

Supermarket Billing Analysis Using Data Analytics for Customer Purchase Insights

Shubhangi Yadav

Department of Science and Technology, G.H.Raisoni Skill Tech University, Nagpur, India

ABSTRACT

Supermarkets generate a lot of information every day. This information includes what products are sold, how many of these products are sold, at what price they're sold and how people purchase them. Many supermarkets do not make use of this information. They mostly use their billing systems to process sales not to gain knowledge from the information. If they properly examined this information they could learn about the shopping habits of their customers how money they are making and how demand for products changes. This study uses data analysis and machine learning to help supermarkets make decisions. First, we prepare the supermarket data. By understanding these things GS supermarkets can make decisions and run more smoothly. We use tools like Linear Regression and Decision Tree algorithms to predict sales at supermarkets. We check how well these tools work by looking at their accuracy and other measures. We found that by examining billing data from supermarkets can make decisions about inventory prediction, forecast demand for products and plan promotions. Supermarket billing data can be very useful. It is about turning this data into something so supermarkets can make decisions and stay ahead of other supermarkets. By using data analysis and machine learning supermarkets can understand what is working for supermarkets and what is not and make changes. This can lead to sales, happier customers and a more sustainable business, for supermarkets.

Keywords: Supermarket Billing Analysis, Retail Data Analytics, Sales Forecasting, Machine Learning, Inventory Optimization, Customer Behaviour Analysis.



This is an open-access article under the [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) license

1. Introduction

Supermarkets generate an amount of data daily from their computer billing systems. Every time a customer makes a purchase from the supermarket the system records details such as what they bought from the supermarket how much they bought from the supermarket and how much it cost them. The question is, how did they pay for the items they bought from the supermarket. This vast amount of data from the supermarket is an opportunity for analysis and business insights for the supermarket.

Traditionally most supermarket billing systems were used for recording sales and making invoices for the supermarket. However, with the increasing competition in retail there is a growing need to utilize billing data from the supermarket for more than daily operations.

Analysing billing records from the supermarket can provide information about how customers buy products from the supermarket what products are in demand at the supermarket how much money the supermarket makes and when people shop at the supermarket. By employing data-driven

strategies supermarkets can improve operations at the supermarket manage inventory effectively at the supermarket increase revenue at the supermarket and gain a competitive advantage, in retail.

1.1 Motivation

The retail industry is growing fast and supermarkets are competing. This means that making decisions based on data is no longer a choice it is necessary. Modern supermarkets use computer systems to handle thousands of transactions every day. Each transaction has a lot of information about the products, customers, prices and how people pay. Many retail businesses only use this data to record transactions and do basic accounting. They are not using the data to get insights.

One of the reasons for this research is to turn raw data from billing into useful information for businesses. Supermarkets do not make a lot of profit. Even small improvements in how they plan their inventory price their products or target their customers can make a big difference. By looking at the billing data retailers can see which products are selling well which ones are not moving and reduce waste from having much stock or products that are expiring. This helps them manage their stock better and make their supply chain more efficient.

Another reason for this research is to understand how customers behave. In a market keeping customers is just as important as getting new ones. Supermarkets sell more during holidays and special occasions. If they do not analyze the data they might run out of stock. Have too much. By using the billing history to forecast sales supermarkets can prepare for the demand. Make sure their customers are happy. Also many retailers are now using technology, which means they have access to a lot of data and tools to analyse it. With the help of machine learning and data visualization medium-sized supermarkets can use intelligent billing analysis systems. This research is motivated by the opportunity to use these technologies to improve how retailers manage their businesses.

Finally, this study is motivated by the desire to support efficient retail operations. Analysing billing data efficiently not increases profits but also reduces waste improves the customer experience and helps businesses make smarter decisions. In a time when data's an asset using billing data effectively is crucial for long-term success. So, the Supermarket Billing Analysis project is motivated by the need to use data intelligently improve how businesses operate strengthen customer relationships and enable growth, in the retail sector based on data.

