A PROJECT REPORT

on

"Annual Sales Analysis and Recommendations"

Submitted to

KIIT Deemed to be University

In Partial Fulfilment of the Requirement for the Award of

BACHELOR'S DEGREE IN
INFORMATION TECHNOLOGY

BY

AASTHA SINHA 2105087

UNDER THE GUIDANCE OF
SOVAN KUMAR SAHOO

SCHOOL OF COMPUTER ENGINEERING
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
BHUBANESWAR, ODISHA - 751024
May 2020

Acknowledgements

I am profoundly grateful to **SOVAN KUMAR SAHOO** of **Affiliation** for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion.

I would like to express our sincere gratitude to Senior Leadership Team of ACME Inc for entrusting us with this critical analysis and for their valuable guidance and support throughout the project.

I am deeply thankful to our sir, for their expert guidance, insightful feedback, and unwavering encouragement that played a pivotal role in shaping this project and enhancing our analytical skills.

I would also like to extend our appreciation to the Data Science and Analytics team at ACME Inc for their cooperation and for providing the necessary resources and datafiles required for the analysis.

Furthermore, we acknowledge the contributions of our peers and colleagues who provided valuable insights, feedback, and support at various stages of the project.

AASTHA SINHA

ABSTRACT

The increasing volume of transactional data in retail and e-commerce sectors necessitates effective analysis techniques to derive actionable insights. This project focuses on analyzing sales transactions spanning multiple years to uncover trends, patterns, and key performance indicators. Utilizing Python libraries such as Pandas, Matplotlib, and Seaborn, the study merges and cleanses datasets related to customers, products, shipping destinations, and transaction details. The primary objectives include understanding yearly sales trends, sales distribution by product category and market, the impact of discounts on profit margins, and regional sales variations. Additionally, the project aims to identify top-performing and underperforming products based on sales volume and profit. The findings from this analysis can assist businesses in optimizing their product offerings, pricing strategies, and marketing efforts to enhance profitability and customer satisfaction.

Keywords: Sales Analysis, Data Visualization, Pandas, Matplotlib, Seaborn, E-commerce, Retail, Profit Margin, Discount Impact, Product Performance.

Contents

Chapter 1

Introduction

This project aims to analyze sales transactions data to derive meaningful insights and patterns that can assist businesses in enhancing profitability and optimizing strategies. With the advent of big data and analytics, there is a growing need to process and interpret large datasets efficiently. This project focuses on utilizing Python programming and data visualization libraries to conduct a comprehensive analysis of sales data spanning multiple years. The report outlines the methodology, findings, and recommendations based on the analysis.

