**COVER PAGE:**

# DSC 682-001 - Fall 2024:

# Data Visualization For Business Analytics

# PROJECT TITLE:

**Sales Analysis and Profit Forecasting for Adidas Products**

**GROUP MEMBERS:**

**1.BhavyaSree Pasupuleti**

**2.Raghunath Reddy Bandaru**

**3. Vineeth Kumar Reddy Vaddi**

**4.Venkata Reddy Naru**

**Link to your project’s published dashboard:**

**https://public.tableau.com/views/Book2-project-progressfinal\_17293704025800/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display\_count=n&:origin=viz\_share\_link**

**Background and Motivation:**

We choose the adidas sales dataset for this project to analyze sales trends and evaluate product performance across various regions. This dataset provides key metrics such as units sold, sales revenue, and sales locations, which are essential for identifying successful products and marketing strategies. By uncovering insights from this data, we aim to enhance marketing efforts, optimize inventory management, and compare Adidas's performance with competitors, ultimately supporting strategic planning and businessgrowth.

**Research Questions:**

1. What is the trend of total sales over time?

2. How do sales vary by region and state?

3. How does performance vary between different sales channels (sales method) across regions?

4.Which products are the best-selling?

5.What are the busiest months or periods in terms of units sold?

**Questions added By Professor:**

6.What are the key factors driving high sales volume across different regions and cities? This question could explore which factors (e.g., price sensitivity, product preferences, or sales method) contribute to the high sales volumes.

7.How do different sales methods impact operating profit margins across various regions?

8.Explore relationships between various factors (e.g., price per unit, region, sales method) and total sales or operating profit to identify key variables of success.

**New Questions:**

1**.**What are the top four states which has high profit margin as region wise

2.What are best selling products by gender wise/comparison between men and women sales.

**DATA:**

**Dataset Source:**

The dataset used for this project is publicly available on Kaggle and is titled as Adidas Sales Dataset. An Adidas sales dataset is a collection of data that includes information on the sales of Adidas products. This type of dataset may include details such as the number of units sold, the total sales revenue, the location of the sales, the type of product sold, and any other relevant information.

Source Link: https://www.kaggle.com/datasets/heemalichaudhari/adidas-sales-dataset

**The dataset includes the following variables:**

**Retailer :** Represents the business or individual that sells Adidas products directly to consumers.

**Retailer ID :** A unique identifier assigned to each retailer in the dataset.

**Invoice Date :** The date when a particular invoice or sales transaction took place.

**Region :** Refers to a specific geographical area or district where the sales activity or retail

operations occur.

**State :** Represents a specific administrative division or territory within a country.

**City :** Refers to an urban area or municipality where the sales activity or retail operations

are conducted.

**Product :** Represents the classification or grouping of Adidas products.

**Price per Unit :** The cost or price associated with a single unit of a product.

**Units Sold :** The quantity or number of units of a particular product sold during a specific

sales transaction.

**Total Sales :** The overall revenue generated from the sales transactions.

**Operating Profit :** The profit earned by the retailer from its normal business operations.

**ProfitMargin:** The percentage of revenue that turns into profit, calculated as (Operating Profit / Total Sales) \*100 which measures the profitability of each sale.

**Sales Method :** The approach or channel used by the retailer to sell its products or services.

**ETL ON DATA:**

**Cleaning & Transform The Data:** This step includes cleaning and transforming the data to fit for accurate analysis

**Step1:** Transformed the Date columns into Date Type

adidas$'Invoice Date' <- mdy(adidas$'Invoice Date')

**Step2:** Remove Dollar Signs and Convert Columns to Numeric, For columns like "Price per Unit," and "Operating Profit," Remove the dollar signs and convert them into numeric types.

adidas$`Price.per.Unit` <- as.numeric(gsub("[\\$,]", "", adidas$`Price.per.Unit`))

adidas$`Operating.Profit` <- as.numeric(gsub("[\\$,]", "", adidas$`Operating.Profit`))

**Step3:** **#Create a new column for profit margin in R**

adidas <- adidas %>% mutate(Profit.Margin = `Operating.Profit` / `Total.Sales`)

#Round the profit margin upto 3 decimal

adidas <- adidas %>% mutate(Profit.Margin = round(Profit.Margin, 3))

**Step4:** Renamed the Column names

adidas <- adidas %>%

rename(

Retailer\_ID = Retailer.ID,

Invoice\_Date = Invoice.Date,

Price\_per\_Unit = Price.per.Unit,

Units\_Sold = Units.Sold,

Total\_Sales = Total.Sales,

Operating\_Profit = Operating.Profit,

Sales\_Method = Sales.Method,

Profit\_Margin = Profit.Margin

)

**Loading Data:**

Finally ,the transformed data is stored in an excel file .

