IBM Skillsbuild Internship Project

Topic : Analysis of Superstore Dataset

Student Details :

Name: Vani Singh

Skillsbuild Email: singhvani3105@gmail.com

College Name: LDRP Institute of Technology & Research

State: Gujarat

Internship Domain: Data Analytics (DA)

Internship Start & End : 12/06/2023 - 24/07/2023

Organization: DGT

LIST OF

CONTENTS

PROJECT TITLE: ANALYSIS OF SUPERSTORE DATASET

INTRODUCTION:

The goal of this project is to conduct a comprehensive analysis of the Sample Superstore dataset to gain valuable insights into Sales trends and Profitability of the store.

- **Sales Analysis:** Analyze sales metrics, trends, and factors influencing sales fluctuations.
- **Profit Analysis:** Analyze the profit and factors affecting profit of various items in the store.

Based on the analysis of the Superstore dataset, our report identifies areas for improvement and provides data-driven recommendations to optimize store performance and achieve better results.

PROJECT OVERVIEW

PURPOSE:

The purpose of the project is to perform Descriptive Data Analysis on Superstore data to gather relevant insights regarding sales and profit of the superstore -

- Extract, Transform & Load (ETL) the data
- Perform Exploratory Data Analysis (EDA) on the dataset

SCOPE:

The project focused on understanding sales trends and profitability across regions, cities, categories, and subcategories within the Superstore dataset.

- **GOAL:**
    a. Perform Descriptive Analysis of the 'SampleSuperstore' Dataset using Python
    b. Highlight the insights regarding the Sales and Profit of the Superstore as gathered from analyzing the data.
- **Tools Used**: Python and libraries- Numpy, Pandas, Matplotlib, and Seaborn

PROJECT OVERVIEW

OBJECTIVES:

- Understand, clean, and visualize the dataset.
- Analyze sales patterns in regions and cities for highest-selling areas.
- Determine top-selling categories and sub-categories.
- Assess profitability of different products and identify most profitable areas.
- Find cities, states, and regions with maximum profit and sales.
- Provide data-driven insights and recommendations for sales optimization and improved profitability.

END USERS

- **Store managers and executives:** Optimize operations, and drive profitability with analysis insights.
- **Sales and marketing teams:** Enhance marketing strategies, and boost sales with analysis findings.
- **Financial analysts and stakeholders:** Gain insights for informed decisions, cost reduction, and investments.

SOLUTION AND ITS VALUE PROPOSITION

SOLUTION:

- The Superstore dataset was analyzed using Python libraries like NumPy, Pandas, and visualization tools. The project aimed to gain insights into sales trends and profitability, employing statistical techniques and data mining. Valuable information about the store's performance, trends, and areas for improvement was obtained.

VALUE PROPOSITION:

- The analysis offers data-driven recommendations to optimize sales, improve profitability, and inform personalized marketing strategies. Python libraries and visualization tools enhance efficiency and interpretability, aiding decision-making processes.

CUSTOMIZATION OF PROJECT

The project followed the 3 basic steps: Loading and Importing the Dataset, Descriptive statistics, and Data cleaning.

