PREDICTIVE ANALYSIS OF RETAIL SALES: FURNITURE AND HOME FURNISHINGS

OLUWATOBI EKUNDAYO
x19173105
MSc. Data Analytics, National College of Ireland

I. INTRODUCTION

Sales of furniture products in the retail industry has had a gradual growth in consumers demand for furniture products over the years. This increase can be attributed to an increase in consumers personal income, increase in urban population, increase in corporate business establishments or simply customers wants. The benefit of implementing predictive analysis in this domain is to gain competitive advantage in the retail industry by leveraging on analytical insights provided. Better decisions and strategic plans can thus be implemented by the decision makers to increase retail sales of furniture products.

The ***great recession*** [1] which officially took place between December 2007 and June 2009 in USA, had a ripple effect across world causing a global financial crisis. The retail industry was greatly impacted due to this crisis. It took about 4-5 years for the retail sales in furniture to fully recover from this crisis [2]. In January 2020, a pandemic (Coronavirus, known as COVID-19) broke out in Wuhan, China, and spread rapidly across the world [3]. As a result of this global pandemic bring some many business activities have been slowed down and, in some cases, a shutdown. Thus, causing a recession (***coronavirus recession*** or ***the great lockdown***) to resurface. The international monetary funds (IMF) projects that this recession will have more impact than that of the great recession [4][5].

1. BACKGROUND

The Sales and Marketing team have been tasked to come up with strategies and business models to recover from the coronavirus recession. To strategically plan for recovery from the impact of the recession, both teams of IKEA, (a major furniture retail company) requires qualitative and predictive analysis to provide them with sufficient insights.

As suggested by Yucesan et al [6], the cost and benefits of a forecasting model should be evaluated by the decision makers of the business domain before selecting the model for use. In the research they carried out, they applied artificial neural network (ANN) model to monthly sales data for a corporate furniture manufacturing company. Their results showed that bayesian training (a component of ANN model) can be applicable for forecasting a periodic data (such as the monthly sales data used). The research on forecasting retail demand by Fildes et al [7] revealed the retailers need for forecasts. They identified the retailers forecasting needs on a strategic level, tactical level, and operational level. They further examined aggregate sales forecasting on a market-level, characteristics of retail sales data on a product-level, drivers of demand and forecasting demand for new products in retail for managerial decisions

Aras et al [8] carried out sales forecast on furniture retail sales data provided by a global furniture retailer located in Turkey. They implemented five models (ARIMA model, AFRIMA model, ANFIS model, ANN model and state space model) and compared results of the models against each other. Their conclusion was that performance of the individual models was relatively the same. However, combining the time series models could yield a better sales forecast result. The analysis carried out by Ansuj et al [9] on sales behaviour of a medium sized enterprise in brazil involved the use of two time series models, neural network back propagation model and ARIMA model. Their forecast result showed that the accuracy for the back-propagation model was better than that of ARIMA model.

Zhang [10] recommended that a combination of autoregressive integrated moving average (ARIMA) model and artificial neural network (ANN) model will yield a better performance than using the individual models alone. The methodology of model combination is effective in improving the accuracy of the forecasts. The performance of ARIMA model and State Space model was compared in the study by Ramos et al [11] for predicting retail sales. They however concluded that for both multi-step forecasts and one step forecasts, both models will have similar performance in prediction.

2. SCOPE

This project will evaluate business values that can be obtained from insights gained from predictive analysis. Applicable model selected for prediction of retail sales of furniture will be implemented. The predicted results (forecasts) and summary of analysis, will aid planning and decision making for the business strategies to be implemented in the business domain.

Also, the ethical concerns with regards to carrying out predictive analysis in this business domain will be highlighted, to serve as a guide for taking further business decisions.

3. OBJECTIVES

- To predict retail sales of furniture for the next 2 years (24 months)
- To critically analyse the historical data to extract past, current, and futuristic insights of significant business value for the business domain (furniture retail sales).
- Evaluate different models applicable to predicting retail sales of furniture.
- Develop a model with significant accuracy for predicting furniture retail sales

4. RESEARCH QUESTION

- Will there be an increase or decrease in retail sales of furniture products for the next 2 years (24 months), Post-Covid19?

