Name: Drishti Harshit Doshi

Student ID: 2203898

Module Code: CS5803

Module Title: Data Visualisation

Academic Year: 2022-2023

Assignment Title: Visualisation Design Task

Data:

The given dataset is about sales data of Adidas, a sportswear manufacturing company. This dataset has 9649 rows and 13 columns describing different retailers, regions, product details, total sales revenue, operating margins, and sales methods - Adidas company. The Adidas Sales Dataset is available on Kaggle.

https://www.kaggle.com/datasets/heemalichaudhari/adidas-sales-dataset?resource=download

Data Dictionary:

Name	Description	Domain
Retailer	Stores that sell Adidas Merchandise	String [eg Walmart, Footlocker]
Retailer ID	Unique ID of the retailer	String [eg 1128299, 1185732]
Invoice Date	Date of transaction	Ordinal [January 2020 – December 2021]
Region	Region for retailer	String [eg South, North-East]
State	State within that region	String [eg California, Texas]
City	A city within that state	String [eg San Francisco, New York]
Price per unit	Price of each unit	Integer [$7 - $110]
Product	Type of product being sold	String [eg Men's Athletic Footwear, Women's Apparel]
Units Sold	Number of units sold	Integer [0 -1275]
Total Sales	Price of Goods * Number of units sold	Integer [$0 - $825000]
Operating Profit	Total Sales - COGS	Integer [$0 - $390000]
Operating Margin	(Total Sales - COGS)/ Total Sales	Integer [10% - 80%]
Sales Method	Method of payment for that sale	String [Outlet , Online]
Profit Ratio	**The ratio of Operating Profit to the Total amount of Sales** **SUM ([Operating Profit])/SUM([Total Sales])**	**Ratio 0.25 - 0.50**
Rank	**Retailer companies position** **RANK ([Profit Ratio], ('desc'))**	**Integer (1 - 6)**

Persona:

Josh Gayle

Store head at Sports Direct

"I am all about numbers"

DEMOGRAPHICS		BEHAVIOR	
Gender:	Male	Active:	★★★☆☆
Age:	47	Statistical:	★★★★☆
Location:	Texas	Competent:	★★★★☆
Job:	Store head	Management:	★★★☆☆

BIO

Josh Gayle is a Store head at Sports Direct. His goal is to improve the shopping experience for his customers and constantly keep up to date with the latest technologies and trends in the sports industry. Josh has always been passionate about sports and fitness. He holds a master's degree in Sports Management. Josh has been working as a Sports Retailer in the industry for the past 10 years. He enjoys playing football, listening to country music, and campaigning in his free time. Apart from this, Josh also takes part in volunteering for sports charity events.

GOALS	FRUSTRATIONS
● Higher profit margins beneficial for the store	○ Enough stock inventory especially during seasons
● Attract more customers for sales	○ Troublesome customer reviews
● Providing customer with the best possible service and products.	○ Competition from online retailers
● Varied sales methods at the retail for customer convenience	○ Customer churn management.

Questions:

Q1) How do the sales trends vary among the retailers across the different cities?

Q2) How do seasonal fluctuations impact the performance of retailers' businesses?

Requirements:

Here we specify the main requirements in terms of the relationships that must be visualized to answer each question, followed by brief design ideas for representations and interactions.

R1) **To answer Q1** the user needs to see the distribution of sales trends distributed across cities in the USA for different retailers.)

- **Symbol Map** for visualizing the distribution of **Units sold**, **Total Sales**, and **Operating Profit** of **Retailers** across different **Cities.**
- Filter by **Retailer**

R2) **To answer Q1** the user needs to make a comparison of the sales distributed across cities for different retailers.

- **Word Cloud** for visualizing **Operating Profit** and **Total Sales** across **Cities**
- Filter by **Retailer**

R3) **To answer Q2** the user needs to see the sales trends across seasons for different retailers.

- Dual Axis Bar and a Line graph of Total Sales and Operating Profit by **Invoice Date (Quarterly)**
- **Bar graph - Total Sales**
- **Line graph - Operating Profit**
- Filter by **Retailer**
- Filter by **Sales Method**

R4) **To answer Q2** the user needs to compare the sales trends of all products for the years 2020 and 2021 for different retailers.

- **Dumbbell graph** for **Products** and **Total Sales** for the years **2020 and 2021**
- Filter by **Retailers**

R5) **To answer Q2** the user needs to compare the operating profit with total sales for different retailers.

- **Scatter Plot** (Linear Regression) between **Operating Profit** and **Total Sales**
- Filter by **Retailer**

R6) **To answer Q2** the user needs to rank the Retailer company based on their financial performance for 2 years i.e., January 2020 - December 2021.

- Creating Calculated field: **Profit Ratio** = SUM ([Operating Profit])/SUM ([Total Sales])
- Creating Calculated Field: **Rank** = RANK ([Profit Ratio], ('desc'))
- **Text Table** to display **Retailers**, **Profit**, and **Rank**

Retailer and **Sales Method** have been added as **Global Filters** in the dashboard.

R7) Non-Functional Requirements:

- Comprehensible **user interface** for analyzing sales trends for different retailers.
- **Private** database for **indoor** retail shop

R8) Technical Requirements:

- **Higher Screen Resolution** for better display of visuals and statistics with clarity.
- Updated version of **Tableau** software
- **Presentation mode** of dashboard and story

Design:

Prototype Design using Paper Landscape: (CS5703 Week 9 Teaching Materials, 2022)

Adidas Sales Data

Goal: What are the variations among US retailers in terms of sales determinants such as total sales, operating profit, and the number of units sold for different products?

