decision-augmentation

Babazadeh, R., Razmi, J., Rabbani, M. & Pishvaee, M. (2017). An integrated data envelopment analysis–mathematical programming approach to strategic biodiesel supply chain network design problem. *Journal of Cleaner Production*, vol. 147, pp. 694-707.

Bendoly, E. (2013). Real-time feedback and booking behavior in the hospitality industry: Moderating the balance between imperfect judgment and imperfect prescription. *Journal of Operations Management*, vol. 31 (1-2), pp. 62-71.

Bharat, A., N, P. & Reddy, R. (2018). Using Machine Learning algorithms for breast cancer risk prediction and diagnosis. *IEEE Third International Conference on Circuits, Control, Communication and Computing*, vol. 83, pp. 1-4.

Blackburn, R., Lurz, K., Priese, B., Göb, R. & Darkow, I. (2015). A predictive analytics approach for demand forecasting in the process industry. *International Transactions in Operational Research*, vol. 22 (3), pp. 407-428.

Bogetić, S., Đorđević, D., Ćoćkalo, D. & Bešić, C. (2017). The Analysis Of Quality Aspects In Development Of Competitiveness Of Domestic Hotel Enterprises. *International Journal "Advanced Quality"*, vol. 44 (2), p. 1.

Boone, C., Hazen, B., Skipper, J. & Overstreet, R. (2016). A framework for investigating optimization of service parts performance with big data. *Annals of Operations Research*, vol. 270 (1-2), pp. 65-74.

Brandon-Jones, A. & Carey, S. (2015). The impact of user-perceived e-procurement quality on system and contract compliance. *International Journal of Operations & Production Management*, vol. 31 (3), pp. 274-296.

Brandon-Jones, A. & Kauppi, K. (2018). Examining the antecedents of the technology acceptance model within e-procurement. *International Journal of Operations & Production Management*, vol. 38 (1), pp. 22-42.

Bünger, M. (2017). "Blockchain for industrial enterprises: Hype, reality, obstacles and outlook - IoT Agenda". *Internetofthingsagenda.techtarget.com* [online]. Available at: https://internetofthingsagenda.techtarget.com/blog/IoT-Agenda/Blockchain-for-industrial-enterprises-Hype-reality-obstacles-and-outlook

Buhalis, D. & Foerste, M. (2015). SoCoMo marketing for travel and tourism: Empowering co-creation of value. *Journal of Destination Marketing & Management*, vol. 4 (3), pp. 151-161.

Buhalis, D. & Leung, R. (2018). Smart hospitality—Interconnectivity and interoperability towards an ecosystem. *International Journal of Hospitality Management*, vol. 71, pp. 41-50.

Calderon, T. & Onita, C. (2020). "Big Data and the Perceived Expectations Gap in Digital Authentication Processes". *SJSU ScholarWorks* [online]. [Accessed 16 March 2020]. Available at: https://scholarworks.sjsu.edu/acc_fin_pub/65/

Casey, M. & Wong, P. (2020). "Global Supply Chains Are About to Get Better, Thanks to Blockchain". *Harvard Business Review* [online]. [Accessed 8 May 2020]. Available at: https://hbr.org/2017/03/global-supply-chains-are-about-to-get-better-thanks-to-blockchain

Chai, J., Liu, J. & Ngai, E. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications*, vol. 40 (10), pp. 3872-3885.

Chang, L., F. Backman, K. & Chih Huang, Y. (2014). Creative tourism: a preliminary examination of creative tourists' motivation, experience, perceived value and revisit intention. *International Journal of Culture, Tourism and Hospitality Research*, vol. 8 (4), pp. 401-419.

Cheung, C., Wang, W., Lo, V. & Lee, W. (2012). An agent-oriented and knowledge-based system for strategic e-procurement. *Expert Systems*, vol. 21 (1), pp. 11-21.

Cheng M.H., Sedaghat R., Siddavaatam P. (2016) A New Adaptable Construction of Modulo Addition with Scalable Security for Stream Ciphers. In: Chen J., Piuri V., Su C., Yung M. (eds) Network and System Security. NSS 2016. Lecture Notes in Computer Science, vol 9955. Springer, Cham

Chuah, S., Rauschnabel, P., Krey, N., Nguyen, B., Ramayah, T. & Lade, S. (2016). Wearable technologies: The role of usefulness and visibility in smartwatch adoption. *Computers in Human Behavior*, vol. 65, pp. 276-284.

Chui, M. & Malhotra, S. (2020). "AI adoption advances, but foundational barriers remain". *McKinsey & Company* [online]. Available at: https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain

Claveria, O., Monte, E. & Torra, S. (2015). A new forecasting approach for the hospitality industry. *International Journal of Contemporary Hospitality Management*, vol. 27 (7), pp. 1520-1538.

Collins, G., Wynants, L., Timmerman, D., Steyerberg, E. & Van Calster, B. (2019). Predictive analytics in health care: how can we know it works?. *Journal of the American Medical Informatics Association*, vol. 26 (12), pp. 1651-1654.

Creswell, J. & Creswell, J. (2008). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches*. 3rd edn. Lincoln:Sage.

Croom, S., Brandon-Jones, A.: Impact of e-procurement: experiences from implementation in the UK public sector. J. Purchasing Supply Manag. **13**(4), 294–303 (2007)

Daghfous, A. & Belkhodja, O. (2019). Managing Talent Loss in the Procurement Function: Insights from the Hospitality Industry. Sustainability, vol. 11 (23), p. 6800.

De Simoni, G. (2019). "Digital Platforms: The Role Data and Analytics Plays in Their Success". *International Research Organisation Webinars* [online]. Available at: http://www.International Research Organisation.com/webinar/3895194

Deer, M. (2020). "NLP, AI, Deep Learning, and Machine Learning: A Brief Guide". *Towards data science* [online]. [Accessed 9 February 2020]. Available at: https://towardsdatascience.com/nlp-ai-deep-learning-and-machine-learning-a-brief-guide-f225a1f6340b

Delafenestre, R. (2019). New business models in supply chains: a bibliometric study. *International Journal of Retail & Distribution Management*, vol. 47 (12), pp. 1283-1299.

Delen, D. & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems*, vol. 55 (1), pp. 359-363.

Denizci Guillet, B. & Mohammed, I. (2015). Revenue management research in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, vol. 27 (4), pp. 526-560.

Dhaouadi, R., Benmiled, A. & Ghédira, K. (2014). Ontology based Multi Agent System for Improved Procurement Process: Application for the Handicraft Domain. *Procedia Computer Science*, vol. 35, pp. 251-260.

Dzeng, R. & Wang, P. (2017). C-Negotiation Game: An educational game model for construction procurement and negotiation. *Automation in Construction*, vol. 75, pp. 10-21.

Edghiem, F. & Mouzughi, Y. (2018). Knowledge-advanced innovative behaviour: a hospitality service perspective. *International Journal of Contemporary Hospitality Management*, vol. 30 (1), pp. 197-216.

Elliot, B. & Andrews, W. (2020). "A Framework for Applying AI in the Enterprise". *International Research Organisation* [online]. Available at: https://www.International Research Organisation.com/doc/3751363/framework-applying-ai-enterprise

Er-Kara, M., Oktay Fırat, S. & Ghadge, A. (2019). A Data Mining-Based Framework for Supply Chain Risk Management. *SSRN Electronic Journal*.

Fan, T., Chang, X., Gu, C., Yi, J. & Deng, S. (2014). Benefits of RFID technology for reducing inventory shrinkage. *International Journal of Production Economics*, vol. 147, pp. 659-665.

Fang, K., Jiang, Y. & Song, M. (2016). Customer profitability forecasting using Big Data analytics: A case study of the insurance industry. *Computers & Industrial Engineering*, vol. 101, pp. 554-564.

Fosso Wamba, S., Gunasekaran, A., Dubey, R. & Ngai, E. (2018). Big data analytics in operations and supply chain management. *Annals of Operations Research*, vol. 270 (1-2), pp. 1-4.

Giotopoulos, I., Kontolaimou, A., Korra, E. & Tsakanikas, A. (2017). What drives ICT adoption by SMEs? Evidence from a large-scale survey in Greece. *Journal of Business Research*, vol. 81, pp. 60-69.

Gomezelj, D. (2016). A systematic review of research on innovation in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, vol. 28 (3), pp. 516-558.

Goudarzi, F. (2019). Travel Time Prediction: Comparison of Machine Learning Algorithms in a Case Study. *20th International Conference on High Performance Computing and Communications*, vol. 4 (20), p. 1.

Govindan, K., Cheng, T., Mishra, N. & Shukla, N. (2018). Big data analytics and application for logistics and supply chain management. *Transportation Research Part E: Logistics and Transportation Review*, vol. 114, pp. 343-349.

Greasley, A. (n.d.). *Simulating business processes for descriptive, predictive, and prescriptive analytics*. Frankfurt:Walter de Gruyter GmbH & Co.

Grossman, R. (2018). A framework for evaluating the analytic maturity of an organization. International Journal of Information Management, vol. 38 (1), pp. 45-51.

Grover, P. & Kar, A. (2017). Big Data Analytics: A Review on Theoretical Contributions and Tools Used in Literature. *Global Journal of Flexible Systems Management*, vol. 18 (3), pp. 203-229.

Gunasekaran, A., Kumar Tiwari, M., Dubey, R. & Fosso Wamba, S. (2016). Big data and predictive analytics applications in supply chain management. *Computers & Industrial Engineering*, vol. 101, pp. 525-527.

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S., Childe, S., Hazen, B. & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, vol. 70, pp. 308-317.

Hadian, H., Chahardoli, S., Golmohammadi, A. & Mostafaeipour, A. (2019). A practical framework for supplier selection decisions with an application to the automotive sector. *International Journal of Production Research*, vol. 58 (10), pp. 2997-3014.

Hassanien, A., Dale, C. & Clarke, A. (2011). *Hospitality business development*. London [u.a.]:Routledge.

Hazen, B., Boone, C., Ezell, J. & Jones-Farmer, L. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, vol. 154, pp. 72-80.

Hazen, B., Skipper, J., Ezell, J. & Boone, C. (2016). Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering*, vol. 101, pp. 592-598.

Helo, P., Gunasekaran, A. & Rymaszewska, A. (2017). *Designing and Managing Industrial Product-Service Systems*. Cham:Springer International Publishing.

Hemmington, N. (2007). From Service to Experience: Understanding and Defining the Hospitality Business. *The Service Industries Journal*, vol. 27 (6), pp. 747-755.

Hennelly, P., Srai, J., Graham, G. & Fosso Wamba, S. (2019). Rethinking supply chains in the age of digitalization. *Production Planning & Control*, vol. 31 (2-3), pp. 93-95.

Hersent, O., Boswarthick, D. & Elloumi, O. (2011). *The Internet of Things*. Hoboken:John Wiley & Sons.

Higgins, L. (2009). Community music and the welcome. *International Journal of Community Music*, vol. 1 (3), pp. 391-400.

Hofmann, E. (2015). Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. *International Journal of Production Research*, vol. 55 (17), pp. 5108-5126.

Holtgrewe, U., Gasparri, S., Ilsøe, A., Kirov, V., Ramioul, M., Sharma, A., Tassinar, A. & Vereycken, Y. (2020). "Shaping Industrial Relations in a Digitalising Services Industry -

Challenges and Opportunities for Social Partners". *Shaping industrial relations in a digitalising services industry – Service Markets in Europe* [online]. [Accessed 16 March 2020].

Hongdao, Q., Bibi, S., Khan, A., Ardito, L. & Khaskheli, M. (2019). Legal Technologies in Action: The Future of the Legal Market in Light of Disruptive Innovations. *Sustainability*, vol. 11 (4), p. 1015.

"How Predictive Analytics Will Change the Supply Chain of Tomorrow | Transforming Data with Intelligence". (2019). Available at: https://tdwi.org/articles/2018/01/19/adv-all-how-predictive-analytics-will-change-supply-chain-of-tomorrow.aspx

Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z. & Zhang, H. (2019). Deep Learning with Long Short-Term Memory for Time Series Prediction. IEEE Communications Magazine, vol. 57 (6), pp. 114-119.

