Unveiling Market Analysis: Analysing Spending Behaviour and Identifying Opportunities for Growth

DATA ANALYTICS

EXTERNSHIP PROJECT REPORT

By

PARTH GOYAL (**20BCE0621**)

SWATI RAI (**20BCE0996**)

RIA GOEL (**20BCE2001**)

AISWARYA NS (**20BCE2829**)

VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

Vellore Institute of Technology, Vellore

June 2023

1. INTRODUCTION

1.1 Overview

Market research is essential during the planning phases of any start-up. On a basic level, market research is the collection and analysis of data related to a business's target market. Market research can entail everything from information on competitors' products to the interpretation of demographic data related to potential customers.

The main purpose of market research is to gain an understanding of customer needs and wants in an effort to reveal potential business opportunities. When you have a clear picture of what your target market is and what it wants, you can more effectively design your marketing mix to engage that demographic.

Here we are analysing the dataset and making useful visualisation to interpret meaningful outcomes helpful for business and better understanding for the user.

1.2 Purpose

- With our visualisation we are able to visualise vast amounts of data at a glance and in a better way. It helps to understand the data better to measure its impact on the business and communicates the insight visually to internal and external audiences.

- Our visualisation can also help businesses identify which factors affect customer behaviour; pinpoint areas that need to be improved or need more attention; make data more memorable for stakeholders understand when and where to place specific products; and predict sales volumes

- It helps to explore, represent and monitor any information by using visual elements like charts ,graphs and maps and provide an accessible way to see and understand trends, outliers and patterns in data.

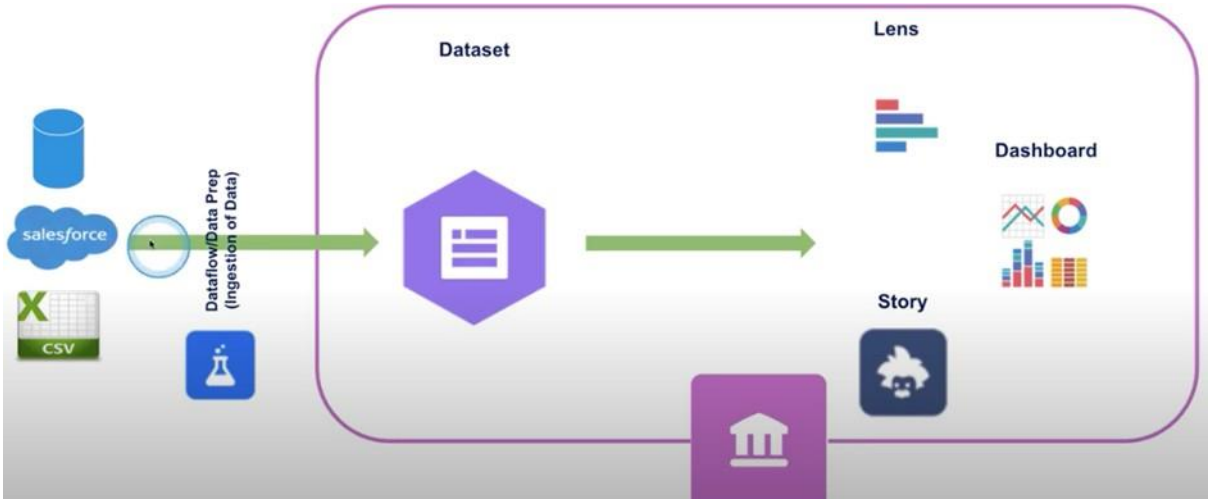
2. LITERATURE SURVEY

Paper Title	Author(s)	Year	Observations
Structural Model for Analysis of Key Performance Indicators for Sustainable Manufacturer-Supplier Collaboration: A grey-DEMATEL based Approach	Kannan Govindan, Aditi, Jyoti Dhingra Darbari, Arshia Kaul, P.C. Jha	2021	With growing competition in the market and dire need for sustainability, it has become imperative for companies to build long-term relationships with their supply chain partners through sustainable collaboration. This study focuses on identification of KPIs for an Indian home appliance company through exhaustive discussions involving multiple decision makers. A Grey based Decision Making Trial and Evaluation Laboratory (DEMATEL) model is proposed in the study for analysing the importance levels among the fifteen KPIs based on multiple stakeholder perspectives.
The measurement of customer satisfaction towards the service quality at xyz wholesale by using fuzzy service quality method	Chalis Fajri Hasibuan	2020	The service quality is the most important factor for the company, in order for the company to keep running and developing. The ways to improve service quality are to find out and understand the customers' desires. This research was conducted at XYZ wholesaler to find out the quality service level and propose the attributes that need to be improved the quality of services. It was concluded that the grocery location was the attribute rated by the customers better than the other attributes and that customer' assessments towards the assurance dimension was better than the others dimension.