5. HYPOTHESIS

- The impact of coronavirus recession will result in poor sales performance for **only** the next six months.

6. DATASET

- The first piece of dataset termed "Dataset 1", contains 322 records of time series seasonal data. This dataset has a periodic series information of annual retail sales of furniture from 1992 to 2020.
  Source:
  https://fred.stlouisfed.org/series/RSFHFS
- The second piece of dataset termed "Dataset 2", contains 3694 records and 14 attributes of retail sales of IKEA furniture products for the year 2020.
  Source:
  https://www.kaggle.com/ahmedkallam/ikea-sa-furniture-web-scraping

II. ETHICAL CONCERNS

Reviewing ethics based on historical data is key for the business, to prevent violating laws that could impact the business revenue generation and public image.

a) Confidentiality

Information provided in confidence should be respected. The sensitivity of information should be protected diligently. The global impact of the EUs General Data Protection Regulation (GDPR) [12] is significant and must be taken seriously. The GDPR concerns both EU and Non-EU Business enterprises or organizations who operate with EU customers. Compliance to the ethics of this regulatory body should be adhered to.

b) Fairness

Collection of data should not be biased. There should be no discrimination [13]. Also, the treatment of humans equally should be considered in the implementation of a strategy based on derived analysis and insights. Avoid engaging in sales or marketing strategy that conflicts with interest of other competitors. Any sales tactics such as "price gouging" and "bait-and-switch" [14], commonly used in retailing should not be encouraged or used.

c) Integrity and Honesty

Strategies to be implemented to increase sales should not be misleading or over-exaggerate quality, function, or purpose of products. Commitments and actions taken based on insights from the analysis should be done truthfully [14]. Also, an honest approach should be taken in data processing to produce predictive results.

d) Responsibility

It is required that full responsibility be taken for any strategy and decision made to foster retail sales of furniture products [14]. They should be willing to accept the consequences of such actions. Responsibility must be taken for products that fails to deliver its marketed functions.

e) Accountability

The procedures, methods, technical measures put in place by the organisation should adhere to the policies and must effectively demonstrate all that was done with the use of the data[12][15].

f) Openness or Transparency

The policies by GDPR [12] demands best practices of transparency by ensuring complete information being communicated to the customers about use of the data. Also, for any strategy to be implemented, it is important that the price deals, product description and other relevant details about the furniture products should be stated clearly [14]. Deceptive promotions or misleading information should be avoided.

g) Consent

Customers must expressly be informed before their obtaining their consent irrespective of the channel used [15]. Upon initial offer, the terms & conditions, the right to cancel or optout from any plan and period of offer must be communicated clearly to the customer. Customers consent must be obtained for any billing, additional charge, or participation in any plan.

III. STRATEGY FOR ANALYSIS

To successfully carry out predictive analysis with the sourced dataset, the strategy presented below will be implemented.

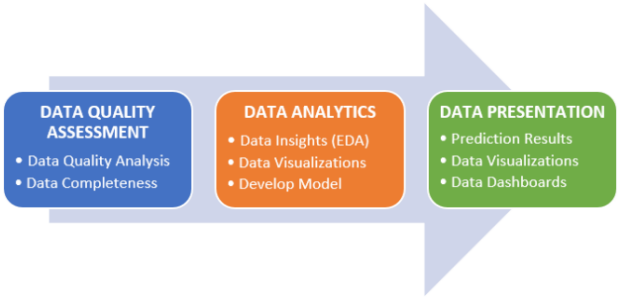

Fig. 1. Strategy to analyse the dataset

1. Data Quality Assessment

This initial stage is very essential for predictive analysis. It requires assessing the quality of the data to be used. In this stage, the completeness (percentage of missing data) of the dataset will be examined. Also, variables suitable for analysis (SFA) will be identified. Quality assessment of the data taken at this stage, will be shared with the business or retail enterprise, to provide them with information about the quality of their

data. This can help them improve on their channel of data collection and/or acquire more relevant data to carry out further analysis, to provide business insights.

