Analysing Customer Behaviour to Optimize Stock Levels

A Proposal report for the BDM capstone Project

Submitted by

Name: VIJAYAGEETHA.V

Roll number: 23F1000049

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Declaration Statement

I am working on a Project Title “Analysing Customer Behaviour to Optimize Stock Levels”. I extend my appreciation to Reliance Digitals, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analysed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively, and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate: vijayageetha.v

Name: VIJAYAGEETHA.V

Date:06-08-2024

1 Executive Summary and Title:

The project, titled "Analysing Customer Behaviour to Optimize Stock Levels," aims to delve into the purchasing behaviours of customers at Reliance Digital, a major electronics retailer. Utilizing data collected over a span of two years from 2022 to 2024, the project seeks to identify patterns and trends that can inform better inventory management decisions. The project presents a detailed explanation of the chosen analysis methods, with a focus on justifying their suitability over alternative methods. This decision is based on the specific objectives and characteristics of the dataset. The analysis has yielded valuable insights and findings, that are presented through graphs and other visual representations, facilitating a clear understanding of identified trends and patterns.

The analysis involves a comprehensive approach that includes descriptive statistics to summarize central tendencies, exploratory data analysis to uncover initial insights, and data visualization to clearly present the findings. By understanding which products are most frequently purchased, identifying seasonal trends, and analysing customer preferences, the project provides actionable insights that will help optimize stock levels.

2 Proof of Originality of Data:

Data Collection:

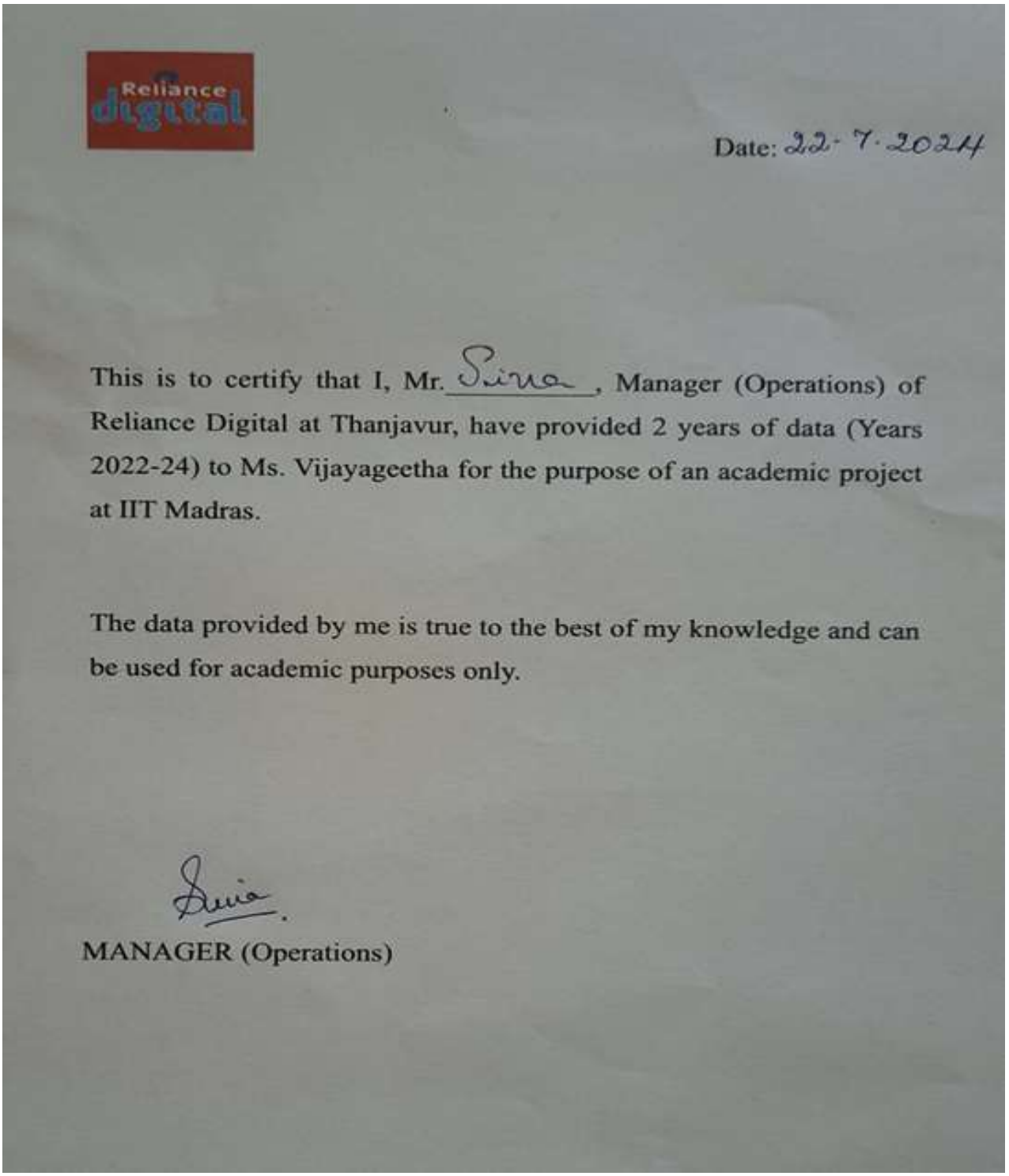
● I have included link for the raw data that I have collected: DATASET

