Optimizing Retail Operations Using Data Analytics to Gain a Competitive Advantage and Increase Sales

Deirdre Gouta Mmolotsa, Reneilwe Keoagile, Letso Manyanda and Loago Rantao

Date: April 28, 2024

Contents

Declaration

We, the undersigned authors of this research paper entitled "Optimizing Retail Operations Using Data Analytics to Gain a Competitive Advantage and Increase Sales," hereby declare that:

1. This paper is the result of our original research and has not been submitted for publication elsewhere.

2. Any assistance received during the course of this research has been duly acknowledged.

3. The data and analysis presented in this paper accurately represent our findings.

4. We take full responsibility for the content and integrity of this research paper.

We understand that any violation of academic integrity or ethical standards would undermine the credibility of this paper and the reputation of Botswana Accountancy College.

Signed:

Deirdre Gouta Mmolotsa	D.G.M
Reneilwe Keoagile	R.K
Letso Manyanda	L.M
Loago Rantao	L.R

Date: April 28, 2024

Acknowledgements

Abstract

The retail environment is highly competitive, in order to survive organizations need to leverage data analytics to gain a competitive advantage and increase sales. This study explores the application of data analytics in optimizing the retail operations to enhance competitiveness and boost sales. Leveraging the regression analysis predictive modeling technique, this research aims to investigate the efficiency of sales forecasting to guide strategic decision making. By analysing the historical sales and other important variables, the regression model provides valuable insights into future sales trends, which enables retailers like Walmart Inc. to anticipate the demand, optimize inventory management, refining pricing strategies, and enhancing the overall operational efficiency. This paper highlights the potential of predictive modeling in shaping retail strategies, illustrating how Walmart Inc. and similar organizations can harness data analytics to gain a competitive edge in the dynamic retail environment.

List of Figures

Chapter 1

Introduction

The retail environment is highly competitive, businesses are faced with increasing challenges that can make or break them. With e-commerce on the rise, customer preferences are changing, hence there is a need for retailers to adapt to the changing environment to survive, by optimizing their retail operations to gain a competitive advantage and increase sales. The main purpose of this research project is to gain actionable insights that can be derived from data analytics implemented by Walmart Inc. to enhance efficiency, effectiveness, and the overall performance of all the various departments in an organisation. Optimising retail operations requires more than just inventory management and streamlining processes; it represents a strategic approach that retailers need to use to thrive in this competitive landscape. Achieving optimisation of retail operations requires thorough market research to understand changing customer preferences, market trends and the emergence of advanced technologies. It is a must for retailers to harness the power of data analytics, and predictive modelling, to deliver personalized shopping experiences across both the physical and digital channels.

1.1 Background of the Study

A common belief among executives and investors is that consumer data may be leveraged to obtain a competitive advantage that cannot be matched. Having more consumers means you can collect more data, which when combined with machine learning techniques for analysis, enables you to provide a better product that draws in more business. After that, you can gather even more information and, in the end, displace your rivals like companies with substantial network effects. Some of those companies may be Walmart Inc., Amazon and E-bay which are mainly retail operation-based companies. Retail operations companies can harness the power of data analytics by collecting client data and using it to create bet-

ter goods and services, but in the past, the procedure was cumbersome, sluggish, and impossible to scale. Focus groups, customer surveys, and data analysis were necessary for automobiles, consumer packaged goods businesses, and numerous other conventional manufacturers. However, the sales data was frequently not associated with specific consumers, and only a small number of customers' data was gathered because surveys and focus groups were costly and time-consuming.

Furthermore, as of today, the retail industry is very much customer-driven. Which means data is a pretty big deal. Retail businesses can discover a great deal about their target market, their products, and the state of the market by utilising consumer data. It's similar to unlocking a hidden door into people's purchasing patterns. Data is the source of all those insights!

Any information pertaining to your company's operations, such as sales, inventory, and pricing, is referred to as retail data. Thus, gathering, examining, and summarising this data constitute the analytical process. The goal is to get information on all the different aspects of your company, from supply chains and inventory levels to customer behaviour and loyalty, in order to enhance every aspect of your business operations. Believing that your company is operating smoothly and being aware of your shortcomings in general are insufficient. Having accurate insights into the performance of various supply chain or customer journey segments is what matters most. The key to solving these puzzles and improving a number of factors, such as pricing and consumer loyalty, is retail data.

As stated in the introduction, the main purpose of this research project is to gain actionable insights that can be derived from data analytics implemented by Walmart Inc. to enhance efficiency, effectiveness, and the overall performance of all the various departments in an organisation. This amounts to the need for Walmart to to adopt cutting-edge data analytics approaches after realising the shortcomings of conventional methodologies. This requirement stems from the understanding that out-of-date tactics impede operational optimisation, producing less-than-ideal results and passing up chances to increase revenue. Globalisation, the rise of online shopping, and technical advancements all highlight how important it is for Walmart to implement new technologies like artificial intelligence, machine learning, and big data. Moreover, by figuring out the best prices for various products, the business may boost sales and profitability. Walmart gains from this data-driven strategy, which also gives customers competitive prices and improves their shopping experiences. Walmart leverages big data in a large way for supply chain optimisation. Walmart can find inefficiencies and bottlenecks by examining data on suppliers, logistics, and inventory levels. The business can save expenses and streamline processes by using data to inform decision-making. This

7

guarantees that products are always available, which lowers customer churn and enables Walmart to give customers lower pricing. [16]

The ultimate objective of this study is to give Walmart the knowledge and resources it needs to improve customer experiences, streamline retail processes, boost sales, and maintain its position as the industry leader.

1.2 Problem Statement

The old techniques of retail management, which rely on intuition and historical data, have proven inadequate to traverse the rapidly altering market trends and change consumer preferences in the highly competitive and dynamic retail industry, as demonstrated by Walmart Inc. This shortcoming makes it more difficult to optimize operations, which leads to less-than-ideal results and lost chances to increase revenue. The necessity for Walmart to implement cutting-edge technologies like big data, machine learning, and artificial intelligence is further highlighted by the growth of internet shopping, globalisation, and technical breakthroughs. Additionally, big data is being used more and more to customise online buying. For instance, data-driven algorithms are used by online merchants to suggest products to customers based on their past purchases to add to their baskets both before and after they check out. [27] A significant time, financial, and resource investment is necessary for globalisation. Many shops have had severe defeats when they have ventured into uncharted and uncertain areas because they did not completely understand the special difficulties presented by disparate political, economic, and cultural contexts. As they enter new growth markets, retailers like Walmart Inc. can set and meet realistic goals for increasing sales and profitability by conducting thorough due diligence on both the potential and limitations. [13]

To maintain and improve its market position, Walmart Inc. must optimise its retail operations, which is a critical problem that this study project aims to solve. The issue at hand concerns the difficulties presented by antiquated retail management techniques, which impede the organization's capacity to adjust to evolving market conditions and seize new prospects. The modern retail environment, which is marked by heightened rivalry, waning brand loyalty, and an increasing emphasis on price over quality by consumers, is too demanding for the traditional approaches.

1.3 Aims

The goal is to offer Walmart Inc. meaningful data analytics insights that it can use to boost profitability and obtain a competitive edge. To verify these findings, the study will also assess the possible advantages and difficulties of putting data analytics into practice.

1.4 Objectives

Specific Objectives of this project are:

• Data Collection and preparation: obtaining a suitable dataset which entails variables such as invoice ID, product line, cost of goods sold, unit price etc. that relate to Walmart Inc.'s operations and ensuring it's data quality and consistency.

• Exploratory Data Analysis: Conducting a fact finding/investigational data analysis to identify patterns, trends, and outliers in the data, providing insights into the underlying factors impacting Walmart Inc.'s retail operations and sales performance.

• Predictive Modelling Development: Developing predictive models using machine learning algorithms and statistical techniques to forecast consumer demand, optimize inventory levels and sales forecasting accuracy.

• Implementation Strategy formulation: A comprehensive strategy for implementing data analytics solutions in Walmart Inc.'s retail operations addressing key considerations such as technology to be used/integrated, organizational change management, collaborations, and resource allocation.

• Development of a framework: A framework that guides the design and implementation of the project can be used by Walmart Inc. to improve efficiency and profitability of its retail operations, their inventory management, customer experience, marketing efficiency, including the use of sales forecasting to maximize profits and minimize costs.

• Framework Validation: The effectiveness and viability of the proposed framework will be done through benchmarking to compare the performance of the data analytics framework against industry benchmarks, best practices, or competing solutions available in the market.

• Documentation and Knowledge Transfer: Detailed documentation and training materials (A business report) will be created to facilitate the adoption and integration of data analytics practices within Walmart Inc.'s personnel, ensuring that stakeholders are equipped with the necessary skills and knowledge to leverage data effectively for decision making.

1.5 Significance of the study

The study's importance stems from its crucial role in tackling the urgent issue statement recognised in the retail sector, with Walmart Inc. serving as a prime example. The conventional methods of retail management, which depend on gut feeling and past performance, have not been sufficient in adjusting to the quickly shifting consumer tastes and market trends. Retail management oversees all of a retail organization's everyday operations, including strategy development, inventory control, customer support, marketing, and sales. A thorough approach is required to ensure profitable operation and boost efficiency. [18] This study

acknowledges the restriction of old methods of retail management and attempts to close the gap by promoting the use of state-of-the-art data analytics methods. Retail businesses can make data-driven decisions, increase operational effectiveness, improve consumer experiences, and spur business growth with the use of data analytics. Retailers can obtain a competitive advantage in the ever-changing and fiercely competitive retail sector by harnessing the power of data. [23]

By taking this approach, the research hopes to provide Walmart with the knowledge and resources it needs to make wise decisions and streamline its business practices in the face of a constantly changing and fiercely competitive retail market.

This study fills a vacuum in the literature about the transformative potential of data analytics for operational efficiency and performance enhancement, which greatly enhances knowledge in data analytics and retail management. It combines the fields of data analytics with retail management, providing insights on how they interact within Walmart Inc. Empirical analysis is used to provide useful insights to improve retail operations, adding to the body of knowledge in the sector and directing future research endeavours.

Furthermore, the study's findings have significant implications for the theoretical frameworks of retail management and data analytics. Through an investigation of the revolutionary potential of data analytics in improving retail operations, the study seeks to validate and advance current theories on the integration of sophisticated technology in business decision-making processes. The empirical information obtained from the research helps to evolve and improve theoretical models through the effective application of data-driven approaches in the retail sector and their ability to improve operational efficiency and performance. The research also elucidates the interplay between consumer behaviour, market trends, and operational strategy, offering perspectives that improve theoretical understanding of retail dynamics in the digital age.

The report's conclusions provide useful information for improving retail management practices, especially at Walmart Inc. Retailers can use advanced data analytics techniques like predictive modelling and exploratory data analysis to gain a deeper understanding of customer behaviour, market trends, and operational inefficiencies. By using this data to inform their pricing, marketing, and inventory management decisions, merchants may increase their operational effectiveness and profitability. Furthermore, the study offers a thorough process for utilising data analytics solutions for retail operations, managing organisational transformation, and integrating technology. To ensure sustained success in the cutthroat retail industry, retailers are urged to cultivate a culture of data-driven decision-making.