TABLE I
QUALITY ASSESSMENT 1

Quality Assessment 1				
	Variables	Missing Data	Count	SFA
DATASET 1	Date	No		Yes
	Sales (millions in $)	No		Yes
DATASET 2	item_id	No		No
	link	No		No
	name	No		Yes
	category	No		Yes
	designer	Yes	143	Yes
	price	No		Yes
	old_price	No		No
	sellable_online	No		Yes
	other_colors	No		Yes
	short_description	No		No
	dimension	Yes	1382	No
	depth	Yes	1463	No
	height	Yes	988	No
	width	Yes	589	No

Table I assesses the variables of the sourced dataset to check for missing data or values. The count of the missing data for the corresponding variable is measured. Also, in preparation for analysis, the suitable variables for analysis (SFA) are identified.

TABLE II
QUALITY ASSESSMENT 2

Quality Assessment 2		
		Count
DATASET 1	Years of Sales Data (1992-2020)	29
	Months of Sales Data (1992-2020)	341
DATASET 2	Levels of Category	17
	Number of Designers	218
	Total Number of Sold Products	3694
	Products with Designers	3551

Table II examines the counts and levels of key variables to provide information about the dataset before carrying out predictive analysis.

2. Data Analytics

This stage involves carrying out an exploratory data analysis to extract data insights. Hence, some data visualizations will be created to understand and interpret the data. Some business insights identified at this stage will also be shared with the business management as shown in this document. The last step in this stage will involve developing a model using a selected technique for prediction.

3. Data Presentation

This final stage will involve analysing and evaluating the predicted results. After justification, a summary of the predicted results is prepared for presentation. Also, as part of the data presentation, a data dashboard will be prepared to harmonize the visualizations prepared for data insights. This dashboard will be interactive in nature and will aid communicating business insights of value to the business.

IV. PRELIMINARY VISUALIZATIONS

One of the main aims of predictive analysis is to gain both past and futuristic insights of the data and extract relevant information to make decision about actions to take for the future. Data insights and visualizations are one of the main steps under the data analytics stage for this project.

Developed preliminary visualizations and data insights are highlighted below.

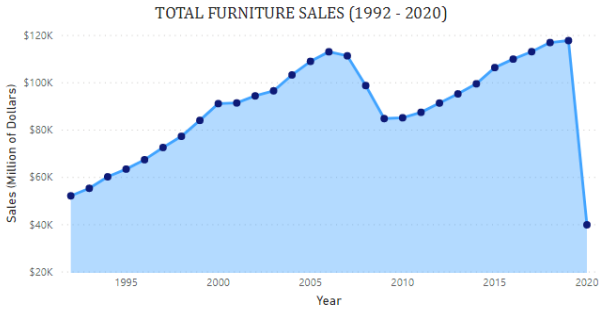

Fig. 2. Total furniture retail sales from January 1992 to May 2020

Fig. 2. Shows the trend in retail sales of furniture captured from January 1992 to May 2020. The trend shows some relative increase in retail sales except for the decline between 2007 and 2010. This decline is attributed to the great recession in US, between 2007 and 2009 [2].

The retail sales of furniture gradually recovered and by 2015, it began to maintain a steady increase. However, by 2020 the trend shows a great decline in sales. This decline is attributed to the current coronavirus regression. The trend also shows that the drop in sales is the worst seen since 1992 for retail sales in furniture.

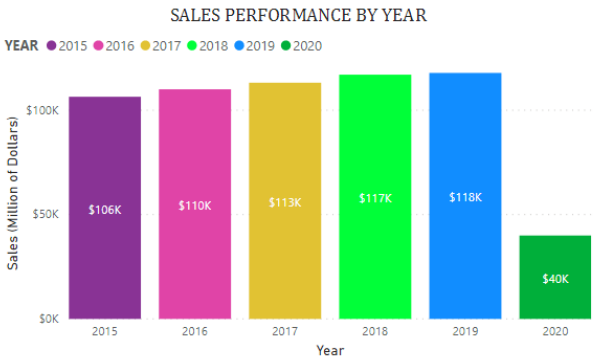

Fig. 3. Bar chart of Sales from 2015 to 2020

Fig. 3. Shows the performance of retail sales of furniture from 2015 to 2020. Sales performance for the last 5 years is compared with current sales performance in year 2020. The sales performance for the captured years shows that in the history of the retail sales of furniture, 2018 and 2019 peaked in sales.