Hyndman, R. & Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, vol. 22 (4), pp. 679-688.

Incorporated, I. (2017). "Newly Licensed, Independent Pharmacy-Owned Indy Health Insurance To Launch Medicare-D Plan This Fall". *Prnewswire.com* [online]. Available at: https://www.prnewswire.com/news-releases/newly-licensed-independent-pharmacy-owned-indy-health-insurance-to-launch-medicare-d-plan-this-fall-301019362.html

Issabayeva, S., Yesseniyazova, B. & Grega, M. (2019). Electronic Public Procurement: Process and Cybersecurity Issues. NISPAcee Journal of Public Administration and Policy, vol. 12 (2), pp. 61-79.

Ittmann, H. (2015). The impact of big data and business analytics on supply chain management. *Journal of Transport and Supply Chain Management*, vol. 9 (1).

Ivanov, S., Webster, C. & Seyyedi, P. (2020). "Consumers' attitudes towards the introduction of robots in accommodation establishments". *Hrcak.srce.hr* [online]. [Accessed 16 March 2020]. Available at: https://hrcak.srce.hr/206163

James K, S. (2018). "Agent Based Fraud Detection And Reporting In Public E-Procurement | International Journal of Current Research". *Journalcra.com* [online]. Available at: http://journalcra.com/article/agent-based-fraud-detection-and-reporting-public-e-procurement

Jamshidi, D. & Hussin, N. (2016). Forecasting patronage factors of Islamic credit card as a new e-commerce banking service. *Journal of Islamic Marketing*, vol. 7 (4), pp. 378-404.

Jang, Y., Chen, C. & Miao, L. (2019). Last-minute hotel-booking behavior: The impact of time on decision-making. *Journal of Hospitality and Tourism Management*, vol. 38, pp. 49-57.

Jarmin, R. & O'Hara, A. (2016). BIG DATA AND THE TRANSFORMATION OF PUBLIC POLICY ANALYSIS. *Journal of Policy Analysis and Management*, vol. 35 (3), pp. 715-721.

Jin, J., Gubbi, J., Marusic, S. & Palaniswami, M. (2014). An Information Framework for Creating a Smart City Through Internet of Things. *IEEE Internet of Things Journal*, vol. 1 (2), pp. 112-121.

Johnson, M.: A study of e-market adoption barriers in the local government sector. J. Enterp. Inf. Manag. **25**(6), 509–536 (2012)

K, T. (2017). "Supply Chain Management vs. Procurement". *Lean Supply Solutions - Innovative Supply Chain Solutions* [online]. [Accessed 12 March 2020]. Available at: http://www.leansupplysolutions.com/blog/supply-chain-management-vs-procurement/

K, T. (2020). "Supply Chain Management vs. Procurement". *Lean Supply Solutions - Innovative Supply Chain Solutions* [online]. [Accessed 9 February 2020]. Available at: http://www.leansupplysolutions.com/blog/supply-chain-management-vs-procurement/

Kalbaska, N., Janowski, T., Estevez, E. & Cantoni, L. (2017). When digital government matters for tourism: a stakeholder analysis. *Information Technology & Tourism*, vol. 17 (3), pp. 315-333.

KAREN, G. (2020). "Assessing the Fit of Regression Models - The Analysis Factor". *The Analysis Factor* [online]. [Accessed 1 May 2020]. Available at: https://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/

Kaur, H. & Singh, S. (2018). Heuristic modeling for sustainable procurement and logistics in a supply chain using big data. *Computers & Operations Research*, vol. 98, pp. 301-321.

Kazandzhieva, V. & Santana, H. (2020). "E-tourism: Definition, development and conceptual framework". *Hrcak.srce.hr* [online]. [Accessed 16 March 2020]. Available at: https://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=335260

Kim, D. & Kim, S. (2017). The Role of Mobile Technology in Tourism: Patents, Articles, News, and Mobile Tour App Reviews. *Sustainability*, vol. 9 (11), p. 2082.

King, C. (1995). What is hospitality?. *International Journal of Hospitality Management*, vol. 14 (3-4), pp. 219-234.

Kissinger Amayi, F. (2020). "Procurement Oversight Agencies And Growth Of Mses In Trans Nzoia County, In Kenya". *Pdfs.semanticscholar.org* [online]. [Accessed 16 March 2020].

KiKoetsier, J. (2017). "Blockchain Beyond Bitcoin: How Blockchain Will Transform Business in 3 to 5 Years". *Inc.com* [online]. Available at: https://www.inc.com/john-koetsier/how-blockchain-will-transform-business-in-3-to-5-years.html

Kohli, P. & Arora, S. (2019). Application of Machine Learning in Disease Prediction. *International Conference on Computing Communication and Automation (ICCCA)*, vol. 8, pp. 1-4.

Korte, D. (2013). Business Intelligence in the Hospitality Industry. *International Journal of Innovation, Management and Technology*.

Kothari, T., Hu, C. & Roehl, W. (2018). *Adopting e-Procurement technology in a chain hotel: An exploratory case study* [online].

Król, K. & Zdonek, D. (2020). Analytics Maturity Models: An Overview. *Information*, vol. 11 (3), p. 142.

Kshetri, N. (2018). Blockchain’s roles in meeting key supply chain management objectives. *International Journal of Information Management*, vol. 39, pp. 80-89.

Lamba, K. & Singh, S. (2017). Big data in operations and supply chain management: current trends and future perspectives. *Production Planning & Control*, vol. 28 (11-12), pp. 877-890.

Lee, Y. & Min, O. (2017). Comparative Analysis of Machine Learning Algorithms to Urban Traffic Prediction. *ICTC 2017*.

Lis-Gutiérrez, J., Gaitán Angulo, M., Bouza, C., Balaguera, M., Lis-Gutiérrez, M. & Aguilera-Hernandez, D. (2020). "Text Mining Applied to Literature on Sustainable Supply Chain (1996–2018) An Analysis Based on Scopus". *Papers.ssrn.com* [online]. [Accessed 16 March 2020]. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3375836

Liu, P. & Yi, S. (2018). Investment decision-making and coordination of a three-stage supply chain considering Data Company in the Big Data era. *Annals of Operations Research*, vol. 270 (1-2), pp. 255-271.

Mamgain, S., Kumar, S. & Manjari Nayak, K. (2018). Car Popularity Prediction: A Machine Learning Approach. *Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pp. 1-5.

Mariani, M., Baggio, R., Fuchs, M. & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, vol. 30 (12), pp. 3514-3554.

Massa, S. & Testa, S. (2009). A knowledge management approach to organizational competitive advantage: Evidence from the food sector. *European Management Journal*, vol. 27 (2), pp. 129-141.

Mathew, E. (2019). Big Data Analytics in E-procurement of a Chain Hotel. *Advances in Internet, Data and Web Technologies* [online]. Vol. 29, pp. 295-308. Available at: https://link.springer.com/chapter/

Mergel, I., Rethemeyer, R. & Isett, K. (2016). Big Data in Public Affairs. *Public Administration Review*, vol. 76 (6), pp. 928-937.

Moktadir, M., Ali, S., Paul, S. & Shukla, N. (2019). Barriers to big data analytics in manufacturing supply chains: A case study from Bangladesh. *Computers & Industrial Engineering*, vol. 128, pp. 1063-1075.

Nandola, K., Koshal, M. & Koshal, R. (1982). Forecasting Restaurant Food Sales. *Cornell Hotel and Restaurant Administration Quarterly*, vol. 23 (2), pp. 92-96.

Neuhofer, B., Buhalis, D. & Ladkin, A. (2015). Smart technologies for personalized experiences: a case study in the hospitality domain. *Electronic Markets*, vol. 25 (3), pp. 243-254.

Nguyen, B. & Simkin, L. (2017). The Internet of Things (IoT) and marketing: the state of play, future trends and the implications for marketing. *Journal of Marketing Management*, vol. 33 (1-2), pp. 1-6.

Nikiel, S. (2019). New business models for Cultural and Creative Institutions. *Management*, vol. 23 (2), pp. 124-137.

Özdemir, A. (2018). "Curriculum development of SCM master's degree program in Turkey". *Urn.fi* [online]. [Accessed 16 March 2020]. Available at: http://urn.fi/URN:NBN:fi-fe2018110247071

Noone, B. & McGuire, K. (2013). Pricing in a social world: The influence of non-price information on hotel choice. *Journal of Revenue and Pricing Management*, vol. 12 (5), pp. 385-401.

O'Marah, K. (2020). "Circular Economy Is The Answer". *Forbes* [online]. [Accessed 8 May 2020]. Available at: https://www.forbes.com/sites/kevinomarah/2017/10/26/circular-economy-is-the-answer/

Padhi, S. & Pati, R. (2017). Quantifying potential tourist behavior in choice of destination using Google Trends. *Tourism Management Perspectives*, vol. 24, pp. 34-47.

Pahwa, K. & Agarwal, N. (2019). Stock Market Analysis using Supervised Machine Learning. *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing*,.

Panicker, V., Reddy, M. & Sridharan, R. (2016). Ant colony optimisation algorithm for location routing decisions in a supply chain. *International Journal of Value Chain Management*, vol. 7 (4), p. 285.

Pauleen, D. & Wang, W. (2017). Does big data mean big knowledge? KM perspectives on big data and analytics. *Journal of Knowledge Management*, vol. 21 (1), pp. 1-6.

Perkins, B. & Wailgum, T. (2020). "What is supply chain management (SCM)? Mastering logistics end to end". *CIO* [online]. [Accessed 12 March 2020]. Available at: https://www.cio.com/article/2439493/what-is-supply-chain-management-scm-mastering-logistics-end-to-end.html

Perner, P. (2018). Machine Learning and Data Mining in Pattern Recognition. Springer:New York.

Plinere, D. & Borisov, A. (2011). A Negotiation-Based Multi-Agent System for Supply Chain Management. *Scientific Journal of Riga Technical University. Computer Sciences*, vol. 45 (1).

Podlogar, M. (2007). *E-procurement success factors*. 3rd edn. e-Procurement in Emerging Economies Theory and Cases, Editors Pani and Agrahari. IDEA Group Publishing.

Pougel, F., Bonnel, S. & Beier, G. (2020). *E-Gov_and_Sustainability_a_Literature_Review* [online]. [Accessed 16 March 2020].

Pramanik, H., Kirtania, M. & Pani, A. (2019). Essence of digital transformation—Manifestations at large financial institutions from North America. *Future Generation Computer Systems*, vol. 95, pp. 323-343.

R., Radhakrishnan., Aithal, S. & Lm, M. (2018). "Business Strategy of Top Indian Company: L&T InfoTech.". *10.5281/zenodo.1302770* [online]. [Accessed 16 March 2020].

Rajkumar, C., Kavin, L., Luo, X. & Stentoft, J. (2016). Doctoral dissertations in logistics and supply chain management: a review of Nordic contributions from 2009 to 2014. *Logistics Research*, vol. 9 (1).

Raschka, S., Julian, D. & Hearty, J. (2016). *Python deeper insights Into machine learning*. 2nd edn. Birmingham B3 2PB, UK.:Packt Publishing Ltd.

Reavie, V. (2020). "Council Post: Do You Know The Difference Between Data Analytics And AI Machine Learning?". *Forbes* [online]. [Accessed 12 March 2020]. Available at:

https://www.forbes.com/sites/forbesagencycouncil/2018/08/01/do-you-know-the-difference-between-data-analytics-and-ai-machine-learning/#4121ef135878

Rekik, R., Kallel, I., Casillas, J. & Alimi, A. (2016). Using multiple criteria decision making approaches to assess the quality of web sites. *International Journal of Computer Science and Information Security* [online]. Vol. 14 (7). [Accessed 26 January 2020]. Available at: http://sites.google.com/site/ijcs

Roespinoedji, R., Mulyati, Y., Istambul, R. & Mutalib, N. (2019). How the Hotel Website Management Influence Hotel Supply Chain Management and Tourism Industry? A Case of Malaysian. *International Journal of Supply Chain Management*, vol. 8 (1), pp. 231-242.