#save your data

getwd()

write.csv(adidas,"modified\_adidas\_data.csv")

**DESIGN OF VISUALIZATION:**

**Types Of Visualizations Considered”**

"In my project, we used a variety of visualizations to represent the Adidas sales data effectively:

* **Bar Charts** were used to compare total sales and operating profit across different retailers. This helped identify the top-performing retailers and regions.
* **Line Charts** depicted the trend of sales over time, highlighting peak periods of sales activity.

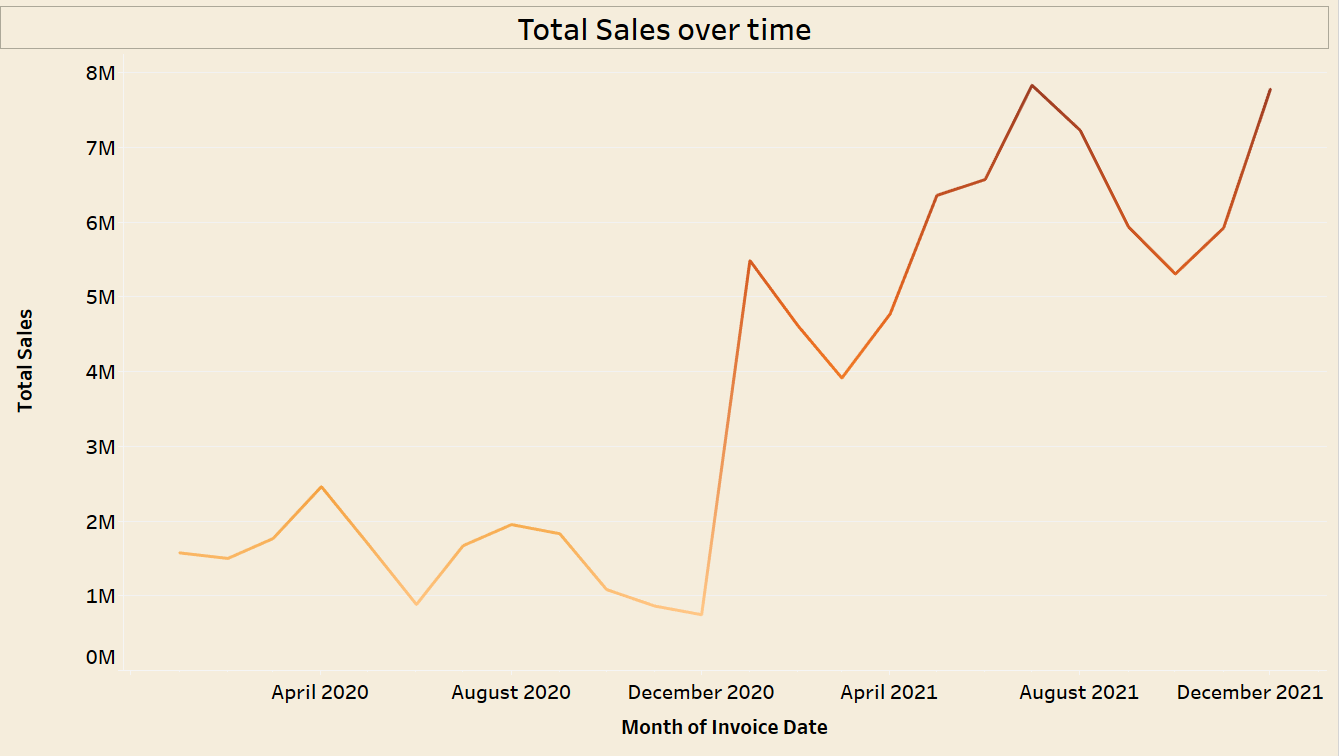
**Design Principles Used:**

Some important design principles might include:

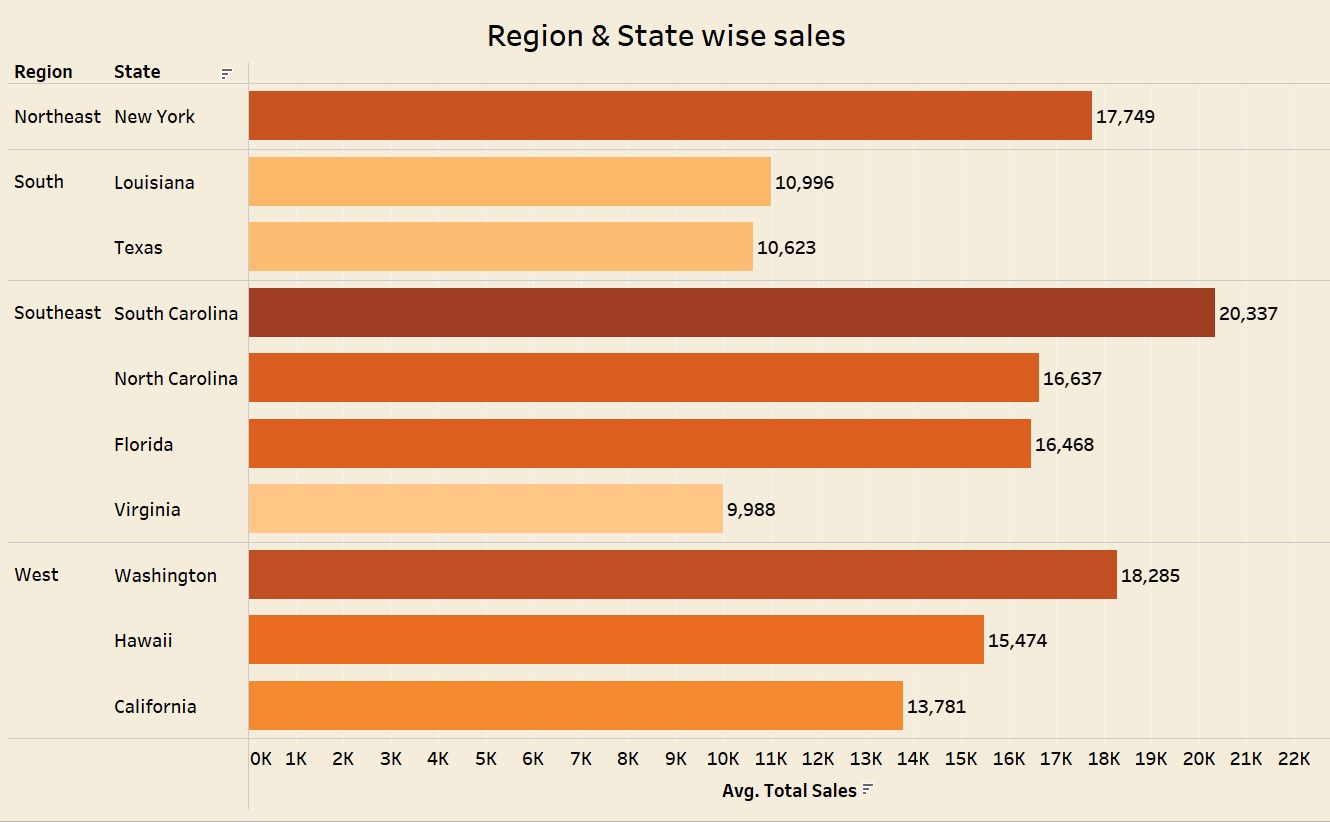
* **Simplicity**: Avoid clutter and unnecessary details. To Ensure the visualizations are clear and easy to interpret by the audience.
* **Color Usage**: Used color purposefully to highlight key trends, regions, or categories.
* **Consistency**: Maintained consistent use of colors, fonts, and labels across all charts to create a cohesive look.
* **Clarity**: The charts are labeled properly, including axes, legends, and titles, so the viewers can easily understand the context of the data.

