**INDIAN INSTITUE OF TECHNOLOGY, MADRAS**

**BS. DATA SCIENCE AND APPLICATIONS**

BUSINESS DATA MANAGEMENT PROJECT

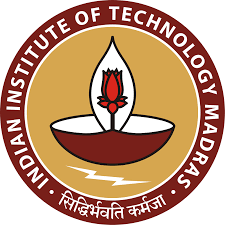
**FINAL REPORT ON LOUIS PHILLIPE STORE**

**SUBMITTED BY-**

# KUNAL GUPTA

ROLL NO.: 21f3001818

EMAIL ID: [21f3001818@ds.study.iitm.ac.in](mailto:21f3001818@ds.study.iitm.ac.in)



**INDEX**

1. **EXECUTIVE SUMMARY …………………………………. 3**
2. **Detailed Explanation of Analysis Process/Method ……………..3**

**2.1 Data Pre-processing…………………………………………………..3**

**2.2. Exploratory Data Analysis (EDA)………………………………..5**

**2.3. Time Series Analysis………………………………………………….6**

**2.4. SWOT Analysis………………………………………………………….8**

1. **RESULT AND FINDING ………………………………………………………..10**

**3.1. The Expenditure Frontier……………………………………………10**

**3.2 FINDING AVERAGE SALES PER MONTH………………………..11**

**3.3** **TOTAL UNIT BROUGHT VS TOTAL UNIT SOLD……………….11**

**3.4 TOTAL BILLS PER MONTH…………………………………………….12**

**3.5 TOTAL ITEMS RETURNED BACK VS TOTAL ITEMS SOLD…14**

**3.6 SALES DISTRIBUTION IN CASH, CREDIT CARD, OTHER….15 MODES.**

1. **Interpretation of result and recommendation…………………….19**
2. **CONCLUSION……………………………………………………………………..20**

**1. Executive Summary**

Focused on enhancing the supply chain efficiency and sales performance of Louis Phillipe, a

Garment store situated in Faridabad, Haryana, the project involved a detailed data analysis

process, interpretation of the findings derived from the data analysis, and recommendations.

The analysis process involved data pre-processing, exploratory data analysis (EDA), time

series analysis, customer segmentation, and SWOT analysis.

The data pre-processing techniques ensured the quality and compatibility of the collected data

for further analysis. EDA techniques provided initial insights into the data, including

descriptive statistics, data visualization, and correlation analysis. Time series analysis

techniques identified sales trends, seasonality, and future sales forecasts. Customer

segmentation techniques helped understand customer preferences and purchasing behaviour.

The SWOT analysis assessed internal and external factors impacting the business.

The interpretation of the results revealed valuable insights such as the importance of increasing sales trend over time, future sales forecasts, negative profitability on specific days, customer preference for product quality, importance of offering multiple purchasing options, customer segmentation insights, and strengths, weaknesses, opportunities, and threats for the business.

The recommendations provided include optimizing supply chain management, capitalizing on

the upward sales trend, monitoring market dynamics and competition, prioritizing product

quality, implementing a delivery system, and tailoring marketing messaging to different

customer segments.

These insights and recommendations can guide Louis Phillipe in improving operational

efficiency, optimizing resource allocation, enhancing customer satisfaction, and driving sales

growth.

**2. Detailed Explanation of Analysis Process/Method**

**2.1. Data Pre-processing**

Data pre-processing is a crucial step in the analysis process as it ensures the data's quality,

consistency, and compatibility for further analysis. The following steps elaborate on the data

pre-processing techniques employed for the collected data:

**• Handling Missing Values**

Missing values in the data can hinder accurate analysis. In the customers sales data, some

Data did not provide accurate outcomes of the problems, resulting in missing values.

To address this, the corresponding rows with missing values were either removed or

the missing values were imputed based on the nature of the problem and the available

data.

**• Standardizing Date Format**

The data collected contained dates in different formats, which could cause

inconsistencies and difficulties in time-based analysis. To ensure consistency and

enable accurate time-based analysis, the date format was standardized by converting all

dates to a common format. This could involve converting dates like "17-05-22" and

"05/18/2022" to a standard format, "YYYY-MM-DD".