Ruiz-Molina, M., Servera-Francés, D., Arteaga-Moreno, F. & Gil-Saura, I. (2018). Development and validation of a formative scale of technological advancement in hotels from the guest perspective. *Journal of Hospitality and Tourism Technology*, vol. 9 (3), pp. 280-294.

Russom, P. (2020). "TDWI Best Practices Reports | Transforming Data with Intelligence". *Transforming Data with Intelligence* [online]. [Accessed 12 March 2020]. Available at: https://tdwi.org/research/list/tdwi-best-practices-reports.aspx

S K, P., T N, M., T V, M., Singh, A. & Suhas, C. (2018). Class Result Prediction using Machine Learning. *International Conference On Smart Technology for Smart Nation*, pp. 6-11.

S. Padhi, S. (2020). "Nuance of Government Procurement Ethics in India". *Iimk.ac.in* [online]. [Accessed 16 March 2020]. Available at: https://www.iimk.ac.in/websiteadmin/FacultyPublications/Working%20Papers/228fullp.pdf

Samie, F., Bauer, L. & Henkel, J. (2019). From Cloud Down to Things: An Overview of Machine Learning in Internet of Things. *IEEE Internet of Things Journal*, vol. 6 (3), pp. 4921-4934.

Sanchez-Rodriguez, C., Martinez-Lorente, A. & Hemsworth, D. (2020). "The EFQM excellence model as enabler of e-procurement adoption and the effect on performance". *AIS Electronic Library (AISeL)* [online]. Available at: https://aisel.aisnet.org/ecis2012/180/

SATRIA PRAMDANA, H. (2019). Behavior Analysis Of Goods / Services Providers In Using Electronic Procurement System (E-Procurement). *Journal of Government and Politics*, vol. 10 (1).

Siami-Namini, S. & Namin, A. (2018). "Forecasting Economics and Financial Time Series: ARIMA vs. LSTM". arXiv.org [online]. Available at: https://arxiv.org/abs/1803.06386

Gerald, S., Borkar, D., Mayuram, R. & Carey, M. (2020). "Have Your Data and Query It Too | Proceedings of the 2016 International Conference on Management of Data". *Dl.acm.org* [online]. [Accessed 8 May 2020]. Available at: https://dl.acm.org/doi/10.1145/2882903.2904443

Ghaffarianhoseini, A., AlWaer, H., Ghaffarianhoseini, A., Clements-Croome, D., Berardi, U., Raahemifar, K. & Tookey, J. (2017). Intelligent or smart cities and buildings: a critical exposition and a way forward. *Intelligent Buildings International*, vol. 10 (2), pp. 122-129.

Shi, X., Liao, Z.: Managing supply chain relationships in the hospitality services: an empirical study of hotels and restaurants. Int. J. Hospitality Manag. **35**, 112–121 (2013)

Sibertin-Blanc, C., Therond, O., Mazzega, P. & Monteil, C. (2019). *The Entity-Process Framework for Integrated Agent-Based Modeling of Social-Ecological Systems*. 42nd edn. Cham:Springer International Publishing.

Soliman, M. & Karia, N. (2016). "Enterprise Resource Planning (ERP) Systems in the Egyptian Higher Education Institutions: Benefits, Challenges and Issues". *Pdfs.semanticscholar.org* [online]. [Accessed 16 March 2020].

Song, H., Qiu, R. & Park, J. (2019). A review of research on tourism demand forecasting: Launching the Annals of Tourism Research Curated Collection on tourism demand forecasting. *Annals of Tourism Research*, vol. 75, pp. 338-362.

Srai, J. & Lorentz, H. (2019). Developing design principles for the digitalisation of purchasing and supply management. *Journal of Purchasing and Supply Management*, vol. 25 (1), pp. 78-98.

Stalidis, G., Karapistolis, D. & Vafeiadis, A. (2015). Marketing Decision Support Using Artificial Intelligence and Knowledge Modeling: Application to Tourist Destination Management. *Procedia - Social and Behavioral Sciences*, vol. 175, pp. 106-113.

Stefanovic, N. (2014). Proactive Supply Chain Performance Management with Predictive Analytics. *The Scientific World Journal*, vol. 2014, pp. 1-17.

Subramanian, G. (2015). *Python Data Science Cookbook*. 1st edn. Birmingham B3 2PB, UK.:Packt Publishing.

Sutono, A. (2019). Supply chain management: implementation issues and research opportunities in tourism industry. *Uncertain Supply Chain Management*, pp. 427-438.

Tang, M., Fung, M., Kulendran, N., King, B. & Yap, M. (2016). Research Note: Using Demand Determinants to Anticipate Fluctuations in Hotel Occupancy. *Tourism Economics*, vol. 22 (1), pp. 179-187.

Toktaş-Palut, P. & Ülengin, F. (2010). Modeling a Supply Chain as a Queuing System. *IFAC Proceedings Volumes*, vol. 43 (17), pp. 242-247.

Tom, K. (2020). "Supply Chain Management vs. Procurement". Lean Supply Solutions - Innovative Supply Chain Solutions [online]. [Accessed 9 February 2020]. Available at: http://www.leansupplysolutions.com/blog/supply-chain-management-vs-procurement/

Treiblmaier, H. (2018). The Impact of the Blockchain on the Supply Chain: A Theory-Based Research Framework and a Call for Action. *SSRN Electronic Journal*,.

Ulutas, A., Shukla, N., Kiridena, S. & Gibson, P. (2015). A utility-driven approach to supplier evaluation and selection: empirical validation of an integrated solution framework. *International Journal of Production Research*, vol. 54 (5), pp. 1554-1567.

Valeri, M. & Fadlon, L. (2018). Co-Evolutionary Prospects in Tourism. *International Journal of Business and Management*, vol. 13 (9), p. 95.

Verhoef, P., Kooge, E. & Walker, N. (2016). *Creating value with big data analytics*. London:Routledge.

Waller, M. & Fawcett, S. (2013). Click Here for a Data Scientist: Big Data, Predictive Analytics, and Theory Development in the Era of a Maker Movement Supply Chain. *Journal of Business Logistics*, vol. 34 (4), pp. 249-252.

Wang, G., Gunasekaran, A., Ngai, E. & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, vol. 176, pp. 98-110.

Wen, H., Li, S., Li, W., Li, J. & Yin, C. (2017). Comparison of four machine learning techniques for the prediction of prostate cancer survivability. *Institute of Information Science and Engineering*, (3).

"What Are Companies' Legal Obligations Around Coronavirus?". (2020). [Accessed 12 March 2020]. Available at: https://hbr.org/2020/03/what-are-companies-legal-obligations-around-coronavirus

Witell, L., Kowalkowski, C., Perks, H., Raddats, C., Schwabe, M., Benedettini, O. & Burton, J. (2019). Characterizing customer experience management in business markets. *Journal of Business Research*,.

Wokabi, V. & Fatoki, O. (2019). The Use and Benefits of E-Technology Business Applications. *International Journal of Business and Management*, vol. VII (1).

Wooldridge, M. & Jennings, N. (1995). Intelligent agents: theory and practice. *The Knowledge Engineering Review*, vol. 10 (2), pp. 115-152.

Wu, C., Patil, P. & Gunaseelan, S. (2018). Algorithms for Multiple Regression on Black Friday Sales Data. *IEEE explore*, vol. 14 (1), pp. 17-21.

Yeom, H., Kim, J. & Chung, C. (2020). LSTM Improves Accuracy of Reaching Trajectory Prediction From Magnetoencephalography Signals. IEEE Access, vol. 8, pp. 20146-20150.

Yu, A., Yevu, S. & Nani, G. (2020). Towards an integration framework for promoting electronic procurement and sustainable procurement in the construction industry: A systematic literature review. *Journal of Cleaner Production*, vol. 250, p. 119493.

Zhang, J., Yang, X. & Appelbaum, D. (2015). Toward Effective Big Data Analysis in Continuous Auditing. *Accounting Horizons*, vol. 29 (2), pp. 469-476.

Zhan, Y. & Tan, K. (2020). An analytic infrastructure for harvesting big data to enhance supply chain performance. *European Journal of Operational Research*, vol. 281 (3), pp. 559-574.

APPENDICES

I. Descriptive Analytics

A. Buyer Summary

```
Buyer_df=Buyer_df.groupby(['Supplier_Name','Category_Name']).sum()
Buyer_df.sort_values('PurchaseAmount',ascending=False,axis=0,inplace=True)
Buyer_df.head(10)
```

B. Spend Summary

```
# get a list of the columns
col_list = list(total_spend2018)
# use this handy way to swap the elements
col_list[0], col_list[1] = col_list[1], col_list[0]
# assign back, the order will now be swapped
total_spend2018.columns = col_list
explode=(.1,0,.1,0,.1,0,.1,0,.1,0,.1,0,.1,0,.1,0,.1,0,.1,0)
total_spend2018.plot(kind='pie',subplots=True,
                            figsize=(15, 6),
                            autopct='%1.1f%%',
                            startangle=90,
                            shadow=True,
                            labels=None,
                            pctdistance=1.12,
                            explode=explode
                            )



plt.title('2018 top 20 products ')
plt.axis('equal')
plt.legend(labels=total_spend2018.SupplierName, loc='upper left')
plt.show()
```

C. Spend report (yearly)

```
import pandas as pd
def mapper(month):
    return month.strftime('%m')

data['Month'] = data['Gendate_time'].apply(mapper)
d_f=data[['Month','Purchasetype','Categoryname','Extension']]
d_f_1=d_f[(d_f['Purchasetype']=='Beverage')]
d_f_2=d_f[(d_f['Purchasetype']=='Food')]
d_f_3=d_f[(d_f['Purchasetype']=='General')]
d_f_frames=[d_f_1,d_f_2,d_f_3]
result=pd.concat(d_f_frames)
result
```

D. Spend by Purchase type

```
from matplotlib import pyplot
import pandas as pd

ax = tx.plot(kind="bar",figsize=(20, 10))


plt.title('2019 products')
plt.xlabel('Top 10 Products of 2019')
plt.ylabel('Total Spent')

plt.xticks(
    rotation=45,
    horizontalalignment='right',
    fontweight='light',
    fontsize=12
)
```

II. Diagnostic Analytics

A. Pie Chart

```
labels=group.index
group.plot(kind='pie',figsize=(16, 9),subplots=True,
                             autopct='%1.1f%%',
                             startangle=90,
                             shadow=True,
                             labels=labels,
                             pctdistance=0.7,
                            labeldistance=1.05)
plt.legend(labels='_nolegend_', loc='upper left')
plt.title('Spend by category for the year ')
plt.axis('equal')
plt.show()
```

B. Time Series

```
from matplotlib import pyplot
from sklearn.preprocessing import StandardScaler
d_f1=dt[['Month','Occupancy']]
d_f2=dt[['Month','Extendhomeamt']]
d_f1.groupby('Month').mean().plot(grid=True, title="2019 Total Occupancy- by mont
gend=True)
d_f1.groupby('Month').sum().plot(grid=True, title="2019 Total Spend- by month", 
True)
```

C. Trend Analysis

```
import pandas as pd
d_f=dt[['Month','BuyercompID','Occupancy']]
result = d_f.groupby(['Month','BuyercompID']).mean().unstack().fillna(0)
result.plot(grid=True,figsize=(14,8), title="2019 Average Occupancy by Property", legend=True)
plt.legend(bbox_to_anchor=(1.04,1), loc="upper left")
```

D. Stacked Bar Chart

```
import numpy as np
import matplotlib.pyplot as plt

df=dt[['Month','Categoryname','Extendhomeamt']]
df.columns=['Month','Categories','Total_Spend']
group = df.groupby(['Categories','Month']).sum()
group=group.sort_values('Total_Spend',ascending=False)
```

```
group=group.reset_index().pivot(index='Categories', columns='Month', values='Total_Spend')

group=group.append(pd.Series(group.sum(),name='Total'))\
  .assign(Total=group.sum(1))\
  .set_value('Total','Total',group.values.sum())
group=group.sort_values('Total',axis=0,ascending=False)
group=group.head(25)
group
```

```
group= group.drop(['Total'],axis=1)
```

```
group.plot(kind='barh',stacked=True,figsize=(15,20),width=0.4)
plt.savefig(r'C:\Users\admin\Documents\_PhD\PhDRawData\2018\Top25Category2018.jpg')
```

