STUDENT DETAILS

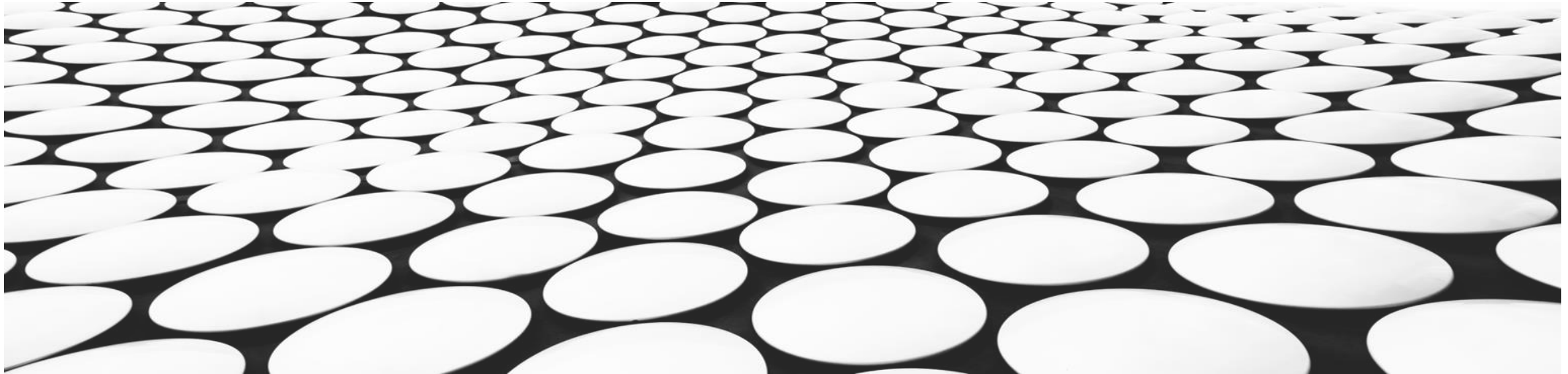
NAME: JAYESH NISHIKANT INAMDAR

SKILLSBUILD EMAIL ID: JAYESHINAMDAR316@GMAIL.COM

COLLEGE NAME : DEOGIRI INSTITUTE OF ENGINEERING AND MANAGEMENT STUDIES, AURANGABAD

COLLEGE STATE : MAHARASHTRA

DOMAIN: DATA ANALYTICS START DATE: 12/06/2023 END DATE: 24/07/2023



EXPLORATORY DATA ANALYSIS OF SUPERSTORE DATABASE

- Exploratory Data Analysis Using Data Analytics, work on the Superstore Dataset to discover insights and record observations of the statistics calculated and graphical representations, to produce solutions to check with our hypotheses and assumptions for the company's actions and policies in order to increase its productivity and profits and minimize losses.

AGENDA

- Super Store is a small retail business located in the United States. They sell Furniture, Office Supplies and Technology products and their customers are the mass Consumer, Corporate and Home Offices.
- The data set contains sales, profit and geographical information of Super Store. Our task is to analyze the sales data and identify weak areas and opportunities for Super Store to boost business growth via answering questions like :-

- Which Category is the Best Selling and most profitable One?
- Which one is the most profitable region?
- Which City is the one with highest sales?

PROJECT OVERVIEW

- Many companies nowadays often neglect the necessity of proper stock management, observation of regular trends and buying patterns of customers and fail to understand its importance.
- The data gathered via the efforts of the IT department and their technology has to be used by Data Analysts in order to recognize patterns, gain insights and draw predictions and suggests solutions which can be beneficial to the enterprise in the future. Such Technology is often used in higher institutions like business corporations, Stock marcescent.
- In the Superstore, Data Analysis can help identify the trends in the purchases of customers, Which area or city are the most consumers belong to, which cities or branches contribute to the profits of the company.

WHO ARE THE END USERS OF THIS PROJECT?