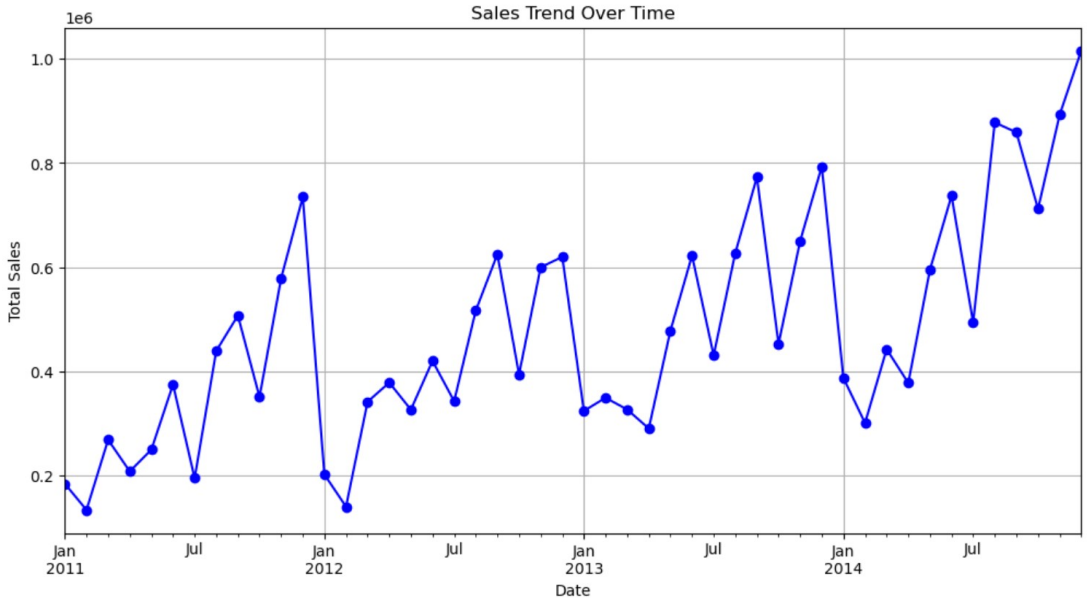


Figure 1.1: Overview of Sales Analysis Process

- Consumer represents 52% of the total Customer Base.
- Inconsistency in the sales as it fluctuates significantly througout the period .
- There are notable trough around january .
- Overall during the years 2011-2014 , sales has increased significantly

Chapter 2

Basic Concepts/ Literature Review

This section provides an overview of the fundamental concepts related to the tools and techniques utilized in the project for analyzing and visualizing sales transaction data.

2.1 Python Programming Language

Python is a high-level programming language known for its simplicity and readability. In this project, Python is used for data manipulation, analysis, and visualization. Libraries such as Pandas, Matplotlib, and Seaborn are extensively used for handling data and generating visualizations.

2.2 Pandas

Pandas is a powerful open-source data analysis and manipulation library for Python. It provides data structures like DataFrame and Series, which are essential for cleaning, transforming, and analyzing the sales transaction datasets in this project.

2.3 Matplotlib

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It is used in this project to generate various types of plots, including line plots, bar charts, and scatter plots, to visualize sales trends, distributions, and relationships between variables.

2.4 Seaborn

Seaborn is a data visualization library based on Matplotlib that provides a high-level interface for drawing attractive and informative statistical graphics. It is utilized in this project to create more advanced and visually appealing plots, including heatmaps, pair plots, and distribution plots, to enhance the analysis and interpretation of the sales transaction data.

2.5 Data Cleaning and Preprocessing

Data cleaning and preprocessing are crucial steps in the data analysis process to ensure the accuracy and reliability of the analysis results. Techniques such as handling missing values, removing duplicates, and transforming data are performed using Pandas to prepare the sales transaction datasets for analysis.

2.6 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is an approach to analyzing data sets to summarize their main characteristics, often using statistical graphics and other data visualization methods. In this project, EDA is conducted using Python libraries like Pandas, Matplotlib, and Seaborn to explore and understand the sales transaction data, identify patterns, and uncover insights that can be useful for further analysis and decision-making.

Chapter 3

Problem Statement

You are an Analytics Engineer working for ACME Inc, and are tasked with analyzing their sales over the past few years, and providing recommendations on improving their revenue next year. You will be presenting your findings to the Senior Leadership Team (note: non-technical audience). Refer to the yearly sales datafiles.

Requirement Specifications

The primary objective of this project is to analyze ACME Inc's sales data from the past few years to identify sales trends, patterns, and insights. The analysis aims to provide actionable recommendations to the Senior Leadership Team for improving revenue in the upcoming year.