The findings of this study have implications for how policies are made within the retail sector, particularly about creating a supportive regulatory framework for the adoption and innovation of new technologies. According to the research, the retail industry must adopt cutting-edge technologies like big data analytics, machine learning, and artificial intelligence due to the rise of online shopping, the rise of globalisation and technological improvements. Furthermore, the research results of this study are significant regarding the way policies are made within the retail sector, particularly about creating a supportive regulatory framework for the implementation and innovation of emerging technologies. As to the research, the retail industry must adopt cutting-edge technologies like big data analytics, machine learning, and artificial intelligence due to the rise of online shopping, globalisation, and technological improvements. [1]

1.6 Structure of the Report

The report first starts with a literature review after the introduction, the literature review is chapter 2. A thorough examination of the state of retail operations optimisation is offered by the literature review, which aims to boost sales and provide a competitive edge. It identifies important trends, best practices, and cutting-edge techniques in the retail industry by referencing scholarly works, industry publications, and empirical studies. Moreover, provides businesses with useful information and suggestions that go beyond inventory control and sales forecasting. It highlights how important it is to use data analytics methods like business intelligence, big data analytics, and predictive modelling to improve decision-making and boost operational effectiveness. The review also emphasises how critical it is to solve issues like data availability, accuracy, and the shortage of trained workers. Retailers such as Walmart Inc. can enhance their competitive edge, streamline their operations, and manage the intricacies of the contemporary marketplace by implementing data-driven strategies.

The report then has the methodology right after the literature review as Chapter 3. With a focus on creating a competitive edge and boosting sales, the methodology chapter presents a thorough plan for optimising Walmart Inc.'s retail operations. Data gathering, exploratory data analysis (EDA), predictive modelling, creation of implementation strategies, framework development, validation, and documentation are just a few of the phases it includes. The study's methodology makes use of secondary data from Kaggle, notably the WalmartSalesData.csv dataset, which is in line with the project's goal of offering useful information for strategic decision-making. The selection of secondary data is based on its affordability, temporal efficiency, and availability of pertinent information. Excel and Power BI are used for data analysis, which includes descriptive statistics and

visualisations to find trends, patterns, and performance benchmarks. Extensive assessments of validity, limits, and reliability are required to evaluate the quality of data, with exploratory data analysis (EDA) being a key component. Forecasting customer demand and optimising inventory levels are the goals of predictive modelling development, and a plan for implementation is developed to guarantee that data analytics solutions are successfully adopted. Stakeholder engagement is facilitated by knowledge transfer and documentation, and a framework is created and validated to direct project implementation. Throughout, restrictions about the dataset's duration are discussed along with ethical issues about participant consent and data privacy. All things considered, the methodology offers a methodical way to maximise retail operations, utilising data-driven insights for strategic choice-making and long-term company expansion. Finally, for results, we have the outcomes and findings that were found using linear regression for variables picked from the dataset. The results aligned with the objectives outcomes and deliverables for the study.

Chapter 2

Literature Review

The literature review seeks to explore the current state of the research surrounding the optimization of retail operations within the context of gaining competitive advantage and increased sales. By synthesizing insights from academic literature, industry reports, and empirical studies, this review aims to identify the key trends, best practices, and emerging strategies in the retail sector. It will examine how retailers optimize their retail operations to enhance their market positioning, improve customer satisfaction, and drive sustainable revenue growth in a highly competitive environment. The review is not only limited to inventory management and sales forecasting, by critically evaluating the different methodologies, the review aims to provide actionable insights and practical recommendations for retailers seeking to optimize their retail operations to gain a competitive advantage and increase sales. With the use of data analytics, retailers can use a pivotal strategic approach to navigate the highly competitive retail environment through understanding retail operations management practices and provide valuable insights for practitioners, researchers, and stakeholders alike.

A pricing strategy is purposeful pricing by channels and customers to maximize value perception and business results to increase customer engagement and loyalty [14]. It is essential for every retailer to have a pricing strategy as competition is tough in the retail landscape. Key-Value Observing (KVO) and Key-Value Coding (KVC) provide powerful mechanisms for handling data in a flexible manner. [14] claims that the new digital retail era has several trends in the way customers are pricing their products including cross-channel customer decision journeys; responding at the right time and on the right channels, price transparency; the use of price comparison shopping engines, dynamic pricing; changing prices of online products 3-4 times a day, personalization; tailored recommendations

based on their purchase history and big data; the use of real-time data updates with sophisticated tools. These trends represent the emerging practices and dynamics within the retail pricing sector. Retailers keeping their prices low to keep up with its competitors is not always the best pricing strategy but rather, retailers need to transform their approaches to select Key-Value Items (KVI) through leveraging new data sources by using media platforms like Twitter to get reviews instead of relying on traditional data and refreshing KVIs more frequently to have real-time data. Walmart Inc. can use these strategies together to respond swiftly to market fluctuations, competitive threats and changing customer behaviour that influences its sales. However, a KVI pricing strategy implies linking the prices for inexpensive items with prices for rather expensive yet also cross-dependent items in a retailer's portfolio [15].

The layout design increases the sales of stores by navigating consumer behaviour [28]. Grouping products arranging the substitute goods and complementary goods in the same area, leads customers along specific paths to visit as many stores as possible. There are four main store layout types being the grid layout (long aisles are parallel to one another), freedom layout (asymmetric arrangement with aisles having products of different sizes and products), racetrack k layout (offers major aisle control traffic at store entrance) and the circulation layout (traffic loop around entire store). The use of AI to generate a conceptual design that uses the Sense-Think-Act-Learn. "Sense" to collect raw data like the CCTV camera. "Think" to process data collected. "Act" to use the knowledge and insights from the second phase. According to [6] the process operates as a continuous learning cycle. However, the AI technologies face limitations in roofed venues due to microwave signal attenuation and scattering by construction materials used for roofs and walls [11]. Walmart can use these AI technologies to improve their store layout which will encourage the buying behaviour hence increasing sales and staying ahead of the competition.

Inventory management can provide a more comprehensive understanding of optimizing retail operations using modern machine-learning techniques that leverage extensive datasets, including historical sales, weather data, store locations and special days to enhance inventory decisions. Retailers dealing with perishable goods face difficulties managing inventory due to frequent stockouts and substitution effects within product categories. Additionally, the unknown true demand distributions of products further complicate decision-making. The proposed data driven approach that uses modern machine learning techniques is empirically evaluated using real world datasets. The findings indicate that a data-driven approach outperforms a model-based benchmark. This proves that data driven approaches

lead to a more effective inventory decisions in real world scenarios which leads to better inventory management especially for the complex inventory management challenges faced by retailers dealing with perishable goods and multiple products [19]. If inventory management is well implemented right, stock levels are kept at optimum levels, and sales increase, however poor inventory management strategies can result in undesirable stock levels, late deliveries, and loss of revenue [24].

Regression analysis serves as a crucial component in Walmart Inc.'s sales forecasting model. This statistical technique allows analysts to explore and quantify the relationships between Walmart's sales and various economic indicators. Unlike qualitative methods that rely on subjective insights, regression analysis provides a data-driven approach by identifying the causal links between predictors, like oil [9]. By utilizing regression analysis, Walmart can optimize its operations, enhance customer satisfaction, and stay ahead of its competition. Additionally, regression analysis enables Walmart to anticipate major economic shifts by identifying leading indicators and confirm patterns and trends using lagging indicators. This approach empowers both practitioners and policymakers within Walmart to make informed decisions and adapt to changing market conditions effectively [25].

There is no absolute framework to guide organisations through social media competitive intelligence due to data privacy and sharing issues, which makes it difficult for organisations to compare their internal data with their competitor's data. However, social media data can be divided into two categories being social media data-at-rest and social media data-in-motion which is readily available to the public hence a comparison between organisations and their competitors can be made [9]. With the use of text mining or content analysis, Walmart Inc. can dig through the negative and positive comments about their products and services to identify areas of improvement [10].Overall, regression analysis plays a pivotal role in Walmart's sales forecasting efforts, enabling the company to leverage economic indicators and quantitative methods to drive strategic decision-making and achieve sustainable growth.

Many organisations are aware that data is a key asset and through acquiring valuable data they can gain competitive advantage [3]. According to [8] data needs to be collected from various sources and integrated to gain valuable business insights. Gathering information sounds very easy to implement, it is rather difficult and time consuming as data accuracy issues can cause bottlenecks in the project. By implementing this Walmart Inc. will be able to gather vast amounts of data and using data mining tools valuable insights can be extracted from the data which will ensure Walmart Inc. gains competitive advantage and increases their

sales.

Big data analytics has more volume, variety, velocity, veracity, data sources, analytical techniques, tools and platforms, storage, processing power and insight generation. Walmart Inc. uses predictive analytics to forecast, optimize inventory levels and manage logistics across its massive networks of stores. Through getting to know their customers and improving valuation, Walmart Inc can optimize marketing spending using customer data, predict the customers journey life. Predictive marketing can help Walmart Inc. understand buying behaviours of customers using Business Intelligence tools and techniques. This can help them improve and customer relations to ensure loyalty which will lead to more sales [2]. Although Business Intelligence has many benefits it is a very complex concept with several shortcomings making it unattractive and expensive to maintain [17]. White [29] claims that Business Intelligence users have a difficult time using Business Intelligence applications and that these applications are time consuming. According to [7] Business Intelligence tools must be easy to use simultaneously providing flexibility and significant power to gain valuable insights which can help in making informed decisions which will to increased sales and gaining competitive advantage.

Many strategies such as pricing and promotions strategies, inventory management, store layout design and sales forecasting are all ways in which an organisation can optimize their retail operations to gain competitive advantage and increase their sales. Though these are good strategies an organisation can implement, product exposure is a very effective strategy that is overlooked. Research shows that over two million grocery store shoppers see products displayed for promotion twice as much as products displayed in the inner aisles of the store [20]. There are several topologies to identify a business strategy, Miles and Snow [26] suggests a topology that consists of four business strategies consisting of prospector strategy, analyser strategy, defender strategy and reactor strategy. The first three increase operational performance while the reactor slows down operational performance. If the strategies stated above are implemented by Walmart Inc. can increase sales and gain competitive advantage.

This project has limitations which can result in the opposite of the outcomes expected. One of these limitations is data accuracy and availability. Data privacy and sharing regulations limit the amount of data that can be accessed, this can lead to Inaccurate forecasts which will drive the decision making in the wrong directions hence sales will not increase, nor will Walmart Inc. be able to gain competitive advantage. Due to the limitations in the forecasting techniques demand forecasting errors may occur, resulting in increased costs and operational

complexity. The strategies recommended are expensive however the results are not guaranteed as factors like lack of skilled personnel can hinder the deployment of the forecasting and pricing strategies can result in missed revenue targets and market share loss.