- The End Users of this project are supposed to be the Owners, The branch heads and the Marketing and Service Department of the Superstore chain,
- Owners are likely not people who are expert in reading charts or the interpretation of statistical analysis, thus our analysis will be in layman terms and easy to understand.
- The Branch heads ,can observe the trends seen in the analytics and stock up the stores' wares and inventories accordingly.
- The Service and the Marketing Department of the Superstore now can know about how and where to boost their publicity and what changes are to be expected in the revenues in the upcoming years.

YOUR SOLUTION AND ITS VALUE PROPOSITION

- Our solution for the "Exploratory Data Analysis of Superstore Database" project provides comprehensive data-driven insights and actionable recommendations to optimize the Superstore's business operations, improve customer experience, and increase overall profitability.
- By leveraging data analysis techniques and methodologies, we can offer valuable solutions to the challenges faced by the Superstore, enabling them to make well-informed decisions and stay ahead in the competitive retail market.
- By making data-driven decisions, the Superstore can enhance operational efficiency, improve customer satisfaction, increase revenue, and maintain a competitive edge in the dynamic retail market. Our value proposition lies in empowering the Superstore with actionable insights, enabling them to unlock the full potential of their data and drive sustainable growth.

HOW DID YOU CUSTOMIZE THE PROJECT AND MAKE IT YOUR OWN

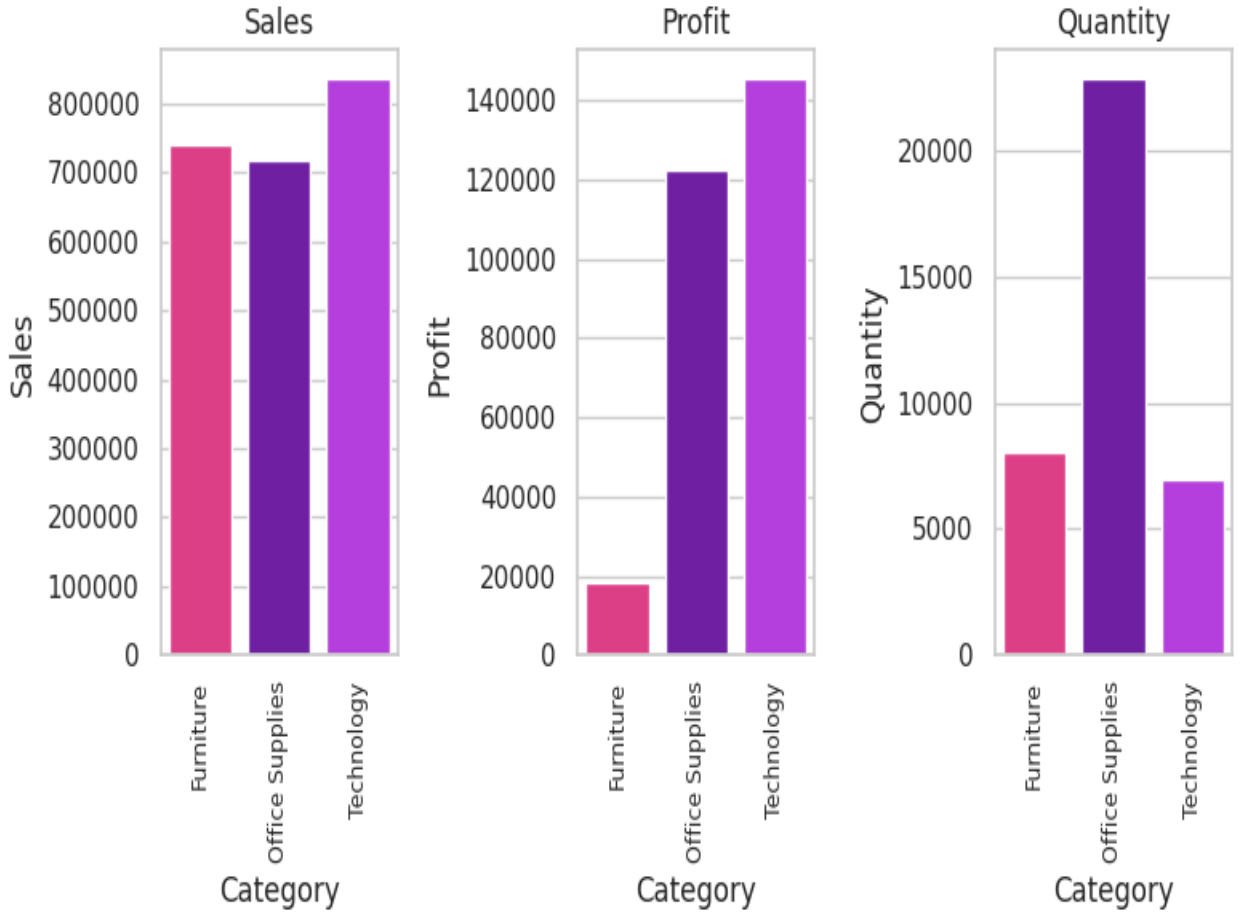
- In the Exploratory Data Analysis of the Superstore Dataset, Firstly I have separated the activity into two sections first part being data cleaning and refining where the data with null and “NaN” values, blank data as well as duplicate data rows are removed from the dataset. And the Data Analysis part where the data is analyzed and are then represented graphically.(Preprocessing is done before this by importing libraries and datasets).
- For Data cleaning, common operations of pandas library are used like info(),sum(),drop_duplicates() and etc.
- For Data Analysis part, the libraries pandas(to manipulate rows and columns of dataset), matplotlib and seaborn(to plot graphs and show them) are used often.
- Here, the given data is adjusted and split into subgroups and then are operated on accordingly and the converted into graphical representation of an appropriate setting which accentuates the insights gained and observations to be recorded.