The Effects of Agglomeration on Customer Traffic & Commercial Real Estate Values: Evidence from Grocery Store Openings	Franklin Qian, Qianyang Zhang, Xiang Zhang	2023	This paper investigates the driving forces and effects of agglomeration in the local non-tradable service sector on the value of commercial real estate (CRE). The spillovers of demand are strongest between new grocery stores and existing grocery stores and businesses in wholesale and retail. Landlords capitalised on the benefits of agglomeration and increased the rents in newly-signed leases by 23.8% in a half-mile radius from the real grocery store openings in the first two years after the openings. Preliminary analysis suggests that CRE values capitalise on the benefits of agglomeration economies caused by grocery anchors.
Community-Based Recommendations to Improve Customer Turnout at a Non-Profit Grocery Store	Leena Daniel, Sarah Hinman, Bengucan Gunen, Kaitlyn Harper, Lisa Poirier, Joel Gittelsohn	2020	Salvation Army's first non-profit grocery store in Baltimore City has so far failed to attract a large customer base. The paper explored these reasons for low store usage: prices are not low and food quality concerns. In light of these responses, authors made recommendations, including: make displays showing price differences between DMG Foods and other competitors and improve and maintain the quality of produce and meat in store. Additionally, store management should negotiate with wholesalers and distributors to obtain competitive wholesale prices.
Comparing Prophet and Deep Learning to ARIMA in Forecasting Wholesale Food Prices	Lorenzo Menculini, Andrea Marini, Massimiliano Proietti, Alberto Garinei, Alessio Bozza, Cecilia Moretti, Marcello Marconi	2021	This paper examines different techniques to forecast sale prices applied by an Italian food wholesaler, as a step towards the automation of pricing tasks usually taken care by human workforce. They consider ARIMA models and compare them to Prophet, a scalable forecasting tool by Facebook based on a generalised additive model, and to deep learning models exploiting Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs). ARIMA models and LSTM neural networks perform similarly for the forecasting task under consideration, while the combination of CNNs and LSTMs attains the best overall accuracy, but requires more time to be tuned. On the contrary, Prophet is quick and easy to use, but considerably less accurate.

3. THEORETICAL ANALYSIS

3.1 Block Diagram



3.2 Hardware and Software Designing

Hardware Requirements:

- Processor: Intel Core i5 or higher (or equivalent AMD processor)
- RAM: 8 GB or higher (more RAM is beneficial for larger datasets and complex visualisations)
- Hard Disk Space: 1.5 GB of free disk space for installation
- Screen Resolution: Minimum 1366 x 768 pixels (higher resolutions are recommended for better viewing and design experience)
- Internet Connection: Required for accessing online resources and sharing dashboards

Software Requirements:

- Operating System: Tableau Desktop is available for Windows and macOS
- Tableau Desktop: Tableau Desktop is the authoring tool for creating dashboards. It can be downloaded from the Tableau website
- Web Browser: Tableau dashboards can be viewed in web browsers like Google Chrome, Mozilla Firefox, or Microsoft Edge
- Data Sources: Tableau supports various data sources, including databases, spreadsheets, and cloud services

4. EXPERIMENTAL INVESTIGATIONS

Link to Data Set: **https://www.kaggle.com/datasets/thedevastator/analyzing-customer-spending-habits-to-improve-sa**

Total Number of Visualisations: 11
Total Unique Visualisations: 7
Total number of Calculated fields: 4
The dataset contains a hierarchy of product categories and product subcategories which is represented by the sunburst chart showing the distribution of the product categories among customers. The size of each ring is based on the metric quantity which can be helpful to gain insight in the quantity of each product subcategory.