1.2 Contribution

The Supermarket Billing Analysis project makes some contributions to research and practical retail management. The main contribution of this study is creating a framework. This framework turns supermarket billing data into business intelligence. Unlike billing systems that record transactions this research shows how transactional data can be processed, analysed and interpreted to support strategic decision-making for the Supermarket Billing Analysis project.

One major contribution is designing an end-to-end data analysis pipeline for supermarket environments. The study outlines steps like this:

- ✓ Collecting data from POS systems
- ✓ Cleaning and preprocessing the data
- ✓ Transforming the data
- ✓ Doing exploratory data analysis
- ✓ Building models

This structured methodology can serve as a reference model for medium-scale retail businesses seeking to implement data analytics solutions for Supermarket Billing Analysis. The Supermarket Billing Analysis project also combines analytical techniques. These include:

- ✓ Analytics to measure sales performance and revenue trends for Supermarket Billing Analysis
- ✓ Analytics to understand sales fluctuations in Supermarket Billing Analysis
- ✓ Market Basket Analysis to identify product associations in Supermarket Billing Analysis
- ✓ Predictive analytics for sales forecasting and demand planning in Supermarket Billing Analysis

By integrating these approaches, the Supermarket Billing Analysis research provides a complete analytical model for Supermarket Billing Analysis. The study contributes to inventory management strategies for Supermarket Billing Analysis. It analyses billing records to identify moving and slow-moving products. This helps supermarkets optimize stock levels. Reduce wastage. Good inventory management reduces costs. Prevents revenue loss due to expired goods in Supermarket Billing Analysis. From a customer perspective the Supermarket Billing Analysis project provides insights into purchasing behaviour and spending patterns. By analysing transaction frequency, average bill value and product preferences the study supports marketing strategies and loyalty programs for Supermarket Billing Analysis. This improves customer satisfaction and long-term retention which're critical in competitive retail markets.

The Supermarket Billing Analysis project also implements forecasting models for demand prediction. It uses billing data to predict future sales trends. This helps supermarkets prepare for demand and festive sales spikes in Supermarket Billing Analysis. The project shows how modern data analytics tools like Python libraries SQL databases and visualization platforms can be used. These tools help build cost- retail intelligence systems without large-scale enterprise infrastructure for Supermarket Billing Analysis. The Supermarket Billing Analysis research bridges the gap between data mining concepts and real-world supermarket operations. It translates techniques into retail strategies. This practical alignment enhances the relevance of data analytics research in business for Supermarket Billing Analysis. The Supermarket Billing Analysis project promotes a data-driven culture in businesses. It demonstrates improvements in decision-making, profitability and operational efficiency. This encourages supermarkets to adopt thinking as part of their strategic planning process.

2. Related Work

The area of analytics and supermarket billing analysis is important these days. It is getting a lot of attention from researchers and companies because supermarket billing analysis and analytics are crucial for them. Many studies have looked at how data from sales can be used to make stores run better get customers to come back and make money. Supermarket billing analysis is a part of this as it helps stores understand what sells well and how customers buy things.

- This book talks about data mining methods, such as finding patterns grouping things and predicting what will happen all of which're essential for supermarket billing analysis.
- These methods are used a lot in supermarket billing analysis to find patterns in amounts of sales data, which's vital for understanding customer behaviour and improving sales.

Another important book is "Data Mining: The Textbook" by Charu C. Aggarwal. This book discusses ways of analysing data and using machine learning both of which're critical for supermarket billing analysis and retail analytics. The ideas of customer segmentation, frequent pattern mining and time-series forecasting are very useful for understanding how customers buy things and what sells well which is the goal of supermarket billing analysis. Market Basket Analysis has been studied a lot in research. It is highly relevant to supermarket billing analysis. The Apriorism algorithm, which was introduced by Rakesh Agrawal and Ramakrishnan Srikant helps find items that are often bought together which's useful for supermarket billing analysis. This information can be used to sell things to customers, bundle products and put the products on shelves, all of which're important for supermarket billing analysis and retail analytics. Many

supermarkets use these techniques to increase the amount customers spend, which's a key objective of supermarket billing analysis.