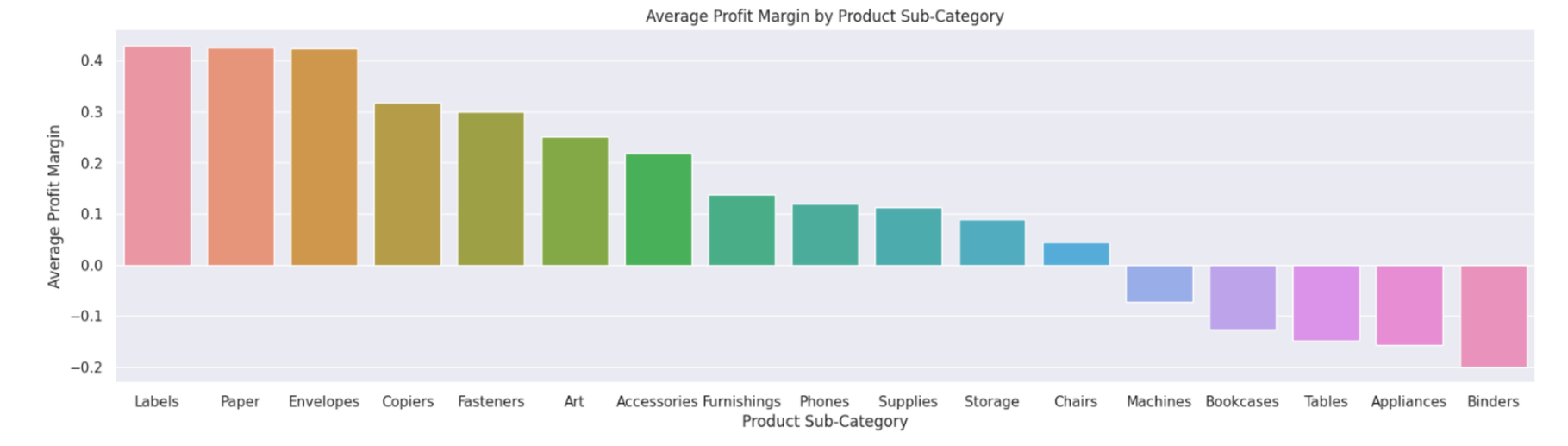
USE OF ADVANCED VISUALIZATION TOOLS:

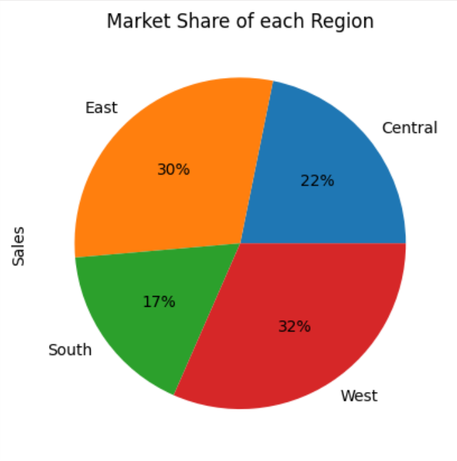
- This project showcases its uniqueness by utilizing advanced visualization libraries like Matplotlib and Seaborn. Through visually appealing representations, it presents insights from data analysis in a clear and understandable manner, ensuring end users grasp the information effectively and easily.

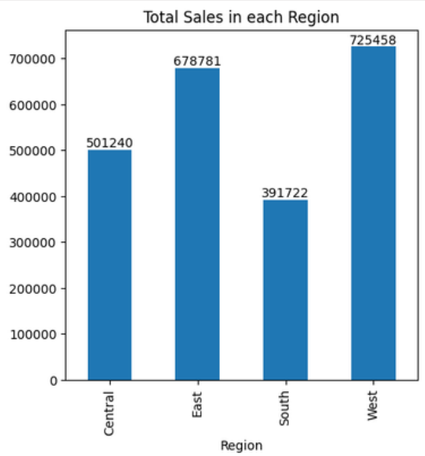
DESCRIPTIVE ANALYSIS & EDA:

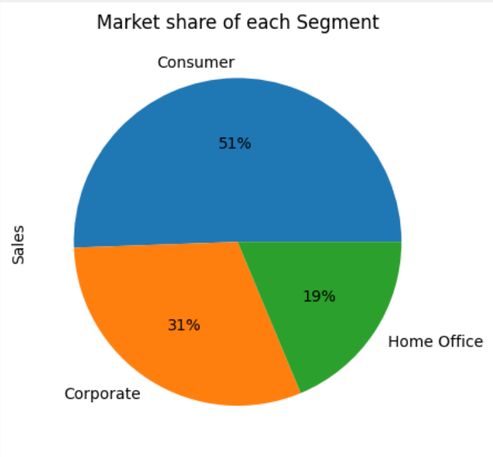
- This project employs Descriptive Analytics and Exploratory Data Analysis techniques to highlight key insights on the Sales and Profit of the Superstore. Utilizing advanced visualization libraries, it presents visually appealing representations to facilitate clear understanding for end users, aiding in decision-making and performance comprehension.