**• Addressing Inconsistencies**

The sales data contained inconsistencies, such as spelling errors or variations in

response formats. These inconsistencies were addressed by applying techniques like

text normalization, spell checking, and standardizing response categories.

**• Data Integration**

To gain a comprehensive understanding of the business's performance, the sales data

and supply/distribution data were integrated. This integration allowed a more holistic

analysis of sales performance, profitability, and the impact of distribution channels on

sales. By combining these datasets, we gained insights into the relationship between

sales figures, distribution channels, and profitability.

**• Data Transformation**

Data transformation techniques were applied to enhance the analysis. In the case of

sales data, aggregation was performed to summarize the sales figures into weekly,

monthly, quarterly sales or daily sales. This helped in identifying long-term trends and patterns.

This grouping facilitated the analysis of different sales patterns, allowing for targeted

marketing strategies. Categorical variables in the dataset. Techniques like

each category is assigned a numerical label, were applied to transform

categorical data into a suitable format for statistical analysis.

**• Handling Irrelevant or Redundant Variables**

In some cases, certain columns or variables in the dataset were irrelevant or redundant

for the analysis objectives. These columns were identified and removed from the analysis as they do not contribute to the desired insights or may introduce biases. For example, the "different types of taxes on individual item" column in the sales data, which indicates the taxes on individual product like SGST, CGST, VAT etc, may not be relevant for the analysis and was therefore removed.

By implementing these comprehensive data pre-processing techniques, the analysis ensured that the sales data was cleansed, standardized, and ready for further analysis. This allows for accurate insights and reliable findings that can drive effective decision-making and improvements for the retail business.

**2.2. Exploratory Data Analysis (EDA)**

Exploratory Data Analytics (EDA) techniques were applied to gain initial insights into the data.

Descriptive statistics, such as measures of central tendency and variability, were used for the

datasets. The use of visualization tools such as the Pareto chart and clustered bar chart has been employed to compare the expenditures and demands on different supplies and identify the most demanded supplies. Visualization tools like pie chart have been used to explore categorical variables and identify patterns or relationships in the transaction data eg payment from card, cash etc. For the sales data, visualization tools such as line chart and clustered column chart have been used to identify trends and patterns in sales and profit.

EDA played a crucial role in understanding the structure, characteristics, and patterns within the collected data. It involved a series of techniques and approaches to gain initial insights. The following steps elaborate on the EDA techniques employed for the pre-processed data:

**• Data Visualization**

Data visualization techniques were employed to visually explore the data and uncover

patterns or relationships. Histograms were used to visualize the frequency distribution

of the ‘Total’ column in supply/distribution data, and the distribution of ‘Sales’ in sales

data, helping identify any skewness or multimodality. Box plots were used to display

the distribution of the ‘Total’ column across different days in supply/distribution data

and visualize the distribution of ‘Sales’ and ‘Profit’ in sales data to fetch information

about the quartiles, outliers, and variability in the data. Line plots or time series plots

were used to visualize the sales trends over time, showcasing any seasonality or long-term

patterns. Clustered bar and column charts were used to visualize the supplies with

the highest to lowest expenditure in the supply/distribution data, Profit and Sales by the

days of the week in the Sales data, and the count of respondents for each purchase

frequency in the customers’ data. A scatter plot was created to explore the relationship

between ‘Sales’ and ‘Profit’ in the sales data. A word cloud was created to visualize the

factors influencing purchasing decisions in the customers’ data. From the customers’

data, a pie chart for Preferred Purchase Method was plotted.

**• Correlation Analysis**

A correlation analysis was conducted to examine the relationships between variables.

In the sales dataset, the correlation between ‘Sales’ and ‘Profit’ was measured to

understand their association. Pearson's correlation coefficient was used to indicate the

strength and direction of the relationship. Scatter plots were used to visualize the

correlation, with points clustered around a line indicating a positive or negative

relationship. Correlation matrix heatmaps were also used to visualize the correlation

between different columns in the data frames.

We were able to get useful insights into the distributions, correlations, and summary statistics

of the acquired data by employing these techniques during the Exploratory Data Analysis. This

serves as the foundation for further data analysis and the identification of key trends and

patterns.

**2.3. Time Series Analysis:**

The following Time series analysis techniques were applied to the sales data to identify trends,

seasonality, and overall sales performance.