	A	B	C	E	F	G	H
1	Sl.NO	Appliance Name	Price	EMI Mode	Tax	Warranty Years	Customer Rating
2	1	Fan	5887	Paid	588.71	5	3
3	2	Fan	8231	Paid	823.06	2	2.3
4	3	Laptop	159019	Paid	15901.88	5	4.3
5	4	Water Purifier	11081	Paid	1108.08	1	3.9
6	5	Headphones	6901	Paid	690.14	2	2.1
7	6	Mobile Phone	77561	Paid	7756.08	3	2.4
8	7	Digital Camera	249419	Paid	24941.92	2	4.8
9	8	Security Camera	36232	Paid	3623.19	2	3
10	9	Projector	38379	Paid	3837.94	4	4.5
11	10	Smartwatch	39043	EMI	3904.3	3	3
12	11	Laptop	142122	Paid	14212.15	1	4.1
13	12	Speaker	8510	EMI	850.95	4	4.7
14	13	Security Camera	33191	Paid	3319.08	1	3.9
15	14	Headphones	20720	Paid	2071.97	4	3.6
16	15	Smart Home Device	16562	Paid	1656.18	4	1.8
17	16	Air Conditioner	26610	Paid	2661.01	1	2.5
18	17	Smart Home Device	10485	Paid	1048.45	5	3.6
19	18	Digital Camera	216396	Paid	21639.58	4	3.1
20	19	Water Purifier	14827	Paid	1482.67	2	1.9
21	20	Headphones	25713	Paid	2571.33	5	1.8
22	21	Mixer Grinder	5893	Paid	589.27	4	3.6

Sheet1

Fig 1.0

● **Letter from the organization:**



Reliance digital

Date: 22-7-2024

This is to certify that I, Mr. Siva, Manager (Operations) of Reliance Digital at Thanjavur, have provided 2 years of data (Years 2022-24) to Ms. Vijayageetha for the purpose of an academic project at IIT Madras.

The data provided by me is true to the best of my knowledge and can be used for academic purposes only.

Siva.

MANAGER (Operations)

Fig 1.1

● Here are some pictures of the Store:

Fig: 1.2

Fig: 1.3

3 Metadata:

- **Data Source:** The dataset was provided by Reliance Digital, a leading electronics retailer in Thanjavur, Tamil Nadu, with over 15+ electronic products ranging from televisions to smartphones.

- **Data Collection Period:** Data was collected over a period of 2 years, from 2022 to 2024.

- **Data Collection Method:** The data was systematically recorded through Reliance Digital's point-of-sale (POS) systems and stored on their server.

- **Data Format:** The data was collected in CSV format, ensuring compatibility with various data analysis tools and ease of use.

- **Data Quality:** The initial dataset collected from Reliance Digital's point-of-sale systems had several inconsistencies, including missing values. A thorough data cleaning process was implemented, involving the removal of duplicate records and imputation of missing values using median values for numerical data and mode for categorical data.

- **Data Fields:** The data fields are
    - Appliance Name
    - Price
    - Brand Name
    - Amount Paid
    - EMI Mode
    - Tax
    - Warranty Years
    - Customer Ratings

- **Purpose of Data:** The primary goal of this dataset is to analyse customer purchasing behaviour with the aim of optimizing stock levels, ensuring product availability while minimizing overstock.