E. Line graph

```
import pandas as pd
d_f=dt[['Month','BuyercompID','Occupancy']]
result = d_f.groupby(['Month','BuyercompID']).mean().unstack().fillna(0)
result.plot(grid=True,figsize=(14,8), title="2019 Average Occupancy by Property", legen
d=True)
plt.legend(bbox_to_anchor=(1.04,1), loc="upper left")
```

III. Machine Learning Algorithm for Finding Purchase Duration

A. Import data

```
import pandas as pd
import os
from datetime import date, datetime
pd.set_option('display.max_columns', 500)
data = None

for file in os.listdir("./"):
    if file.endswith(".csv"):


        fileData = pd.read_csv(file,encoding='cp1252')

        if data is None:
            data = fileData
        else:
            data = data.append( fileData )
```

B. Import Libraries

```
#import libraries
from __future__ import division
from __future__ import unicode_literals
from datetime import datetime, timedelta,date
import pandas as pd
%matplotlib inline
from sklearn.metrics import classification_report,confusion_matrix
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns

from sklearn.cluster import KMeans

#do not show warnings
import warnings
warnings.filterwarnings("ignore")

#import plotly for visualization
import chart_studio.plotly as py
import plotly.offline as pyoff
import plotly.graph_objs as go

#import machine learning related libraries
from sklearn.svm import SVC
from sklearn.multioutput import MultiOutputClassifier
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.ensemble import RandomForestClassifier
from sklearn.linear_model import LogisticRegression
import xgboost as xgb
from sklearn.model_selection import KFold, cross_val_score, train_test_split

#initiate plotly
pyoff.init_notebook_mode()
```

C. Date conversion to datetime data type

```
import pandas as pd
dt=data
dt['Recdate'] = pd.to_datetime(dt['Recdate'],format='%d/%m/%Y',infer_datetime_format=Tr
ue)
dt
```

Adding 'Month' column to the data frame

```
def mapper(month):
    return month.strftime('%m')
dt['Month'] = dt['Recdate'].apply(mapper)
```

Filtering data frame to extract information for UAE

```
H_UAE = H_data.query(("Country=='UNITED ARAB EMIRATES '")).reset_index(drop=True)
```

D. Cut off entire data into two [9 month and 3 months] for finding purchase duration

```
H_9m = H_UAE[(H_UAE.Recv_Date <= date(2019,9,30)) & (H_UAE.Recv_Date >= date(2019,1,1
))].reset_index(drop=True)
H_3m = H_UAE[(H_UAE.Recv_Date >= date(2019,10,1)) & (H_UAE.Recv_Date <= date(2019,12,31
))].reset_index(drop=True)
H_9m.sort_values(['PropertyID','Recv_Date'])
```

Selecting unique buyers

```
H_user = pd.DataFrame(H_data['PropertyID'].unique())
H_user.columns = ['PropertyID']
H_user
```

E. Finding latest purchase date and earliest purchase date from the cutoff point

```
#create a dataframe with customer id and first purchase date in H_3m
H_3m_first_purchase = H_3m.groupby(['PropertyID']).Recv_Date.min().reset_index()
H_3m_first_purchase.columns = ['PropertyID','MinPurchaseDate']

#create a dataframe with customer id and last purchase date in tx_6m
H_9m_last_purchase = H_9m.groupby(['PropertyID']).Recv_Date.max().reset_index()
H_9m_last_purchase.columns = ['PropertyID','MaxPurchaseDate']

#merge two dataframes
H_purchase_dates = pd.merge(H_9m_last_purchase,H_3m_first_purchase,on=['PropertyID'],ho
w='left')

H_purchase_dates
```

F. Finding purchase duration

```
#calculate the time difference in days:
H_purchase_dates['PurchaseDuration'] = (H_purchase_dates['MinPurchaseDate'] - H_purchas
e_dates['MaxPurchaseDate']).dt.days

#merge with H_user
H_user = pd.merge(H_user, H_purchase_dates[['PropertyID','PurchaseDuration']],on='Prope
rtyID',how='left')

#print H_user
H_user.head
```

```
#fill NA values with 999
H_user = H_user.fillna(999)
H_user
```

G. Adding Initial feature

```
#find the recency in days and add it to H_user
H_max_purchase['Recency'] = (H_max_purchase['MaxPurchaseDate'].max() - H_max_purchase[
'MaxPurchaseDate']).dt.days
H_user = pd.merge(H_user, H_max_purchase[['PropertyID','Recency']], on='PropertyID',how
='left')

H_user
```

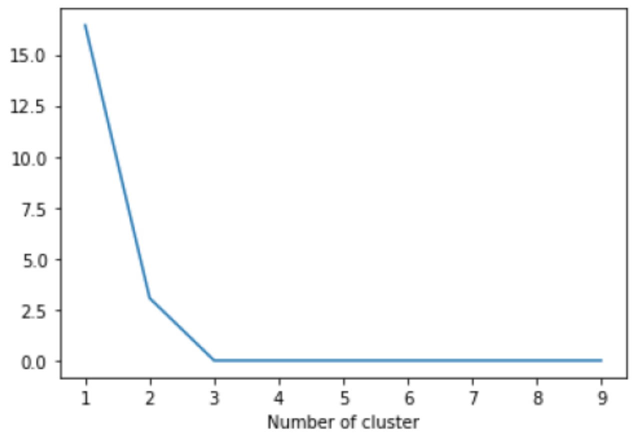
H. Plot Features
[similar code is used for spend, frequency, occupancy]

```
plot_data = [
    go.Histogram(
        x=H_user['Recency']
    )
]

plot_layout = go.Layout(
        title='Recency'
    )
fig = go.Figure(data=plot_data, layout=plot_layout)
pyoff.iplot(fig)
```

I. Elbow method for making cluster
[similar code is used for recency, spending, frequency and occupancy]

```
sse={}
H_Recency = H_user[['Recency']]
for k in range(1, 10):
    kmeans = KMeans(n_clusters=k, max_iter=1000).fit(H_Recency)
    H_Recency["clusters"] = kmeans.labels_
    sse[k] = kmeans.inertia_
plt.figure()
plt.plot(list(sse.keys()), list(sse.values()))
plt.xlabel("Number of cluster")
plt.show()
```



Creating cluster [similar code is used for creating all clusters]

```
kmeans = KMeans(n_clusters=3)
kmeans.fit(H_user[['Recency']])
H_user['RecencyCluster'] = kmeans.predict(H_user[['Recency']])
```

Order clusters [similar code is used to order all clusters]

```
def order_cluster(cluster_field_name, target_field_name,df,ascending):
    new_cluster_field_name = 'new_' + cluster_field_name
    df_new = df.groupby(cluster_field_name)[target_field_name].mean().reset_index()
    df_new = df_new.sort_values(by=target_field_name,ascending=ascending).reset_index(d
rop=True)
    df_new['index'] = df_new.index
    df_final = pd.merge(df,df_new[[cluster_field_name,'index']], on=cluster_field_name)
    df_final = df_final.drop([cluster_field_name],axis=1)
    df_final = df_final.rename(columns={"index":cluster_field_name})
    return df_final
```

In [38]:

```
H_user = order_cluster('RecencyCluster', 'Recency',H_user,False)
H_user.groupby('RecencyCluster')['Recency'].describe()
```

J. Finding overall score

```
#building overall Gradeation
H_user['OverallScore'] = H_user['RecencyCluster'] + H_user['FrequencyCluster'] + H_user
['SpendCluster']+H_user['OccupancyCluster']

H_user
```

Assign grades for the score

```
#assign Grade names
H_user['Grade'] = 'Low-Value'
H_user.loc[H_user['OverallScore']>=np.percentile(H_user['OverallScore'], 25),'Grade'] =
'Mid-Value'
H_user.loc[H_user['OverallScore']>=np.percentile(H_user['OverallScore'], 75),'Grade'] =
'High-Value'
H_user
```

K. Scatter plot between each features.
[Similar code is used for plotting all other combinations of features.]

```
# plot Spend vs Frequency
H_graph = H_user.query("Spend < Spend.max() and Frequency < Frequency.max()")

plot_data = [
    go.Scatter(
        x=H_graph.query("Grade == 'Low-Value'")['Frequency'],
        y=H_graph.query("Grade == 'Low-Value'")['Spend'],
        mode='markers',
        name='Low',
        marker= dict(size= 7,
            line= dict(width=1),
            color= 'blue',
            opacity= 0.8
           )
    ),
        go.Scatter(
        x=H_graph.query("Grade == 'Mid-Value'")['Frequency'],
        y=H_graph.query("Grade == 'Mid-Value'")['Spend'],
        mode='markers',
        name='Mid',
        marker= dict(size= 9,
            line= dict(width=1),
            color= 'green',
            opacity= 0.5
           )
    ),
        go.Scatter(
        x=H_graph.query("Grade == 'High-Value'")['Frequency'],
        y=H_graph.query("Grade == 'High-Value'")['Spend'],
        mode='markers',
        name='High',
        marker= dict(size= 11,
            line= dict(width=1),
            color= 'red',
            opacity= 0.9
           )
    ),
]

plot_layout = go.Layout(
        yaxis= {'title': "Spend"},
        xaxis= {'title': "Frequency"},
        title='Grades'
    )
fig = go.Figure(data=plot_data, layout=plot_layout)
pyoff.iplot(fig)
```

L. Adding more features

```
### Steps to find difference between last 3 purchases
#create a dataframe with PropertyID and Invoice Date
H_day_sequence = H_9m[['PropertyID','Recv_Date']]
#convert Invoice Datetime to day
H_day_sequence['Recv_day'] = H_9m['Recv_Date'].dt.date
H_day_sequence = H_day_sequence.sort_values(['PropertyID','Recv_Date'])
#drop duplicates
H_day_sequence = H_day_sequence.drop_duplicates(subset=['PropertyID','Recv_day'],keep=
'first')
H_day_sequence
```

Finding difference between last 3 purchases

```
#shifting last 3 purchase dates
H_day_sequence['P1Recv_Date'] = H_day_sequence.groupby('PropertyID')['Recv_day'].shift(
1)
H_day_sequence['P2Recv_Date'] = H_day_sequence.groupby('PropertyID')['Recv_day'].shift(
2)
H_day_sequence['P3Recv_Date'] = H_day_sequence.groupby('PropertyID')['Recv_day'].shift(
3)
```

```
H_day_sequence['P1Diff'] = (H_day_sequence['Recv_day'] - H_day_sequence['P1Recv_Date'])
.dt.days
H_day_sequence['P2Diff'] = (H_day_sequence['Recv_day'] - H_day_sequence['P2Recv_Date'])
.dt.days
H_day_sequence['P3Diff'] = (H_day_sequence['Recv_day'] - H_day_sequence['P3Recv_Date'])
.dt.days
H_day_sequence
```

Finding mean and standard deviation of the day difference

```
H_P_Diff = H_day_sequence.groupby('PropertyID').agg({'P1Diff': ['mean','std']}).reset_i
ndex()
H_P_Diff.columns = ['PropertyID', 'P_DiffMean','P_DiffStd']
H_P_Diff
```