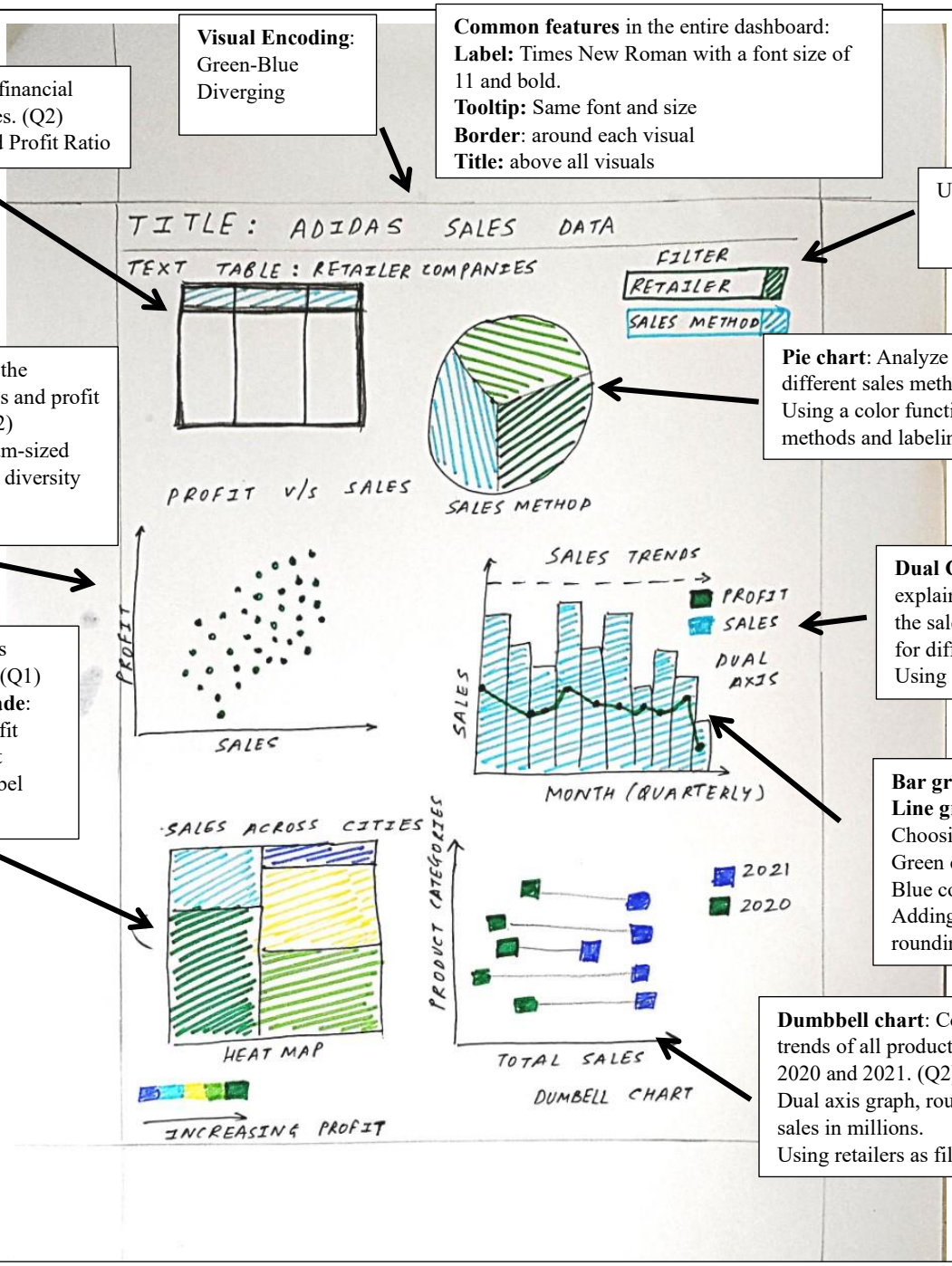
Q1) How do sales trends vary among retailers across different cities?
Q2) How do seasonal fluctuations impact the performance of retailers' businesses?

Visual Encoding: Green-Blue Diverging

Common features in the entire dashboard:
Label: Times New Roman with a font size of 11 and bold.
Tooltip: Same font and size
Border: around each visual
Title: above all visuals

Text Table: Compare the financial performances of companies. (Q2)
Calculated field: Rank and Profit Ratio

Use of **Global Filter**:
1) Retailer
2) Sales Method



Scatter plot: To analyze the relationship between sales and profit for different retailers. (Q2)
Using green color, medium-sized filled dots to show visual diversity in the dashboard.

Pie chart: Analyze sales distribution for different sales methods. (Q2)
Using a color function for 3 different sales methods and labeling the total sales using

Dual Combination chart: To explain seasonal variations in the sales trends among retailers for different products. (Q2)
Using Quarterly invoice date

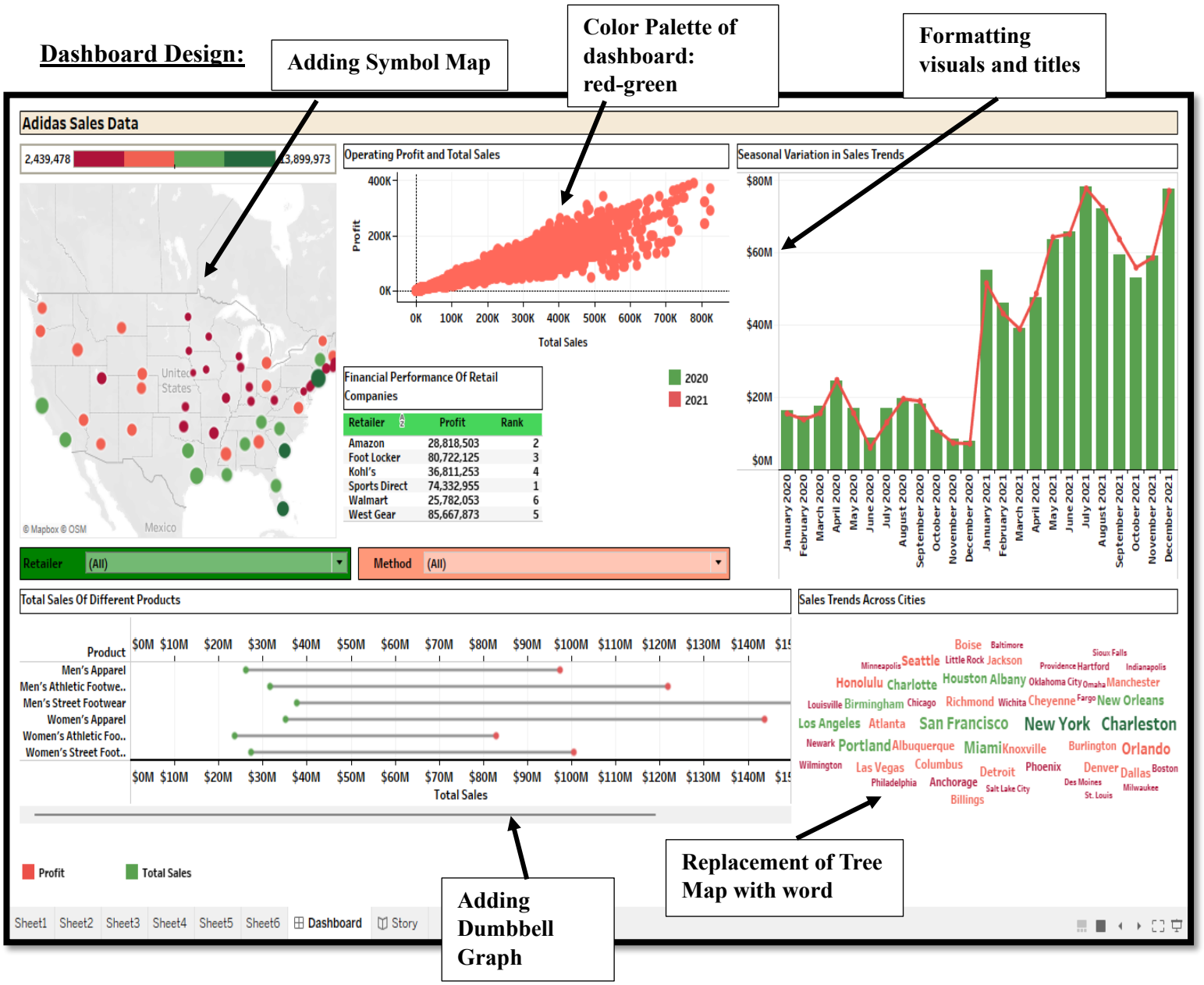
Tree Map: To show sales trends in different cities. (Q1)
Using **Stepped color shade**:
Dark Green: Highest Profit
Dark Blue: Lowest Profit
Labeling Cities using label function.