COLUMNS	Data Type	Description
Appliances Name	Text	Electronic appliances
Price	Numerical (Accounting)	Retail price of the appliance
Brand Name	Categorical	Manufacturer of the appliance
Amount Paid	Numerical (Accounting)	Total amount paid for the appliance
EMI Mode	Categorical	Mode of EMI payment
Warranty Years	Numerical	Number of years the appliance is under warranty
Tax	Numerical (Accounting)	Tax amount applied to the purchase

Table No: 1.0

• Descriptive Statistics:

- Descriptive statistics were applied to the sales data to determine the central tendency of various key variables. The mean, median, and mode were calculated for key variables such as "Price" and "Amount Paid" to determine typical spending on electronic appliances.

- Measures of dispersion, including the range, variance, and standard deviation, were computed to assess the variability in prices and the associated financial risks. This information is essential for subsequent analyses and decision-making processes, particularly in inventory management.

Descriptive Statistics for Purchase data:

Column1	Price	Amount Paid	Tax	Warranty years	Customer Rating
Mean	48327	24374	4832.4	3	3
Standard Error	854.069	553.044	85.0432	0.023	0.019
Mode	9856	3007	1173.1	5	2.6
Median	31208	11618	3120.7	3	3
First Quartile	12308.2	4129.4	12308	2	2
Third Quartile	65555.1	30626.07	6555.15	4	4
Variance	729435.6	306299.2	7292.63	2.02	1.35
Std Deviation	854.06	553.44	80.96	1.42	1.16
Skewness	2.02	2.18	2.02	-0.016	0.014
Range	298968	276697	29896.8	4	4
Minimum	1019	890	101.8	1	1
Maximum	299987	277587	29998.6	5	5
Sum	177941152	89744716	17794116	11066	11031
Count	3682	3682	3682	3682	3682

Table No :1.1

4 Detailed Analysis:

a) Collection:

For analysing customer behaviour, data related to the sales of electronic appliances from 2022 to 2024, Excel sheets, pivot tables, and charts were utilized for the analysis.

Following actions were taken after Data collection:

- **Data cleaning was prioritized,** focusing on removing any rows or columns with missing or invalid entries to ensure the integrity of the analysis.
- **Data visualization followed,** utilizing various charts such as bar graphs and pie charts to effectively depict sales trends, customer preferences, and the performance of different electronic appliances over the two-year period.
- **Data analysis was performed** to uncover patterns and trends, such as identifying the most popular product categories, detecting seasonal sales fluctuations, and examining the correlations between customer demographics and purchasing decisions.

b) Analysis:

For analysing the customer behaviour in the electronic dataset, I followed a structured process to gain insights and understand patterns. Here's a detailed overview of the analysis process:

1. **Descriptive Statistics**: I began by calculating basic descriptive statistics to get an initial understanding of the dataset. This included measures such as the mean, median, mode, and standard deviation for key variables. This step helped in summarizing the central tendencies and variability within the data.

2. **Exploratory Data Analysis (EDA)**: I conducted an Exploratory Data Analysis to uncover initial insights and patterns. This involved generating summary statistics, visualizing distributions, and identifying potential anomalies. EDA helped in forming hypotheses about customer behaviour and guided the direction for more detailed analysis.

3. **Data Cleaning**: Before diving deeper into analysis, I cleaned the dataset by handling missing values, removing duplicates, and addressing any outliers. This ensured the accuracy and reliability of the analysis results.

4. **Correlation Analysis**: I assessed the relationships between different variables using correlation analysis. This analysis helped identify which factors are strongly related.

5. **Group By Analysis**: I performed group by analysis to summarize data across different segments. This involved aggregating data to calculate summary statistics like average purchase value and frequency for each group. Pivot tables were used to present these summaries effectively.

6. **Basic Hypothesis Testing**: To test specific hypotheses about customer behaviour, I conducted basic hypothesis tests. For example, I used t-tests to compare purchase behaviour between different age groups or regions and chi-square tests to examine the relationship between categorical variables such as customer preferences and product types.

7. **Frequency Analysis**: I performed frequency analysis to count occurrences of various customer behaviours and attributes. This helped in identifying the most common patterns and preferences.

5 Results and Findings:

Here, I have presented pictorial representations and graphs to clearly illustrate the trends and patterns observed in the analysis on a quarterly period basis.