In conclusion the literature review provides a comprehensive overview of the significance, challenges, and emerging trends in the retail landscape. As evidenced by the research reviewed, the strategies optimize retail operations, lead to strategic decision making and enhance customer satisfaction in the overly competitive retail environment. It highlights the importance of using predictive modelling techniques, big data analytics and business intelligence to get accurate and reliable insights that will lead to informed decision making. By adopting data driven approach to sales forecasting, retailer can gain valuable insights on demand patterns, inventory management, and pricing opportunities, therefore improving the overall business performance and profitability.

However, the challenges and limitations associated with the optimization initiatives have also been addressed in the literature review. From data accuracy and availability issues to lack of skilled personnel. Solutions to these challenges can be achieved using advanced technologies and models. Optimization techniques are important as they can help retailers like Walmart Inc. to navigate the complexities of the modern marketplace and deliver exceptional value to their customers hence increasing sales and gaining competitive advantage.

Chapter 3

Methodology

This chapter outlines the comprehensive approach undertaken in the data analysis project aimed at optimizing retail operations to gain a competitive advantage and increase sales for Walmart Inc. The methodology encompasses several key stages, including data collection and preparation, exploratory data analysis (EDA), predictive modelling development, implementation strategy formulation, framework development, framework validation, and documentation and knowledge transfer.

Research Approach and Research Design: The overall research approach for this project is to utilize secondary data to provide actionable insights that can be implemented by Walmart Inc. to gain a competitive advantage and increase profitability. The rationale for using secondary data is that it provides a cost-effective and time-efficient method for obtaining relevant data, which is essential for achieving the project's objectives. A datset containing secondary data will be sourced from Kaggle. This approach aligns with the research objectives and questions by providing a comprehensive dataset that can be used to conduct a literature review, exploratory data analysis, predictive modelling, and framework development. The secondary data will enable the project to identify patterns, trends, and outliers in Walmart Inc.'s retail operations and sales performance, which can be used to develop predictive models and formulate an implementation strategy. The use of secondary data also allows for the validation of the proposed framework through benchmarking against industry benchmarks, best practices, or competing solutions available in the market [21]. The secondary data will be analyzed using quantitative data analysis and qualitative data analysis, which will provide a comprehensive understanding of the underlying factors impacting Walmart Inc.'s retail operations and sales performance. In conclusion, the use of secondary data is essential for achieving the project's objectives by providing a comprehensive

18

dataset that can be used to conduct a literature review, exploratory data analysis, predictive modelling, and framework development. The secondary data will enable the project to identify patterns, trends, and outliers in Walmart Inc.'s retail operations and sales performance, which can be used to develop predictive models and formulate an implementation strategy. Secondary data also allows for the validation of the proposed framework through benchmarking against industry benchmarks, best practices, or competing solutions available in the market.

Data Source Description: The primary data source, WalmartSalesData.csv, obtained from Kaggle, includes variables relevant to Walmart Inc.'s operations, aligning with the project's objectives. It contains data about Sales transactions, customer demographics, store layout, and inventory levels.

Rationale for Using Secondary Data: Secondary data is appropriate for this study due to its cost-effectiveness, time efficiency, and access to a large volume of relevant information. As emphasized by[30], using secondary data is more convenient than gathering primary data, which typically is the most time-consuming and costly component of the research. By leveraging pre-existing data, sourced from datasets such as WalmartSalesData.csv obtained from Kaggle, provide comprehensive insights into various aspects of Walmart Inc.'s operations, including sales transactions, customer demographics, store layout, and inventory levels. Instead of starting from scratch and trying to gather all the data , the information that's already out there can be used. This saves time and effort without worrying about having to find people to take part in the study, whether the collected data is enough, or worrying about people dropping out of the study. Secondary data facilitates a thorough analysis of historical trends and patterns, facilitating the development of data-driven strategies to optimize retail operations and increase sales. Purpose: The use of secondary data is justified by its relevance to the research objectives, enabling a comprehensive analysis of retail operations and sales performance.

Data Selection Criteria: The dataset was selected based on criteria such as completeness, relevance, and availability of key variables. However, the limitation of the dataset covering a period of three months is acknowledged.

Data Analysis Procedures: Data Analysis was done using Power BI and Excel. Excel was used to obtain the summary of descriptive statistics. They provide insights into the central tendency, variability, and distribution of relevant variables under investigation (Unit Price, Total Sales, Cost of Goods Sold, Gross Income, and Rating). Subsequently, visualizations done on Power BI are used in comparative analysis to identify the impact of independent variables on dependent variables. Comparative/ Benchmarking analysis is another important method used to compare different segments of the dataset or compare Walmart Inc.'s perfor-

19

mance with industry benchmarks or competitors. This analysis helps to identify areas of strength and weakness and benchmark performance against industry standards [4]. Moreover, qualitative methods such as content analysis may be applied to analyse textual data within the dataset, such as customer feedback or product reviews. Content analysis involves systematically categorizing and interpreting qualitative data to identify recurring themes or patterns, providing additional insights into customer preferences and behaviors. The data analysis procedure aims to interpret the secondary data comprehensively, providing valuable insights into retail operations and sales performance for Walmart.

Evaluation of Data Quality: To ensure the integrity of the analysis, the reliability, validity, and limitations of the secondary data are thoroughly assessed. Exploratory Data Analysis (EDA) is conducted to uncover patterns, trends, and outliers in the data [12]. Valuable insights into sales and customer behavior are gained through the utilization of various techniques such as visualizations, statistical summaries, and segment analysis. However, EDA plays a crucial role in assessing data quality, identifying anomalies, and generating hypotheses for further analysis [12]. It may also have limitations, such as its inability to detect certain complex relationships within the data.. Therefore, a comprehensive assessment of data quality involves considering the strengths and weaknesses of EDA alongside other evaluation methods, ensuring a thorough understanding of the dataset's reliability and validity.

Predictive Modelling Development: The project proceeds with the development of predictive models using statistical techniques such as Linear regression. Leveraging historical data, these models provide valuable insights for strategic decision-making across different aspects of Walmart Inc's retail operations. These models aim to forecast consumer demand, optimize inventory levels, and enhance sales forecasting accuracy. Predictive modeling also plays a crucial role in informing pricing strategies for Walmart Inc. By analyzing historical pricing along various market factors as seen in the dataset, predictive models can provide insights into optimal pricing strategies.

Implementation Strategy Formulation: A comprehensive strategy is devised for implementing data analytics solutions in Walmart Inc.'s retail operations. Key considerations such as technology selection, organizational change management, collaborations, and resource allocation are addressed to ensure successful implementation and adoption of data analytics practices.

Framework Development and Validation: A framework is developed to guide the design and implementation of the project, focusing on areas such as retail operations efficiency, inventory management, and customer experience. The frame-

work's effectiveness and viability are then validated through benchmarking against industry standards and best practices.

Documentation and Knowledge Transfer: Detailed documentation and training materials are created to facilitate the adoption and integration of data analytics practices within Walmart Inc.'s personnel. Stakeholders are equipped with the necessary skills and knowledge to leverage data effectively for decision-making, ensuring the successful implementation and sustainability of data-driven initiatives.

Ethical Considerations: The confidentiality of the data that has been used in this research has been considered. Each customer has been identified by the invoice ID, gender and customer type. It doesn't really contain data which can be deemed as sensitive therefore it is safe to say privacy rights have been protected by the person who initially collected the data. Biases have been avoided during the data analysis process to ensure fairness and equity of the individuals represented in the data. This can be seen in the sample data being used for predictive modeling where an equal number of both genders is included to avoid perpetuating stereotypes or discrimination. Transparency has been maintained throughout the project about the practices involved from accessing the WalmartSalesDataset, pre-processing and analysing it as well as on how it was used in predictive model development and the methodologies used.

Limitations of Secondary Analysis: The data being used in the project is secondary data obtained from Kaggle, however, conducting secondary analysis mean relying on data collected by others. This lack of control over the data collection process can lead to limitations in the depth of analysis. As seen in the Exploratory Data Analysis, it was identified that the dataset covers a specific period of three months. This limits the ability to analyze long term trends or seasonal variations. There is also a risk of misinterpreting or misunderstanding the context in which the data was collected. Without firsthand knowledge of the data collection process and any limitations of biases that may have been present, it may be possible that incorrect conclusions or faulty assumptions about the data are made about the data during analysis. Another highlighted issue is that Secondary Data Analysis in this project is limited to the variables and information contained in the Walmart Salesdataset.csv. The challenge of not being able to collect primary data to suit the nature of the research and address specific research aims/questions is limiting because some hypothesis cannot be made based on that dataset.

When analyzing secondary data, there is a limitation in terms of the scope of available data and certain types of study designs and statistical tests as highlighted by [30]. These studies typically focus on describing, exploring, and finding con-

nections between different things in the data. They rely on observations from past events, so they're often looking backward rather than forward. This means test cause-and-effect relationships wouldn't be possible unlike in a controlled experiment where changes can be made during data collection to see what happens.

Chapter 4

Results

Exploratory Data Analysis (EDA) Report

This report serves as a roadmap for leveraging data analytics to enhance operational efficiency, drive revenue growth, and gain a competitive advantage. By embracing data-driven decision making, Walmart In. can strengthen its position in the competitive retail landscape and pave the way for a more prosperous future. This project focuses on using data analytics to optimize retail operations for Walmart Inc., aiming to increase sales and profitability. The goal is to provide actionable insights from exploratory data analysis (EDA) that Walmart Inc. can implement to enhance performance and gain a competitive edge by increasing sales. Through analysing variables such as total sales, customer segmentation (Customer type and Gender), product lines (for inventory management strategies) and pricing strategies from the Walmart Inc. dataset, the aim is to uncover opportunities for improvement. The raw Walmart Inc. dataset was extracted from Kaggle and opened in excel. It is comprised of numerical and categorical data. There are 17 attributes and 101 observations uniquely identified by Invoice Id, one of the variables. According to [22], the second phase of data processing is data cleaning where values in the dataset are identified and dealt with. At first glance the dataset doesn't seem to consist of any missing values, however it was still transformed and clean to ensure consistency of the data to improve data quality.