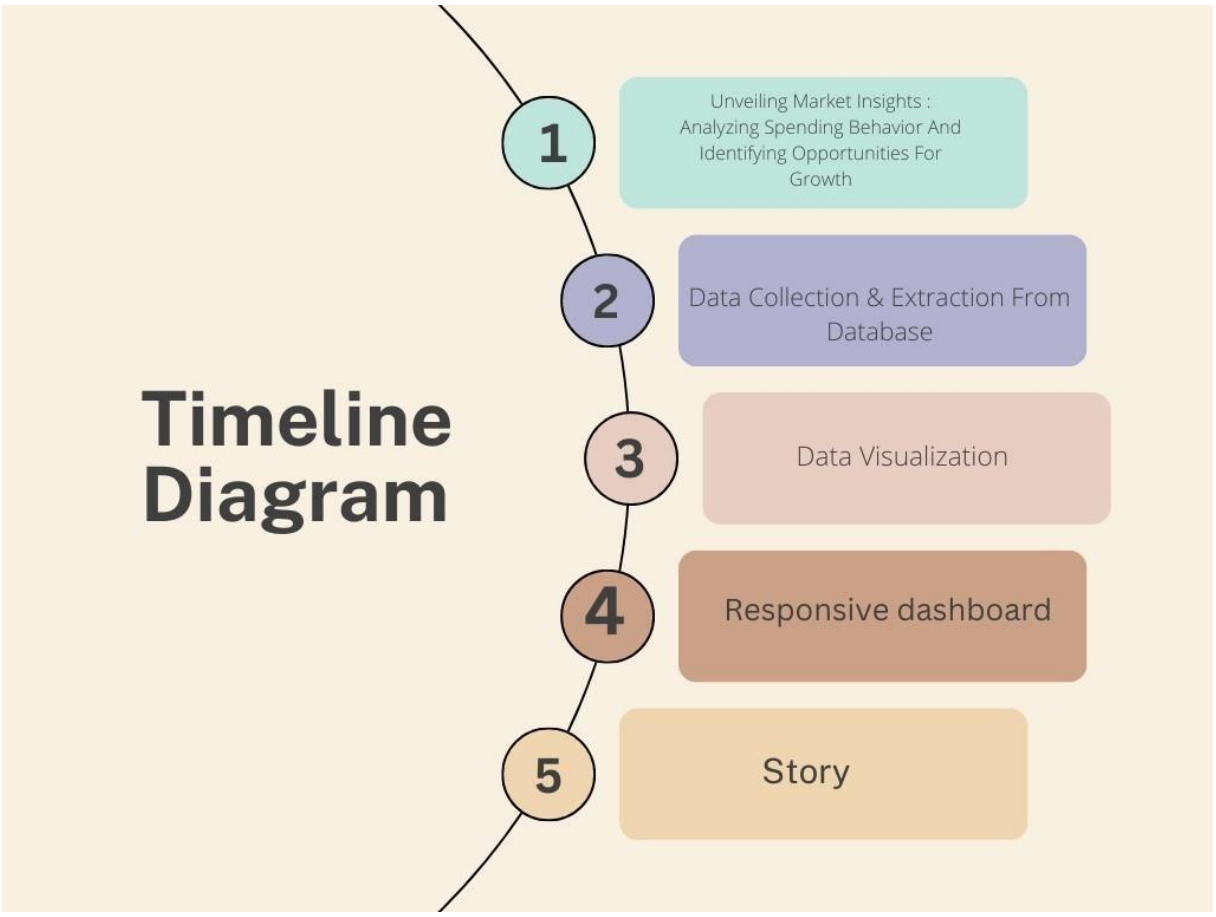
There are various kinds of columns in the used dataset which show the multidimensionality and to represent the same, the best chart would be a heat chart as it is suitable for comparing values across different dimensions and display by using colour encoding.