Bar graph: Total Sales
Line graph: Profit
Choosing a dotted line path
Green color for Profit and Blue color for Sales
Adding dollars to sales, rounding off to millions.

Dumbbell chart: Comparing sales trends of all products for the years 2020 and 2021. (Q2)
Dual axis graph, rounding of total sales in millions.
Using retailers as filters.

Results: The user is able to gain insights into the sales trends for different retail companies in the USA.

Dashboard Design:



Changes in the final implementation of the dashboard:

Firstly, **TreeMap** was replaced with **Word Cloud**. In the process of making a dashboard, I realized that the darker density-filled color of the TreeMap was obstructing the pop-up effect of the city names. Also, the dynamic changing of the Word cloud helps to give better insights about the data and makes them more attractive. As my research question signifies sales trends in different regions of the USA, I was guided by my teacher to add **Symbol Maps**. Later, I came across **Dumbbell graphs**, they were better at revealing differences in a year span. Hence, I tried to portray details about sales of different products for different retailers in the years 2020 and 2021. Due to fitting issues, I had to delete the **Sales Method Pie chart** and it was not useful in answering my questions. Then I tried fitting the visuals as per the prototype in the actual Dashboard. Due to size issues, I had to change the layout of the Dashboard for visual clarity and fitting. The use of a blue-green palette was causing the blue color to be very opaque. Hence, I tried different color combinations using a tableau palette and found the green-red palette better. Initially, I aimed to use **Times new roman** font but later I had to change the formatting to Tableau medium and added borders, color shades to labels, formatting tooltips, and titles.

Implementation:

a) Importing Adidas Sales Data file. Renaming the feature names and changing their data types.(Figure1)

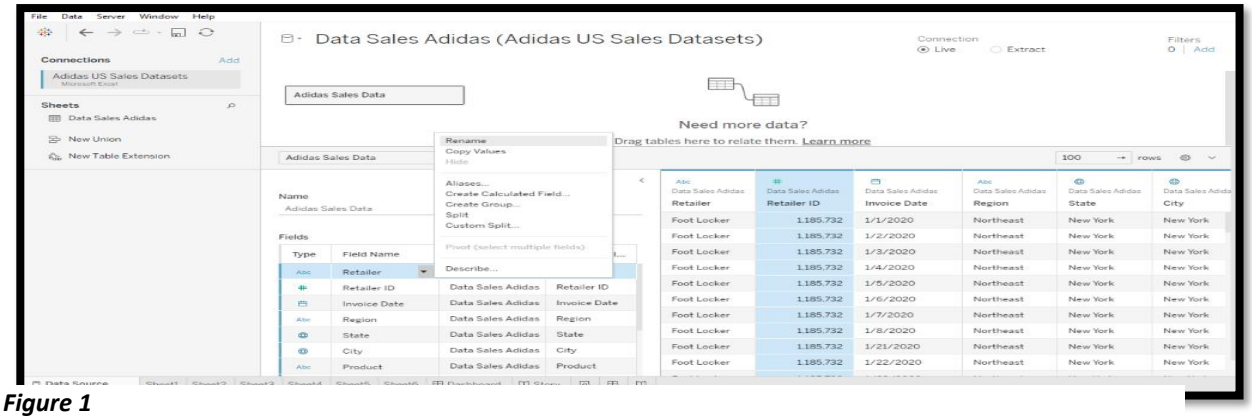


Figure 1

b) Sheet 1: Financial performance of retail companies (Figure2 and Figure 3)
Creating calculated fields by right-clicking on the measures and dimensions column
- Profit Ratio = Operating profit/Total sales
- Rank = RANK ([Profit Ratio],('desc'))

Using "show me" for the suggestion. Selecting text-table for Retailers, Profit, and Rank Formatting the visual by changing the font type to Tableau medium, size to 8, and Bold, Black color and adding a border to the title with a green shade.