**SCREENSHOTS OF VISUALIZATIONS:**

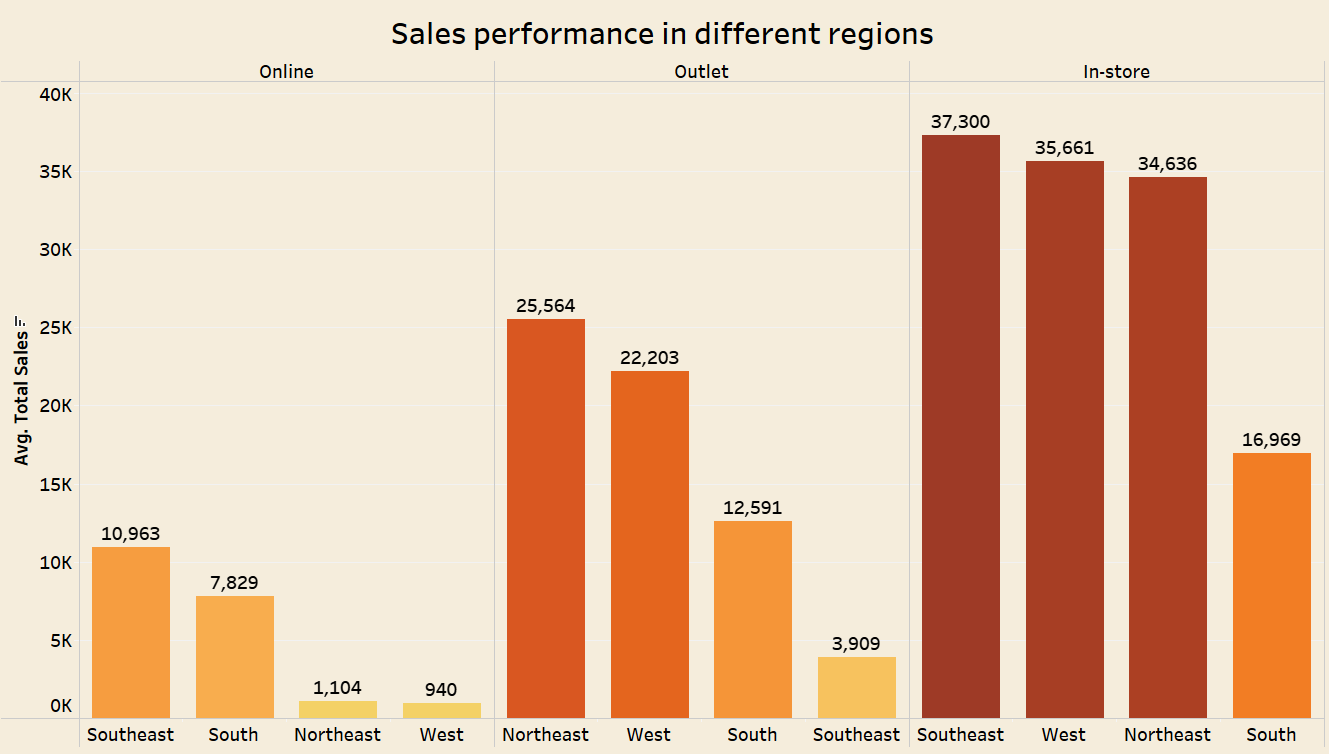
**1.Visualization for Describes Trend of Sales Overtime**

****

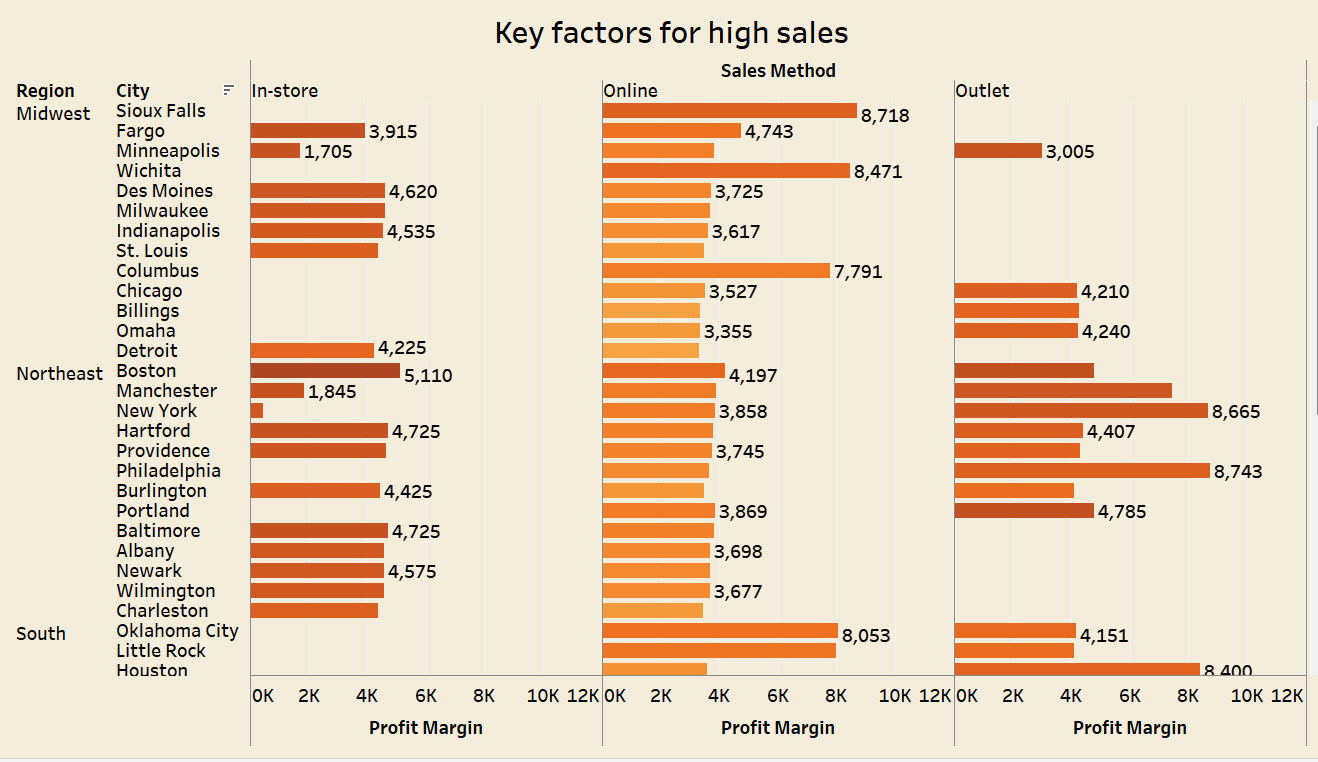
**2.Visualization for Sales across region wise:**