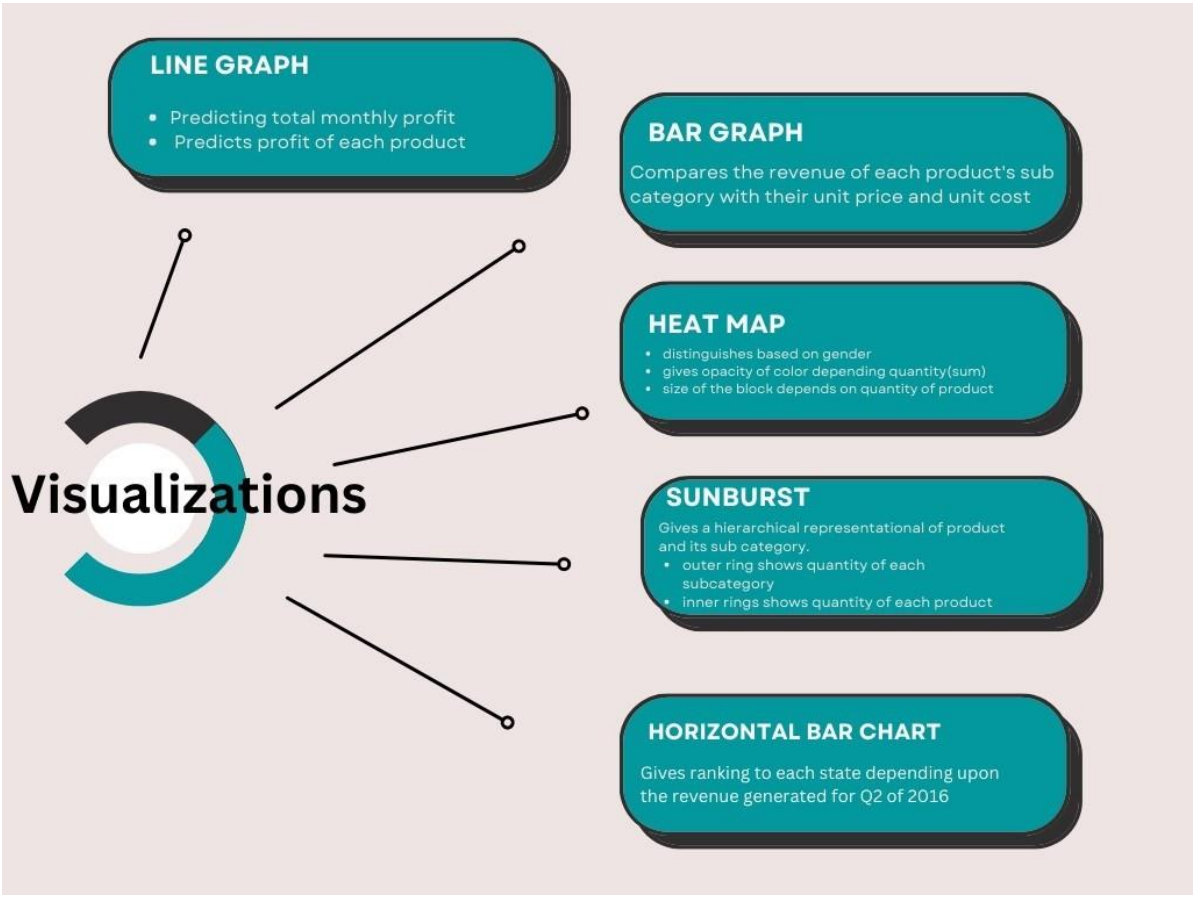
Understanding the dataset, there are a total of 4 quarters of the year 2015 and 2 quarters of the year 2016, and using this timestamp we have shown the variation in the revenue in different states from 2015 to 2016 in quarters. Through this, we observed the evolution of the metrics used and analysed the performance of different states.