Figure below shows a successfully transformed and cleaned Walmart Inc. dataset

Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quantity	Tax 5%	Total	Date	Time	Payment	Cost o
750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	05/01/2019	13:08:00	Ewallet	
226-31-3081	C	Naypyitaw	Normal	Female	Electronic accessories	15.28	5	3.82	80.22	08/03/2019	10:29:00	Cash	
631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	03/03/2019	13:23:00	Credit card	
123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.288	489.048	27/01/2019	20:33:00	Ewallet	
373-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7	30.2085	634.3785	08/02/2019	10:37:00	Ewallet	
699-14-3026	C	Naypyitaw	Normal	Male	Electronic accessories	85.39	7	29.8865	627.6165	25/03/2019	18:30:00	Ewallet	
355-53-5943	A	Yangon	Member	Female	Electronic accessories	68.84	6	20.652	433.692	25/02/2019	14:36:00	Ewallet	
315-22-5665	C	Naypyitaw	Normal	Female	Home and lifestyle	73.56	10	36.78	772.38	24/02/2019	11:38:00	Ewallet	
665-32-9167	A	Yangon	Member	Female	Health and beauty	36.26	2	3.626	76.146	10/01/2019	17:15:00	Credit card	
692-92-5582	B	Mandalay	Member	Female	Food and beverages	54.84	3	8.226	172.746	20/02/2019	13:27:00	Credit card	
351-62-0822	B	Mandalay	Member	Female	Fashion accessories	14.48	4	2.896	60.816	06/02/2019	18:07:00	Ewallet	
529-56-3974	B	Mandalay	Member	Male	Electronic accessories	25.51	4	5.102	107.142	09/03/2019	17:03:00	Cash	
365-64-0515	A	Yangon	Normal	Female	Electronic accessories	46.95	5	11.7375	246.4875	12/02/2019	10:25:00	Ewallet	
252-56-2699	A	Yangon	Normal	Male	Food and beverages	43.19	10	21.595	453.495	07/02/2019	16:48:00	Ewallet	
829-34-3910	A	Yangon	Normal	Female	Health and beauty	71.38	10	35.69	749.49	29/03/2019	19:21:00	Cash	
299-46-1805	B	Mandalay	Member	Female	Sports and travel	93.72	6	28.116	590.436	15/01/2019	16:19:00	Cash	
656-95-9349	A	Yangon	Member	Female	Health and beauty	68.93	7	24.1255	506.6355	11/03/2019	11:03:00	Credit card	
765-26-6951	A	Yangon	Normal	Male	Sports and travel	72.61	6	21.783	457.443	01/01/2019	10:39:00	Credit card	
329-62-1586	A	Yangon	Normal	Male	Food and beverages	54.67	3	8.2005	172.2105	21/01/2019	18:00:00	Credit card	
319-50-3348	B	Mandalay	Normal	Female	Home and lifestyle	40.3	2	4.03	84.63	11/03/2019	15:30:00	Ewallet	
300-71-4605	C	Naypyitaw	Member	Male	Electronic accessories	86.04	5	21.51	451.71	25/02/2019	11:24:00	Ewallet	
371-85-5789	B	Mandalay	Normal	Male	Health and beauty	87.98	3	13.197	277.137	05/03/2019	10:40:00	Ewallet	
273-16-6619	B	Mandalay	Normal	Male	Home and lifestyle	33.2	2	3.32	69.72	15/03/2019	12:20:00	Credit card	
636-48-8204	A	Yangon	Normal	Male	Electronic accessories	34.56	5	8.64	181.44	17/02/2019	11:15:00	Ewallet	
549-59-1358	A	Yangon	Member	Male	Sports and travel	88.63	3	13.2945	279.1845	02/03/2019	17:36:00	Ewallet	
227-03-5010	A	Yangon	Member	Female	Home and lifestyle	52.59	8	21.036	441.756	22/03/2019	19:20:00	Credit card	

DESCRIPTIVE SUMMARY OF STATISTICS

Unit Price		Total Sales		Cost of Goods		Gross income		Rating	
Mean	55.67213	Mean	322.966749	Mean	307.58738	Mean	15.379369	Mean	6.9727
Standard Error	0.837833713	Standard Error	7.775577022	Standard Error	7.405311449	Standard Error	0.370265572	Standard Error	0.054346281
Median	55.23	Median	253.848	Median	241.76	Median	12.088	Median	7
Mode	83.77	Mode	829.08	Mode	789.6	Mode	39.48	Mode	6
Standard Deviation	26.49462835	Standard Deviation	245.8853351	Standard Deviation	234.1765096	Standard Deviation	11.70882548	Standard Deviation	1.718580294
Sample Variance	701.9653313	Sample Variance	60459.59802	Sample Variance	54838.63766	Sample Variance	137.0965941	Sample Variance	2.953518228
Kurtosis	-1.218591428	Kurtosis	-0.08188476	Kurtosis	-0.081884758	Kurtosis	-0.081884758	Kurtosis	-1.151586839
Skewness	0.007077448	Skewness	0.892569805	Skewness	0.892569805	Skewness	0.892569805	Skewness	0.009009649
Range	89.88	Range	1031.9715	Range	982.83	Range	49.1415	Range	6
Minimum	10.08	Minimum	10.6785	Minimum	10.17	Minimum	0.5085	Minimum	4
Maximum	99.96	Maximum	1042.65	Maximum	993	Maximum	49.65	Maximum	10
Sum	55672.13	Sum	322966.749	Sum	307587.38	Sum	15379.369	Sum	6972.7
Count	1000	Count	1000	Count	1000	Count	1000	Count	1000

figure 4.Line chart showing Average total sales by Date.

Interpretation of the Descriptive Summary of Statistics

Unit Price: With an average of approximately $55.67, the unit price of products sold exhibits moderate variability, indicated by a standard deviation of around $26.49. The dataset encompasses a diverse range of product pricing, with unit prices ranging from $10.08 to $99.96. Symmetrical distribution with a slight positive skewness suggests that most unit prices cluster around the mean.

Total Sales: The average total sales amount per transaction stands at approximately $322.97, reflecting significant variability with a standard deviation of around $245.89. A wide range of total sales, spanning from $10.68 to $1042.65, indicates varying transaction volumes. Positive skewness in the distribution suggests the presence of transactions with exceptionally high sales amounts. Cost of Goods Sold (COGS): On average, the cost of goods sold per transaction is approximately $307.59, exhibiting moderate variability with a standard deviation of around $234.18. A wide range of COGS, from $10.17 to $993, suggests varying costs associated with producing goods. Like total sales, the distribution of COGS

is positively skewed, indicating the presence of transactions with exceptionally high COGS values.

Gross Income: The average gross income per transaction is approximately $15.38, with moderate variability reflected by a standard deviation of around $11.71. A relatively wide range of gross income, from $0.51 to $49.65, suggests varying levels of profitability. Positive skewness in the distribution indicates the presence of transactions with exceptionally high gross income values.

Rating: Customers provide an average rating of approximately 6.97, with moderate variability indicated by a standard deviation of around 1.72. Ratings range from 4 to 10, reflecting a range of customer satisfaction levels. The distribution of ratings appears to be approximately symmetrical, with a slight negative kurtosis, suggesting a relatively normal distribution.

Overall, these interpretations offer insights into the characteristics and variability of the variables in the dataset. Having gained valuable insights into the characteristics and variability of the variables in the dataset through the summary of descriptive statistics, the focus shifts to the analysis of visualizations. By delving deeper into the data through visual exploration, the aim is to uncover actionable patterns, trends, and relationships that will inform strategic decision-making and contribute to achieving this project's overarching goal.

Analysis of Visualizations

Average Total sales by Date
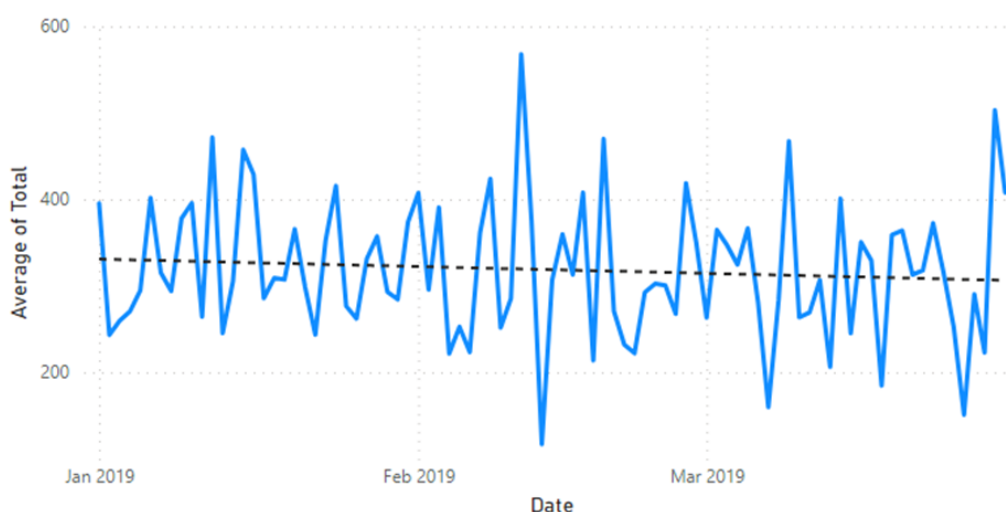
This visualization shows the trend of Average total sales over a period of 3

25

months.

1. Recent Trend: The recent uptrend in Average total sales, particularly between Tuesday, February 26, 2019, and Saturday, March 2, 2019, corresponds to an increase in sales volume by 96.89 units. This trend mirrors the overall average sales amount per transaction, suggesting a period of heightened sales activity during this timeframe.

2. Steep Trend: Similarly, the steep upward trend in Average total sales observed during the same period indicates a significant rise in sales, further corroborating the positive sales trend identified. This steep increase in sales aligns with the range of total sales, indicating a notable spike in sales volumes during this period.

3. Long Trend: The prolonged upward trend in Average total sales between Sunday, January 6, 2019, and Wednesday, January 16, 2019, with a rise of 27.33 units, reflects a sustained period of growth in sales. This long-term trend is consistent with the average sales amount per transaction, indicating a gradual but steady increase in sales over the specified period.

4. Anomalies: The anomalies detected in Average total sales highlight instances where sales deviated significantly from the expected range.
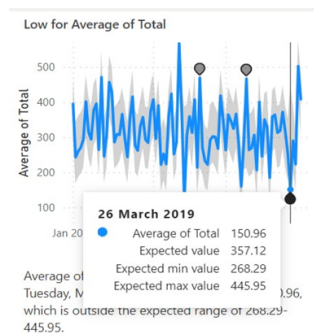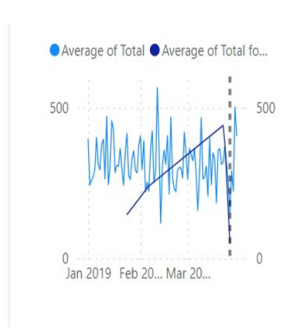
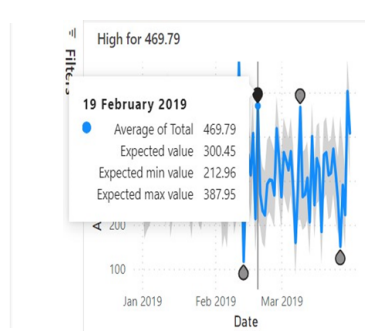

Figure 4.1: fig 5

Figure 4.2: fig 5

Figure 4.3: fig 6

For example, as illustrated in figure 5 and 6, the unexpectedly low sales on Tuesday, March 26, 2019, and unusually high sales on Tuesday, February 19, 2019, suggest potential anomalies in sales performance. These anomalies may warrant further investigation to understand the underlying factors contributing to these deviations from the expected range.