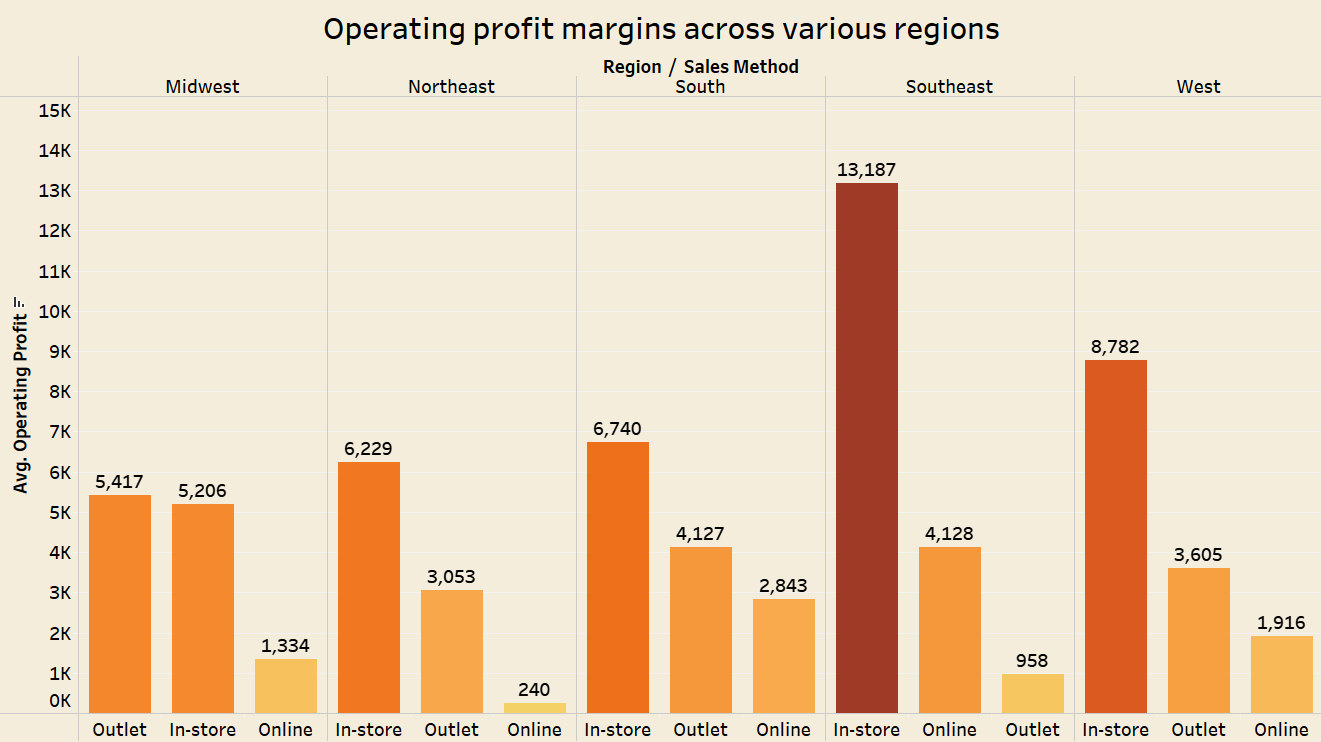
**3.Visualizatio For Sales Performance across regions:**