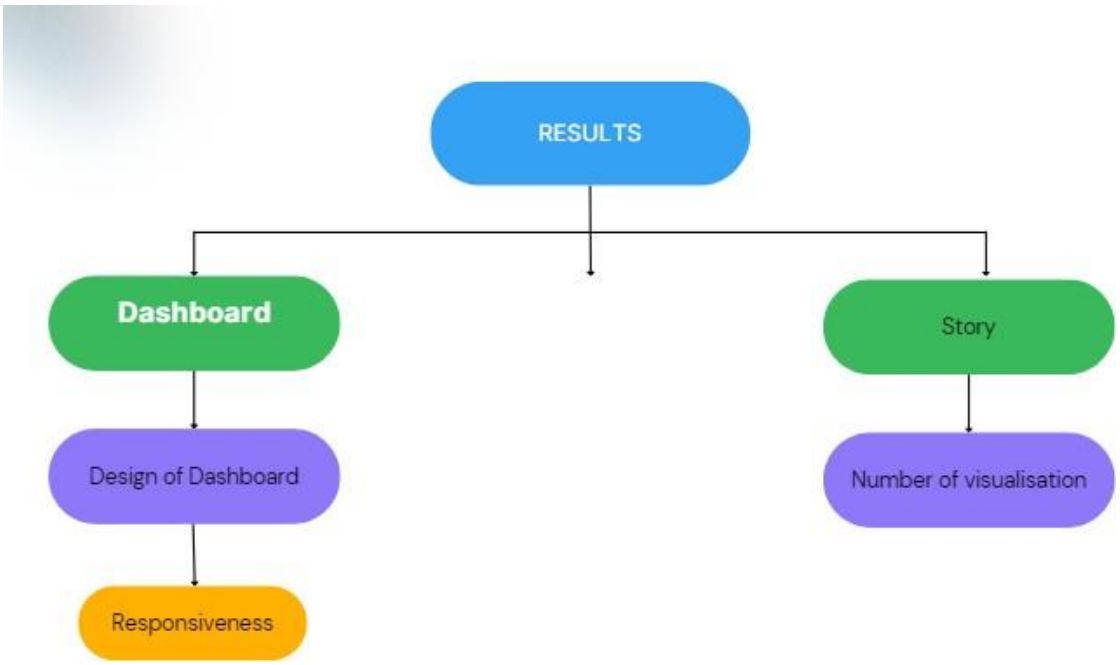
The forecast chart, predicts the values based on the past trend, i,e historical data trends. It is useful for making projections and anticipating future performance. By using this feature of Tableau, we have shown an estimated profit in the year 2016 of each product subcategory.