Understanding customer behavior is also very important for supermarket billing analysis as it helps stores tailor their offerings to customer needs. Research in marketing such as the concepts presented in "Marketing Management" by Philip Kotler and Kevin Lane Keller shows that customer segmentation and personalized marketing are crucial for supermarket billing analysis. These ideas support the use of sales data to group customers based on how they buy things how much they. What they like to buy, all of which are essential for supermarket billing analysis. Newer methods use machine learning models like Random Forest and Neural Networks to make predictions, which's useful for supermarket billing analysis.

Most research focuses on ideas or large companies, but supermarket billing analysis is important for all stores regardless of size. This research tries to fill that gap by combining established theories, customer behavior models and forecasting techniques into a system for supermarket billing analysis. It builds on work to create a practical scalable and cost-effective framework that is suitable for modern retail operations with a focus on supermarket billing analysis. So the research that has been done provides a foundation for supermarket billing analysis.

Retail analytics and supermarket billing analysis are the focus of this research. It aims to provide a useful tool, for retail analytics and supermarket billing analysis.

3. Research Methodology

3.1 Problem Statement

In stores today supermarkets get a lot of information from the bills they make. Every time a customer buys something the store records what the customer bought how much of it the price, any discounts, tax, the time and how they paid. This information is very useful for the supermarket. If they have much of something it costs them money to store it and it might go bad especially if it is food. If they do not have enough of something they miss out on sales and customers get upset. Supermarkets also have a time predicting how busy they will be on weekends, holidays or special events. Another issue is that supermarkets do not understand when people like to shop and how they like to pay. If they do not know how people like to pay, they cannot make it easier for them to buy things. A lot of supermarkets do not have the tools they need to look at all the information they get from bills. They only use software to make invoices. If they can do this, they can run their supermarkets better manage their inventory make money and be more competitive. Supermarket billing data is very important. Supermarkets need to start using it in a way. By using supermarket billing data supermarkets can make a lot of improvements.

3.2 Data Preprocessing

Data preprocessing is a step when we do research. It is not in the format. The preprocessing stage has steps:

3.2.1 Removing Duplicate Records

Sometimes we get transactions in the billing data. This can happen because of mistakes in the system when we enter the billing data. These duplicate records can make our sales analysis of the billing data wrong. They can make it look like we sold more than we did. To fix this problem with the billing data:

- We looked at the dataset for Invoice IDs in the billing data.
- We found the duplicate rows with the transaction details in the billing data.

- We removed the records. Kept the original valid transaction in the billing data. This way we made sure that each transaction in the billing data was unique and accurate.

3.2.2 Handling Values

Sometimes the billing data is missing some information. This can happen because the transaction was not complete or because of system failures. Fields like product category or unit price in the billing data might be empty. We did the following:

- If some important fields like Invoice ID or Total Amount were missing, we removed those records to keep the billing data accurate.
- For the missing values we used techniques to fill them in. For example, we used the value for categories and the average value for numbers in the billing data.
- We checked to make sure that the Total Amount was correct by multiplying the Unit Price and Quantity in the billing data. By handling the missing values, we made the billing data complete and more accurate.

3.2.3 Converting Date and Time

The Date and Time attribute was in a format. We needed to convert it into a format that we could analyse for the billing data. We did the following:

We converted the date-time column into a format for the billing data. * We broke down the timestamp into parts like:

- ✓ Year
- ✓ Month
- ✓ Day
- ✓ Hour
- ✓ Day of the week

This way we could analyse the sales trends over time. Find out when people shop the most from the billing data.

3.2.4 Creating New Features

We created variables from the existing attributes to make the analysis of the billing data better. We did this by:

- Creating a month-sales indicator to see how sales change each month in the billing data.
- Creating a day-of-week feature to compare sales on weekdays and weekends in the billing data.
- Creating an hour-of-day feature to find out when people shop the most from the billing data.
- Calculating the transaction value by dividing the revenue by the number of invoices in the billing data.
- Calculating the revenue per category by adding up the sales for each product category in the billing data. These new features helped us understand how people buy things and made our predictions more accurate from the billing data. Overall data preprocessing made sure that the billing dataset was clean and accurate. It helped us get results, from our analysis. Made our predictions more accurate. Billing data needs to be clean and structured so that we can analyse it and make predictions. Data preprocessing helps us do this for the billing data. We need to preprocess the billing data to make sure it is accurate and reliable.