This chart in figure 7 compares the average total sales across different branches and product lines.
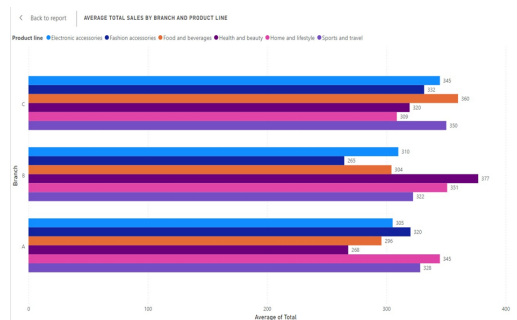
Figure 4.4: fig 7

1. Top-Performing Branches and Product Lines: Branch B appears to have the highest average total sales across all product lines, particularly in the categories of Health and beauty at \$377 and Home and Lifestyle at \$351. Among the product lines, Food and Beverages consistently show high average total sales across all branches.

2. Areas Requiring Attention: Branch A shows relatively lower average total sales compared to Branches B and C, particularly in the categories of Health and beauty and Home and Lifestyle. Fashion accessories in Branch B and Food and Beverages in Branch C also exhibit lower average total sales compared to other product lines within the same branches.

3. Potential Strategies for Improvement: To replicate success in underperforming areas, such as Branch A and certain product lines, strategies like cross-selling or product bundling could be explored. Leveraging the success of high-performing product lines, such as Food and Beverages, across all branches could contribute to overall sales growth.

4. Correlations Between Branches and Product Lines: Branch B consistently performs well across various product lines, indicating a potential correlation between the branch's operational efficiency and sales performance. The relatively lower average total sales in Branch A and C may suggest areas for improvement in marketing strategies or product offerings specific to those branches.

5. Actionable Insights: Identifying top-performing branches and product lines enables Walmart Inc. to allocate resources effectively and focus on strategies that capitalize on existing strengths. Analysing underperforming areas provides opportunities for targeted interventions to enhance sales performance and drive overall profitability.

Average Total sales by Customer type

315.... (44...)

Customer type
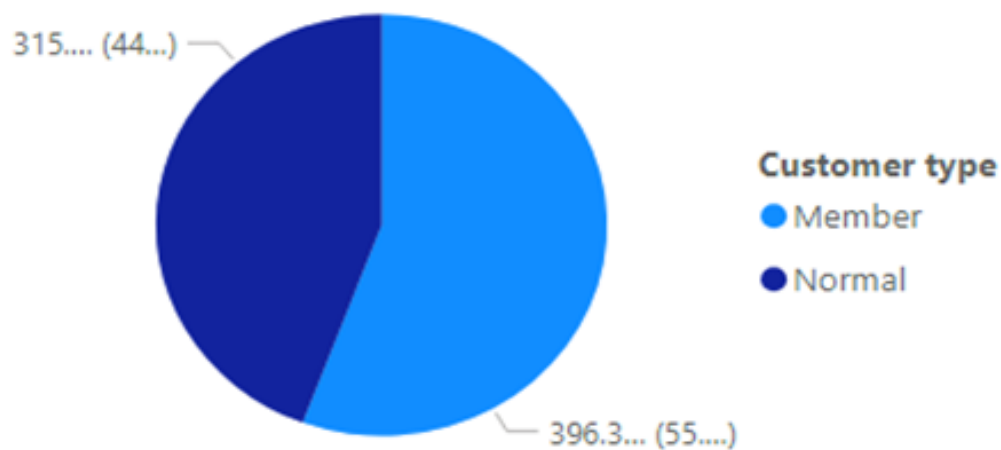● Member
● Normal

396.3... (55....)

fig:8.Pie chart showing Average Total

The pie chart in Figure 8 displays the distribution of average total sales between different customer types. The percentage of average total sales contributed by members and non-members allows for a comparison of their respective shares in the overall sales revenue.

1. Member vs. Non-Member Spending: Members account for approximately 44% of the total pie chart, with an average total sales value of $315.12.

Non-members represent the majority, constituting around 55% of the pie chart, with an average total sales value of $396.30.

2.Purchasing Behaviour Differences: The higher average total sales value for non-members suggests that they may make larger purchases compared to members, despite members accounting for a smaller portion of the total sales.

Members, on the other hand, contribute a significant portion of sales but may have more frequent, smaller transactions compared to non-members.

Overall, by analysing the distribution of total sales by customer type and its relationship to summary statistics and customer segmentation, Walmart Inc. can gain valuable insights into purchasing behaviour, customer preferences, and opportunities for enhancing customer engagement and satisfaction.
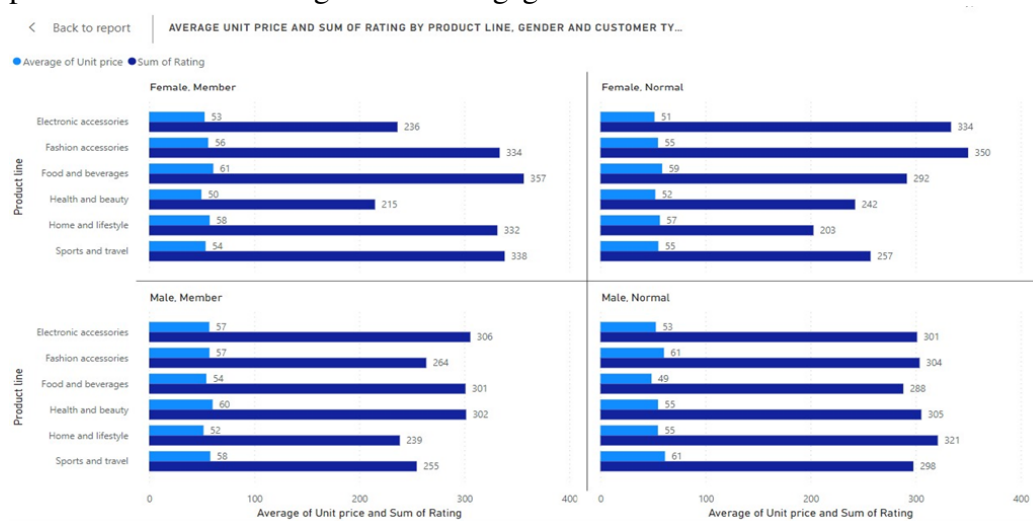
Clustered column chart

Pricing Strategies: Across different product lines, there are variations in the average unit prices. For instance, "Food and Beverages" and "Home and Lifestyle" have relatively higher average unit prices compared to "Electronic accessories" and "Health and beauty. "Analysing pricing strategies involves understanding how these variations align with customer preferences and market demands. For example, higher-priced products may target a niche market segment or offer premium quality, while lower-priced items may aim for broader accessibility.

Pricing Strategies Analysis:

Examining the average unit prices across various product lines reveals intriguing insights:

"Electronic accessories" boast average unit prices ranging from $51 to $57. "Fashion accessories" exhibit prices ranging from $55 to $61. "Food and Beverages" showcase prices ranging from $59 to $61. "Health and beauty" products vary from $50 to $60. "Home and Lifestyle" items range from $52 to $58. "Sports and Travel" products span prices from $54 to $61.

Customer Satisfaction Levels: The sum of ratings provides insights into customer satisfaction levels across different product lines, genders, and customer

types. Comparing the sum of ratings across product lines allows us to identify which categories consistently receive higher ratings, indicating greater customer satisfaction. For instance, "Fashion accessories" and "Sports and Travel" demonstrate relatively higher sum of ratings across all customer segments, suggesting higher satisfaction levels in these product categories.

Customer Satisfaction Levels Analysis: Analysing the sum of ratings sheds light on customer satisfaction across various product lines: "Fashion accessories" and "Food and Beverages" generally boast higher customer satisfaction ratings. "Health and beauty" products also show relatively high satisfaction levels.

Relationship between Pricing, Product Quality, and Customer Satisfaction: While this analysis suggests a potential correlation between higher-priced products and customer satisfaction levels, it's important to note that this observation is based on the data available and may not fully capture the complexities of customer preferences and perceptions. Further research and validation are necessary to confirm the relationship and understand the underlying factors influencing customer satisfaction.

Recommendations

Adjusting pricing strategies based on customer demographics could further enhance this relationship. For example: Offering competitive pricing on products with high satisfaction ratings could attract more customers and foster loyalty.

Tailoring pricing strategies to specific customer segments, such as offering discounts or promotions to incentivize purchases among certain demographics, could drive sales and strengthen customer relationships. Values in the Analysis: Average unit prices and sum of ratings for each product line, segmented by gender and customer type, provide the foundation for this analysis. For example:

"Fashion accessories" have average unit prices ranging from $55 to $61, with corresponding sum of ratings ranging from 264 to 350 across different customer segments.

By leveraging these insights, Walmart Inc. can refine its pricing strategies to align with customer preferences, enhance product value perception, and ultimately drive sales and loyalty in the competitive retail market.

The chart in Figure 10 compares the average cost of goods sold and average total sales across different branches. This visualization offers insights into the cost-effectiveness of sales in each branch, shedding light on potential areas for improvement and optimization.

From the data presented:

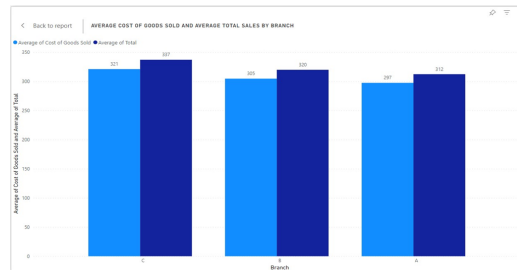Branch A has an average cost of goods sold of $321 and average total sales of $337.

Figure 4.5:

Branch B shows an average cost of goods sold of $305 and average total sales of $320.

Branch C indicates an average cost of goods sold of $297 and average total sales of $312.

Analysis

Branch A demonstrates robust sales performance, with an average unit price of approximately $55.67 and an average total sales amount per transaction of approximately $322.97. However, the average cost of goods sold, approximately $307.59, suggests that a significant portion of the sales revenue is allocated towards covering the costs of goods sold.

In comparison, Branch B exhibits slightly lower average total sales compared to Branch A, but also lower average cost of goods sold. With similar average unit price and total sales figures, Branch B may be more cost-effective in managing its inventory and production costs, potentially resulting in improved profitability.

Branch C demonstrates the lowest average total sales among the three branches, accompanied by a relatively low average cost of goods sold. While the cost of goods sold is relatively low, indicating efficient cost management, the lower sales figures raise questions about sales performance and potential opportunities for revenue growth.

While Branches A and C appear to have a relatively favourable cost-effectiveness ratio, Branch B stands out with slightly lower average total sales compared to its average cost of goods sold. This suggests that while sales revenue is generated, it may not be sufficient to cover the associated costs effectively.

To enhance profitability and increase sales in Branch B, exploring cost-saving measures or operational improvements is crucial. Optimizing inventory management practices or negotiating better supplier deals could help reduce costs. Additionally, analysing sales strategies specific to Branch B may reveal opportunities

for targeted marketing efforts or product diversification to stimulate sales growth and improve cost-effectiveness.

In conclusion, this Exploratory data analysis has unearthed valuable insights into Walmart Inc.'s retail operations, underscoring the pivotal role of data analytics in driving profitability and operational efficiency.

Key Findings

The line chart depicting average total sales by date revealed notable trends in sales performance over time, highlighting periods of growth and potential anomalies.