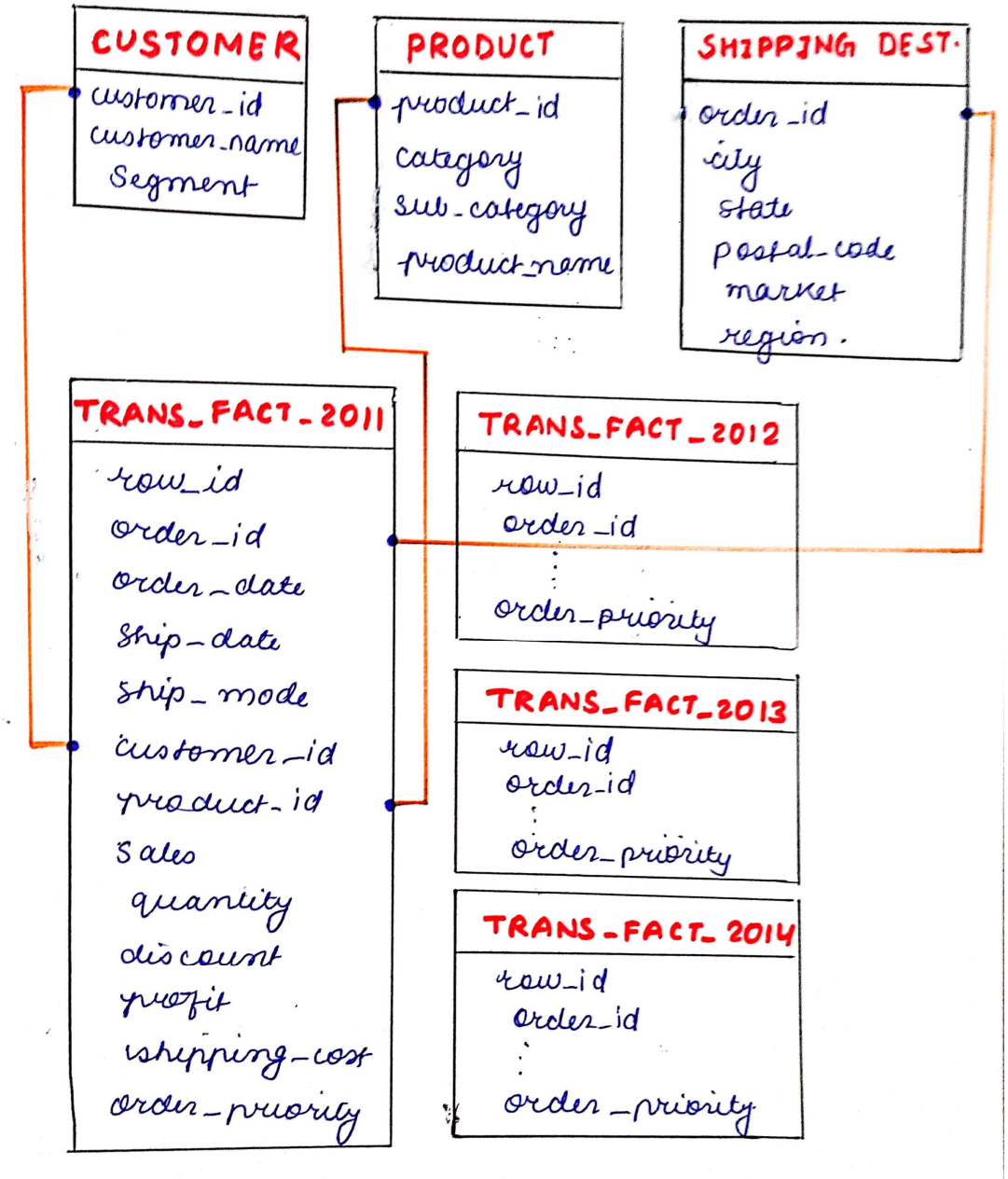
The project will focus on the following key areas:

3.1 Project Planning

The project will be divided into the following phases:

1. **Data Collection:** Gathering the yearly sales datafiles provided by ACME Inc.
2. **Data Cleaning and Preprocessing:** Cleaning and preprocessing the sales data to ensure accuracy and reliability.
3. **Exploratory Data Analysis (EDA):** Conducting EDA to explore and understand the sales trends, patterns, and relationships within the data.
4. **Data Analysis and Insights:** Analyzing the sales data to derive actionable insights and recommendations for improving revenue.
5. **Reporting and Presentation:** Compiling the findings into a comprehensive report and presenting the insights and recommendations to the Senior Leadership Team.

6. **Dependencies of the datasets are as follows:**

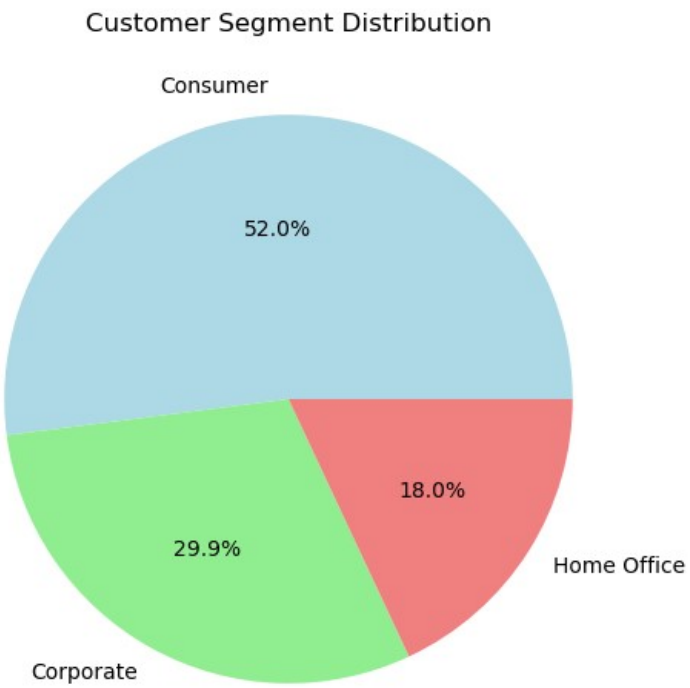


3.2 Project Analysis

The project requires the following software and libraries:

1. Python: Programming language for data analysis and visualization.
2. Pandas: Data manipulation and analysis library.
3. Matplotlib: Plotting library for creating static visualizations.
4. Seaborn: Statistical data visualization library for creating advanced visualizations.
5. Jupyter Notebook: Interactive environment for executing Python code and documenting the analysis process.

The project will utilize the provided yearly sales datafiles in CSV format for analyzing ACME Inc's sales data.



• Consumer represents 52% of the total Customer Base.

Chapter 4

Implementation

The implementation phase of the project involves the execution of the planned steps using the selected tools and techniques. Below are the details of the methodology and procedures followed for analyzing ACME Inc's sales data and deriving actionable insights.

4.1 Methodology OR Proposal

The following methodology is adopted for analyzing ACME Inc's sales data:

1. **Data Loading:** Load the yearly sales datafiles provided by ACME Inc into Pandas DataFrames for analysis.

2. **Data Cleaning and Preprocessing:** Clean and preprocess the sales data to handle missing values, remove duplicates, and ensure data consistency.

3. **Exploratory Data Analysis (EDA):** Explore and visualize the sales data to identify trends, patterns, and insights related to revenue, products, and regions.

4. **Data Analysis and Insights:** Analyze the sales data to derive insights and trends that can help in formulating recommendations for improving revenue.

5. **Reporting and Presentation Preparation:** Compile the analysis findings into a comprehensive report and prepare the presentation slides for the Senior Leadership Team.

4.2 Testing OR Verification Plan

Data validationis done by verifying the correctness of the loaded data by checking the data types, missing values, and duplicates using :

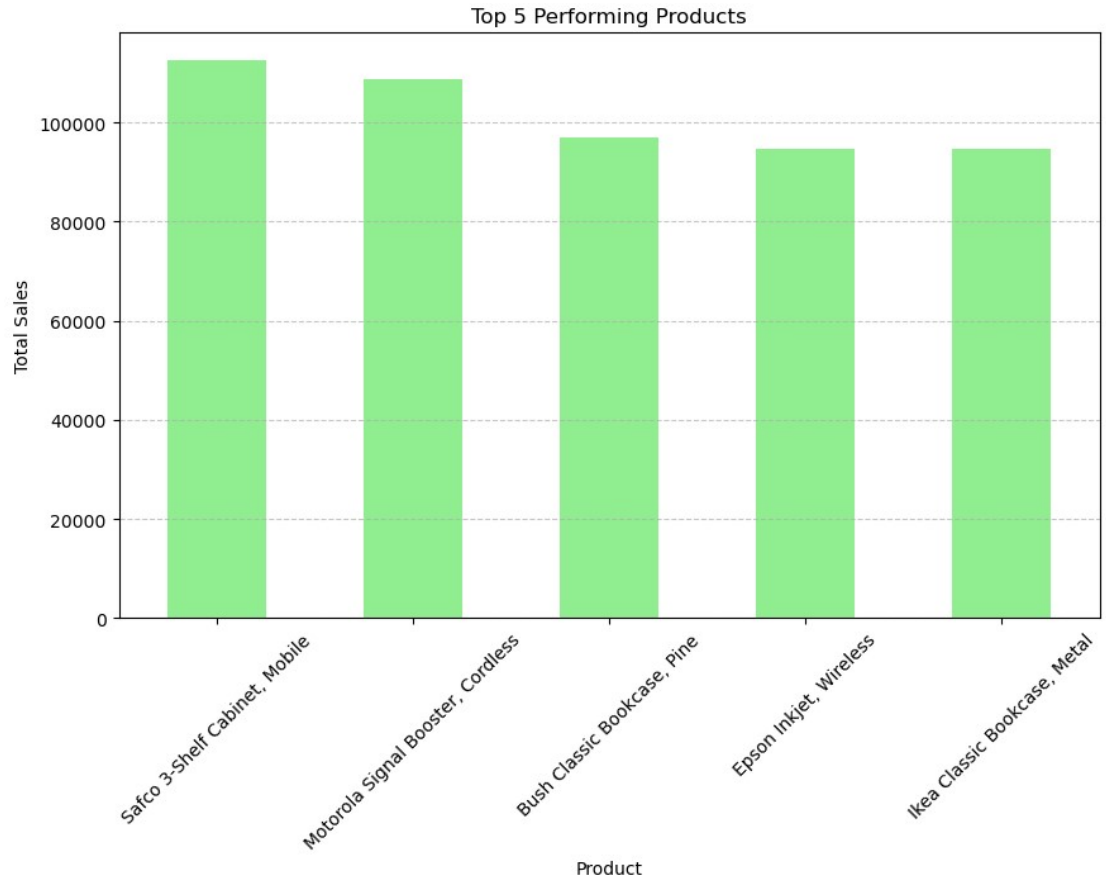
df.isnull().sum()