4. Research Methodology

The Supermarket Billing Analysis project is done step by step.

- We start by collecting data from supermarkets. This data is then processed.
- After processing we analyze it to understand what it means. The main goal of the Supermarket Billing Analysis project is to take information from supermarket bills and turn it into something useful. This useful information can help store owners make decisions. The goal is to use supermarket data to make sure our decisions are fact-based. The Supermarket Billing Analysis project is important. It helps us understand what is happening in the supermarket. We use Supermarket Billing Analysis to make decisions. The project relies on Supermarket data to provide insights. This is why the Supermarket Billing Analysis project is crucial, for supermarkets. It provides an understanding of Supermarket operations.

4.1. Data Description

The Supermarket Billing Analysis project uses data from a supermarket's Point of Sale system. This data shows what customers buy every day, how well products sell, and total sales over a period like six months or a year. The Supermarket Billing Analysis project stores this data in a table format using CSV files, Excel sheets, or SQL databases. This makes it easy to analyze the Supermarket Billing Analysis data using statistics and machine learning.

- The main data used in the Supermarket Billing Analysis project is the Transaction Dataset. The Transaction Dataset has records of every purchase made by customers at the supermarket.
- Each record in the Transaction Dataset has a Transaction ID, date, and time of purchase at the supermarket.
- The Transaction Dataset also has product details like Product ID, Product Name, and Category of the products bought at the supermarket.
- The Transaction Dataset shows how much of each product was bought, which is the Quantity of the products.
- The price per unit of each product is also included in the Transaction Dataset, which is the Unit Price of the products.

Any discount given to customers at the supermarket is included in the Transaction Dataset, which is the Discount Applied. The total amount paid by customers at the supermarket is also in the Transaction Dataset, which is the Total Bill Amount. The way customers pay for products at the supermarket, such as cash, card, or digital payment, is also tracked in the Supermarket Billing Analysis project. If available, a Customer ID is included to see if customers buy again at the supermarket and how loyal they are to the supermarket. The Supermarket Billing Analysis project also uses a Product Dataset. The Product Dataset has information about the supermarket's products.

The Product Dataset includes:

- ✓ Product ID of the products at the supermarket
- ✓ Product Name of the products at the supermarket
- ✓ Cost Price of the products at the supermarket
- ✓ Selling Price of the products at the supermarket
- ✓ How much stock is available of each product at the supermarket, which is the Available Stock Quantity.

By comparing the cost and selling price of products at the supermarket, we can see how profit each product makes for the supermarket. This helps in understanding which products make a profit for the supermarket and in evaluating how well prices are set for the products at the supermarket. The Product Dataset also helps in managing stock at the supermarket. It shows which products sell fast.

at the supermarket and need to be restocked and which ones sell slowly at the supermarket and might take up much space or go to waste at the supermarket.