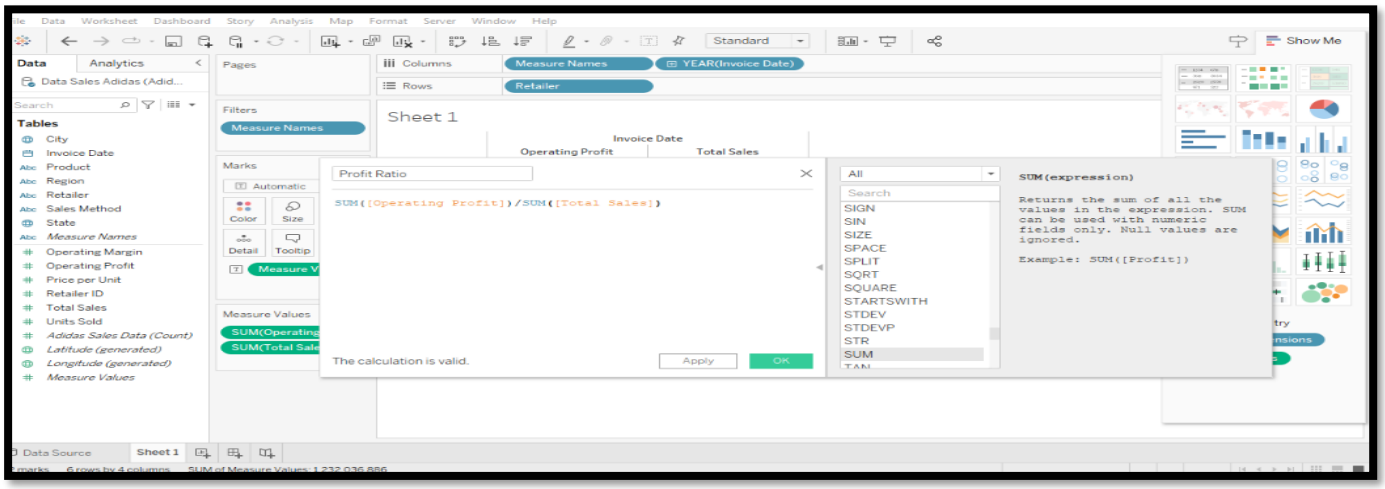


Figure 2

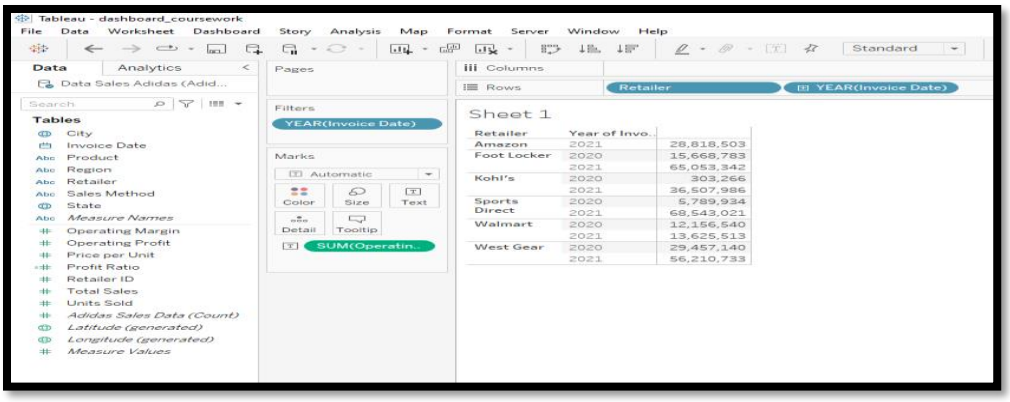


Figure 3

c) Sheet 2: Sales trends across cities (Figure 4)
Creating a heatmap by dragging total sales to rows and cities to columns.
Selecting size from marks and changing the type to text. Creating a word cloud of cities for different sales prices and profit. Adding cities into the color option (marks). Selecting a red-green palette (stepped color). Editing the tooltip by providing state, city, total sales, and profit. (CS5703 Week 3 Teaching Materials, 2022)

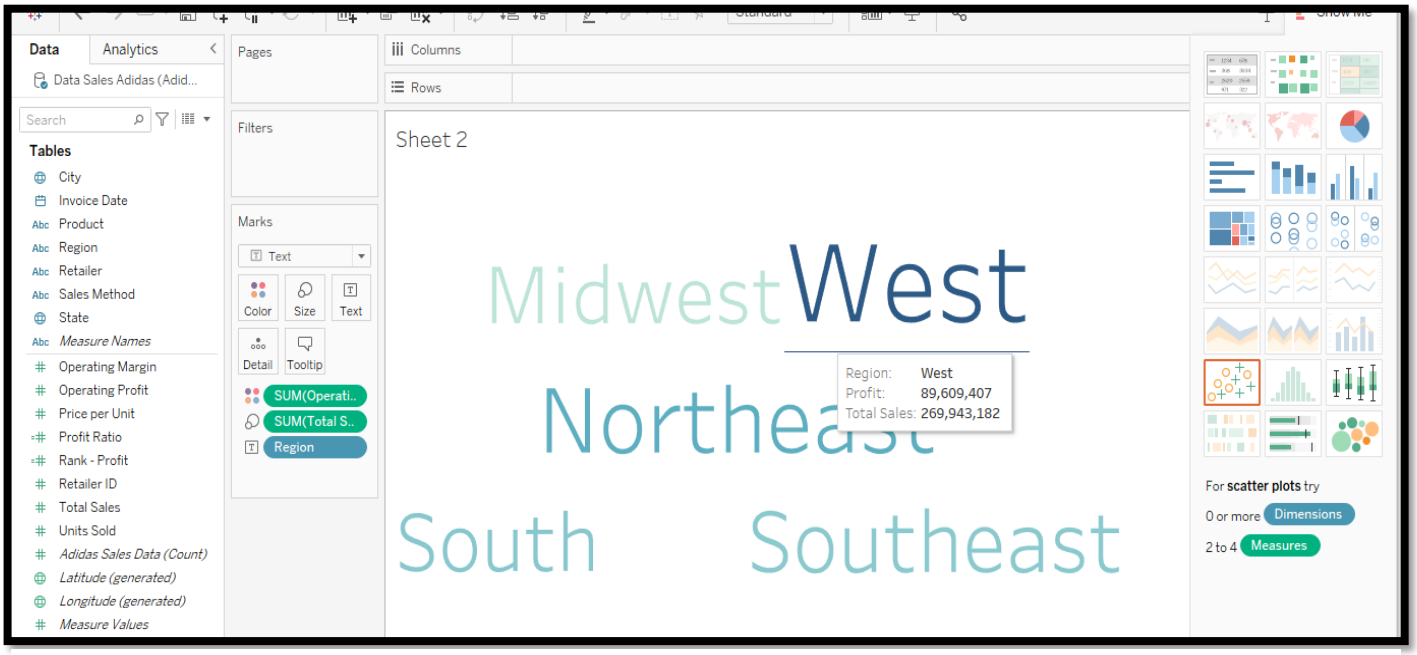


Figure 4

d) Sheet 3: Total sales of different products: Creating a dumbbell chart (Figure 5)
Adding Products into rows. Adding Total Sales into columns two times. Right-clicking sales and selecting Dual-Axis chart. (Tableau)
Adding invoice order in the filter, selecting years 2020 and 2021 from the filter option. Selecting the sum of sales (1) from marks and dropping the order date (year) from the filter option to color. Editing the legend 2020 and 2021 by assigning red and green colors from the palette. Selecting the sum of sales (2) and clicking on the Line option instead of automatic. A zigzag line is formed connecting all data points. Adding invoice order(year) option from the filter to Path. Synchronizing the axis and removing the 2nd dual axis from the visual.