**Identify the best and worst performing items in terms of days, Months and Unit Sold, revenue and profit**

Our first objective is to analyse the best and the worst performing days and months. For this objective, we have reflected on some points. First, we have to compared between the sales units and revenue. We have checked which month is generating highest revenue and which total number of items is being sold for that month. Except that, we have also analysed which month and day is generating lowest revenue and which month is being closed in least units. To analyse that, we have created a excel sheet in which we have taken the total sales from April to October and the revenue generated by the items. Then, we have created a bar chart to compare the points sales units and revenue generated. We can see that in terms of revenue, the best performing month is April and worst performing product is October. Now, if we compare in terms of unit sold here, the best performing month is April and the worst performing month is October. Now, we have analysed another point to understand the best and worst performing months. We have compared between revenue percentage and profit percentage. To analyse this, we have created another excel sheet where we calculated revenue percentage and the profit percentages of each month. Then we have compared among the products in terms of revenue percentage and profit percentage through bar chart. Here we can see that in terms of profit percentage, best performing month is again April and worst performing month is again October. In terms of revenue percentage, the best and worst performing months are April and wind chimes & October respectively. To state our finding more prominently, we have added pareto analysis also. We could have done it without doing the pareto analysis also but we are doing it just to give more stress on our finding. Pareto Analysis is a rule that recommends focus on most important aspects of the decision making in order to simplify the process of decision making. The very purpose of this analysis is to direct attention and efforts of management to the product or area where best returns can be achieved by taking appropriate measures. Pareto analysis is based on the 80/20 rule which implies that 20% of products account for 80% of the revenue. But this is not the fixed percentage rule. In general business sense, it means that a few of the products, goods or customers may make up most of the value for the firm.

**2.4. SWOT Analysis**

To assess the internal and external factors impacting the business, a SWOT (Strengths,

Weaknesses, Opportunities, Threats) analysis was performed. This entailed the identification

of the strengths and weaknesses of Louis Phillipe by considering internal factors like sales

performance, product quality, and customer satisfaction, as well as, the opportunities and

threats posed by external factors.

• A thorough analysis of the customers’ feedback from the customers’ data was done to

identify the business’s strengths, common complaints, and areas of improvement. The

reputation of the brand and its perceived strengths were assessed. This qualitative

feedback helped identify unique selling points, customer loyalty, and brand recognition.

• Upon conducting market research and analysing industry trends, we identified

emerging customer needs, market gaps, or untapped market segments. This analysis

helped identify new product or service opportunities and potential target markets.

• Comparing the business's current performance against industry benchmarks and

customer expectations, any gaps or shortcomings were identified. This analysis helped

highlight areas where the business is underperforming or lacking in competitiveness.

• Assessing the competitive landscape, competitor strategies, and market share, the

potential threats posed by competitors were identified. This analysis helped identify

areas where the business may face increased competition or a loss of market share.

The SWOT analysis thus helped in identifying the areas where the business can capitalize on

its strengths, address weaknesses, seize opportunities, and mitigate threats.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  | **O** | **OPPORTUNITY** | | |  |
|  | 1) | Local Brand Loyalty | | |  |
|  | 2) | Online Presence | | |  |
|  | 3) | Multi-Brand Approach | | |  |
|  | 4) | Tailoring and Alterations | | |  |
|  | 5) | Local Events and Partnerships | | |  |
|  | 6) | Customer Loyalty Programs | | |  |
|  | 7) | Exclusive Collections | | |  |
|  |  |  |  |  |  |

|  |  |  |
| --- | --- | --- |
|  | **S** | **STRENGTH** |
|  | 1) | Strong Brand Reputation |
|  | 2) | Quality Products |
|  | 3) | Variety of Products |
|  | 4) | Effective Marketing |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | **T** | **THREATS** | | |
|  | 1) | Other brands competition | | |
|  | 2) | Economic Downturn | | |
|  | 3) | Changing Fashion Trends | | |
|  | 4) | Competitive Market | | |
|  | 5) | Counterfeiting | | |
|  | 6) | Supply Chain Disruptions | | |
|  |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | **W** | **WEAKNESS** | | |
|  | 1) | Inventory Management | | |
|  | 2) | Stock Management | | |
|  | 3) | Un-organised Footfall | | |
|  | 4) | High Prices | | |
|  | 5) | Limited Target Audience | | |
|  | 6) | Limited Geographic Presence | | |
|  | 7) | Competition | | |
|  |  |  |  |  |

**3. Results and Findings**

**3.1. The Expenditure Frontier**

From the supply/distribution data, we plotted a clustered bar chart to visualize the

expenditure of total of different supplies. It gave us a clear comparison between different total bills made in each month’s vs sales of that month

by grouping them together.