Based on the recovered growth trend in sales since 2010, it is expected that 2020 should also have an increase in sales (following the growth trajectory). Unfortunately, due to the coronavirus recession, retail sales of furniture will most likely have a great loss in sales.

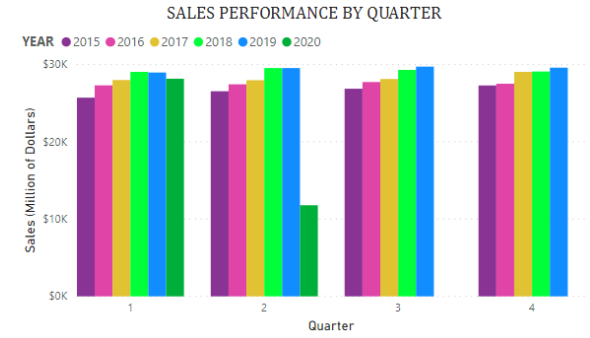

Fig. 4. Bar chart comparison of sales by year and by quarter

Fig. 4. Shows a drill down to the quarter level to examine the sales performance from 2015 to 2020. The comparison shows that in the first quarter (Q1) for Year 2020, the sales dropped, compared to steady increase seen since 2015. The second quarter (Q2) reveals the extent of the impact of the coronavirus recession on retail sales of furniture.

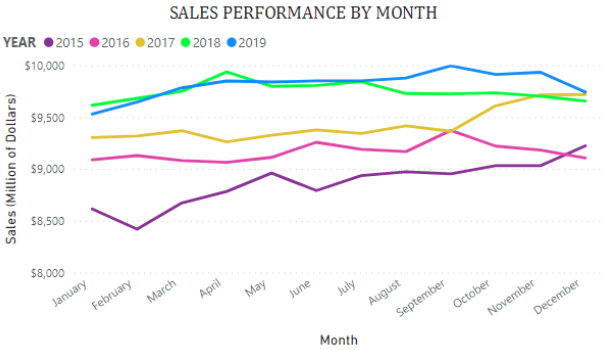

Fig. 5. Line chart comparison of sales by year and by month

Fig. 5. Shows a drill down to the month level to examine critically the sales performance from 2015 to 2019. The comparison of the sales performance by year on a monthly level for the last five years presents a concise overview of monthly sales performance.

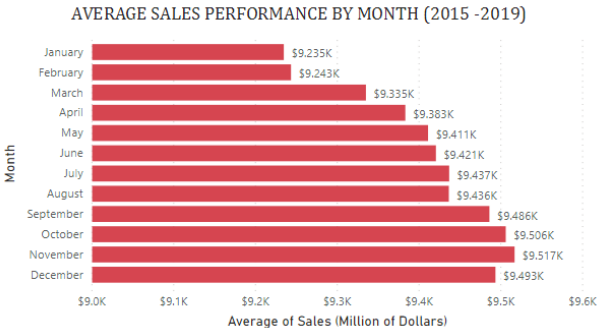

Fig 6: Bar chart of average sales performance by month (2015 to 2019)

Fig. 6. Shows a drill down to the month level to examine critically which month performs better in sales from 2015 to 2019. By taking the average of sales for each month across a span of 5 years (2015 – 2019), it is shown that more sales are being made from September to December. Also, the peak of retail sales of furniture occurs in the month of November.

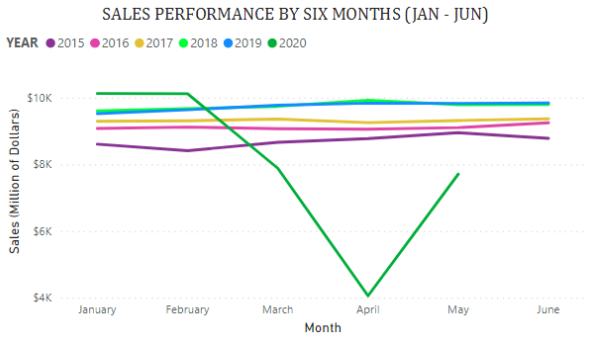

Fig 7: Line chart comparison for the first six months of Year 2015 to Year 2020

Fig. 7. Shows a drill down to six months (January to June) to examine critically the sales performance from 2015 to 2020. This visual shows the drastic drop in sales from February to April 2020 when the pandemic was at its peak. With the “Lockdown” gradually being eased, we can see a gradual recovery in sales.