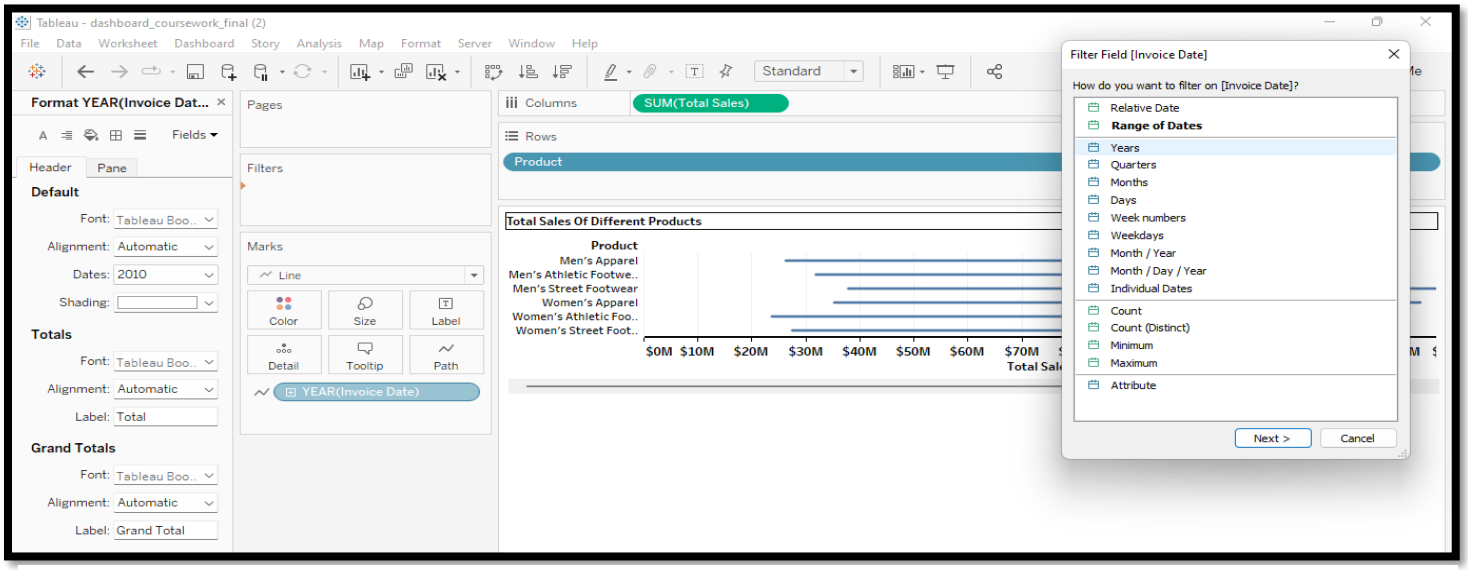


Figure 5

e) Sheet 4: Map (Figure 6)

Selecting cities and sales from columns. Choosing Symbol Map from show me. Adding price quantity into shape and profit into color. Selecting a red-green stepped color palette for the legend.

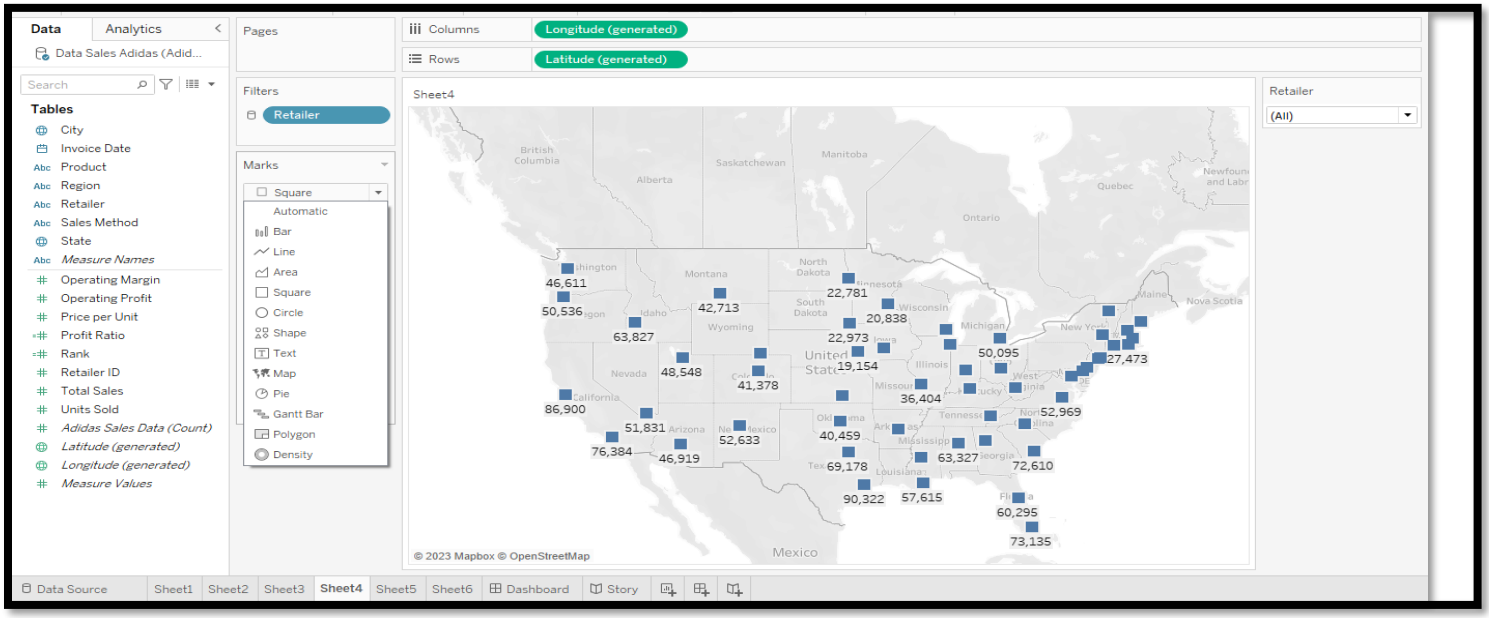


Figure 6

f) Sheet 5: Seasonal variation in the sales trends (Figure 7)
Dragging invoice order into column
Right-clicking and selecting quarterly (month)
Selecting Dual-Axis graph. (sum) in total sales (1): selecting the bar option instead of automatic
Selecting (sum) Total sales (2): selecting line option. Adding a dotted option from the path to the line graph.
Choosing 2 different colors red and green for the bar and line graph. Sync Synchronizing dual axis and removing the second axis. Formatting the labels.