Overall, these charts provide a range of visualisation options to gain insights from your data. They allow for comparisons, trend analysis, pattern identification, and forecasting, enabling you to make data-driven decisions based on the provided columns.

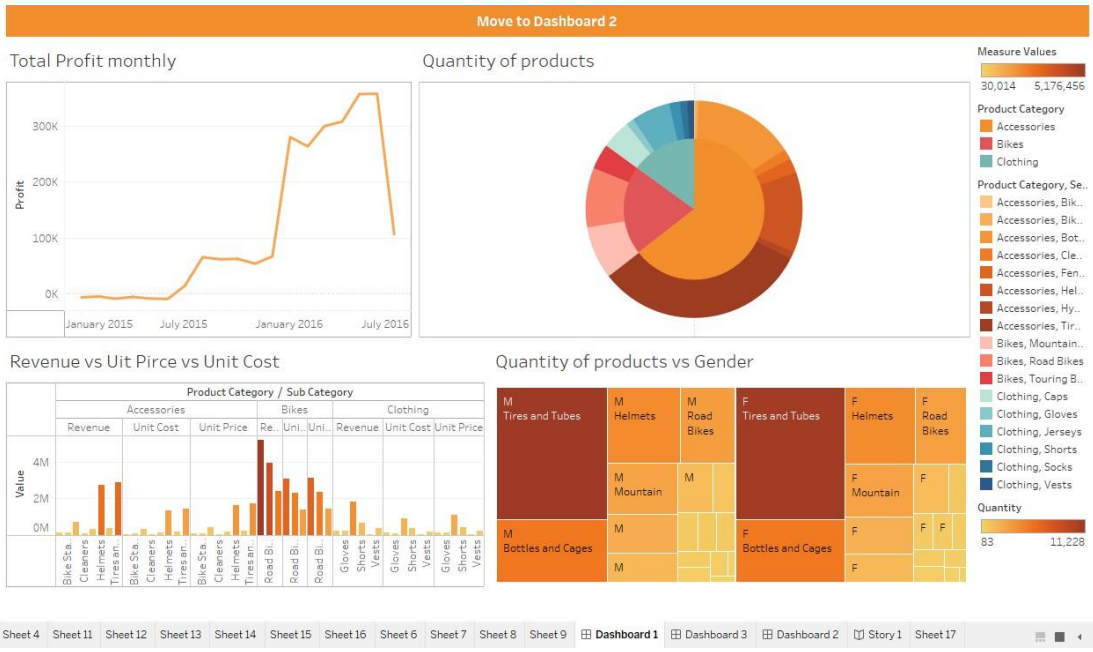
5. FLOWCHART

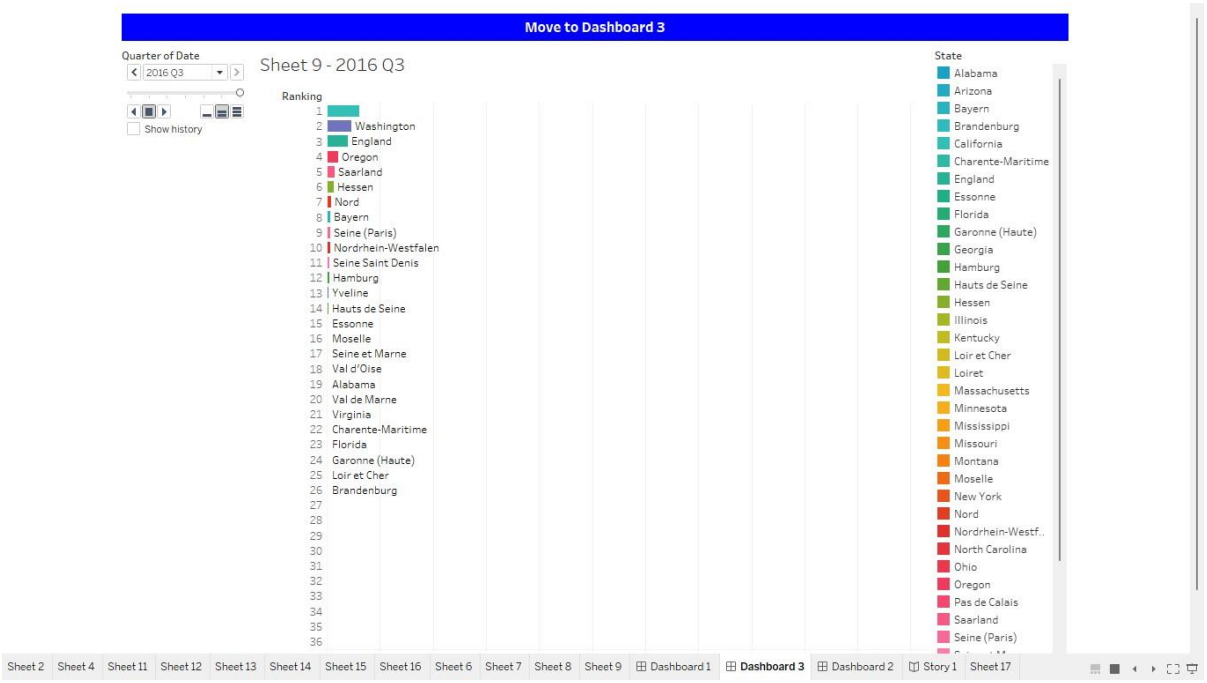

Timeline Diagram
1
Unveiling Market Insights : Analyzing Spending Behavior And Identifying Opportunities For Growth
2
Data Collection & Extraction From Database
3
Data Visualization
4
Responsive dashboard
5
Story