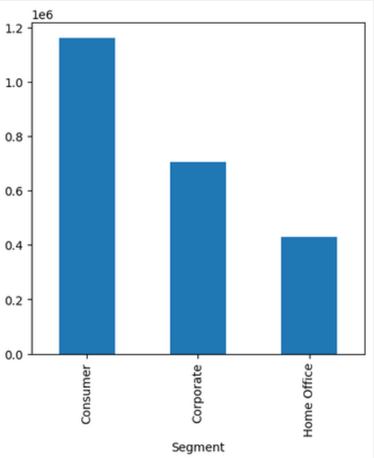
MODELLING & INSIGHTS



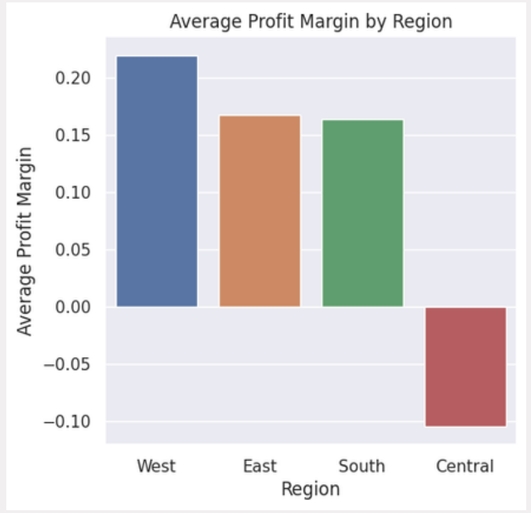


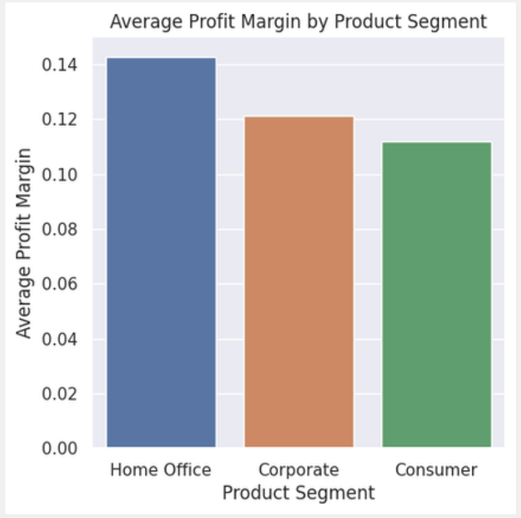


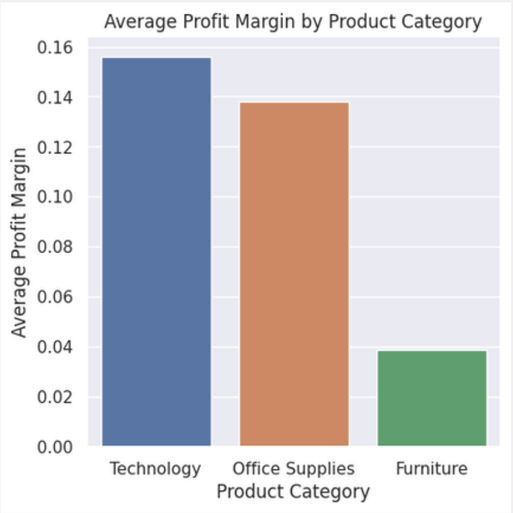


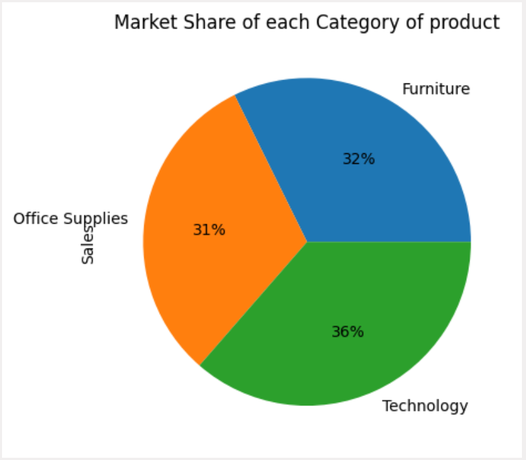


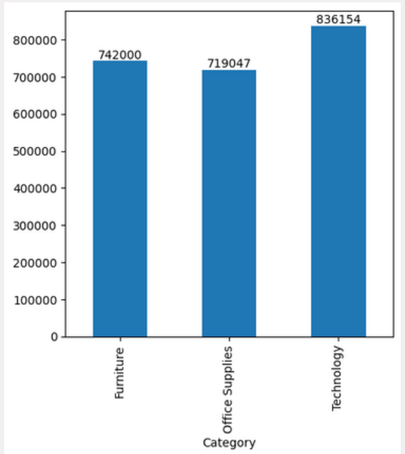
MODELLING & INSIGHTS

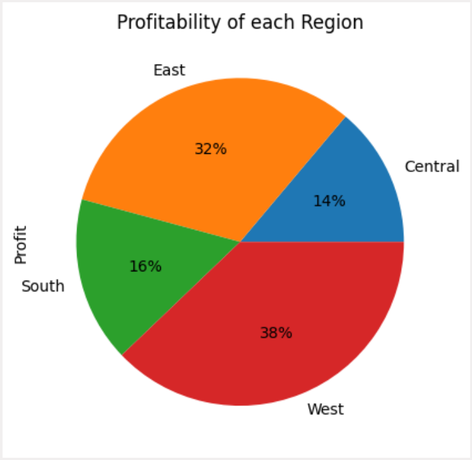


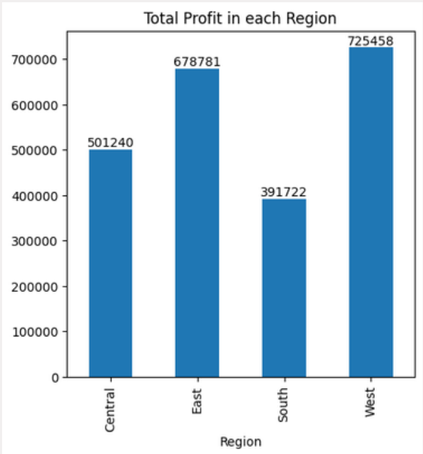




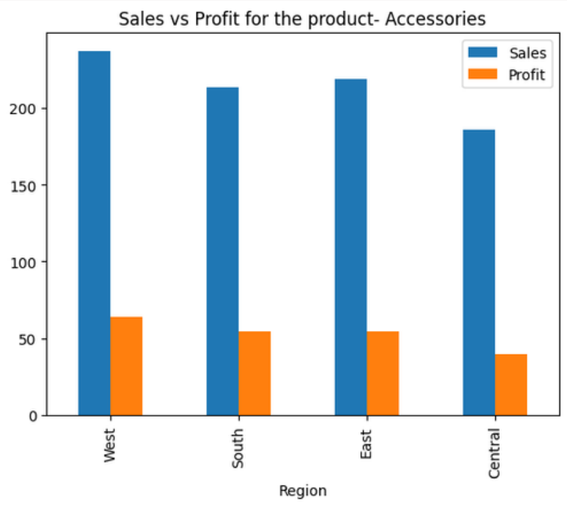


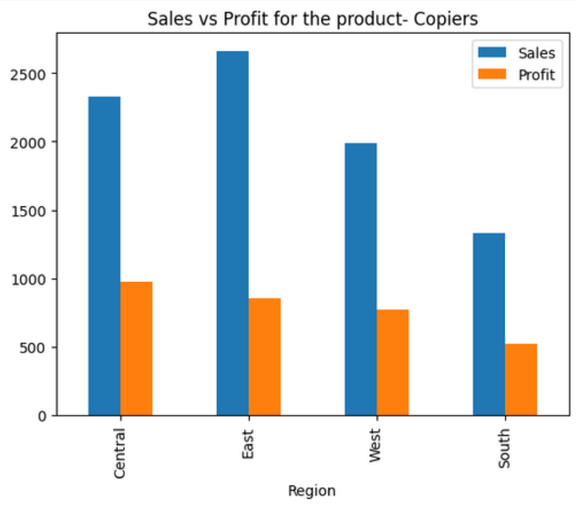


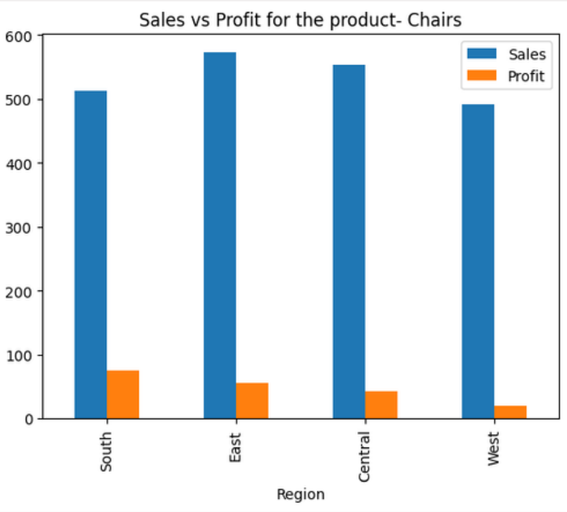


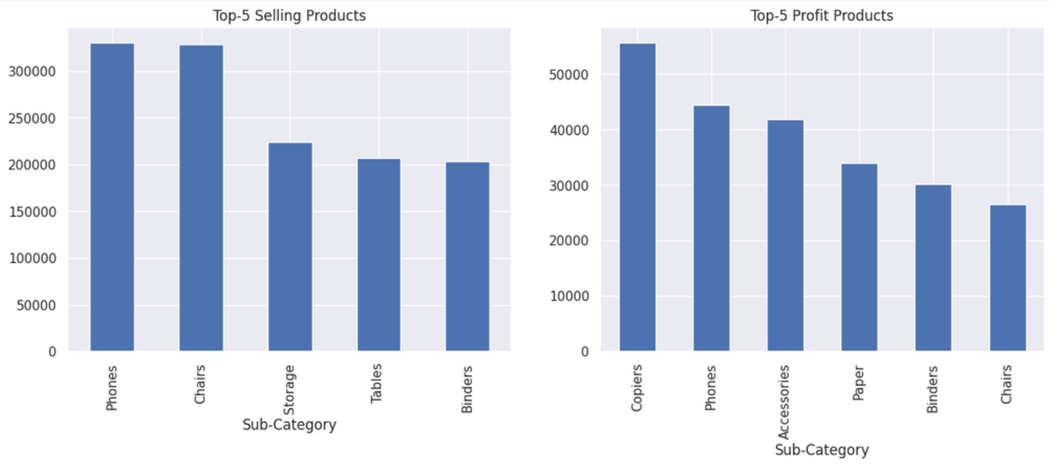


MODELLING & INSIGHTS









RESULTS

KEY INSIGHTS:

- he Sales and Profit Analysis reveals that the Technology Category, particularly Copiers, Fasteners, Accessories, and Phones, yields the highest profit margins, while non-Chair products in the Furniture Category result in losses. Home Office is the most profitable segment, followed by Corporate, and sales in the West Region are the most lucrative, followed by the East, with the Central Region experiencing the highest losses.

ACTIONABLE INSIGHTS:

- To enhance profitability, the Superstore should invest in upgrading Technology Sector production to meet market demand for profitable products consistently. Additionally, efforts must be made to improve profitability in the Furniture Category through cost reduction or finding ways to make other products profitable. Increasing production of Home Office segment products will meet demand and boost overall profit.

RESULTS

CUSTOMIZED SOLUTIONS:

- Customized solutions involve prioritizing resources and investments in cities like Los Angeles (West Region) and New York City (East Region) due to their significant sales contributions. Focusing on production capacities in these cities and others like Seattle will lead to higher profits from product sales. Implementing these actions will ensure improved profitability and sales optimization for the Superstore.

LINKS

GITHUB LINK:

- https://github.com/vanii3105/Superstore-Data-Analysis

DATASET LINK:

- https://www.kaggle.com/datasets/bravehart101/sample-supermarket-dataset

REFERENCE:

- https://github.com/alairdata/Superstore_Analysis
- https://github.com/Sachinnavgale/-The-Spark-Foundation-Tasks