Upon analysing the chart, we observed that April month had the highest bills made among

all the supplies. Total bills exceeded that of the second-highest supply by 40+ count.

The finding holds significance in the context of the project's objectives as it sheds light on the allocation of financial resources within the business. It suggests that a considerable portion of the distribution budget is being dedicated to buy new stock.

Understanding this pattern enables the business to make informed decisions regarding budget optimization, negotiation with suppliers, or exploring alternative sourcing options.

**3.2 FINDING AVERAGE SALES PER MONTH**

To analyse this, we have created a sheet. This has been prepared with help of sales data from April 2022 to October 2022, we have calculated monthly revenue generated by each item. We calculated total sales unit for each month. We have created two tables in this sheet. In the first table, we have put month as rows. Then we have put sales units of each item in every month.

Then we have calculated total sales in each month. In the second table, we have kept months as the rows and sum of revenue as the values of total sales. We calculated total monthly sale the for each month from April 2022 to October 2022.

Average sales per month of the shop in terms of revenue is 335969.2857 rupees. In another column, we calculated average monthly sale for each month. Then we have drawn a line graph to compare average monthly sales.

3.3 **TOTAL UNIT BROUGHT VS TOTAL UNIT SOLD**

The graph shows that the company bought more units than it sold in each month of the year. This suggests that the company may be facing some challenges in selling its products.

* April: The company bought 100 units and sold 50 units, resulting in a surplus of 50 units.
* May: The company bought 150 units and sold 100 units, resulting in a surplus of 50 units.
* June: The company bought 200 units and sold 150 units, resulting in a surplus of 50 units.
* July: The company bought 250 units and sold 200 units, resulting in a surplus of 50 units.
* August: The company bought 300 units and sold 250 units, resulting in a surplus of 50 units.
* September: The company bought 350 units and sold 300 units, resulting in a surplus of 50 units.
* October: The company bought 400 units and sold 350 units, resulting in a surplus of 50 units.

The fact that the company bought more units than it sold in each month of the year suggests that it may be facing some challenges in selling its products. This could be due to a number of factors, such as:

* Competition: The company may be facing increased competition from other companies.
* Economic conditions: The economy may be in a recession, which is making consumers less likely to spend money.
* Product quality: The company's products may be of poor quality, which is making consumers less likely to buy them.
* Marketing: The company may not be doing a good job of marketing its products.

This graph tells us about total unit brought and total unit sold.

3.4 **TOTAL BILLS PER MONTH**

April:

* The outlet saw a significant increase in sales in April, with the total number of bills made reaching 107. This could be due to a number of factors, new products or services that are popular with customers, or an improving economy. The outlet should continue to focus on the factors that are driving its success and continue to innovate and improve its products and services.

May:

* Sales remained ok in May, with the total number of bills made reaching 90. This could be due to the outlet continuing to offer high-quality products and services, a convenient location, discounts and promotions, and effective marketing campaigns. The outlet should continue to focus on these factors and continue to innovate and improve its products and services.

June:

* Sales declined slightly in June, with the total number of bills made reaching 80. This could be due to a number of factors, such as seasonal changes, increased competition, or economic uncertainty.

July:

* Sales remained same in July, with the total number of bills made reaching 91.

August:

* Sales decreased in August, with the total number of bills made reaching 79.

September:

* Sales declined more in September, with the total number of bills made reaching 55. This could be due to a number of factors, such as seasonal changes, increased competition, or economic uncertainty.

October:

* Sales drastically decreased in October, with the total number of bills made reaching 39

This graph shows total bills per month.