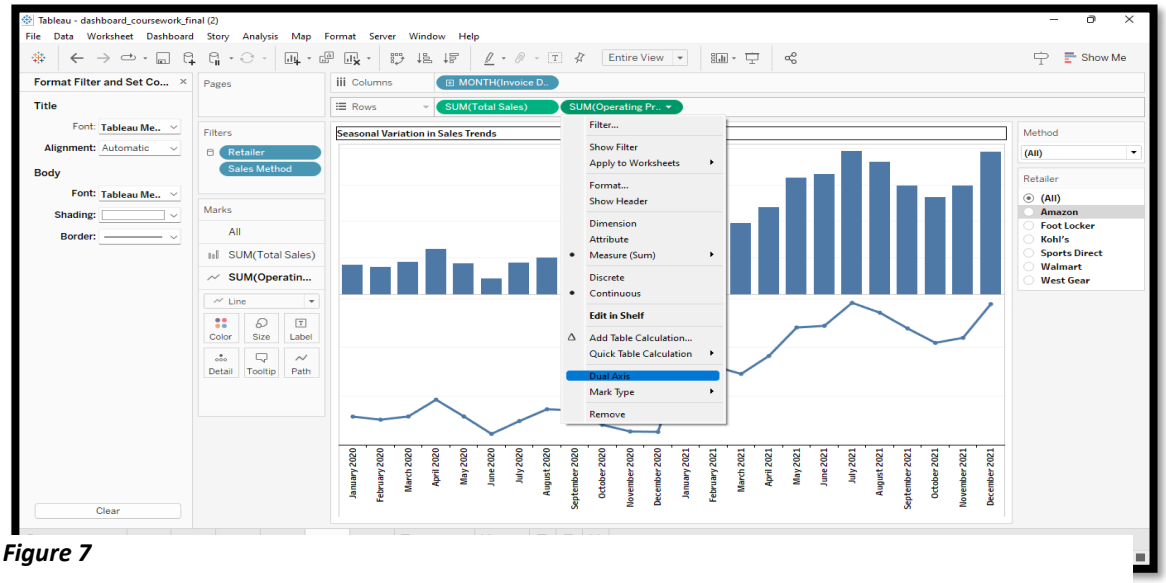


Figure 7

g) Sheet 6: Operating profit and Total sales (Figure 8)
Selecting total sales and operating profit using ctrl. Choosing the scatterplot option. Choosing filled dot options from shapes. Changing the color to red. Adding labels.

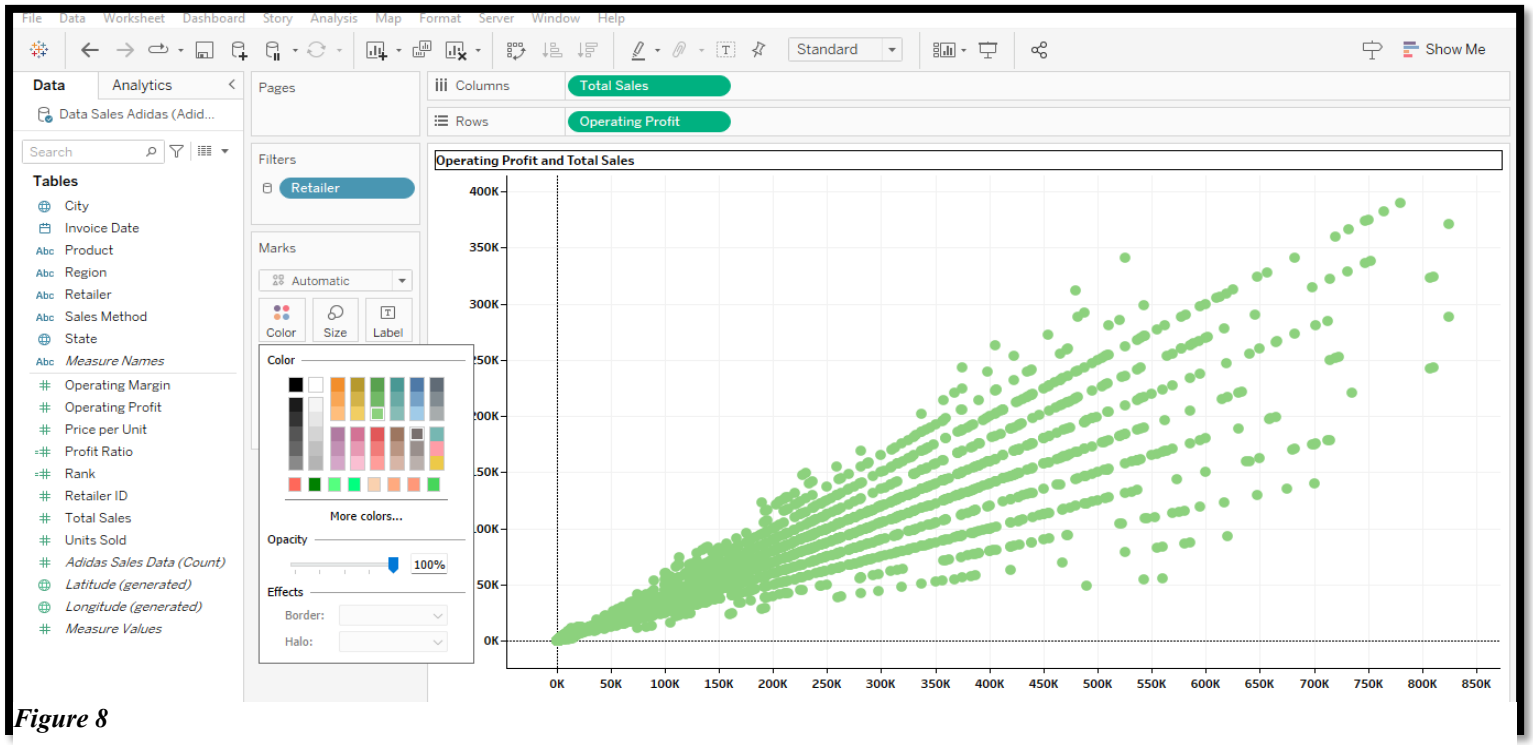


Figure 8

h) Making Dashboard: (figure 9)

Selecting dashboard and dropping sheets into vertical and horizontal (objects) Adjusting the size of visuals. Changing the size of the dashboard to automatic and view to standard. Global filter: adding retail and sales method as filters. Selecting activity from the dashboard. Clicking on filters. Choose all options in source and target. Formatting and adding shade to filter. Selecting multiple lists drop-down with (all) included. Adding URL for interaction: Adding a title to the dashboard. Formatting labels. Saving the dashboard.