M. Adding label

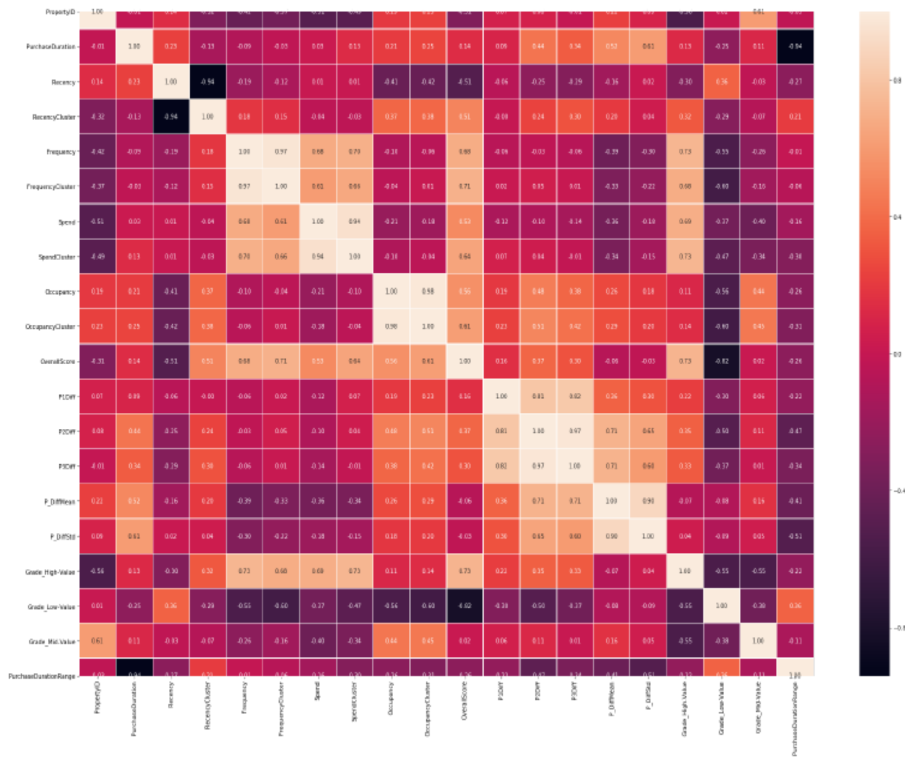
```
H_Class['PurchaseDurationRange'] = 2
H_Class.loc[H_Class.PurchaseDuration>=np.percentile(H_user['PurchaseDuration'], 25),'Pu
rchaseDurationRange'] = 1
H_Class.loc[H_Class.PurchaseDuration>=np.percentile(H_user['PurchaseDuration'], 75),'Pu
rchaseDurationRange'] = 0
```

N. Create Heat map

```
corr = H_Class[H_Class.columns].corr()
plt.figure(figsize = (30,20))
sns.heatmap(corr, annot = True, linewidths=0.2, fmt=".2f")
```

Out[76]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x221762c92c8>
```



O. Split into test and train data

```
X, y = H_Class.drop('PurchaseDurationRange',axis=1), H_Class.PurchaseDurationRange
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=4
2)
```

P. Baseline Model

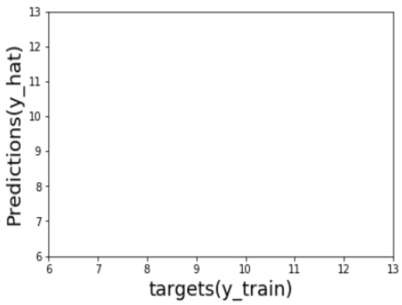
```
from sklearn.preprocessing import StandardScaler
scaler=StandardScaler()
scaler.fit(inputs)
inputs_scaled=scaler.transform(inputs)
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(inputs_scaled, targets,test_size=0.2, random_state=42)
```

```
from sklearn.linear_model import LinearRegression
linereg=LinearRegression()
linereg.fit(X_train,y_train)
```

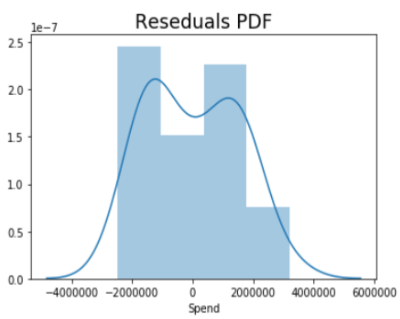
```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

```
y_hat=linereg.predict(X_train)
plt.scatter(y_train,y_hat)
plt.xlabel('targets(y_train)',size=18)
plt.ylabel('Predictions(y_hat)',size=18)
plt.xlim(6,13)
plt.ylim(6,13)
plt.show()
```




```
import seaborn as sns
sns.distplot(y_train-y_hat)
plt.title("Reseduals PDF",size=18)
```

```
Text(0.5, 1.0, 'Reseduals PDF')
```




```
y_pred=linereg.predict(X_test)
y_pred
```

```
array([25070191.00203441, 23751719.71792406, 27727607.14335653,
       16559980.60749849, 21354678.70840001, 16666610.89697071,
       19723800.96965973, 21266640.45840126, 21052350.57211722,
       26109601.52127159])
```

```
from sklearn import metrics
print(np.sqrt(metrics.mean_squared_error(y_test,y_pred)))
```

```
2547629.140609191
```

Q. Machine Learning Model Selection

```
models = []
models.append(("LR",LogisticRegression()))
models.append(("NB",GaussianNB()))
models.append(("RF",RandomForestClassifier()))
models.append(("SVC",SVC()))
models.append(("Dtree",DecisionTreeClassifier()))
models.append(("XGB",xgb.XGBClassifier()))
models.append(("KNN",KNeighborsClassifier()))
```

R. Finding Accuracy

```
for name,model in models:
    kfold = KFold(n_splits=2, random_state=22)
    cv_result = cross_val_score(model,X_train,y_train, cv = kfold,scoring = "accuracy")
    print(name, cv_result)
```

```
LR [0.42857143 0.42857143]
NB [0.71428571 0.71428571]
RF [0.71428571 0.57142857]
SVC [0.28571429 0.28571429]
Dtree [0.85714286 0.71428571]
XGB [0.57142857 0.71428571]
KNN [0.28571429 0.28571429]
```

S. Fine Tuning the model

```
xgb_model = xgb.XGBClassifier().fit(X_train, y_train)

print('Accuracy of XGB classifier on training set: {:.2f}'
       .format(xgb_model.score(X_train, y_train)))
print('Accuracy of XGB classifier on test set: {:.2f}'
       .format(xgb_model.score(X_test[X_train.columns], y_test)))
```

```
y_pred = xgb_model.predict(X_test)
```

```
print (classification_report(y_test, y_pred))
```

```
from sklearn.model_selection import GridSearchCV

param_test1 = {
 'max_depth':range(3,10,2),
 'min_child_weight':range(1,6,2)
}
gsearch1 = GridSearchCV(estimator = xgb.XGBClassifier(),
param_grid = param_test1, scoring='accuracy',n_jobs=-1,iid=False, cv=2)
gsearch1.fit(X_train,y_train)
gsearch1.best_params_, gsearch1.best_score_
```

Scenario 1: Duration of purchase for the frequent buyer

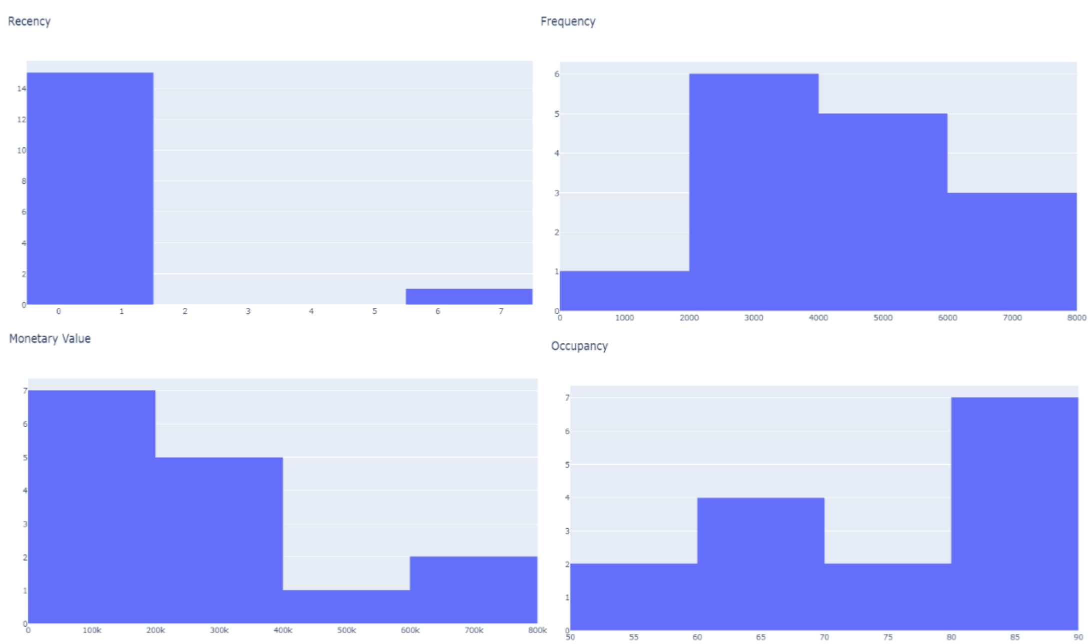
a) Find top 25 frequently bought commodity

```
pn=H_data[['Productname','Extendhomeamt']]
pn=pn.groupby('Productname',as_index=False).agg({'Extendhomeamt': ['sum', 'count']})
pn.sort_values(by=('Extendhomeamt','count'),ascending=False,axis=0,inplace=True)
pn=pn.head(25)
pn
```

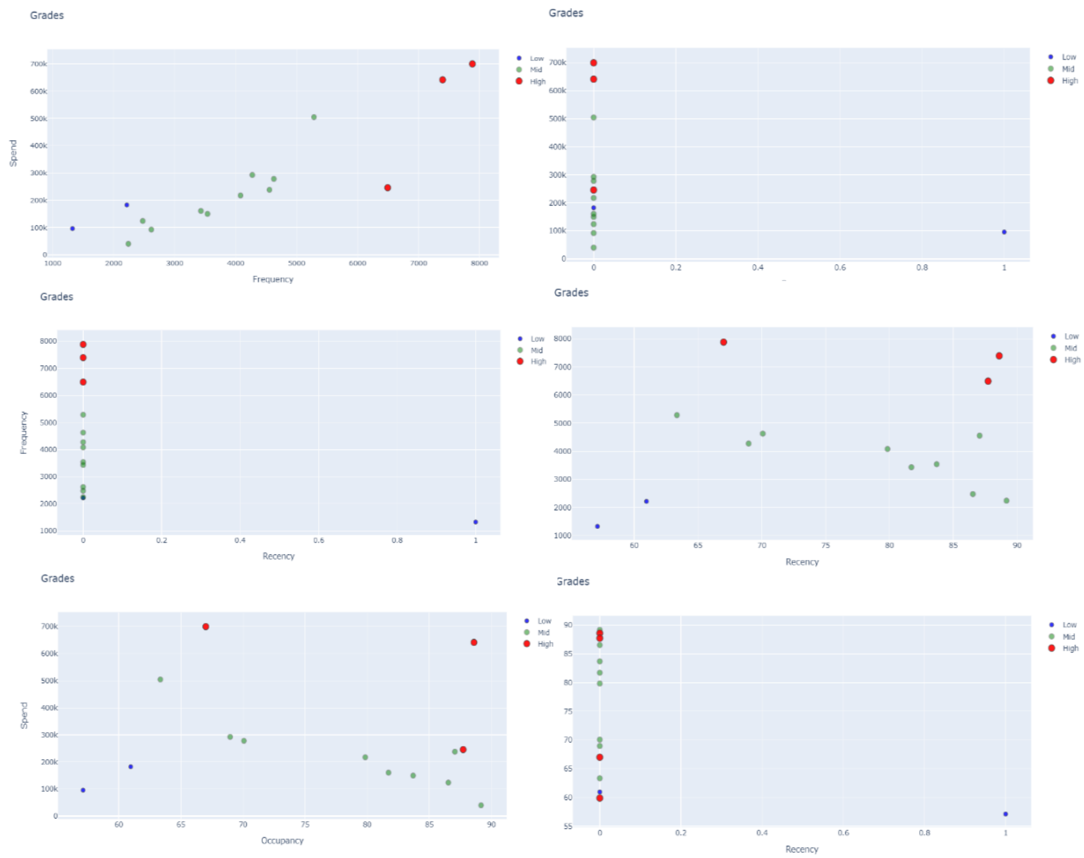
b) Merge this with original data frame.

```
mergedStuff = pd.merge(H_data, pn, on=['Productname'], how='inner')
mergedStuff=mergedStuff.dropna()
mergedStuff
```

c) Histogram for each feature



d) Scatter plot between 2 feature variables



e) Model selection

```
X, y = H_Class.drop('PurchaseDurationRange',axis=1), H_Class.PurchaseDurationRange
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
models = []
models.append(("LR",LogisticRegression()))
models.append(("NB",GaussianNB()))
models.append(("RF",RandomForestClassifier()))
models.append(("SVC",SVC()))
models.append(("Dtree",DecisionTreeClassifier()))
models.append(("XGB",xgb.XGBClassifier()))
models.append(("KNN",KNeighborsClassifier()))
```

```
for name,model in models:
    kfold = KFold(n_splits=2, random_state=22)
    cv_result = cross_val_score(model,X_train,y_train, cv = kfold,scoring = "accuracy")
    print(name, cv_result)
```

```
LR [0.66666667 0.83333333]
NB [0.66666667 0.83333333]
RF [0.66666667 0.83333333]
SVC [0.66666667 0.83333333]
Dtree [0.66666667 0.83333333]
XGB [0.66666667 0.83333333]
KNN [0.66666667 0.83333333]
```