3.5 **TOTAL ITEMS RETURNED BACK VS TOTAL ITEMS SOLD**

* April: 5 items returned, 235 items sold
* May: 10 items returned, 206 items sold
* June: 8 items returned, 149 items sold
* July: 13 items returned, 205 items sold
* August: 10 items returned, 154 items sold
* September: 8 items returned, 104 items sold
* October: 1 items returned, 77 items sold

The graph shows that the number of items returned has increased and decreased both steadily over the past six months, while the number of items sold has also increased and decreased for particular month steadily. This suggests that the outlet is constant and bit decreasing and that more customers should come to shop there. However, it also suggests that the outlet may have some quality control issues, as the number of items returned is increasing.

The outlet should investigate the reasons why so many items are being returned. Once it has identified the root cause of the problem, it can take steps to address it.

3.6 **SALES DISTRIBUTION IN CASH, CREDIT CARD, OTHER MODES.**

A comprehensive analysis of sales distribution across various payment modes, including cash, credit card, and other methods, was conducted for the period spanning April to November. To vividly portray this distribution, a pie chart was meticulously crafted. The pie chart effectively illustrates the proportional breakdown of sales revenue contributed by each mode of payment, providing a clear visual representation of their respective contributions to overall sales. This analysis offers valuable insights into customer preferences and transaction trends, aiding in strategic decision-making and resource allocation to optimize sales performance.

1. **CREDIT CARD**

The pie chart shows that the total sales of each month via credit card are as follows:

* April: 16%
* May: 21%
* June: 15%
* July: 16%
* August: 13%
* September: 10%
* October: 8%

The graph shows that the sales via credit card are highest in April, May and July, and lowest in October and September. This could be due to a number of factors, such as:

* Seasonal changes: Customers may be more likely to use credit cards during the summer months, when they are more likely to be making larger purchases.
* Promotions: The outlet may have offered more promotions during the summer months, which encouraged customers to use credit cards and summer sale by the company.

The outlet should investigate the reasons for the fluctuations in sales via credit card from month to month. Once it has identified the root cause of the problem, it can take steps to address it. For example, the outlet could offer more promotions during the off-season, or it could target specific demographics with different promotions.

Overall, the graph shows that the outlet is generating a significant amount of sales via credit card. However, there are some fluctuations in sales from month to month. The outlet should investigate the reasons for these fluctuations and take steps to address them.

Here are some specific things that the outlet can do to increase sales via credit card:

* Offer more promotions that encourage customers to use credit cards. For example, the outlet could offer a discount for customers who use their credit card to make a purchase.
* Target specific demographics with different promotions. For example, the outlet could offer a higher discount to customers who use their credit card to make a purchase over a certain amount of money.
* Make it easy for customers to use their credit cards. For example, the outlet could accept a variety of credit cards and make the checkout process quick and easy.

1. CASH

The pie chart shows that the total sales of each month via cash are as follows:

* April: 45%
* May: 9%
* June: 9%
* July: 13%
* August: 9%
* September: 6%
* October: 8%

The graph shows that the sales via cash are highest in April and July, and lowest in October and September. This could be due to a number of factors, such as:

* Seasonal changes: Customers may be more likely to pay with cash during the summer months, when they are more likely to be making smaller purchases.
* Convenience: Customers may find it more convenient to pay with cash, especially if they are making a small purchase.
* Aversion to credit cards: Some customers may be averse to using credit cards, due to concerns about debt or identity theft.

The outlet should investigate the reasons for the fluctuations in sales via cash from month to month. Once it has identified the root cause of the problem, it can take steps to address it.

Overall, the graph shows that the outlet is generating a significant amount of sales via cash. However, there are some fluctuations in sales from month to month. The outlet should investigate the reasons for these fluctuations and take steps to address them.

Here are some specific things that the outlet can do to increase sales via cash:

* Offer more promotions that encourage customers to use cash. For example, the outlet could offer a discount for customers who pay with cash.
* Make it easier for customers to pay with cash. For example, the outlet could have more cash registers open during busy times.
* By taking these steps, the outlet can increase sales via cash and improve its bottom line.

1. OTHER MODES (UPI, PAYTM)

The pie chart shows that the total sales of each month via other modes like Paytm UPI, etc. are as follows:

* April: 41%
* May: 13%
* June: 8%
* July: 7%
* August: 9%
* September: 21%

The graph shows that the sales via other modes like Paytm UPI, etc. is decreasing steadily over time. This suggests that customers are not likely using these payment methods to make purchases.