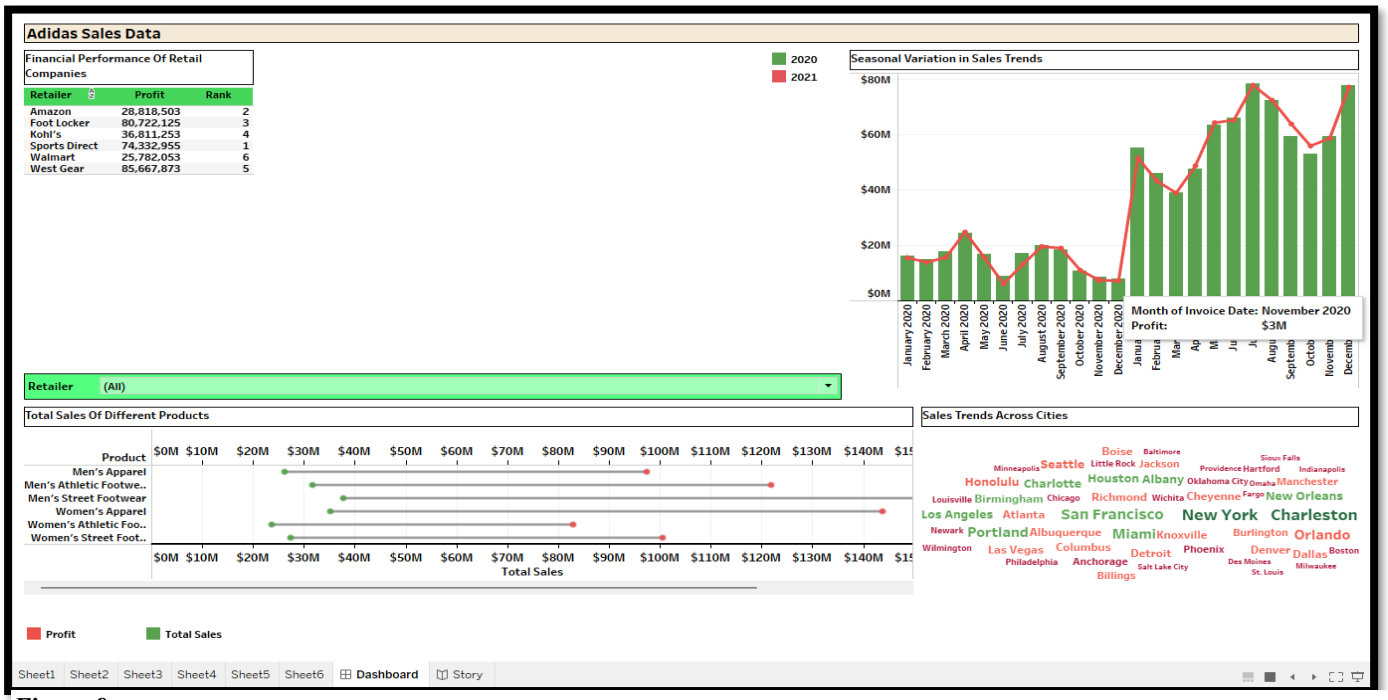


Figure 9

i) Adding Interactions (figure 10)

Selecting dashboard, selecting analysis, adding two interactions URL for the country map Filters for retailers ie keeping sheet 1 containing information about the financial performance of the Retail companies as the source file and the rest of other sheets as the target file.

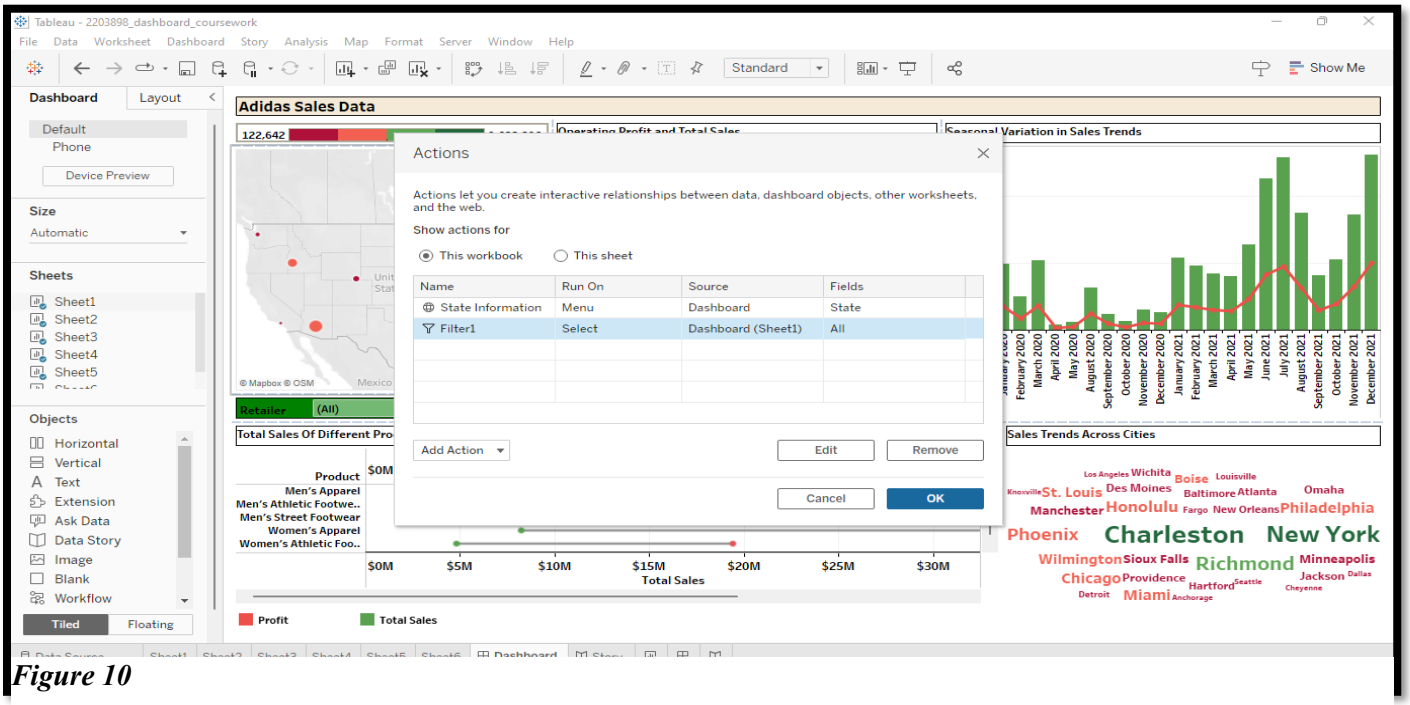
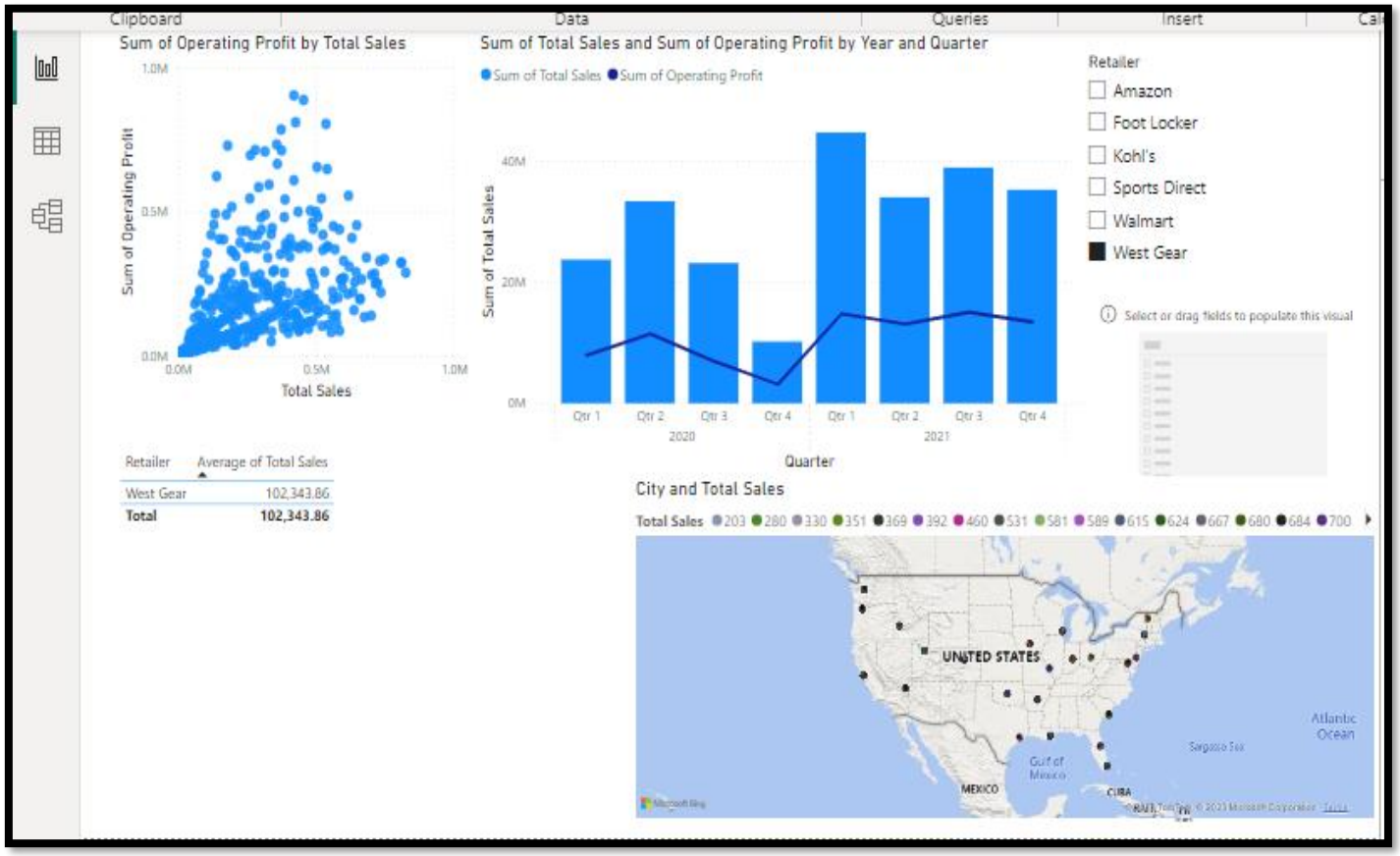


