Ambati Sreeja Rao

**1. Write a brief introduction of your project**:

The dataset contains a wide range of information, including customer details, product categories, order dates, shipping information, and sales metrics. It is commonly used to gain insights into sales performance, customer behavior, and other factors that can influence business decisions. The data from this particular store spans over 1457 days with sales totaling at $2,297,201 and a prof it of $286,397.

**2. Domain:** superstore dataset

**3. The data file**



**4. Data source**: https://www.kaggle.com/code/teddolk/superstore-sales-analysis

**5. Number of records and columns:** Columns: 21**,** Rows: 9994

* Row ID => Unique ID for each row.
* Order ID => Unique Order ID for each Customer.
* Order Date => Order Date of the product.
* Ship Date => Shipping Date of the Product.
* Ship Mode=> Shipping Mode specified by the Customer.
* Customer ID => Unique ID to identify each Customer.
* Customer Name => Name of the Customer.
* Segment => The segment where the Customer belongs.
* Country => Country of residence of the Customer.
* City => City of residence of of the Customer.
* State => State of residence of the Customer.
* Postal Code => Postal Code of every Customer.
* Region => Region where the Customer belong.
* Product ID => Unique ID of the Product
* Category => Category of the product ordered.
* Sub-Category => Sub-Category of the product ordered.
* Product Name => Name of the Product
* Sales => Sales of the Product.
* Quantity => Quantity of the Product.
* Discount => Discount provided.
* Profit => Profit/Loss incurred.

**6. Visualization tools selected:**  tableau and excel

**7. Data cleaning strategies if any:** No. I did not clean the data. The data is accurate.

**8.Cleaned dataset, if any. You could link your data set in the MS Word document:** The data is accurate.

**9. The goals of your project and the complete screenshots of charts generated for each goal.**

Goal 1: Analyzing profit and quantity across different regions (Central, East, South, and West).

Story: The bar chart represents profit and quantity of every region. The West region has the highest profit (108,418) and quantity (12,266). The central region has the lowest profit i.e. 39,706 and the south has the lowest quantity.

A screenshot of a computer

Description automatically generated

Goal 2: Analyzing discount for every state.

Story: This Packed bubble chart represents discount for every state. Texas has the highest discount i.e 364.6 and New Hampshire has the least discount i.e 0.3.

A screenshot of a computer

Description automatically generated

Goal 3: Comparing average quantity for cities.

Story: Lollipop chart represents average quantity for cities and compares the average of quantity within the cities. Santa Fe has highest average quantity with 7.0 and lowest is Kissimmee and Abilene with 2.0

A screenshot of a computer

Description automatically generated

Goal 4: Comparing profits for every shipping mode(standard class, second class, first class and same day).

Story: This tree map shows the profits for each shipping mode and compares the profit with other shipping modes. Standard class has the highest profit with164,089 and lowest profit shipping mode is same day with 15,892.

A screenshot of a computer

Description automatically generated

Goal 5: Comparing sum of sales in states of united states.

Story: The choropleth map represents the sum of sales in states. This choropleth shows the states with similar colors where the highest sales have brighter color and lowest sales have lighter color. So, California has the highest sales i.e 457,688 and is North Dakota with 920 sales.

A map of the united states

Description automatically generated with medium confidence

Goal 6: The goal is to analyze and compare the average values of discount, sales, and profit across different regions (Central, East, South, and West).

Story: The Central region had the highest average discount at 0.240352992, indicating a more aggressive pricing strategy. The South region had the highest average sales of 241.8036451. The East region exhibited the highest average profit at 32.13580758.

A screenshot of a computer

Description automatically generated

**11. Use Dashboards & storyboards to represent your charts and findings**

**Dashboard:**

A screenshot of a computer

Description automatically generated with medium confidence

**Storyboard:**

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence

**12. Overall conclusion for your analysis:**

1. Technology and Home Office have on average a high profit margin, while Furniture are being sold at low margin or even at a loss for some it's sub-categories.

2. Over 50% of their profit is being made in California and New York.

3. Their sales in Texas (which is their 3rd highest state in sales) is on aggregate making a loss.