4.3 Result Analysis OR Screenshots

The analysis results and visualizations generated during the implementation phase are presented below:

- Discount has a significant impact on Profit .
- By looking at category level sales we can see that sales is growing at an increasing rate for all other categories except for furniture in the year 2013-14 in Growth percentage by segment over time graph .
- Among Office Suppliers, Technology and Furniture , The category Technology reaches the highest sales
  but Office Suppliers contribute to maximum profit .
- Maximum number of orders prefer Standard Class Shipping Mode Preferences.
- Sales maximizes with Low shipping cost .
- Critical Orders have more profit margin.

Top 5 Performing Products

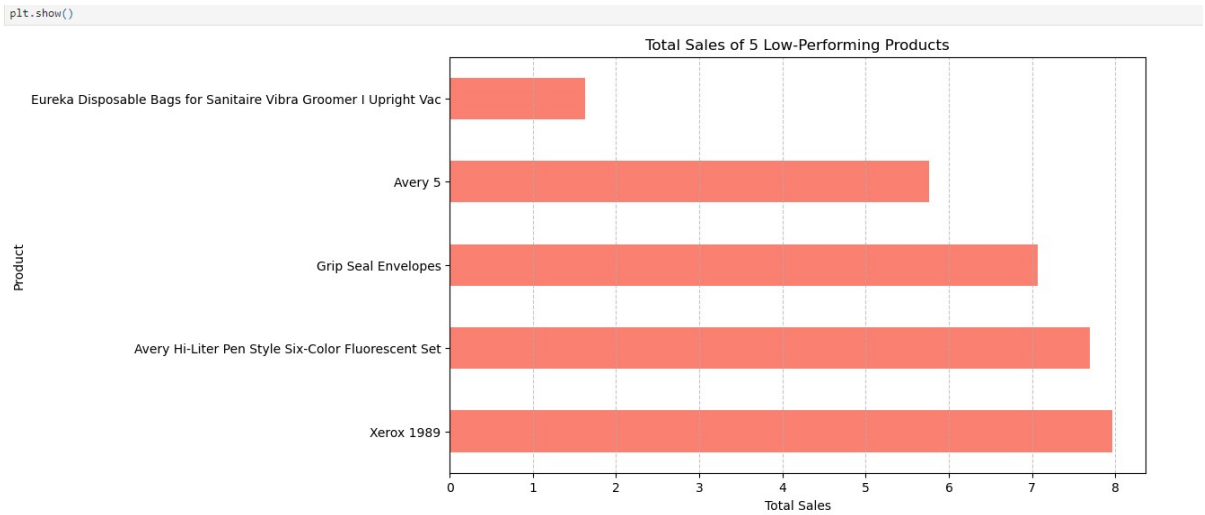


• Sales for 3-shelf Cabinet Mobile , Motorola single booster cordless , Bush Classic Bookcase Pine , Epson Inkjet Wireless and Ikea Classic Bookcase Metal are highest , based on revenue collection that is greater than 90,000 total sales
• It therefore most likely maintains a loyal customer base and meets ongoing market demand.

Recommendations :-

• Product quality and customer satisfaction levels must be maintained to sustain its performance.
• Allocate resources towards targeted marketing campaigns to promote top-performing products.
• Invest in research and development to introduce new features or product variations that align
with consumer preferences.
• Regularly update existing products to maintain relevance and competitiveness in the market.
• Develop customer loyalty programs to incentivize repeat purchases of top-performing products.