Branch and product line comparisons from the clustered bar chart underscored variations in sales performance across different segments, indicating areas of strength and opportunities for improvement.

The pie chart illustrating the distribution of average total sales by customer type shed light on the contribution of member and non-member customers to overall sales, providing insights into customer segmentation strategies.

Analysis of pricing strategies revealed variations in average unit prices across product lines, prompting considerations of how pricing aligns with customer preferences and market dynamics.

Comparisons of average cost of goods sold and average total sales across branches highlighted differences in cost-effectiveness and profitability, signalling areas for optimization.

Actionable Recommendations

Utilize insights from sales trends to inform inventory management and promotional strategies, capitalizing on periods of growth and addressing anomalies promptly.

Tailor marketing efforts and product offerings based on branch and product line performance, optimizing resources to maximize returns.

Implement targeted customer segmentation strategies to cater to the distinct preferences and behaviours of member and non-member customers, fostering loyalty and driving sales.

Refine pricing strategies to align with market demands and customer expectations, balancing premium offerings with accessible options to capture diverse consumer segments.

Enhance operational efficiency by identifying and addressing cost inefficiencies, leveraging data-driven insights to streamline processes and improve profitability.

By leveraging these insights and recommendations, Walmart Inc. can unlock new opportunities for growth, strengthen its competitive position, and ultimately

achieve its goal of optimizing retail operations to drive profitability and success in the dynamic retail landscape.

Predictive Modeling

Walmart comprehends the relationships between variables and provides well-informed forecasts based on past data by using regression analysis, a type of predictive modelling. Regression models has been used at Walmart to examine the relationships between sales forecasting, inventory control, and consumer behaviour.

By employing regression analysis, Walmart aims to uncover meaningful patterns and trends within its vast datasets, enabling the company to make more accurate predictions about future outcomes. Through the examination of historical sales data, demographic information, and other relevant factors, regression models provide valuable insights into consumer preferences, market trends, and the impact of various external factors on retail performance.

This report will delve into the application of regression analysis within Walmart's predictive modelling framework. This will look on how these models have enhanced decision-making processes, optimized inventory levels, and ultimately driven profitability in the dynamic retail landscape. By shedding light on the methodologies, findings, and implications of regression analysis at Walmart, the report aims to illustrate the transformative potential of data-driven approaches in shaping the future of retail operations.

Data Presentation and Analysis

To analyse the relationship between unit price and total sales revenue at Walmart Inc, linear regression was conducted to examine how changes in unit price may affect sales revenue. The regression analysis aimed to quantify the impact of unit price on total sales revenue, providing insights into pricing strategies and revenue optimization.
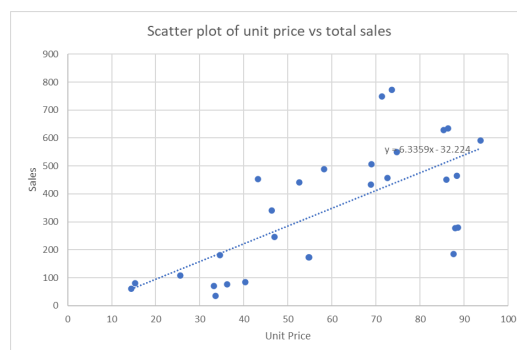

Figure 4.6: Scatter plot of Unit price and Total

Positive Relationship: The analysis indicated a positive relationship between unit price and total sales revenue. As unit price increases, total sales revenue tends to increase as well. This suggests that customers are willing to spend more on products with higher unit prices, contributing to higher overall revenue for Walmart.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.716830627
R Square	0.513846148
Adjusted R	0.476449698
Standard E	160.3935881
Observati	29

ANOVA

	df	SS	MS	F	gnificance F
Regression	2	706979	353490	13.7405	8.5E-05
Residual	26	668879	25726.1		
Total	28	1375858			

	Coefficients	andard Err	t Stat	P-value	Lower 95%	Upper 95%	ower 95.0%	pper 95.0%
Intercept	-265.7317208	176.53	-1.50531	0.1443	-628.594	97.1301	-628.594	97.1301
Unit^price	-0.088645247	0.05962	-1.48694	0.14906	-0.21119	0.0339	-0.21119	0.0339
Unit price	16.366308	6.86241	2.38492	0.02466	2.26042	30.4722	2.26042	30.4722

Figure 4.7: Summary output

Statistical Significance: The analysis reveals a notable positive correlation between unit price and total sales revenue, as evidenced by the moderate correlation coefficient (Multiple R = 0.7168) and a relatively high coefficient of determination (R Square = 0.5138).

The statistical significance is further supported by the results of the ANOVA test, which indicate that the regression model is statistically significant (F = 13.7405, p ¡ 0.0001). Additionally, the regression coefficients associated with unit price demonstrate statistical significance, with a t-statistic value of 2.3849 and a p-value of 0.0247, indicating that changes in unit price have a meaningful impact on total sales revenue.

Price optimization: how changes in unit price affect total sales revenue is essential for maximizing profitability and meeting customer demand.

Promotional Pricing Optimization: Regression analysis can evaluate the effectiveness of promotional pricing strategies by examining the impact of temporary price reductions or discounts on total sales. This helps Walmart optimize promotional pricing tactics to achieve desired sales objectives without sacrificing profitability.

Evaluating Product Performance: Examining the correlation between unit price and total sales over time yields valuable insights into product success and market trends. Monitoring fluctuations in sales volume in relation to price adjustments

enables Walmart to evaluate the impact of its product positioning, branding, and marketing strategies.

Dynamic Pricing Strategies: Regression analysis enables Walmart to implement dynamic pricing strategies that adjust prices in real-time based on factors such as demand fluctuations, competitor pricing, and inventory levels. This allows for more agile and responsive pricing decisions, optimizing revenue and profit margins.

Bundle Pricing Optimization: Bundling products together at a discounted price can influence purchasing decisions and increase total sales revenue. Analysing the impact of bundle pricing on total sales compared to individual unit pricing helps identify the most effective bundling strategies.

By leveraging regression analysis to understand the relationship between unit price and total sales, Walmart can make data-driven pricing decisions that enhance profitability while meeting customer expectations and market demands.

Cost of goods sold and total Sales.

A visual inspection of the data suggests a negative correlation between COGS and Total Sales. As the cost of goods sold increases, there's generally a corresponding decrease in total sales figures. This aligns with the basic economic principle that higher sales typically necessitate lower production cost on goods. Thus, we observe a trend where higher costs of goods sold tend to coincide with lower total sales.
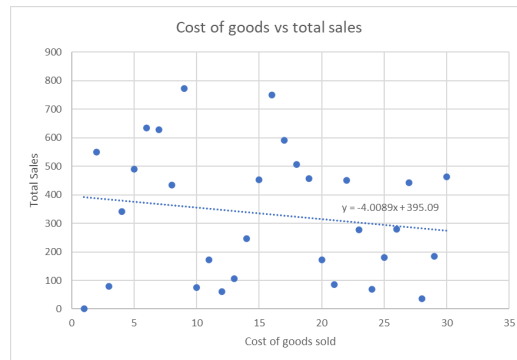

Figure 4.8: Cost of goods sold and total Sales

The regression analysis indicates a perfect correlation (Multiple R = 1) between Cost of Goods Sold (COGS) and Total Sales, with an R-squared value of 1, signifying that all variability in Total Sales can be explained by changes in COGS.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	1
R Square	1
Adjusted R	1
Standard E	5.48E-14
Observatio	29

ANOVA

	df	SS	MS	F	ignificance F
Regression	1	1375858	1375858	4.59E+32	0
Residual	27	8.1E-26	3E-27		
Total	28	1375858			

	Coefficients	andard Err	t Stat	P-value	Lower 95%	Upper 95%	ower 95.0%	pper 95.0%
Intercept	5.68E-14	1.9E-14	2.987434	0.005925	1.78E-14	9.59E-14	1.78E-14	9.59E-14
Cost of Go	1.05	4.9E-17	2.14E+16	0	1.05	1.05	1.05	1.05

Figure 4.9: output

The ANOVA results further confirm the statistical significance of the regression model, with an extremely high F-statistic value and a p-value of 0. This indicates that the relationship between COGS and Total Sales is not due to chance.

Additionally, both the intercept and the coefficient for COGS are statistically significant, with p-values of 0.0059 and 0 respectively. This suggests that changes in COGS have a significant impact on Total Sales, and for every unit increase in COGS, Total Sales increase by 1.05 units.

Impact of regression Analysis on Price optimisation, inventory turnover and logistics efficiency

Price Optimization: The regression analysis revealed a perfect correlation between Cost of Goods Sold (COGS) and Total Sales. This correlation implies that changes in COGS directly impact Total Sales, and vice versa. Therefore, Walmart can adjust their pricing strategies dynamically in response to fluctuations in COGS. For example, if COGS decrease due to efficiency improvements or cost-saving measures, companies can consider lowering prices to stimulate demand and potentially increase sales volume. Conversely, if COGS rise, businesses may need to adjust prices upward to maintain profit margins.

Inventory Turnover: Efficient inventory management is essential for profitability, as highlighted by regression analysis showing COGS's significant impact on Total Sales. Monitoring these metrics enables businesses to optimize inventory levels, striking a balance between supply and demand. This strategy improves inventory turnover rates, reduces carrying costs, and maximizes sales opportunities, ultimately enhancing overall profitability.

Improving logistics efficiency: Improving logistics efficiency is vital, given the perfect correlation between COGS and Total Sales, indicating a direct link between production costs and revenue. By concentrating on streamlining supply

chain operations, reducing transportation expenses, and boosting delivery speed, Walmart Inc can achieve cost savings, enhance customer satisfaction, and ultimately drive-up total Sales.
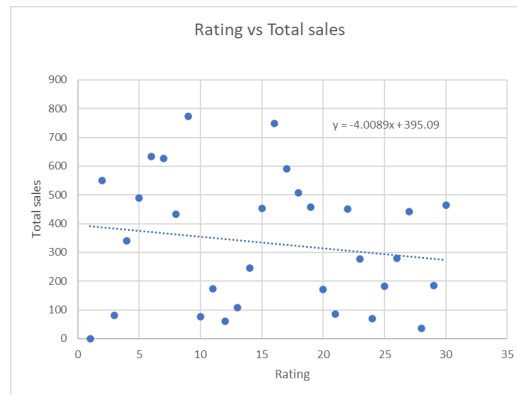
Ratings and Total Sales

Figure 4.10: Scatter plot of ratings and total sales

Figure 4.11: output

The regression analysis reveals a very weak correlation between Ratings and Total Sales, with a Multiple R value of 0.0486 and an R-squared value of 0.0024. This indicates an absence of a linear relationship between Ratings and Total Sales in the dataset. Upon further examination, there is no clear, direct relationship between Ratings and Total Sales. Despite Ratings varying widely, Total Sales figures do not exhibit a discernible pattern or trend based solely on Ratings.

Statistical significance: The coefficient for the Rating variable is not statistically significant, with a p-value of 0.8024. This suggests that changes in Ratings

do not have a meaningful impact on Total Sales, therefore Walmart may not be able to rely solely on Ratings to predict or influence sales performance.