Scenario 2: Duration of purchase for the most spend product

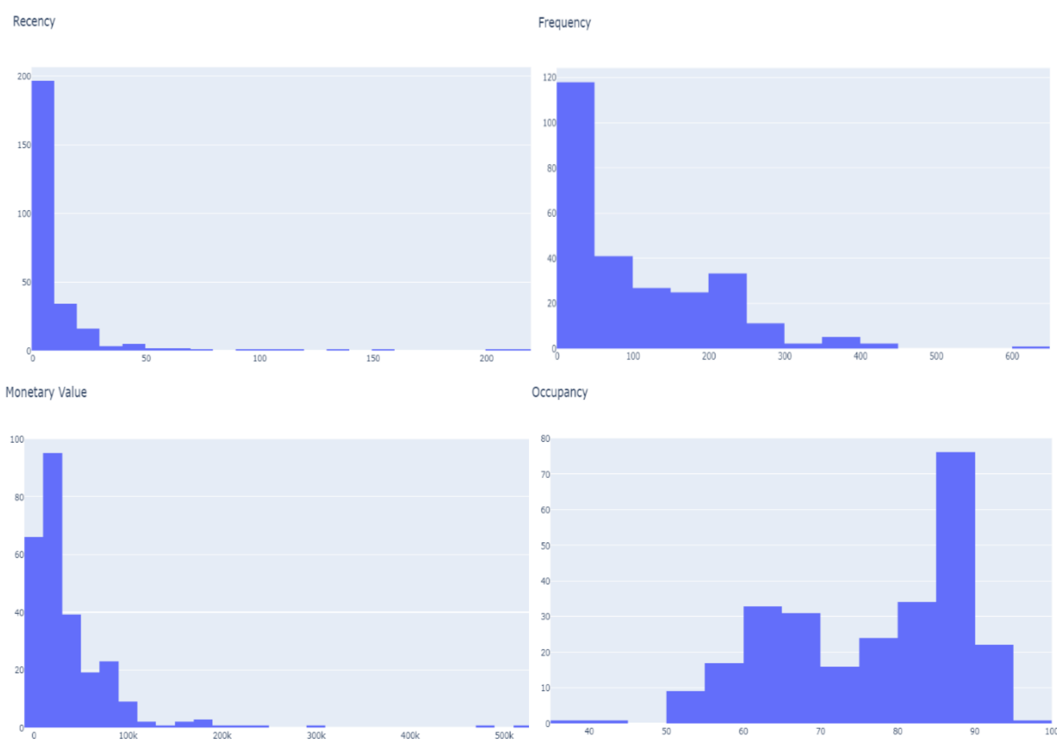
a) Find the top 25 spend products

```
p_n=H_data[['Productname','Extendhomeamt']]
p_n=p_n.groupby('Productname',as_index=False).agg({'Extendhomeamt': ['sum']})
p_n.sort_values(by=('Extendhomeamt','sum'),ascending=False,axis=0,inplace=True)
p_n=p_n.head(25)
p_n
```

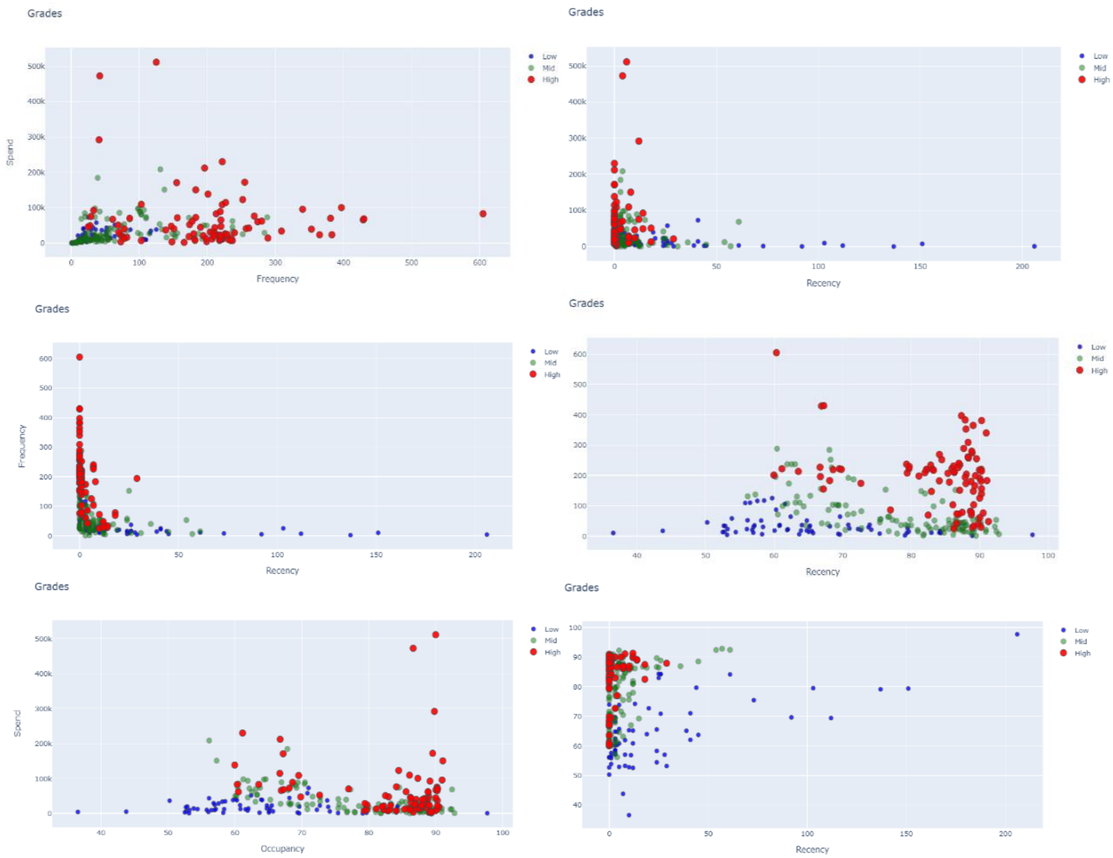
b) Merge this with original data frame

```
mergedStuff = pd.merge(H_data, p_n, on=['Productname'], how='inner')
mergedStuff=mergedStuff.dropna()
mergedStuff
```

c) Histogram for each feature



d) Scatter plot between 2 features



e) Model selection

```
X, y = H_Class.drop('PurchaseDurationRange',axis=1), H_Class.PurchaseDurationRange
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
models = []
models.append(("LR",LogisticRegression()))
models.append(("NB",GaussianNB()))
models.append(("RF",RandomForestClassifier()))
models.append(("SVC",SVC()))
models.append(("Dtree",DecisionTreeClassifier()))
models.append(("XGB",xgb.XGBClassifier()))
models.append(("KNN",KNeighborsClassifier()))
```

```
for name,model in models:
    kfold = KFold(n_splits=2, random_state=22)
    cv_result = cross_val_score(model,X_train,y_train, cv = kfold,scoring = "accuracy")
    print(name, cv_result)
```

```
LR [0.67924528 0.73584906]
NB [0.71698113 0.69811321]
RF [0.67924528 0.66037736]
SVC [0.52830189 0.52830189]
Dtree [0.67924528 0.71698113]
XGB [0.73584906 0.73584906]
KNN [0.56603774 0.49056604]
```

Scenario 3: Duration of most purchase commodity

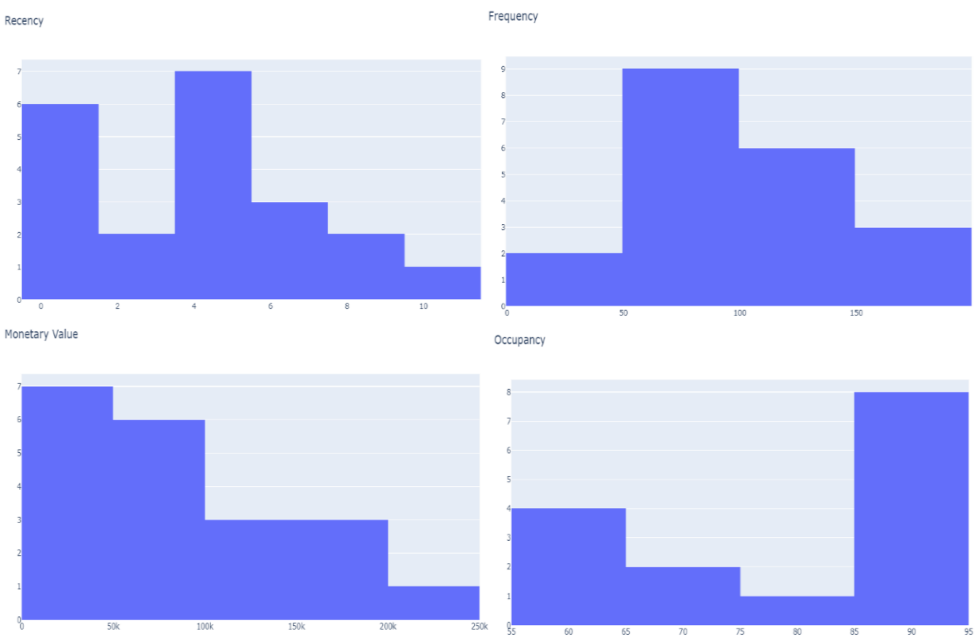
a) Find the most spend product

```
p_n=H_data[['Productname','Extendhomeamt']]
p_n=p_n.groupby('Productname',as_index=False).agg({'Extendhomeamt': ['sum']})
p_n.sort_values(by=('Extendhomeamt','sum'),ascending=False,axis=0,inplace=True)
p_n=p_n.head(25)
p_n
```