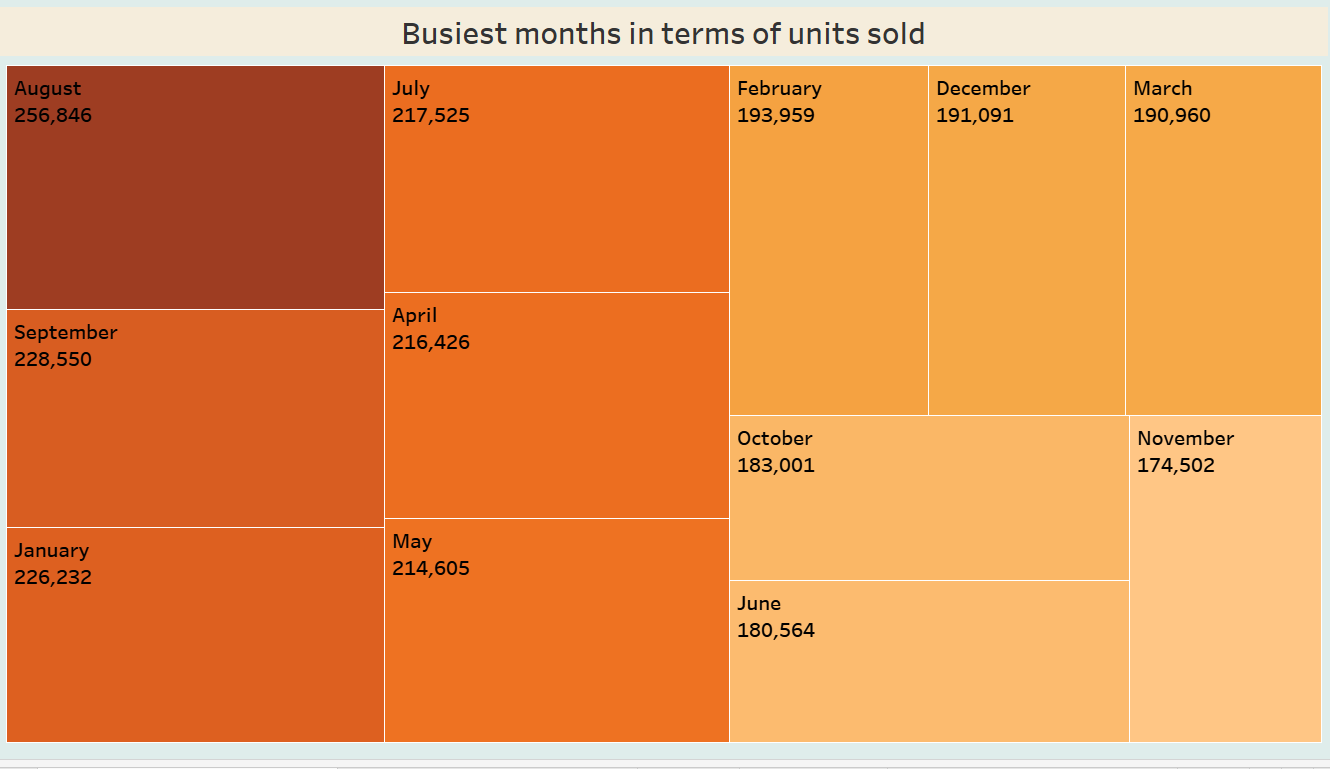
**4.Visualization of Key Factors for High Sales**



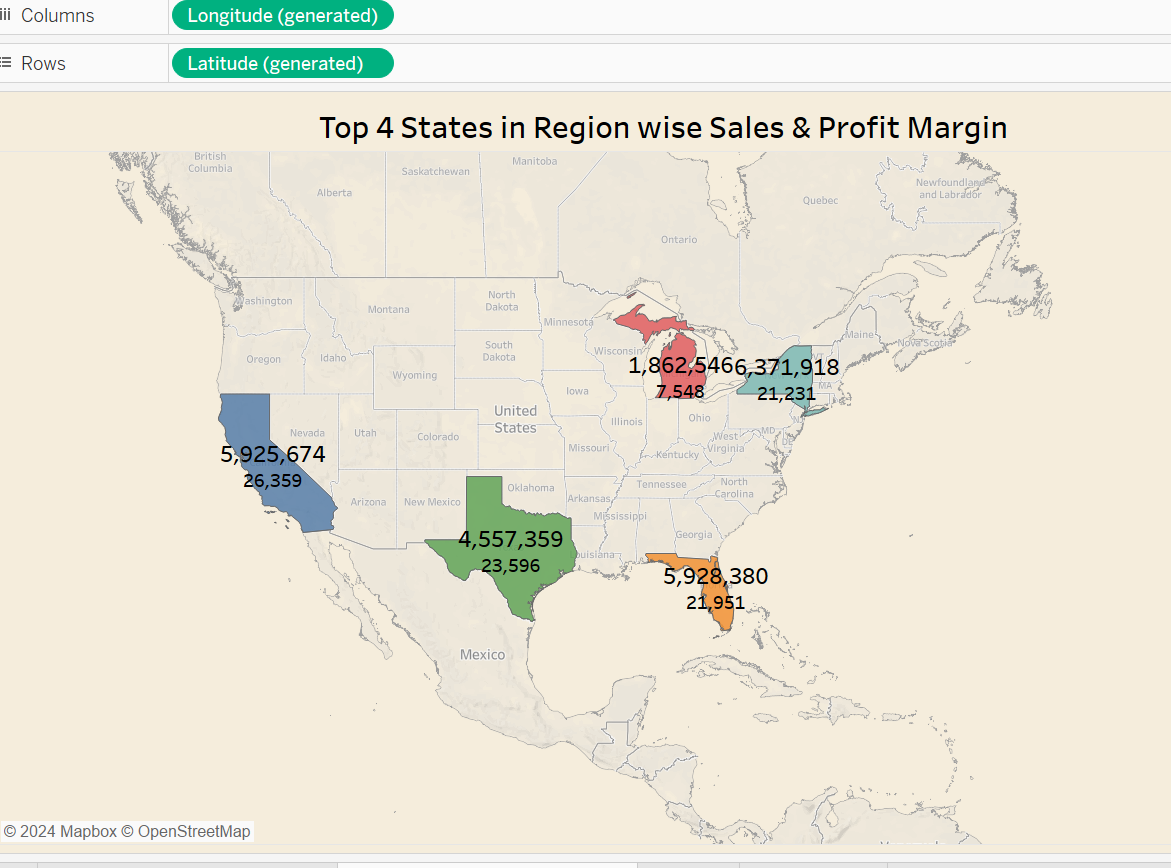
**5.Visualization of Operating Profit margins across various regions:**