4.2. Result Evaluation & Analysis

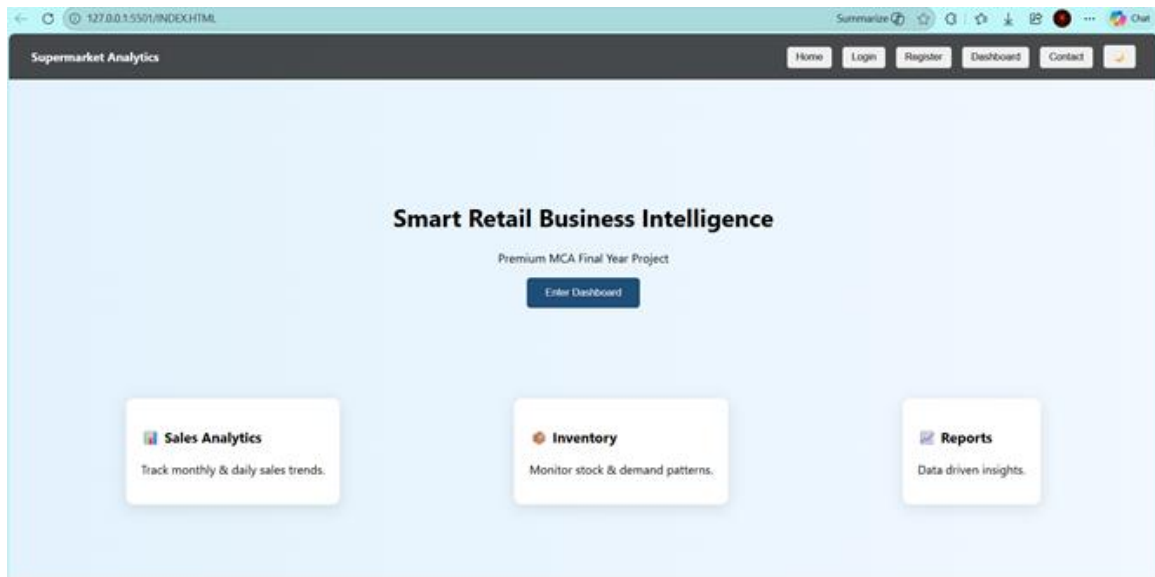


FIG : 1 Smart Retail Business Intelligence – Landing Page UI Design

This picture shows the page of the Supermarket Billing Analytics System. The Supermarket Billing Analytics System was made for a year project. It looks really nice.

The main page has some parts.

- * A menu at the top with options like Home, Login, Register, Dashboard and Contact. It also has a button to switch to mode.
- * A button that says "Enter Dashboard". When you click this button it takes you to the analytics part of the Supermarket Billing Analytics System.
- * Three boxes at the bottom that tell us about the Supermarket Billing Analytics System features:

1. Sales Analytics: This part shows how sales are doing every day and every month for the Supermarket Billing Analytics System.
2. Inventory: This part keeps an eye on how stock we have and what people are buying from the Supermarket Billing Analytics System.
3. Reports: This part gives us information to help us make decisions about the Supermarket Billing Analytics System.

The Supermarket Billing Analytics System looks really clean and easy to use. The background is soft and nice. The buttons are easy to click and use. The Supermarket Billing Analytics System is, on the web. This main page is where you start when you use the Supermarket Billing Analytics System. The Supermarket Billing Analytics System is made to be professional and easy to use.