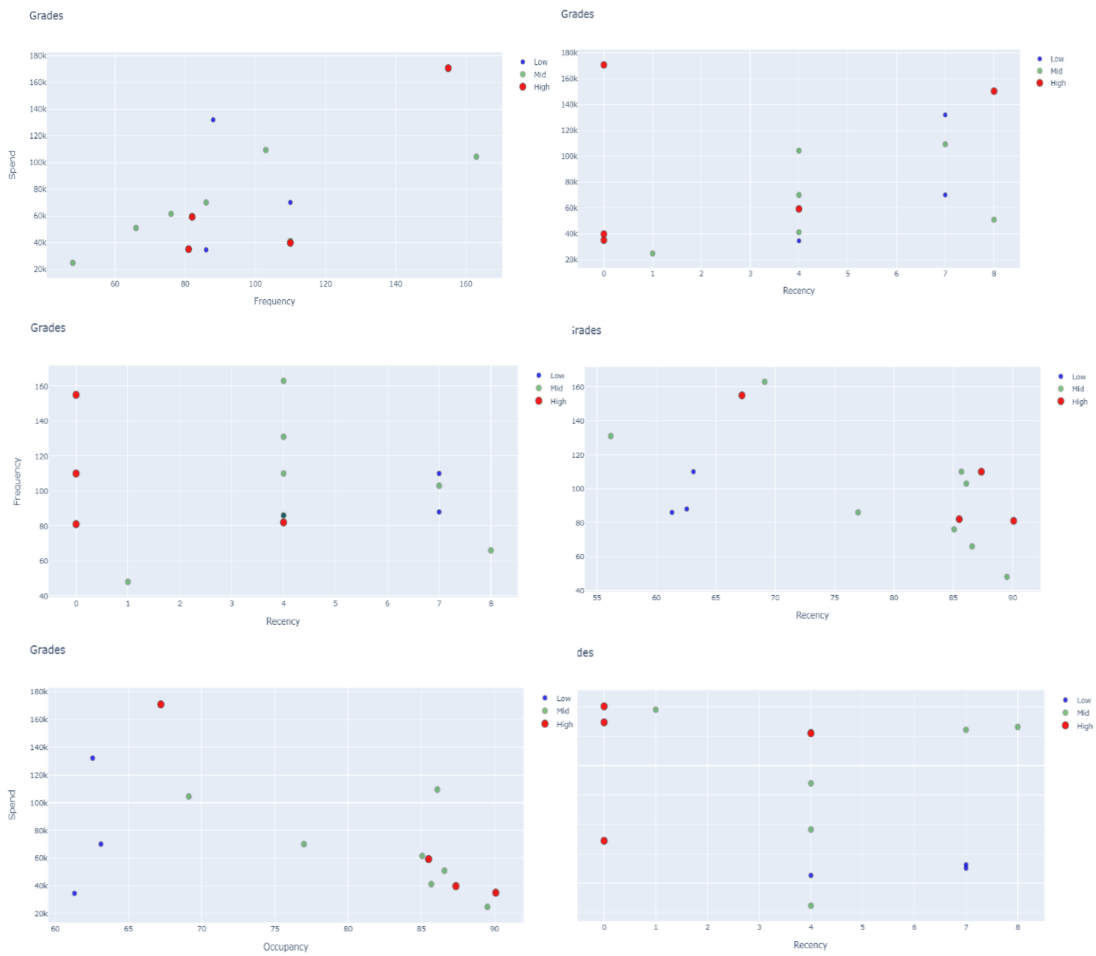
b) Filter data to select records pertaining to the most spend product

```
#H_data=mergedStuff
H_data.columns = ['Recv_Num','Item_Num','ProductName','Recv_Qty','Recv_Date','UnitPrice','PropertyID','Country','OccupancyRate',
H_UAE = H_data.query((("Country=='UNITED ARAB EMIRATES '") and ("ProductName=='WATER MINERAL PLASTIC BOTTLE'"))).reset_index(drc
```

c) Histogram for each feature variables



d) Scatter plot between 2 feature variables



e) Model Selection

```
X, y = H_Class.drop('PurchaseDurationRange',axis=1), H_Class.PurchaseDurationRange
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
models = []
models.append(("LR",LogisticRegression()))
models.append(("NB",GaussianNB()))
models.append(("RF",RandomForestClassifier()))
models.append(("SVC",SVC()))
models.append(("Dtree",DecisionTreeClassifier()))
models.append(("XGB",xgb.XGBClassifier()))
models.append(("KNN",KNeighborsClassifier()))
```

```
for name,model in models:
    kfold = KFold(n_splits=2, random_state=22)
    cv_result = cross_val_score(model,X_train,y_train, cv = kfold,scoring = "accuracy")
    print(name, cv_result)
```

```
LR [0.66666667 0.83333333]
NB [0.66666667 0.66666667]
RF [0.66666667 0.33333333]
SVC [0.66666667 0.66666667]
Dtree [0.66666667 0.33333333]
XGB [0.66666667 0.66666667]
KNN [0.66666667 0.66666667]
```

IV. Machine Learning Algorithm for Demand forecasting

A. Data Wrangling

In [1]:

```
from __future__ import division
from datetime import datetime, timedelta,date
import pandas as pd
%matplotlib inline
import matplotlib.pyplot as plt
import numpy as np
```

In [2]:

```
import warnings
warnings.filterwarnings("ignore")
```

In [3]:

```
import chart_studio.plotly as py
import plotly.offline as pyoff
import plotly.graph_objs as go
```

In [4]:

```
import keras
from keras.layers import Dense
from keras.models import Sequential
from keras.optimizers import Adam
from keras.callbacks import EarlyStopping
from keras.utils import np_utils
from keras.layers import LSTM
from sklearn.model_selection import KFold, cross_val_score, train_test_split
```

```
Using TensorFlow backend.
```

In [5]:

```
pyoff.init_notebook_mode()
```

Data importing, conversion of datatypes, selecting and filtering data is done as same in the earlier part of study.

B. Aggregate spend by month

```
#groupby date and sum the Spend
H_Spend = H_Spend.groupby('date').Spend.sum().reset_index()
H_Spend.head()
```

C. Data Transformation to make stationary data

```
#create a new dataframe to model the difference
H_diff = H_Spend.copy()
```

In [61]:

```
#add previous Spend to the next row
H_diff['prev_Spend'] = H_diff['Spend'].shift(1)
```

Now we have the required dataframe for modeling the difference:

In [62]:

```
H_diff.head()
```

Out[62]:

	date	Spend	prev_Spend
0	2011-07-01	0.00	NaN
1	2011-11-01	678.50	0.00
2	2014-01-01	9732975.30	678.50
3	2014-02-01	10165988.45	9732975.30
4	2014-03-01	11273022.46	10165988.45

In [63]:

```
#drop the null values and calculate the difference
H_diff = H_diff.dropna()
H_diff['diff'] = (H_diff['Spend'] - H_diff['prev_Spend'])
H_diff.head(10)
```

D. Baseline Prediction

A rolling forecast scenario will be used, also called walk-forward model validation.

Each time step of the test dataset will be walked one at a time. A model will be used to make a forecast for the time step, then the actual expected value from the test set will be taken and made available to the model for the forecast on the next time step.

```
from sklearn.metrics import mean_squared_error
# persistence model
def model_persistence(x):
    return x

# walk-forward validation
history = [x for x in train]
predictions = list()
for x in test_X:
    yhat = model_persistence(x)
    predictions.append(yhat)
test_score = mean_squared_error(test_y, predictions)
print('Test MSE: %.3f' % test_score)
```

```
Test MSE: 85043002.467
```

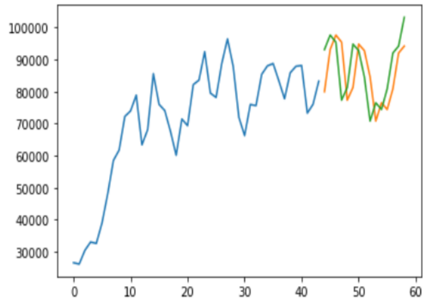
In this case, the error is more than 85,043,002 over the test dataset.

Finally, a plot is made to show the training dataset and the diverging predictions from the expected values from the test dataset.

From the plot of the persistence model predictions, it is clear that the model is 1-step behind reality. There is a rising trend and month-to-month noise in the sales figures, which highlights the limitations of the persistence technique.

```
# plot predictions and expected results
from pandas import concat
from matplotlib import pyplot

pyplot.plot(train_y)
pyplot.plot([None for i in train_y] + [x for x in test_y])
pyplot.plot([None for i in train_y] + [x for x in predictions])
pyplot.show()
```



It assumes nothing about the specifics of the time series problem to which it is applied. This is what makes it so easy to understand and so quick to implement and evaluate.

E. Adding time lags

```
#adding lags
for inc in range(1,13):
    field_name = 'lag_' + str(inc)
    H_supervised[field_name] = H_supervised['diff'].shift(inc)
```

Drop null values

```
#drop null values
H_supervised = H_supervised.dropna().reset_index(drop=True)
```

F. Finding optimum Adjusted R^2

```
# Import statsmodels.formula.api
import statsmodels.formula.api as smf

# Define the regression formula
model = smf.ols(formula='diff ~ lag_1', data=H_supervised)

# Fit the regression
model_fit = model.fit()

# Extract the adjusted r-squared
regression_adj_rsq = model_fit.rsquared_adj
print(regression_adj_rsq)
```

```
0.030465642452062514
```

So what happened above? Basically, we fit a linear regression model (OLS — Ordinary Least Squares) and calculate the Adjusted R-squared. For the example above, we just used lag_1 to see how much it explains the variation in column diff. The output of this code block is:

lag_1 explains 3% of the variation. Let's check out others:

```
# Import statsmodels.formula.api
import statsmodels.formula.api as smf

# Define the regression formula
model = smf.ols(formula='diff ~ lag_1 + lag_2 + lag_3 + lag_4 ', data=H_supervised)

# Fit the regression
model_fit = model.fit()

# Extract the adjusted r-squared
regression_adj_rsq = model_fit.rsquared_adj
print(regression_adj_rsq)
```

```
0.13363362471677465
```

Adding four more features increased the score from 3% to 13%.
How is the score if we use the entire feature set:

```
# Import statsmodels.formula.api
import statsmodels.formula.api as smf

# Define the regression formula
model = smf.ols(formula='diff ~ lag_1 + lag_2 + lag_3 + lag_4 + lag_5 + lag_6 + lag_7 + lag_8 + lag_9 + lag_10 + lag_11 + lag_12
#
# Fit the regression
model_fit = model.fit()

# Extract the adjusted r-squared
regression_adj_rsq = model_fit.rsquared_adj
print(regression_adj_rsq)
```

```
0.7907160335318657
```

```
The result is impressive as the score is 79%. Now we can confidently build our model after scaling our data. But there is one
more step before scaling. We should split our data into train and test sets. As the test set, we have selected the last 6
months' Spend.
```

G. Train-Test-Split

```
#import MinMaxScaler and create a new dataframe for LSTM model
from sklearn.preprocessing import MinMaxScaler
H_model = H_supervised.drop(['Spend','date'],axis=1)
```

```
#split train and test set
train_set, test_set = H_model[0:-6].values, H_model[-6:].values
```

```
H_model.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 47 entries, 0 to 46
Data columns (total 14 columns):
Month     47 non-null object
diff      47 non-null float64
lag_1     47 non-null float64
lag_2     47 non-null float64
lag_3     47 non-null float64
lag_4     47 non-null float64
lag_5     47 non-null float64
lag_6     47 non-null float64
lag_7     47 non-null float64
lag_8     47 non-null float64
lag_9     47 non-null float64
lag_10    47 non-null float64
lag_11    47 non-null float64
lag_12    47 non-null float64
dtypes: float64(13), object(1)
memory usage: 5.3+ KB
```

H. Scaling the data

```
#apply Min Max Scaler
scaler = MinMaxScaler(feature_range=(-1, 1))
scaler = scaler.fit(train_set)
# reshape training set
train_set = train_set.reshape(train_set.shape[0], train_set.shape[1])
train_set_scaled = scaler.transform(train_set)

# reshape test set
test_set = test_set.reshape(test_set.shape[0], test_set.shape[1])
test_set_scaled = scaler.transform(test_set)
```

I. Building LSTM Model

```
X_train, y_train = train_set_scaled[:, 1:], train_set_scaled[:, 0:1]
X_train = X_train.reshape(X_train.shape[0], 1, X_train.shape[1])
```

```
X_test, y_test = test_set_scaled[:, 1:], test_set_scaled[:, 0:1]
X_test = X_test.reshape(X_test.shape[0], 1, X_test.shape[1])
```