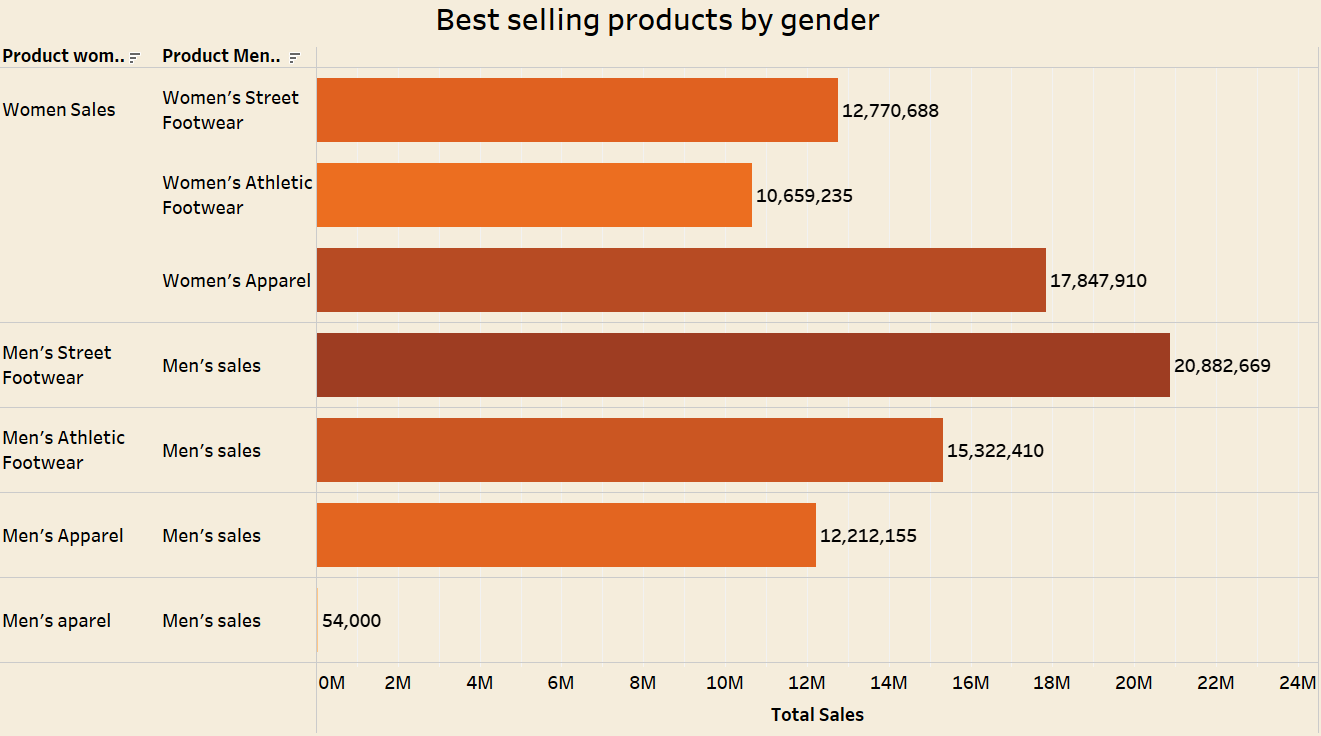
**6.Visualization of which months have the best sales:**



**7. Visualization of top 4 States sales across varios Regions** :



**8. Visualization of Men and Women Sales**:



**IMPLEMENTATION:**

In my visualization, the core functionality revolves around an interactive dashboard that allows users to explore Adidas sales data across different dimensions, such as region, product type, and retailer.