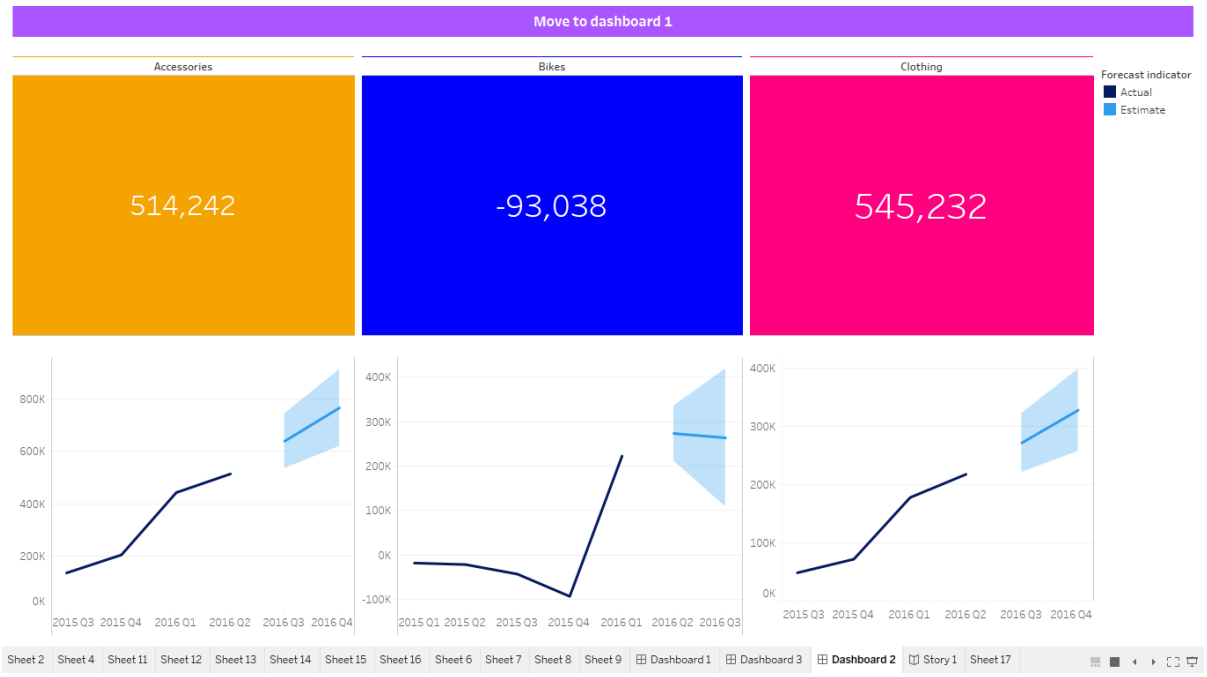

Visualizations
LINE GRAPH
Predicting total monthly profit
Predicts profit of each product
BAR GRAPH
Compares the revenue of each product's sub category with their unit price and unit cost
HEAT MAP
distinguishes based on gender
gives opacity of color depending quantity(sum)
size of the block depends on quantity of product
SUNBURST
Gives a hierarchical representational of product and its sub category.
outer ring shows quantity of each subcategory
inner rings shows quantity of each product
HORIZONTAL BAR CHART
Gives ranking to each state depending upon the revenue generated for Q2 of 2016


RESULTS
Dashboard
Story
Design of Dashboard
Number of visualisation
Responsiveness

6. RESULTS


Move to Dashboard 2
Total Profit monthly
Quantity of products
Measure Values
30,014 5,176,456
Product Category
Accessories
Bikes
Clothing
Product Category, Se..
Accessories, Bik..
Accessories, Bik..
Accessories, Bot..
Accessories, Cle..
Accessories, Fen..
Accessories, Hel..
Accessories, Hy..
Accessories, Tir..
Bikes, Mountain..
Bikes, Road Bikes
Bikes, Touring B..
Clothing, Caps
Clothing, Gloves
Clothing, Jerseys
Clothing, Shorts
Clothing, Socks
Clothing, Vests
Quantity
83 11,228
Profit
300K
200K
100K
0K
January 2015
July 2015
January 2016
July 2016
Revenue vs Uit Pirce vs Unit Cost
Product Category / Sub Category
Accessories
Bikes
Clothing
Revenue
Unit Cost
Unit Price
Value
4M
2M
0M
Quantity of products vs Gender
M Tires and Tubes
M Helmets
M Road Bikes
M Mountain
M Bottles and Cages
F Tires and Tubes
F Helmets
F Road Bikes
F Mountain
F Bottles and Cages
Sheet 2
Sheet 4
Sheet 11
Sheet 12
Sheet 13
Sheet 14
Sheet 15
Sheet 16
Sheet 6
Sheet 7
Sheet 8
Sheet 9
Dashboard 1
Dashboard 3
Dashboard 2
Story 1
Sheet 17





7. ADVANTAGES AND DISADVANTAGES

Advantages:

- With our visualisation we are able to visualise vast amounts of data at a glance and in a better way. It helps to understand the data better to measure its impact on the business and communicates the insight visually to internal and external audiences.

- Our visualisation can also help businesses identify which factors affect customer behaviour; pinpoint areas that need to be improved or need more attention; make data more memorable for stakeholders understand when and where to place specific products; and predict sales volumes

- It helps to explore, represent and monitor any information by using visual elements like charts ,graphs and maps and provide an accessible way to see and understand trends, outliers and patterns in data.
- One can understand interpretation at glance
- Easy to understand

Disadvantages:

- incomplete/null data values give incorrect results
- Incorrect data warehouse gives unreliable visualisation,its future trends cannot be trusted
- Large dataset is difficult to export/import
- Small data set have limitations with respect to analysis and visualisation
- It gives assessment not exactness – While the information is exact in foreseeing the circumstances, the perception of similar just gives the assessment.

8. APPLICATIONS

The results derived from the proposed visualisation dashboard can be applied in various business and economic decision making processes of a company. These visuals can be studied to analyse trends and patterns related to market and customer mood. The organisation can forecast their profit based on previous revenue and sales data findings .The visualisations can be applied to understand demand versus supply economics to address customer needs effectively. The data can be useful to investors and other shareholders to identify product categories which have the highest share in the market. The visualisations can also be applied to understand market dynamics of various regions spread across multiple countries and states.