A weak correlation between ratings and sales indicates a need to improve customer enhancement plan. Walmart Inc. Will have to re-evaluate their priorities, diversify their strategies, segment-specific approaches, data-driven decision making, and align their retail operations with the business goals to drive meaningful improvements in sales performance and overall business outcomes.

Quantity and Total Sales

The multiple R value of 0.8409 suggests a positive correlation between Quantity and Total sales revenue. This indicates that as the Quantity of products sold increases, there is a tendency for Total sales revenue to rise in tandem. Conversely, when the Quantity decreases, Total sales revenue is likely to decline accordingly.
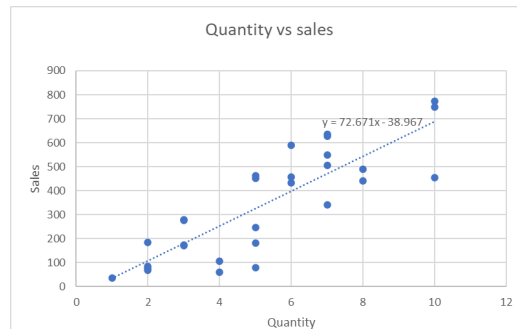

Figure 4.12: Quantity and total sales

Figure 4.13: output

Statistical significance: The ANOVA test evaluate whether the regression model is statistically significant. The F statistic value of 31.38 is compared to a critical

value, and the associated p value (1.16743E-07) indicates the probability of observing such a large F-statistic value if the model were not significant. In this case, the low p-value suggests that the regression model is statistically significant at conventional significance levels.

The coefficient for the Quantity variable (87.231) is statistically significant (p = 0.035), indicating that changes in quantity sold have a meaningful impact on total sales revenue. Specifically, for each unit increase in Quantity, there is an estimated increase in total sales revenue by approximately 87.231 units.

In summary, increasing the quantity of products sold can increase total sales revenue. While focusing on quantity-based strategies is beneficial, it's essential to prioritize the linear relationship between quantity sold and revenue over quadratic effects. By understanding these insights, businesses can make better decisions to improve sales performance and increase revenue

The findings emphasize the crucial role of pricing strategies in influencing total sales revenue. Understanding the positive relationship between unit price and total sales enables Walmart to refine pricing strategies for profitability while meeting customer expectations. Analysis of promotional, bundle, and dynamic pricing strategies offers actionable insights to adapt swiftly to market dynamics.

Efficient inventory management is another area where regression analysis has proven invaluable. The perfect correlation between the cost of goods sold (COGS) and total sales underscores the significance of inventory turnover in driving profitability. By optimizing inventory levels and streamlining supply chain operations, Walmart can minimize costs and maximize sales opportunities.

Furthermore, exploring factors like ratings and quantity sold reveals additional dimensions of Walmart's performance. While ratings may not directly correlate with total sales, understanding this relationship prompts Walmart to adopt a more segmented, data-driven approach to enhance customer satisfaction and drive sales performance improvements.

In conclusion, regression analysis highlights the transformative potential of data-driven approaches in shaping the future of retail operations. By harnessing regression models, Walmart continues to refine strategies, adapt to market dynamics, and maintain competitiveness. As Walmart navigates the complexities of the retail market, regression analysis insights serve as a guidepost for informed decision-making, fostering long-term growth and success.

Chapter 5

Discussions, Conclusions, And Recommendations

DISCUSSION OF FINDINGS

The use of a regression model to examine how the changes in the unit price may affect the sales revenue was conducted using sample data, it was concluded that there is a positive relationship between the two variables of interest being the unit price and the total sales. The model indicated that if there is an increase in the unit price the sales will also increase, which brings increases the sales at Walmart Inc. According to [14] the new digital era has several trends in the way customers are pricing their products including the dynamic pricing: changing the prices of online products 3-4 times a day. This is one of the studies in the literature review that can be implemented by Walmart Inc. to gain a competitive advantage and increase their sales. Walmart Inc. has an online platform, where customers can purchase their products through this e-commerce channel. Dynamic pricing can be implemented using the prices that are predicted to bring higher sales in Walmart Inc., made possible through the regression model to predict the best prices Walmart Inc. can use to position themselves strategically in the market.

Another variable of interest from the Walmart Inc. dataset sourced from Kaggle was the Cost of Goods Sold which is an expense incurred during the production of the products Walmart Inc. has to offer to their customers. The regression model was used to examine how the Cost of Goods Sold influences the performance of the sales using sample data. A negative correlation was identified between the Cost of Goods Sold and the sales, meaning the higher the Cost of Goods Sold the lower the total sales. The Cost of Goods Sold is an expenditure incurred by Walmart Inc. during the production of the goods that they produce.

This indicates a direct impact on the profit margins, as this inverse relationship may be a result of economies of scale, which can make Walmart Inc. adjust their pricing strategies to increase the demand and market share. There is emphasis on good and proper store layout which was made by [28]. The author claims that if the layout design optimizes shelf space, improve product placement, and reduce congestion leading to better inventory management, it will influence the buying behaviour of the customers at Walmart Inc. This will in turn lead to higher sales at Walmart Inc., which will minimize the stock outs and lower the carrying costs, which affect the Cost of Goods Sold if a well-planned layout design is implemented at Walmart Inc.

One of the variables of interest was the ratings, with the use of the regression analysis, it was concluded that there is a weak correlation between the ratings and the total sales. An improvement in the customer enhancement plan can help Walmart re-evaluate their priorities to focus on the other factors that affect the performance of the sales directly through aligning their retail operations with their goals. According to [10] there needs to be an improvement in the data collection so that through text mining Walmart Inc. can be able to dig through their negative reviews and bad ratings as well as positive comments about their products and services to be able to identify key areas. This was not an expected result as; it was assumed that the ratings have a direct impact on the total sales of Walmart Inc.

A positive correlation was identified between the quantity and total sales after testing the sample data using the regression model. This indicates that if the quantity of the goods sold at Walmart increases, the higher the sales. The quantity is a variable in the dataset that represents the quantity of any of the products that a customer purchased online all at one time. This indicates that the inventory at Walmart Inc. for certain products should be high as most of those products are ordered in bulk which will lead to an increase in the sales at Walmart. [19] claims that if the inventory management is planned and implemented right, the stock levels can be kept at an optimum level at Walmart Inc.

There is need for the Key-Value Items to be identified by the management at Walmart Inc., as they can help in improving the pricing strategy. The top performing product line is the Health and Beauty as shown by the analysis of results, as they have higher total sales than any other product line. Walmart Inc. can use the trends showing the best performing product line to determine their KVI. This can enable Walmart Inc. to respond swiftly to the market fluctuations, competitive threats and changing customer behaviours that have an influence on the sales as indicated in the results in the form of visualizations. To attract customers Walmart Inc. can strategically price inexpensive items to entice customers and increase the

41

overall sales [15].

Limitations

Lack of control: There is no control over the design, collection, and formatting of the secondary data. This research relies on the Walmart Inc. data set that was sourced from Kaggle. The dataset does not fully meet the methodological requirements that were stated in the Terms of Reference template, the experimental work was not carried out as the analysis done on this data is secondary data analysis. The experimental work like technology trials was not carried out as it required more data than what was sourced from the Walmart Inc. dataset.

Unavailability of metadata: Secondary data lacks detailed metadata or documentation that describes how the data was collected, processed, and analysed. Without access to this metadata there can be cases where the context of the data is misunderstood. This causes issues when it comes to assessing the quality of the data, reliability, and limitations of the data, if the access to the comprehensive metadata is not given, which was the case when conducting this research.

Limited Scope: The data used when conducting this research was for a period of three months only. This was very limiting as the focus was now on analysis of the data over a period of three months, which is not very comprehensive. It would have been more suitable to identify trends from data that can have the analysis done annually to identify hidden trends and patterns.

Recommendations

The purpose of this section is to propose actionable steps and directions for future research based on the results/findings of the study conducted on Walmart Inc's retail operations. It offers practical recommendations for implementing dynamic pricing strategies, enhancing inventory management practices, prioritizing key-value items improving data collection and analysis, and monitoring market trends. Some of these practical recommendations include:

Implementing pricing strategies: To optimize sales revenue and maintain competitiveness, Walmart Inc. should implement dynamic strategies that adjust unit prices based on market demand (looking at consumer ratings to determine whether customers are satisfied with the prices), competitive pricing (setting prices based on what Walmart's competitors charge) and other relevant factors.

Enhance Inventory Management Practices: A highlighted in [5] effective inventory management is important because it helps to lower the Cost of Goods Sold (a variable from the dataset which as been analysed). This careful monitoring and control of Walmart's inventory levels includes optimizing stock levels, minimizing stockouts and streamlining supply chain processes to ensure efficient inventory turnover to meet customer demand promptly and prioritise customer

42

satisfaction to increase sales.

Prioritizing Key-Value Items: Top-performing product lines can be identified as key-value items. Strategically pricing and promoting them can capitalize on their popularity and drive overall sales. However, these sales are to be monitored continuously to adjust pricing strategies according to their performance in the market.

Enhance Data Collection and Analysis: Investing in advanced data collection and analysis capabilities, including text mining of customer feedback, can provide deeper insights into customer preferences and satisfaction. Leveraging advanced data analytics can identify areas for improvement and implement strategies to enhance customer experience and drive sales.

Monitor Market Trends and Competitor Actions: This was done to some extend when comparing the sales of different branches across the various product lines as explained in the Exploratory Data Analysis. However, this can still be done with Walmart Inc. against its key competitors by monitoring pricing strategies, customer preferences and product offerings to changing market dynamics to help it gain a competitive edge.

Directions for Future Research

Future studies can build upon the findings of this research and address gaps identified by focusing on the following areas or exploring new questions that emerged: Longitudinal Analysis: Conduct a longitudinal analysis using a more extensive dataset covering a more extended period than the three months in the WalmartSalesData.csv. This would allow for a more in-depth examination of long-term trends and seasonal variations in retail operations and sales performance, providing a comprehensive understanding of Walmart Inc.'s dynamics over time.

External Factors Analysis: Investigate the impact of external factors, such as economic conditions, regulatory changes, and competitor actions, on Walmart Inc.'s retail operations and sales performance. Understanding how these external factors influence the company's strategies and outcomes can provide valuable insights for strategic decision-making.

Dynamic Pricing Strategies: Explore the effectiveness of dynamic pricing strategies in response to real-time market fluctuations and customer demand patterns. By delving deeper into how dynamic pricing can optimize pricing decisions and enhance sales performance, future studies can provide Walmart Inc. with actionable insights for pricing strategy refinement.

Customer Segmentation Analysis: Further investigate the role of customer segmentation and personalized pricing strategies in optimizing sales performance

and enhancing customer satisfaction. By tailoring pricing strategies to specific customer segments effectively, future research can help Walmart Inc. better understand customer preferences and behaviours, leading to more targeted marketing and pricing strategies.