The main interactive elements include:

* **Filters** for region, product type, and retailer, allowing users to refine the data and focus on specific areas of interest.
* **Highlights** that emphasize the selected retailer or product when clicked, enabling users to easily compare performance across categories.
* **Tooltips** that provide detailed information about sales, profit, and margins when hovering over a data point, giving users immediate insights without cluttering the visual space.
* **Colours** Used constant coloring for our visualization to make best view for the audience.

**Use Of Generative AI:**

* **Generative AI for Design Guidance**: ChatGPT was used to guide the design process, providing suggestions on effective visualization techniques based on the data type and the insights needed. For example, recommendations on when to use bar charts for categorical data, line charts for trends, or heatmaps for geographic analysis were provided.
* **Layout Optimization**: ChatGPT also helped in optimizing the layout of the dashboard, suggesting where to place filters, how to structure visual hierarchy, and which elements to emphasize (e.g., focusing on profit margin or sales depending on the audience).
* **Enhancing Usability**: Generative AI helped to fine-tune the dashboard by suggesting ways to improve user experience. This included simplifying visuals, choosing appropriate color schemes for accessibility, and ensuring clarity in labels and titles.

**Results and Conclusions:**

**Answers For Research Questions:**

1.**What is the trend of total sales over time?**

Total sales show an upward trend, with certain months or quarters exhibiting stronger sales spikes, possibly tied to promotional events or seasonal demand.

2.**How do sales vary by region and state?**

Sales are higher in the **South** and **Southeast** regions, with states like **Alabama** and **Florida** driving significant sales, particularly in the street footwear category.

3.**How does performance vary between different sales channels (sales method) across regions?**

Outlet sales perform better in terms of profit margins, especially in southern regions, whereas online sales are more consistent across all regions.

4.**Which products are the best-selling?**

**Men's and Women's Street Footwear** are the top-selling products, contributing the most to both units sold and revenue.

5.**What are the busiest months or periods in terms of units sold**

Peak sales occur during **summer months** and promotional periods, with **July** and **August** showing the highest units sold.

6.**What are the key factors driving high sales volume across different regions and cities?**

High sales volume is driven by factors such as **product price**, regional **product preferences**, and the **sales method** used (e.g., outlets in the South).

7.**How do different sales methods impact operating profit margins across various regions?**

**Outlet sales** tend to yield higher operating profit margins in regions like the South, while **online sales** provide consistent, though slightly lower, profit margins across all regions.

8.**Explore relationships between various factors and total sales or operating profit.**

Key variables driving success include **price per unit** (lower prices lead to higher volumes), **sales method** (outlets perform better in certain regions), and **product type** (footwear generates the highest profit and sales).

**What We Learned from the Project:**

* **The importance of regional insights:** Different regions have varying preferences for product types, and understanding this can help in better targeting and inventory management.
* **Profitability analysis:** Analyzing both profit margin and operating profit helps reveal not just the top sellers but also which products are most profitable, even if they sell in lower volumes.
* **Impact of pricing:** Higher unit prices tend to correlate with higher profit margins, but sales volume plays an equally important role in determining total sales and profit, especially for low-priced products.

**Analytical Conclusions:**

* **Retailer Performance**: Sports Direct leads in sales and profit, likely due to its strong focus on street footwear, which is a high-demand product in the South region.
* **Product Profitability**: Footwear, particularly Men's and Women's Street Footwear, contributes significantly to profitability, while apparel products, although still profitable, lag slightly behind.
* **Strategic Opportunities**: There may be opportunities for expanding sales of high-margin products like footwear in regions where they currently underperform or tailoring marketing efforts to regions showing high demand for specific categories.

**References:**

1. **Dataset Reference:**

Adidas Sales Dataset. Retrieved from Kaggle.  
Source Link: https://www.kaggle.com/datasets/heemalichaudhari/adidas-sales-dataset

1. **Visualization Tool:**

Tableau Public for data visualization and insights.  
Source Link: [Tableau Public](https://public.tableau.com/app/discover)

1. **Generative AI:**

OpenAI’s ChatGPT (Versions GPT-3.5 and GPT-4) for suggestions on data analysis, visual design, and interpretation. The tool was used across several phases of the project for insights on data cleaning, feature engineering, and effective visualization strategies. ChatGPT was particularly useful for generating text summaries and enhancing report explanations.  
Retrieved from <https://chat.openai.com>

**Project DashBoard Website:**

https://public.tableau.com/views/Book2-project-progressfinal\_17293704025800/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display\_count=n&:origin=viz\_share\_link