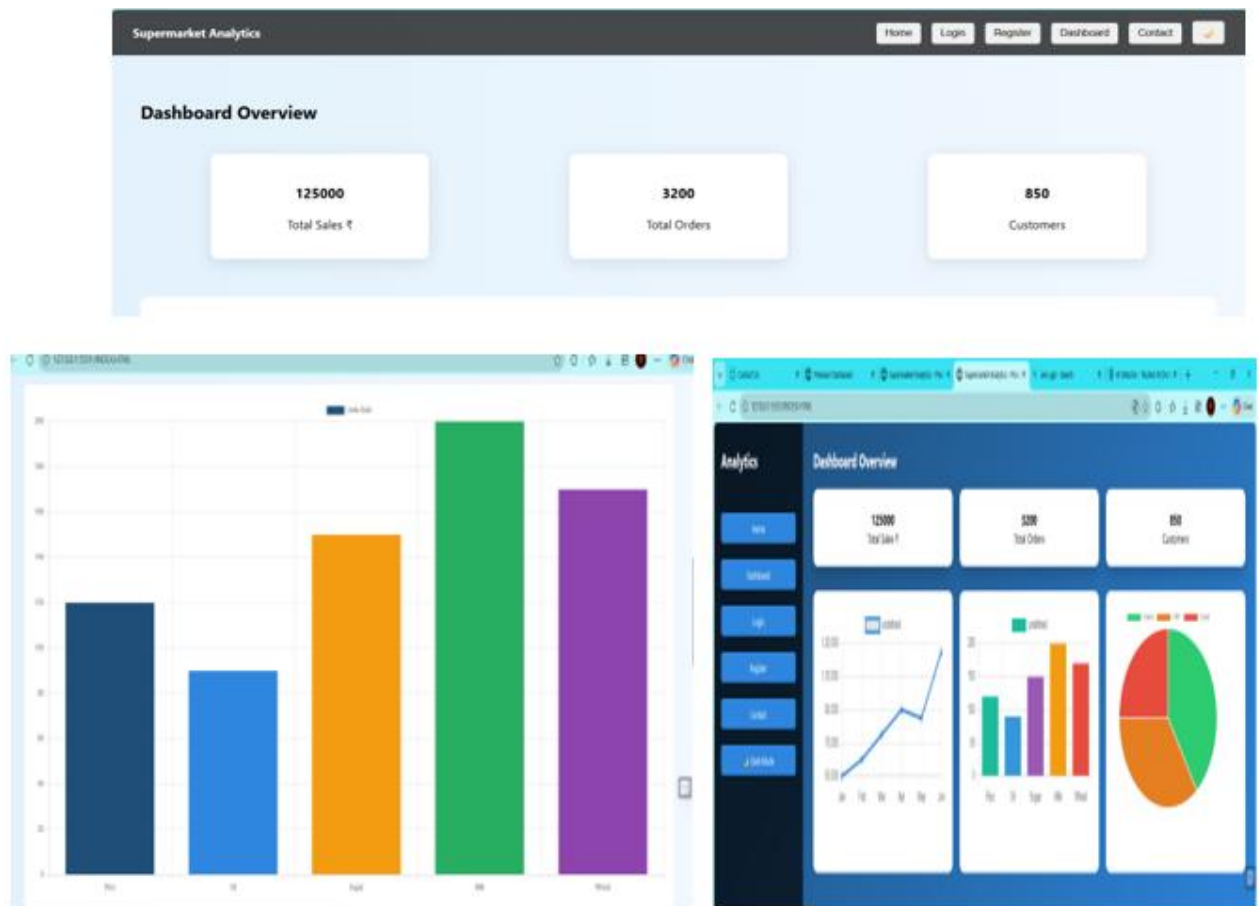


FIG : 2 : Supermarket Sales & Performance Analytics Dashboard

This set of images shows what a supermarket analytics dashboard system looks like. It helps us see how the business is doing by using numbers and interactive charts. The system has two looks one is dark and the other is light. It shows us things like total sales, total orders and how many customers we have. We can also see pictures that show us things like how sales are doing each month which products are selling the most and how people are paying for things.

The supermarket analytics dashboard system gives us a view of what is going on. We can watch sales understand what customers are doing and make decisions based on the information we have. This is really helpful for people who work in supermarkets or stores. The supermarket analytics dashboard system is very useful, for them.