Enhanced Data Collection and Analysis: Invest in advanced data collection and analysis techniques, such as text mining of customer feedback and sentiment analysis, to gain deeper insights into customer preferences and satisfaction. By leveraging advanced analytics methods, future studies can uncover hidden patterns and trends in customer data, enabling Walmart Inc. to enhance its customer experience strategies.

Ethical Considerations in Data Analysis: Further explore ethical considerations in data analysis, ensuring that privacy rights are protected, biases are avoided, and transparency is maintained throughout the research process. By emphasizing ethical practices in data analysis, future studies can uphold fairness, equity, and integrity in research outcomes.

Conclusions

A Walmart Inc. Dataset was collected via the internet. The dataset has variables, Invoice ID, branch, city, customer type, gender, product line, unit price, quantity, tax 5%, total, date, time, payment, cost of goods sold, gross margin percentage, gross income, and rating. The exploratory data analysis report outlines a strategic approach for Walmart Inc. to leverage data analytics in optimizing retail operations, thereby enhancing operational efficiency, driving revenue growth, and gaining a competitive advantage in the dynamic retail landscape. By embracing data-driven decision-making processes, Walmart Inc. can position itself for sustained success and prosperity. The project's focus on utilizing data analytics aims to provide actionable insights derived from exploratory data analysis (EDA), targeting key areas such as total sales, customer segmentation, product lines, and pricing strategies. By analyzing these variables from the Walmart Inc. dataset, the aim is to uncover opportunities for improvement and implement strategies to increase sales and profitability. Ultimately, this road map for leveraging data analytics offers Walmart Inc. a pathway to enhance performance, adapt to evolving market trends, and thrive in the competitive retail environment.

Predictive modelling was used using the statistical technique linear regression to forecast consumer demand, optimize inventory levels, and enhance sales forecasting accuracy. The results of the predictive modelling are as follows in the key findings.

The key findings of the study:

Regarding predictive modelling, Walmart's primary tool for achieving opera-

tional excellence is regression analysis. Regression analysis is a sort of predictive modelling that helps us understand the relationships between variables and produce well-informed projections based on historical data. Walmart has examined the intricate links between inventory control, customer behaviour, and sales forecasting using regression models. Therefore, with the use of regression analysis techniques, there were multiple findings for the study on the implication of regression analysis at Walmart Inc.

There have been three main key findings:

Product unit price and total sales

The first finding was about the unit price per product vs the total sales by Walmart Inc. This finding came in as a positive relationship which indicated that as unit prices increase, total sales revenue increases as well. This implies that consumers are prepared to pay more for goods with greater unit costs, which boosts Walmart's overall income. This finding directly relates to the outcome and deliverable we established as price optimization recommendations. This talks about how the use of data analytics can be implemented in the concept of how price influences sales in Walmart Inc. The higher unit prices in correlation with higher sales respectively indicate that Walmart's competitive pricing strategies attract price-sensitive customers who are willing to spend more. This in turn increases revenue for the organisation as more expensive products are purchased by customers. Furthermore, Walmart has a high-profit margin due to the profitability in selling high-priced goods more than low-priced goods. The statistical significance of the results shows that there is a positive correlation between unit price and total sales revenue. This indicates that indeed, changes in unit price have an impact on total sales, which is a positive impact as it increases profits.

Cost of goods and total sales

This study came in as a negative correlation between cost of goods and total sales. This implies that as the cost of goods increases, there is a corresponding decrease in total sales figures. High total sales of a product may be a sign of significant market acceptability and demand. It may be advised to change the prices of certain products to increase profitability without reducing sales volume. On the other hand, low overall sales of a product could indicate poor demand or problems with price. It may be suggested to alter prices or implement promotional tactics to increase revenue generation and boost sales. Walmart can make well-informed pricing decisions by integrating the cost of goods and total sales data into the price optimisation process. Walmart may enhance its pricing strategy to optimise sales, profitability, and customer satisfaction by utilising data analytics, rival pricing, and customer preferences. By taking a holistic strategy, pricing adjustments are

guaranteed to be grounded on a deep comprehension of demand elasticity, market dynamics, and cost structures, ultimately resulting in sustainable business growth.

Ratings and total sales

This study came in as a very weak correlation between ratings and total sales. This means there is no relationship between the ratings and total sales for Walmart Inc. Walmart can adjust its inventory management optimisation tactics to match customer preferences and demand trends by utilising ratings and total sales data. Demand forecasting, supplier relationship management, and technological solutions like RFID technology and inventory tracking software can all be added to help Walmart increase inventory management efficiency by optimising inventory levels, streamlining operations, and improving overall efficiency. Walmart Inc.'s sales and profitability have improved as a result of this all-encompassing strategy, which guarantees that inventory management procedures stay adaptable to shifting market conditions.

Bibliography

[1] Julian Wright Andre Hagiu. When data creates competitive advantage, 2020.

[2] Omer Artun and Dominique Levin. *Predictive marketing: Easy ways every marketer can use customer analytics and big data.* John Wiley & Sons, 2015.

[3] Haluk Demirkan and Dursun Delen. Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, 55(1):412–421, 2013.

[4] B. K. C. D'Lima. Benchmarking as a measure of competitiveness. *International Journal of Process Management and Benchmarking*, 4(3):342, 2014.

[5] FasterCapital. Inventory management: How effective inventory management can lower your cost of goods sold. https://fastercapital.com/content/Inventory-management–How-Effective-Inventory-Management-Can-Lower-Your-Cost-of-Goods-Sold.html, n.d. Accessed 17 April 2024.

[6] Noemi Ferracuti, Claudia Norscini, Emanuelle Frontoni, Patrizia Gabellini, Marina Paolanti, and Valerio Placidi. A business application of rtls technology in intelligent retail environment: Defining the shopper's preferred path and its segmentation. *Journal of Retailing and Consumer Services*, 47:184–194, 2019.

[7] Jaime Garza. Using business intelligence tools to conduct better margin analytics: A hypothetical case study. *Journal of Performance Management*, 20(3):47, 2007.

[8] Varun Grover, Roger HL Chiang, Ting-Peng Liang, and Dongsong Zhang. Creating strategic business value from big data analytics: A research framework. *Journal of management information systems*, 35(2):388–423, 2018.

[9] Rajeev Gupta, Himanshu Gupta, and Mukesh Mohania. Cloud computing and big data analytics: what is new from databases perspective? In *International conference on big data analytics*, pages 42–61. Springer, 2012.

[10] Wu He, Jiancheng Shen, Xin Tian, Yaohang Li, Vasudeva Akula, Gongjun Yan, and Ran Tao. Gaining competitive intelligence from social media data: Evidence from two largest retail chains in the world. *Industrial management & data systems*, 115(9):1622–1636, 2015.

[11] Hyunwoo Hwangbo et al. A business application of rtls technology in intelligent retail environment: Defining the shopper's preferred path and its segmentation. 2017.

[12] IBM. Ibm exploratory data analysis. https://www.ibm.com/topics/exploratory-data-analysis, 2020. Accessed 20 March 2024.

[13] Harsha Khubwani. From point-of-sale to point-of-smile: Walmart's data strategy enhancing cx and competitive edge, 2023.

[14] Thomas Kilroy, Ian MacKenzie, and Audrey Manacek. Pricing in retail: Setting strategy. Technical report, McKinsey Global Institute, 2015.

[15] V. Kuchkanov. Kvis in retail: How businesses use them to boost sales. 2023.

[16] Pohan Lin. How to use retail data analytics to boost sales, 2024.

[17] L. Marion. Business intelligence: 10 common mistakes. 2008.

[18] Mayank. Globalisation of retailing, 2022.

[19] Sebastian Müller, Jakob Huber, Moritz Fleischmann, and Heiner Stuckenschmidt. Data-driven inventory management under customer substitution. *Available at SSRN 3624026*, 2020.

[20] Olga Pak. *Three essays on optimization and decision-making solutions in grocery retail operations*. PhD thesis, University of South Carolina, 2020.

[21] L. Pederson, V. Van Woerkom, A. Kent, and M. Ravelojaona. Use of secondary data analyses in research: Pros and cons. *Journal of Addiction Medicine and Therapeutics Science*, pages 58–60, 2020.

[22] Ramesh Sharda, Dursun Delen, and Efraim Turban. *Business Intelligence, Analytics and Data Science: A Managerial Approach*. Boston: s.n., 2017.

[23] Deepak singh. How data analytics can help retail companies, 2023.

[24] Alan D Smith. Competitive approaches to new product development: A comparison of successful organizations in an unstable economic environment. *Team performance management: an international journal*, 17(3/4):124–145, 2011.

[25] J. Smith. Leveraging regression analysis for sales forecasting: A case study of walmart. *Journal of Business Analytics*, 2023.

[26] M. Snow et al. *Organisational Strategy, Structure, and Process*. 1978.

[27] Chitkara University. The basics of retail management: Everything you need to know, 2023.

[28] Adam P Vrechopoulos, Robert M O'keefe, Georgios I Doukidis, and George J Siomkos. Virtual store layout: an experimental comparison in the context of grocery retail. *Journal of retailing*, 80(1):13–22, 2004.

[29] C White. Business intelligence: Becoming a smart business. *Retrieved from www. b-eyenetwork. com. Accessed on April*, 16:2005, 2005.

[30] R. Wickham. Secondary analysis research. *Journal of the Advanced Practitioner in Oncology*, 10(4):395–400, 2019.

Appendix A

Appendix: Walmart Case Study

Introduction

This appendix provides additional information and supplementary materials related to the Walmart case study on optimizing retail operations through data analytics.

A: Sample Dataset

- The sample dataset used for the analysis contains 17 attributes and 101 observations, uniquely identified by Invoice ID.

- Attributes include numerical and categorical variables such as total sales, customer segmentation (customer type and gender), product lines, and pricing strategies.

- The dataset was extracted from Kaggle and initially examined for missing values and inconsistencies.

- Data cleaning procedures were applied to ensure data quality and consistency.

B: Regression Analysis Results

- Regression analysis was conducted to examine the relationship between unit price and total sales revenue.

- The analysis revealed a positive correlation between unit price and total sales, indicating that higher unit prices are associated with higher sales revenue.

C: Summary of Key Findings

- A summary of key findings from the exploratory data analysis, regression analysis, and other data-driven insights.

- Insights include recommendations for inventory management, pricing strategies, and customer segmentation to optimize retail operations and increase sales.

D: Visualizations and Charts

- Visual representations of the data, including scatter plots, bar charts, and histograms.

- These visualizations provide additional insights into patterns, trends, and relationships within the dataset.

E: Methodological Details

- Detailed descriptions of the data collection, preparation, and analysis methods used in the project.

- This section provides transparency and reproducibility for the research methodology employed.

F: References

- A list of references cited throughout the project, including academic literature, industry reports, and studies.

- These references support the findings and recommendations presented in the main report.

G: Glossary of Terms

- Definitions of key terms and concepts used throughout the report, ensuring clarity and understanding for readers.

Conclusion

the appendix provides supplementary information and additional insights to support the main findings and recommendations of the Walmart case study. These materials enhance the comprehensiveness and credibility of the analysis, enabling stakeholders to make informed decisions and implement actionable strategies for optimizing retail operations and increasing sales.