Areas For Improvement



• Eureka Disposable Bags,Avery 5,Grip Seal Envelopes, Avery Hi-LiterPen Style Six-Color Fluorescent Set and Xeox are 5 low performing products with less than 8 total sales

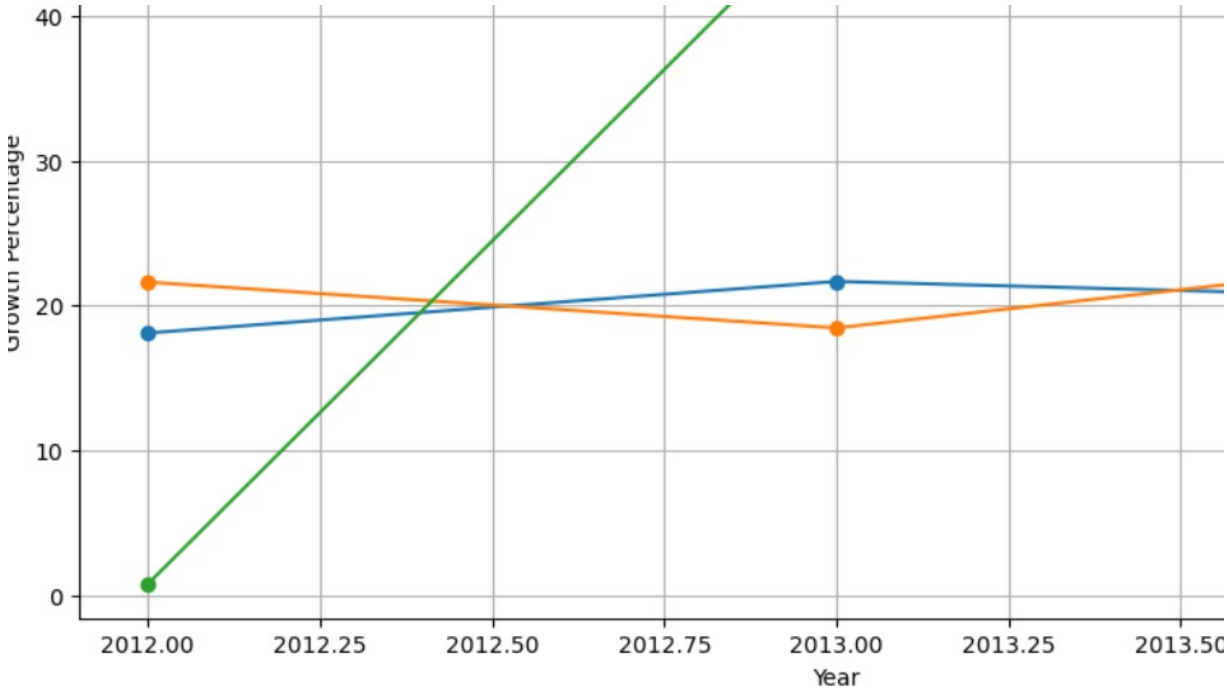
• Affects Brand's reputation as well as increase the probability of financial losses of the company.

Recommendations :-

• Investigate the root causes behind the underperformance of these products. Possible factors could include lack of customer demand, poor product positioning, pricing issues, or inferior product quality.

• Develop targeted marketing campaigns to raise awareness and generate interest in low performing products.

• Review the pricing strategy for low-performing products. It may be necessary to adjust pricing to better align with customer expectations and market competition



Chapter 5

Conclusion

- **This Data Analysis**
  Has provided valuable insights into ACME Inc.'s sales performance and opportunities for revenue improvement. By leveraging the strengths of top-performing products and addressing areas of underperformance, ACME Inc. can drive sustainable growth and maintain a competitive edge in the market.

- **Increased Efficiency**
  Market Expansion Implementing data-driven decision-making processes has reduced cost and improve profitability , thus enhancing efficiency and resource allocation.

- **Market Expansion**
  Explore opportunities for market expansion by targeting high-growth regions or demographic segments. Develop tailored strategies to penetrate new markets and increase market share.

Full Signature of Supervisor:
…….

Full signature of the student:
Aastha Sinha
……