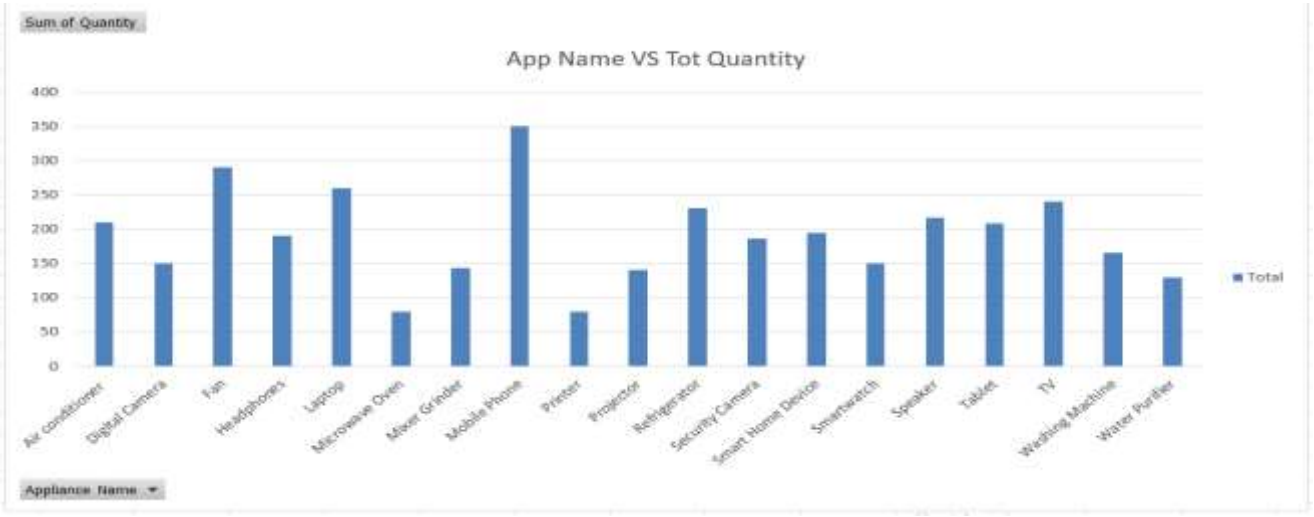


Fig 1.4

- The top revenue-generating product is the **Mobile Phone** with a total of **350 units** sold. The **Fan** follows closely with **290 units,** and **Laptops** rank third with **260 units**

Fig: 1.5

- The top brand by customer rating is **Samsung** with a total score of **737**.5. Apple ranks second with **537,** followed by **Canon** with **277**.2. On the lower end, **Dell** has the lowest rating of **75** among the brands listed.

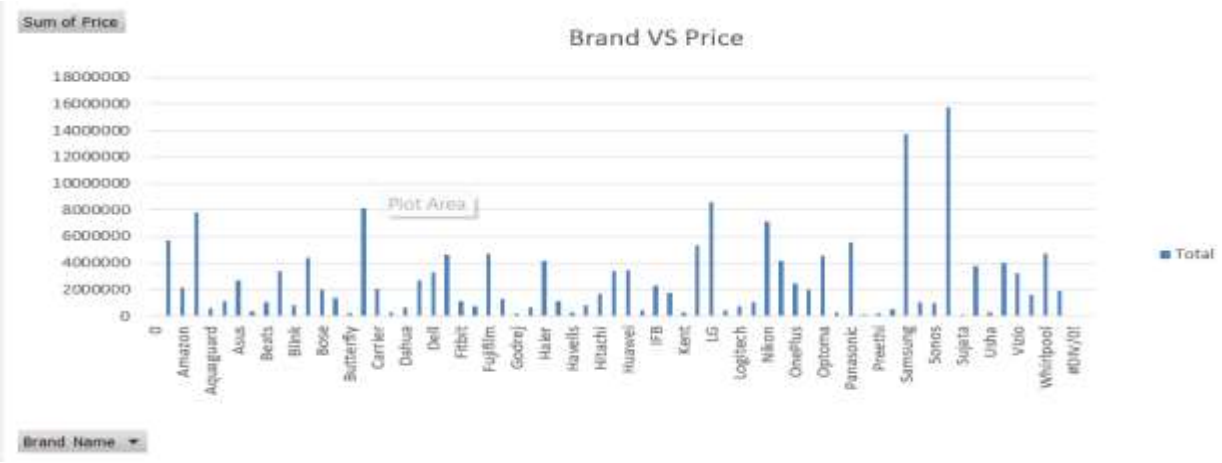


Fig: 1.6

From the above chart, Samsung and Usha have the highest total prices, indicating they offer more expensive products or a larger variety of high-priced items

Fig:1.7

The chart highlights that **Digital Cameras** have the highest total price among the listed appliances, indicating their significant contribution to overall sales or inventory value. **Laptops** and **Projectors** also show substantial total prices, suggesting that they are key products in the inventory