MODELLING

- Here, first we group the columns into subgroups according to sales, profit and quantity.
- Then, we use the data gathered to plot it into bar graphs for the purpose of analysis and observations via the use of libraries like matplotlib and seaborn to use bar plots.
- We see that:-

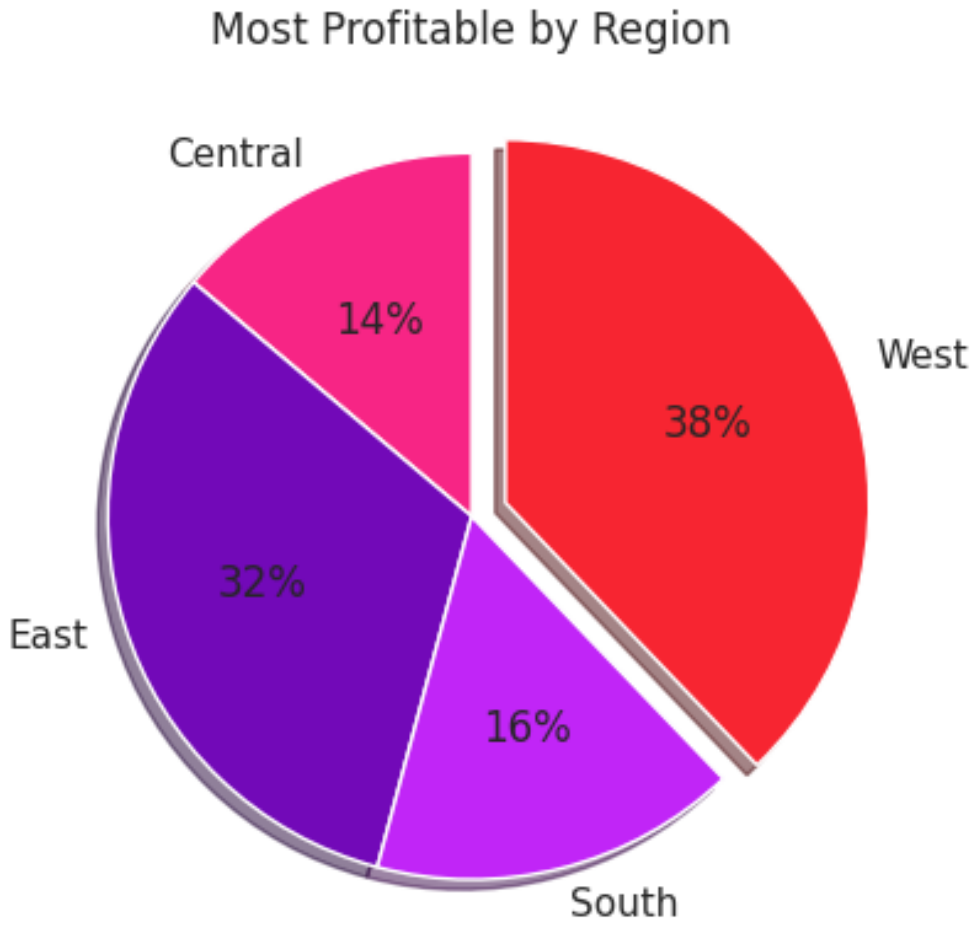
➢ Technology is Best Selling and it's good to know that this category is the Most Profitable too. Only minimal quantity is sold as these products are usually one-off purchases that can last at least 4–5 years.