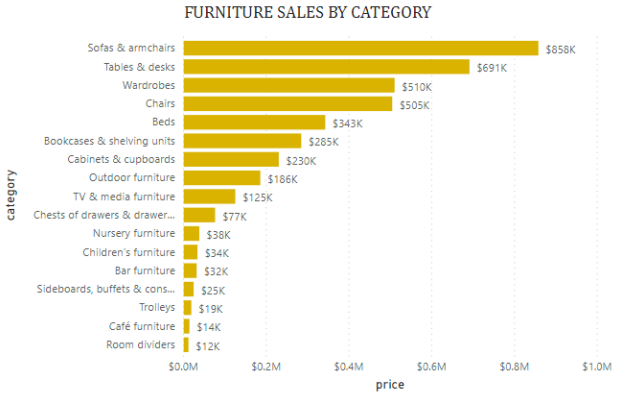

Fig. 8. Stacked bar chart of IKEA furniture sales in Year 2020 by category

Fig. 8. Shows the sales performance of all 17 categories of the furniture products sold so far for the year 2020. This shows immediately the categories with better performance in sales during this pandemic period. The *Sofa & armchairs* and *Tables & desks* category generated a combined sale of $1.55 million out of the total sales of $3.99 million for year 2020.

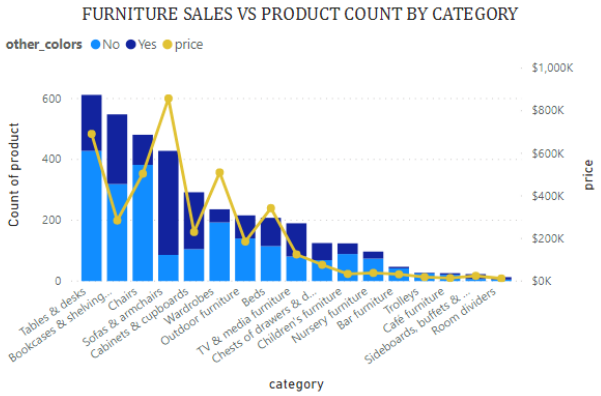

Fig. 9. Line and Stacked column chart showing furnture sales vs count of products sold by category in Year 2020

Fig. 9. Shows the relationship between the count of the products sold and sales generated based on furniture category. Sofas & armchairs category recorded the highest sales ($858,000), however it had the fourth (4th) largest count of products sold.

This can be attributed to the fact that more people stay at home longer than they have ever been. Hence, the need for them to purchase for their comfort in the home. Also, it shows that irrespective of the price tag of the products under this category, there is a good demand.

The furniture category (Tables & desks) recorded the second highest sales ($691,000) in year 2020. It however had the highest count of products sold. This shows there is a very high demand for furniture products in this category, which can be attributed to many businesses and organisations having to transition to work from home (Home-Office).

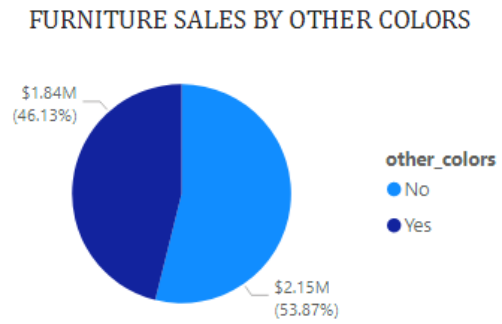

Fig. 10. Pie chart of furniture sales in Year 2020 based on colour preference

Fig. 10 Shows that 53.87% ($2.15 million), of furniture products sold in Year 2020 was of the basic furniture colour (Brown). However, 46.13% ($1.84 million) of the total furniture sales for year 2020 was of products sold in other colours that were made available.

The percentage difference (7.74%) seen in the purchase of furniture based on preference of colour is not so magnanimous. Therefore, it will be in the good interest of the company, to keep making furniture products available in a variety of colours.