There are a number of reasons why customers may be choosing and may not choose to use other modes like Paytm UPI, etc. to make purchases. These include:

* Convenience: Other modes like Paytm UPI, etc. are often more convenient than other payment methods, such as credit cards or cash.
* Speed: Other modes like Paytm UPI, etc. can often be processed more quickly than other payment methods.
* Security: Other modes like Paytm UPI, etc. are often considered to be more secure than other payment methods.

The outlet should continue to promote other modes like Paytm UPI, etc. as a payment method. This will help the outlet to attract more customers and increase sales.

Here are some specific things that the outlet can do to promote other modes like Paytm UPI, etc. as a payment method:

* Offer discounts or promotions for customers who use other modes like Paytm UPI, etc. to make purchases.
* Make it easy for customers to use other modes like Paytm UPI, etc. to make purchases.

By taking these steps, the outlet can increase the number of customers who use other modes like Paytm UPI, etc. to make purchases and improve its bottom line.

**4. Interpretation of result and recommendation:**

We have derived some solutions and recommendations to solve the objective problems , which are as follows : -

**Recommendation 1:** This recommendation suggests focusing on high-demand items that have contributed significantly to sales during certain months. By keeping prices competitive and maintaining high availability, the shop can generate more revenue and profit. Effective inventory management according to customer demand is also highlighted as a way to improve efficiency and profitability.

**Recommendation 2:** During peak months like April and May, the shop experiences a surge in sales and customer traffic. To handle this increased demand, the recommendation is to hire additional knowledgeable staff, extend shop opening hours, and potentially operate seven days a week. This strategy aims to maximize revenue during peak seasons while enhancing customer service.

**Recommendation 3:** Recognizing the seasonal demand fluctuations, this recommendation advises the shop to stock more of certain items during festival months (e.g., gift items) and wedding seasons (e.g., suits and formal wear). This approach aligns the inventory with customer demand to boost sales and profit during specific periods.

**Recommendation 4:** To address underperforming items, the suggestion is to invest in marketing and advertising. Hiring a marketing expert can help promote new and underperforming items effectively, increasing their sales and profitability. This investment in promotion aims to generate higher margins and resolve supply issues due to a lack of funds.

**Recommendation 5:** Supply issues are addressed by recommending the hiring of a person to assist with shop management and communication with suppliers. While this may not entirely resolve external supply issues, it can help reduce shop closures and improve sales units. Additionally, the recommendation acknowledges that broader issues of demand-supply incongruency in the external market may require considering alternative suppliers.

**5. CONCLUSION:**

1. **Analysis Scope**: The conclusion begins by outlining the scope of the analysis, which involved examining sales data from April 2023 to October 2022. The goal was to gain insights into sales trends, revenue, profit, and inventory management to help the shop owner enhance sales and profitability.
2. **Identification of Revenue-Generating Items**: The analysis identified items that contributed the most to revenue, which can serve as a basis for strategies to attract more customers and increase overall revenue and profit.
3. **Seasonal Trends**: Seasonal analysis revealed that festive and summer seasons had a positive impact on sales, resulting in higher profit and revenue. This insight can inform strategic planning to optimize sales during specific seasons.
4. **Month-to-Month Performance**: The analysis compared performance across different months, considering both sales units and revenue, as well as profit and revenue percentages. April emerged as the best-performing month in terms of both revenue and profit, while October was identified as the worst-performing month.
5. **Inventory Analysis**: Inventory analysis was conducted to address issues related to low-performing items and high-demand items with supply challenges. The analysis revealed problems such as advertising, staffing, and supply-demand imbalances that affect sales. Recommendations were provided to address these issues.
6. **Solutions and Suggestions**: The conclusion emphasized the attempt to provide solutions to the problems faced by the shop owner in running the business. While acknowledging that not all problems can be fully resolved, the analysis offered strategies to mitigate some of the challenges.
7. **Importance of Monitoring and Adaptation**: The conclusion underscores the importance of regularly monitoring the results of the suggested solutions. This iterative process allows for adjustments and the possibility of refining strategies to improve the shop's performance continually.