➢ Although Furniture makes similar sales as Technology, it is the least profitable and quantity sold are at a minimum too.

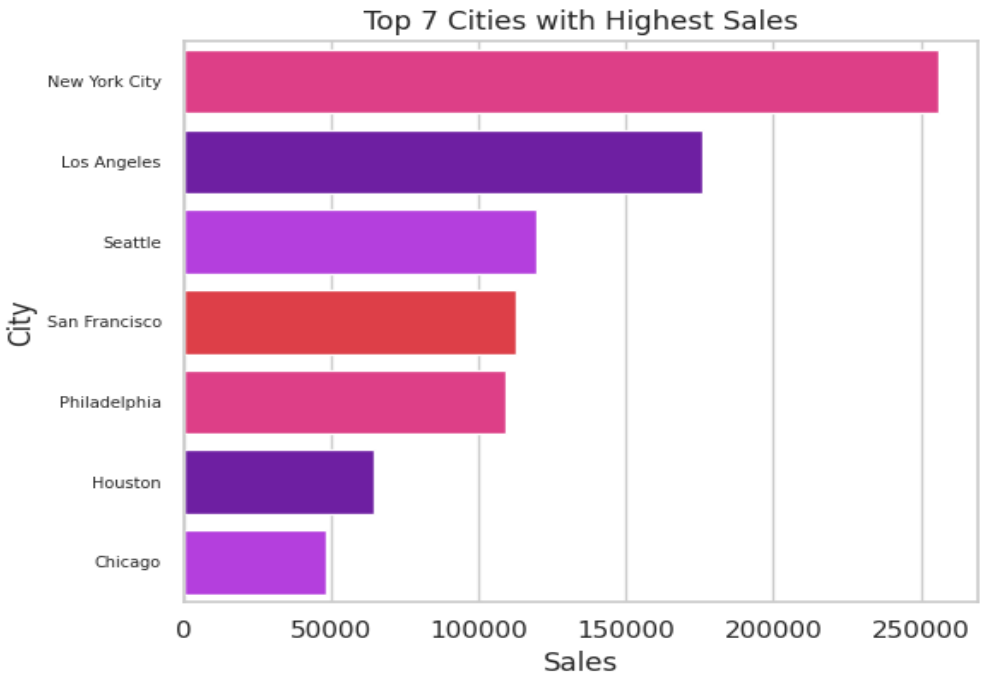


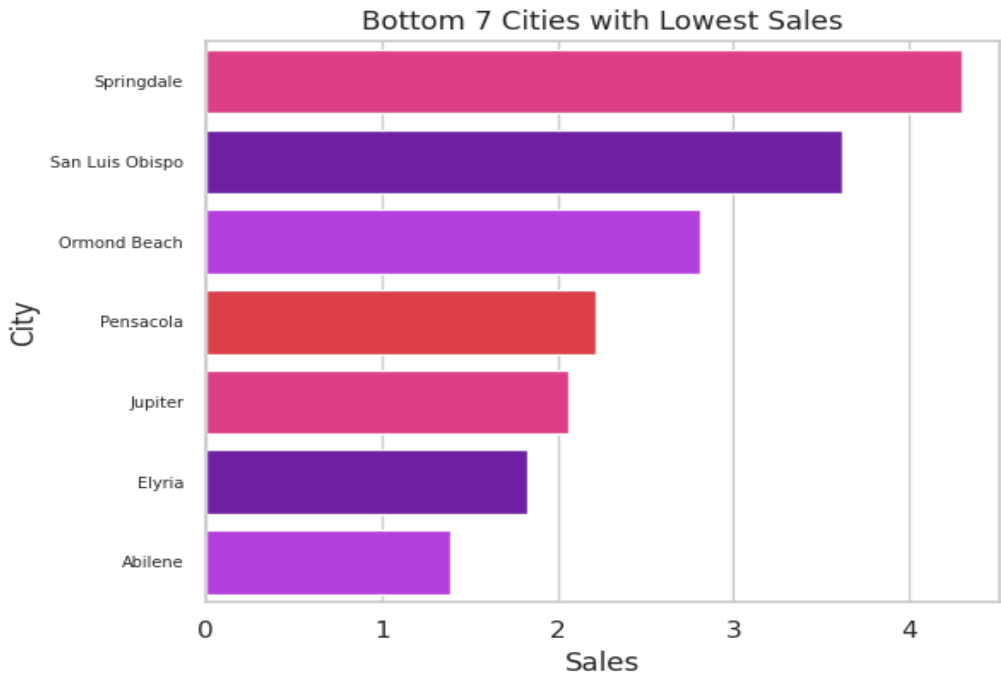
MODELLING

- Here, we simply form subgroups of the columns region and profit to find region with the largest profits.
- We then using the data gathered plot a pie chart to be observed by the user.
- We find the regions in the cities that are responsible for the contribution in the profits for the Superstores.
- East and West region are found to be most profitable.



MODELLING





- Here, firstly we find out the top and bottom 7 cities in the terms of sales revenue generated and the quantity sold via using the library pandas.
- Then, by subgrouping the cities in the data, the bar plot is used to create charts for the first and last 7 cities in terms of sales generated. The libraries used are same as the first two were, but the plotting of bar plot is done with different attributes, dimensions and style to accentuate the conclusions drawn.

RESULTS

- The Data Analysis doe over the Superstore Dataset helps us in analysis of the sales data and identify weak areas and opportunities for Super Store to boost business growth. This has helped us gain insights like:-
- Focusing on Technology sub-category products will benefit more as they are the highest selling and most profitable. Bundling them with the less profitable products such as Bookcases, Table and Chairs(Furniture sub-category products) will help to offset the losses.
- Target the public from profitable areas, especially customers from the East and West region in the Top 10 cities with Highest Sales by introducing special promotions and bundles for mass Consumer and Home Offices and send promotional emails or flyers will boost the profits in the upcoming times.
- We will also need to address the issues in areas like Springdale, Jupiter and Elyria(Cities with lowest Sales) to develop them for better sales performance.

LINKS

- Link to Colab file

https://colab.research.google.com/drive/1LDxBT9bTX9gooxpR19yOWIMQyzPEA_ob?usp=sharing

THANK YOU..