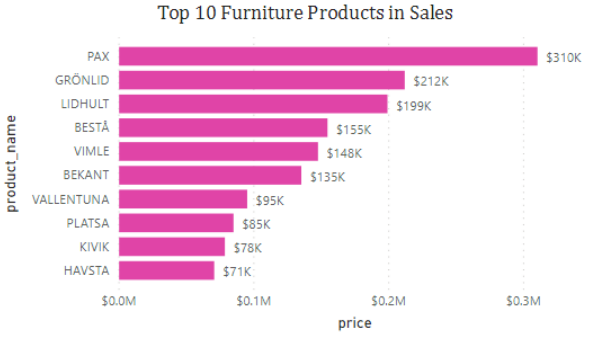

Fig. 11. Bar chart of top 10 furniture products in sale for Year 2020

Fig. 11. Shows the top 10 furniture products sold in year 2020. In the first five months of 2020 (January – May), "*PAX"* yielded $310,000 in sales. The next product "*GRONLID*" yielded $212,000 in sales.

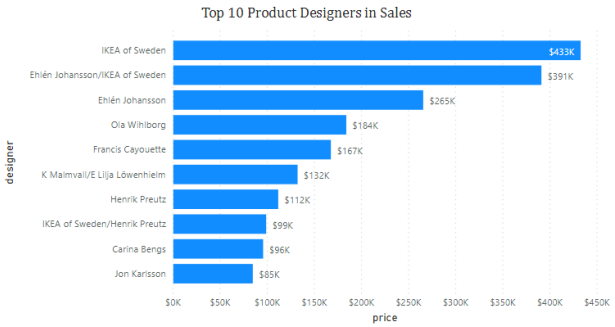

Fig. 12. Bar plot of top 10 furniture designers in sale for Year 2020

Fig. 12. Shows the top 10 furniture designers whose product generated more sales in 2020. The top 3 furniture product designers had sales well-over $200,000. Product designs by "*IKEA of Sweden*" yielded $433,000 in sales. This shows that there is high preference for products by this designer. Also, product designs by the collaboration of "*Ehlén Johansson*" and "*IKEA of Sweden*" made $391,000 in sales, while furniture product designs solely by "*Ehlén Johansson*" made $265,000 in sales.

Fig. 13. Tree map of Top 10 furniture sales in Year 2020 by category

Fig. 13. Provides at a quick glance at the proportion of top 10 furniture category generating the most in sales. Five of the Ten categories are associated with "Home-Office Furniture". This validates some of the emerged industry trends captured by MarketResearch.com [2], the rise of freelancing or working from home (teleworking) as well as demand for luxury furniture.

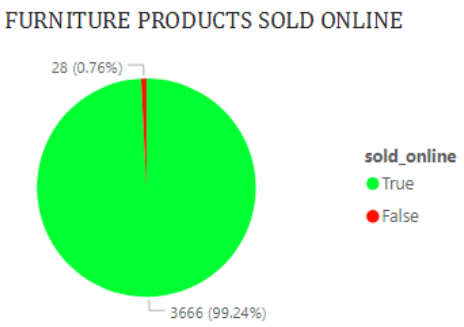

Fig. 14. Pie chart showing count of furniture products sold online

Fig. 14. Shows the percentage and count of total furniture products sold online. 3666 furniture products out of 3694 products was sold online. This shows that 99% of the furniture products was sold online. This is a perfect indication of the use of internet by consumers during this pandemic period. Hence, more strategy can be put in place to maximise the online media space, through use of digital ads, online promotions etc.

V. KEY BUSINESS INSIGHTS

Based on quality assessment and exploratory analysis of the dataset for predictive analysis, some key data insights of business value are highlighted below.

1. Furniture products under the "***Table & desk***" category was purchased the most. Due to the coronavirus pandemic, many businesses and organisations have transitioned to working from home (Home-Office). Hence, a strategic marketing campaign for this category can increase sales of the corresponding and related products.
2. Furniture products under the "***Sofas & armchairs***" category generated the most in sales. This can be attributed to the need for comfort, for the longer time spent at home. Hence, a strategic marketing campaign to appeal to customers need for comfort or luxury, can increase sales of products in this category.
3. Customers preference for furniture products available in other colours should be considered to increase sales. This is because ***46.13%*** of total furniture sales in year 2020 constitutes of furniture products sold in other colours. This percentage is relatively significant.
4. More products designed by Ehlén Johansson, IKEA of Sweden, Ola Wihlborg, and Francis Cayouette should be made available. Also, furniture designs by a collaboration of either of the designers could attract more sales with the right marketing strategy.
5. As part of the plans to recover from the recession, more focus can be given to the top 10 furniture products in sales to sustain the income of the business.