9. CONCLUSION

Based on the analysis of the customer spending behaviour dataset, the visualisations reveal crucial insights regarding the revenue of the products, which are influenced by factors such as customer age and gender. Notably, the tire and tube subcategory emerges as the highest-selling product, suggesting a strong demand. To capitalise on this trend, it is recommended to enhance sales and increase the quantity of these products accordingly.

Furthermore, the forecast charts for different product categories provide valuable information about future profitability. The data indicates that both accessories and clothing categories are projected to yield similar profits. However, a noteworthy finding is that the bike category might experience a significant decline in profit. This emphasises the need for proactive measures to mitigate the anticipated profit decrease in the bike category, while focusing on sustaining and optimising profits for accessories and clothing.

By leveraging these findings and proactively responding to the changing market dynamics, businesses can strategically allocate resources, adapt their product offerings, and tailor their marketing strategies to maximize revenue and profitability.

10. FUTURE SCOPE

Advancements in latest computer technologies have opened a world full of possibilities for data analysis and visualisations. The dashboards can be embedded with artificial intelligence and machine learning logics to identify trends based on previous market fluctuations and provide real world market forecasts. The proposed solution can be bettered with the ability to synchronise real time data into visualisation dashboards as it is generated enabling businesses to make quick decisions. Newer data analysis techniques such as data sonification can also be incorporated to provide another dimension to presentation and understanding of data.

11. PROJECT DEMONSTRATION

S.NO	Name of Visualization	Link for the Visualization
1.	Line Chart	https://drive.google.com/file/d/1cDo9GtNZZUWAorltRKPYWoFEe4tvMplm/view?usp=drive_link
2.	Sunburst Chart	https://drive.google.com/file/d/1Wkt8lLjt2JBMq2MJkvy-jHX3YCizQ3hD/view?usp=drive_link
3.	Forecast Graph	https://drive.google.com/file/d/1CwrTg1lmDmaPhbAbb7ruuiWr1iv6CZ6m/view?usp=drive_link
4.	Text visualization	https://drive.google.com/file/d/1JUSG5nbeUTTUm1v_cIfugRfyNOiRUSsy/view?usp=drive_link
5.	Heat Map	https://drive.google.com/file/d/1nHsn8dgAaOeUVjnF63J2fod1aPRFY1UD/view?usp=drive_link
6.	Bar Chart	https://drive.google.com/file/d/1my4hEMeiXLu-SjyBLZtoQDsCNlevyWsf/view?usp=drive_link
7.	Race Bar chart	https://drive.google.com/file/d/1a2hHr0xj-oiqShvivn3XnAP_s6bxlMOk/view?usp=drive_link
8.	Dashboard	https://drive.google.com/file/d/1Ywu_fE3wgnK-4jPW0bC-Ge5MFIkxkACw/view?usp=drive_link
9.	Story	https://drive.google.com/file/d/1NfrudMst7phaO31BV7Ukhb3IbB7n5nR6/view?usp=drive_link
10.	Flask Deployment	https://drive.google.com/file/d/1ZOJa3Bgr6GawHNd2VlkQ3G9NIkhwOedi/view?usp=drive_link

12. BIBLIOGRAPHY

[1] Govindan, K., Aditi, Dhingra Darbari, J., Kaul, A., & Jha, P. C. (2021). Structural model for analysis of key performance indicators for sustainable manufacturer–supplier collaboration: A grey-decision-making trial and evaluation laboratory-based approach. *Business Strategy and the Environment*, *30*(4), 1702-1722.

[2] Hasibuan, C. F. (2020, December). The measurement of customer satisfaction towards the service quality at xyz wholesale by using fuzzy service quality method. In *IOP Conference Series: Materials Science and Engineering* (Vol. 909, No. 1, p. 012053). IOP Publishing.

[3] Qian, F., Zhang, Q., & Zhang, X. (2023). The Effects of Agglomeration on Customer Traffic & Commercial Real Estate Values: Evidence from Grocery Store Openings.

[4] Daniel, L., Hinman, S., Gunen, B., Harper, K., Poirier, L., & Gittelsohn, J. (2020). Community-Based Recommendations to Improve Customer Turnout at a Non-Profit Grocery Store. *Current Developments in Nutrition*, *4*(Supplement_2), 708-708.

[5] Menculini, L., Marini, A., Proietti, M., Garinei, A., Bozza, A., Moretti, C., & Marconi, M. (2021). Comparing prophet and deep learning to ARIMA in forecasting wholesale food prices. *Forecasting*, *3*(3), 644-662.