J. Fit the LSTM model

```
model = Sequential()
model.add(LSTM(4, batch_input_shape=(1, X_train.shape[1], X_train.shape[2]), stateful=True))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')
model.fit(X_train, y_train, nb_epoch=100, batch_size=1, verbose=1, shuffle=False)
```

```
Epoch 92/100
41/41 [==============================] - 0s 1ms/step - loss: 0.0103
Epoch 93/100
41/41 [==============================] - 0s 1ms/step - loss: 0.0102
Epoch 94/100
41/41 [==============================] - 0s 1ms/step - loss: 0.0101
Epoch 95/100
41/41 [==============================] - 0s 1ms/step - loss: 0.0100
Epoch 96/100
41/41 [==============================] - 0s 2ms/step - loss: 0.0099
Epoch 97/100
41/41 [==============================] - 0s 2ms/step - loss: 0.0097
Epoch 98/100
41/41 [==============================] - 0s 2ms/step - loss: 0.0096
Epoch 99/100
41/41 [==============================] - 0s 1ms/step - loss: 0.0095
Epoch 100/100
41/41 [==============================] - 0s 2ms/step - loss: 0.0094

<keras.callbacks.callbacks.History at 0x21ff63e7b88>
```

The code block above prints how the model improves itself and reduce the error in each epoch:

K. Use the model for prediction

```
y_pred = model.predict(X_test,batch_size=1)
```

```
y_pred
```

```
array([[ 0.2913093 ],
       [ 0.0895184 ],
       [-0.01292165],
       [ 0.4963987 ],
       [ 0.5663966 ],
       [ 1.0344739 ]], dtype=float32)
```

```
y_test
```

```
array([[0.09090909],
       [0.27272727],
       [0.45454545],
       [0.63636364],
       [0.81818182],
       [1.        ]])
```

Results look similar but it doesn't tell us much because these are scaled data that shows the difference. How we can see the actual Spend prediction? First, we need to do the inverse transformation for scaling:

```
#reshape y_pred
y_pred = y_pred.reshape(y_pred.shape[0], 1, y_pred.shape[1])
```

```
#rebuild test set for inverse transform
pred_test_set = []
for index in range(0,len(y_pred)):
    print (np.concatenate([y_pred[index],X_test[index]],axis=1))
    pred_test_set.append(np.concatenate([y_pred[index],X_test[index]],axis=1))
```

L. Inverse Transform

```
#reshape pred_test_set
pred_test_set = np.array(pred_test_set)
pred_test_set = pred_test_set.reshape(pred_test_set.shape[0], pred_test_set.shape[2])
```

```
#inverse transform
pred_test_set_inverted = scaler.inverse_transform(pred_test_set)
```

Second, we need to build the dataframe has the dates and the predictions. Transformed predictions are showing the difference. We should calculate the predicted Spend numbers:

```
#create dataframe that shows the predicted Spend
result_list = []
Spend_dates = list(H_Spend[-7:].date)
act_Spend = list(H_Spend[-7:].Spend)
for index in range(0,len(pred_test_set_inverted)):
    result_dict = {}
    result_dict['pred_value'] = int(pred_test_set_inverted[index][0] + act_Spend[index])
    result_dict['date'] = Spend_dates[index+1]
    result_list.append(result_dict)
H_result = pd.DataFrame(result_list)
```

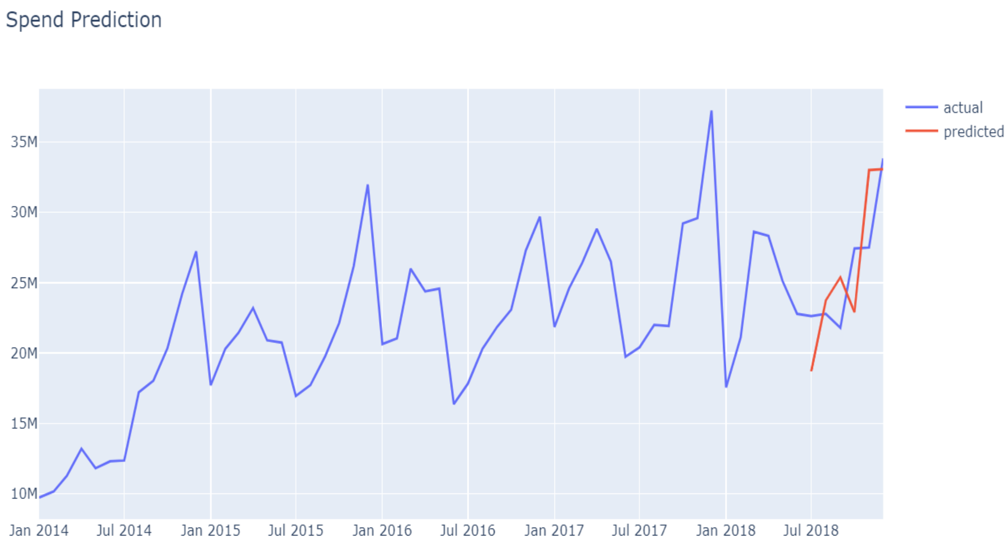
```
H_result
```

M. Plot Actual and Predicted data for Spend forecast

```
#plot actual and predicted
plot_data = [
    go.Scatter(
        x=H_Spend_pred['date'],
        y=H_Spend_pred['Spend'],
        name='actual'
    ),
        go.Scatter(
        x=H_Spend_pred['date'],
        y=H_Spend_pred['pred_value'],
        name='predicted'
    )

]

plot_layout = go.Layout(
        title='Spend Prediction'
    )
fig = go.Figure(data=plot_data, layout=plot_layout)
pyoff.iplot(fig)
```

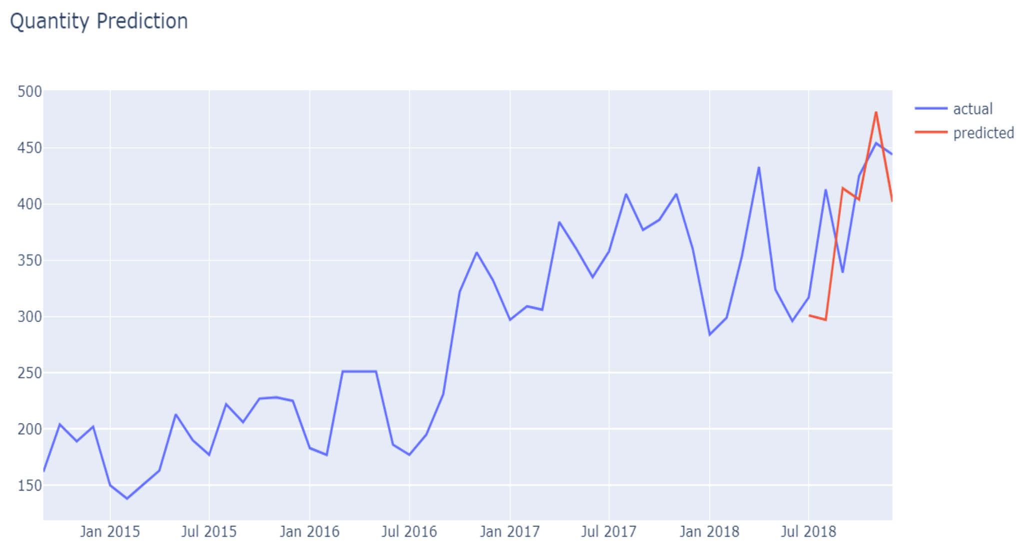


N. Plot Actual and Predicted data for Quantity forecast

```
#plot actual and predicted
plot_data = [
    go.Scatter(
        x=H_Spend_pred['date'],
        y=H_Spend_pred['quantity'],
        name='actual'
    ),
        go.Scatter(
        x=H_Spend_pred['date'],
        y=H_Spend_pred['pred_value'],
        name='predicted'
    )

]

plot_layout = go.Layout(
        title='Quantity Prediction'
    )
fig = go.Figure(data=plot_data, layout=plot_layout)
pyoff.iplot(fig)
```



V. Survey Consent Form

Dear Participant,

The purpose of the survey research is to evaluate the usefulness of the proposed e-procurement system. This is a research project titled **"Integration of Artificial Intelligence and application of Machine Learning in e-procurement of hospitality industry"** conducted by Elezabeth Mathew – Ph.D. student at the British University in Dubai. You are invited to participate in this research project because you are an end-user of the system of the proposed system.

Your participation in this research study is voluntary. You may choose not to participate. If you decide to participate in this research survey, you may withdraw at any time. If you decide not to participate in this study or if you withdraw from participating at any time, you will not be penalized.

The procedure involves filling an online survey that will take approximately 10 minutes. Your responses will be confidential and we do not collect identifying information such as your name, email address, or IP address. The survey questions will be about usefulness, ease of use, system performance about the proposed e-procurement system.

We will do our best to keep your information confidential. All data is stored in a password protected electronic format. To help protect your confidentiality, the surveys will not contain information that will personally identify you. The results of this study will be used for scholarly purposes only and may be shared with the British University in Dubai representatives.

If you have any questions about the research study, please contact Elezabeth Mathew at 2016152110@student.buid.ac.ae . This research has been reviewed and is approved by the Research Ethics Committee of The British University in Dubai. If you wish to discuss any issues relating to this study you can contact me or my supervisor and director of studies Prof. Sherief Abdallah {sherief.abdallah@buid.ac.ae and office telephone number: +971 4 279 1435 }.

If you do wish to participate in the research study, please click the link below.

https://www.surveymonkey.com/r/S2TCDHT

Many Thanks in Advance. Please feel free to forward the survey to anyone related to procurement processes.

Regrads,
Elezabeth

VI. Survey Design

Integration of AI in e-procurement of hospitality industry

⊕ PAGE TITLE

The research proposes a new framework integrated with Artificial Intelligence (AI) in the e-procurement system in the hospitality industry. The system has several agents (like a program) specifically for specific purposes described below. The integrated system will be having interconnectivity and interoperability between all stakeholders of the procurement process. This survey is particularly focusing on the given four agents.

Assumptions: The proposed integrated centralized e-procurement system has standardized policies.

A. Authorization & Security Agent to keep the data encrypted to maintain privacy and confidentiality. This agent limits the person from accessing only the documents/modules that they are authorized for. This includes user identities, group identity, role information, location information, actions, and parameters associated with the actions.

B. Catalog Agent to maintain the system catalogs. Whenever there is a commodity added/edited/deleted by a particular supplier this agent is in charge of notifying the Price regulator. And once the change has been authorized by the Price regulator, the same gets updated in the catalog accessible to the procurement stakeholders. Also, this allows the supplier to get notified of whether their changes have been accepted.

C. Risk Assessment Agent which will monitor the risks such as currency fluctuation rates/pandemic alerts, etc. and augment all related stakeholders to arrange for the mismatch in requirements/resources.

D. Optimization Agent which will continue to monitor decision taken to optimize the requirements and preempt alternative solutions to the given scenario. It uses machine learning on demand forecasting for inventory, spends and purchases duration, etc. The Optimization agent has a continuous improvement feature to its responses in any given scenario based on its previous learnings from real-time responses or outcomes.

You may please proceed to the questionnaire:

A. Demographic details

* 1. Please provide your email addresses:

email:

* 2. What industry do you work for?

- ◯ Hospitality
- ◯ Education
- ◯ Automobile
- ◯ Travel and Tourism

Other (please specify)

* 3. Are you a part of the procurement process? (directly or indirectly)

- ◯ Yes
- ◯ No

* 4. What is the total number of employees in your company (including all locations where your employer operates)?

- ◯ 10-100
- ◯ 100-250
- ◯ 250-400
- ◯ 400-600
- ◯ above 600

B. Importance of the integrated system

* 5. How important is authorization & security agent in your opinion?

Not Important	Less Important	Neutral	Somewhat Important	Very Important
★	★	★	★	★

Please explain why this star is given.

* 6. How important is catalogue agent in your opinion?

Not Important	Less Important	Neutral	Somewhat Important	Very Important
★	★	★	★	★

Whys this star rating is given?

* 7. How important is risk assessment agent in your opinion

Not Important	Less Important	Neutral	Somewhat Important	Very Important
★	★	★	★	★

Please explain why this star rating is given

* 8. How important is Optimization agent in your opinion?

Not Important	Less Important	Neutral	Somewhat Important	Very important
★	★	★	★	★

Please justify your rating

C. Knowledge and recommendation

* 9. Concerning technical work in artificial intelligence, how would you describe your own expertise?

0(None) Average 100(Expert)

* 10. How likely are you to recommend such an integrated system with the agents to your company?

0 (Never) May be 100 (Definitely)