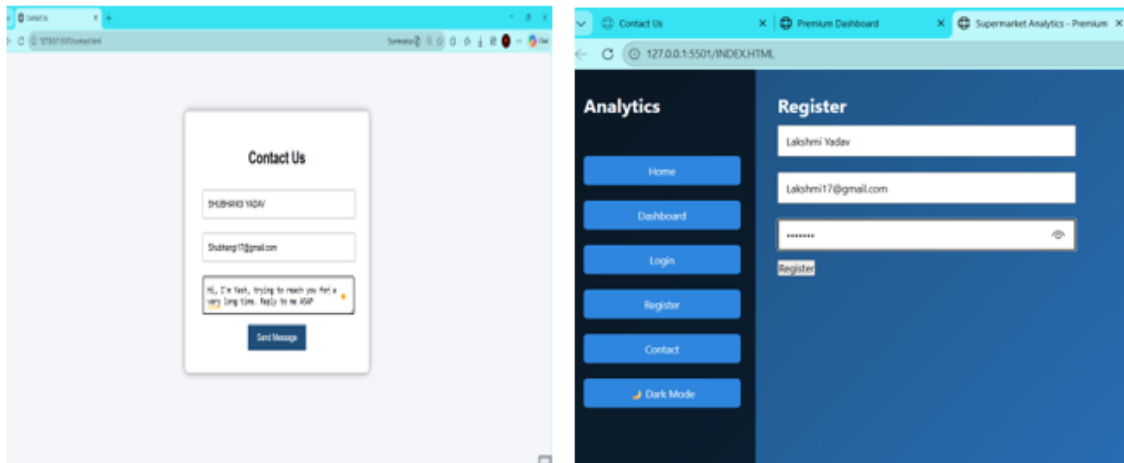


FIG : 3 Contact Us and Register Interface

This image shows the Contact Us page of the Smart Retail Business Intelligence System. I made it as my MCA final year project. The page helps users talk to the system admin or support team directly. The page has: A contact form in the middle with a card design.

The form has spaces to fill in:

- ✓ Your Name
- ✓ Your Email Address
- ✓ Your Message

There's a "Send Message" button to send questions or feedback.

The design is simple. It uses background colors and soft shadows to look modern and easy to use. The form shows how to check what users enter and handle interactions using HTML, CSS and JavaScript. This part helps the Smart Retail Business Intelligence System by letting users and admins communicate. It makes the Smart Retail Business Intelligence System easier to use and look more professional. The Smart Retail Business Intelligence System Contact Us page does its job well. The Contact Us page is really helpful, for getting in touch. It uses the Smart Retail Business Intelligence System to contact them.

4.3 Conclusion

This study looked at supermarket bills to see how we can use data to make business decisions. Usually, bills are just used to record what people buy and how much they pay. This study shows that bills have a lot of useful information that can help us make better choices if we look at them closely. We collected a lot of data from supermarket bills cleaned it up and looked for patterns. We found some things about how well the supermarket does in sales each month and day what kinds of products people like to buy how much money the supermarket makes, when people like to shop and how they like to pay. This information can help the supermarket manage its stock plan promotions and decide how staff to have. We used some computer models like Linear Regression and Decision Tree to try to predict sales based on old bills. These models were pretty good at guessing sales trends. The Decision Tree model was especially good at handling patterns in the data.

This study shows that supermarkets should use their bill data to make decisions. By looking at this data supermarkets can make sure they have the amount of stock cut down on waste lose less money and make customers happier. Using data to make decisions helps supermarkets plan better and compete with stores.

There are some limitations to this study like we did not have much data as we would have liked, and we could not look at as many things as we wanted too. We still found out that looking at

supermarket bills is very important for modern retail. If we can use computer models to predict sales and look at bill data, we can make management better and help supermarkets run more smoothly. In the end looking at supermarket bills is not about processing data it is a tool that helps supermarkets run better. By using computer models and looking at data, we can make supermarkets more efficient, profitable and sustainable. Supermarket Billing Analysis is a part of making this happen.

5. References

1. Jiawei Han, Micheline Kamber, & Jian Pei (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann.
2. Charu C. Aggarwal (2015). Data Mining: The Textbook. Springer.
3. Philip Kotler & Kevin Lane Keller (2016). Marketing Management (15th ed.). Pearson Education.
4. Rakesh Agrawal & Ramakrishnan Srikant (1994). Fast Algorithms for Mining Association Rules in Large Databases. Proceedings of the 20th International Conference on Very Large Data Bases (VLDB).
5. George E. P. Box, Gwilym M. Jenkins, & Gregory C. Reinsel (2015). Time Series Analysis: Forecasting and Control (5th ed.). Wiley.
6. Trevor Hastie, Robert Tibshirani, & Jerome Friedman (2009). The Elements of Statistical Learning. Springer.
7. IBM (2019). Retail Analytics: Driving Insights from Transaction Data. IBM Industry White Paper.
8. McKinsey & Company (2020). The Future of Retail Analytics: Leveraging Data for Competitive Advantage. Industry Report.
9. A. Chaube, "ACO-Enhanced Siamese Networks for Robust Feature Matching in Copy-Move Image Forgery Detection," *2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA)*, Nagpur, India, 2024, pp. 1-6, doi: 10.1109/ICAIQSA64000.2024.10882433.