VI. APPLICABLE TECHNIQUES

Highlighted below are reviewed applicable techniques to carry out predictive analysis for this business domain using the time series data.

1. Autoregressive Integrated Moving Average (ARIMA)
2. Artificial Neural Networks (ANN)
3. State Space Model; also known as ETS Model
4. Multivariate Adaptive Regression Splines (MARS)

Also, highlighted below are the tools that will be used to implement the techniques for predictive analysis.

- **R Programming Language:** This tool will be used to carry out the statistical computing and data modelling.
- **Power BI:** This tool will be used to create data dashboard to consolidate the data visualisations used to provide the data insights.

VII. REFERENCES

[1] D. B. Grusky, B. Western, and C. Wimer eds., "The great recession," Russell Sage Foundation, 2011.

[2] C. Stewart, "5 Top Trends in the Furniture Industry," MarketResearch.com, 2016. [Online]. Available: https://blog.marketresearch.com/5-top-trends-in-the-furniture-industry [Accessed: June 25, 2020].

[3] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K.S. Leung, E.H. Lau, J.Y. Wong, and X. Xing, "Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia," New England Journal of Medicine, 2020. [Online]. Available: https://doi.org/10.1056/NEJMoa2001316 [Accessed on: June 14, 2020]

[4] G. Gita, "The Great Lockdown: Worst Economic Downturn Since the Great Depression," International Monetary Fund, Apr. 14, 2020.

[5] J. Geiff, "The Great Recession Was Bad. The 'Great Lockdown' Is Worse," BloombergQuint, Apr. 2020. [Online] Available: https://www.bloombergquint.com/business/global-great-lockdown-will-dwarf-the-great-recession [Accessed on: June 17, 2020]

[6] M. Yucesan, M. Gul and E. Erkan, "Application of Artificial Neural Networks Using Bayesian Training Rule in Sales Forecasting for Furniture Industry," Drvna industrija: Znanstveni časopis za pitanja drvne tehnologije, vol. 68, no. 3, pp. 219–228, 2017.

[7] R. Fildes, S. Ma, and S. Kolassa, "Retail forecasting: Research and practice," International Journal of Forecasting, 2019.

[8] S. Aras, İpek D. Kocakoç, and C. Polat, "Comparative study on retail sales forecasting between single and combination methods," Journal of Business Economics and Management, vol. 18, no. 5, pp. 803–832, 2017.

[9] A. P. Ansuj, M. E. Camargo, R. Radharamanan, and D. G. Petry, "Sales forecasting using time series and

neural networks," Computers & Industrial Engineering, vol. 31, no. 1–2, pp. 421–424, 1996.
[10] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, vol. 50, pp. 159–175, 2003.
[11] P. Ramos, N. Santos, and R. Rebelo, "Performance of state space and ARIMA models for consumer retail sales forecasting," Robotics and computer-integrated manufacturing, vol. 34, pp.151–163, 2015.
[12] M. Goddard, "The EU General Data Protection Regulation (GDPR): European regulation that has a global impact," International Journal of Market Research, vol. 59, no. 6, pp. 703–705, 2017.
[13] R. Vidgen, G. Hindle and I. Randolph, "Exploring the ethical implications of business analytics with a business ethics canvas," European Journal of Operational Research, vol. 281, no. 3, pp.491–501, 2020.
[14] American Marketing Association, "Codes of Conduct, AMA Statement of Ethics," 2020 [Online]. Available: https://www.ama.org/codes-of-conduct/ [Accessed: June 25, 2020].
[15] ANA Data Marketing & Analytics, "Guidelines for Ethical Business Practice," 2020 [Online]. Available: https://thedma.org/accountability/ethics-and-compliance/dma-ethical-guidelines/ [Accessed: June 25, 2020].