Figure 10

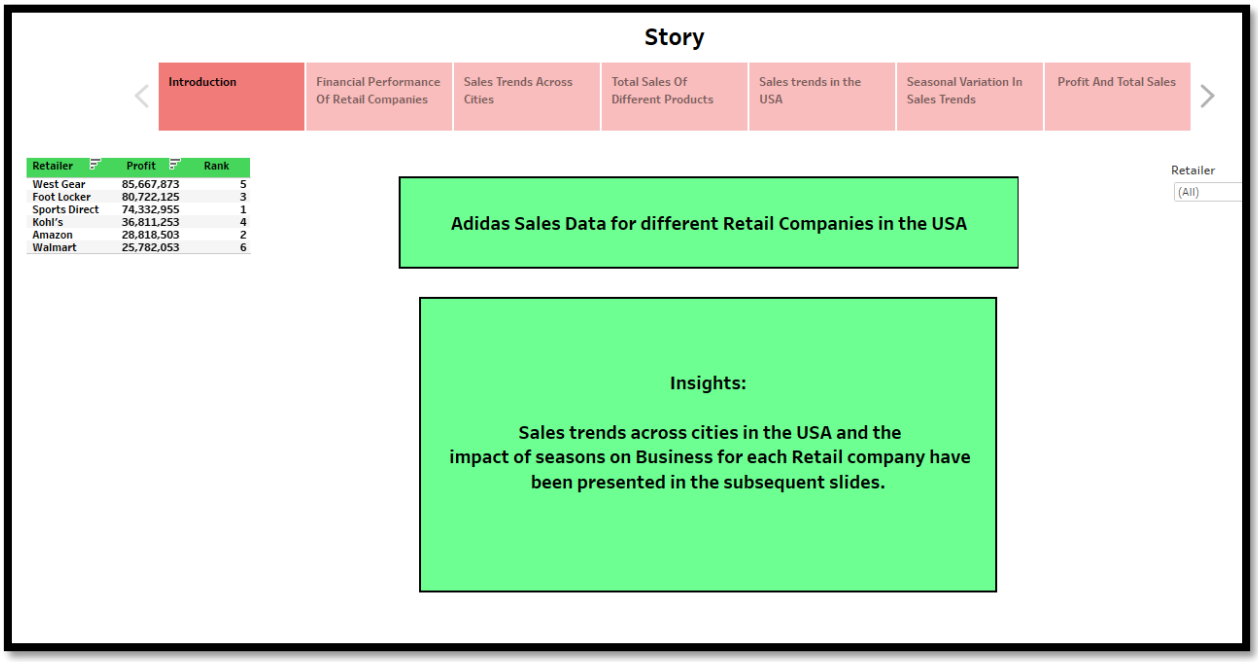
PowerBI implementation:

I tried using PowerBI to explore different Business Intelligence tools. PowerBI was better than Tableau in terms of the User Interface and processing time of the visual. It can be used even by non-programmers due to its design simplicity. PowerBI and Tableau, both had multiple options for diverse visuals. Both had options of connecting to multiple datatypes and servers eg excel, SQL, Jason, Salesforce, etc. Downloading PowerBI was much faster and easier, due to Microsoft's integrity.(Power BI Documentation) Tableau has priced versions and hence cannot be used as an open-sourced platform for data analytics. Both are intuitive and easy to use but aesthetically Tableau is much better than PowerBI. The different color-palette, formatting, shapes, labels, and tooltips option was better user designed in Tableau. Also, I could not make my second calculated field in PowerBI as that option was not there. Machine learning can be performed in both software. The full-screen view was not available in the desktop version of PowerBI. Overall, I realized that PowerBI is useful for getting quicker insights but aesthetically tableau is better.(Datacamp)



Walkthrough:

a) ***Sheet 1:*** The aim of this report is to find answers about the sales trends in the Adidas Sales Data for the period of two years i.e., from 2020 January to 2021 December. The insights from this data will be presented to the manager of the retail shop Sports Direct. Hence, finding sales trends in different regions for each retailer is of competitive advantage for the persona.

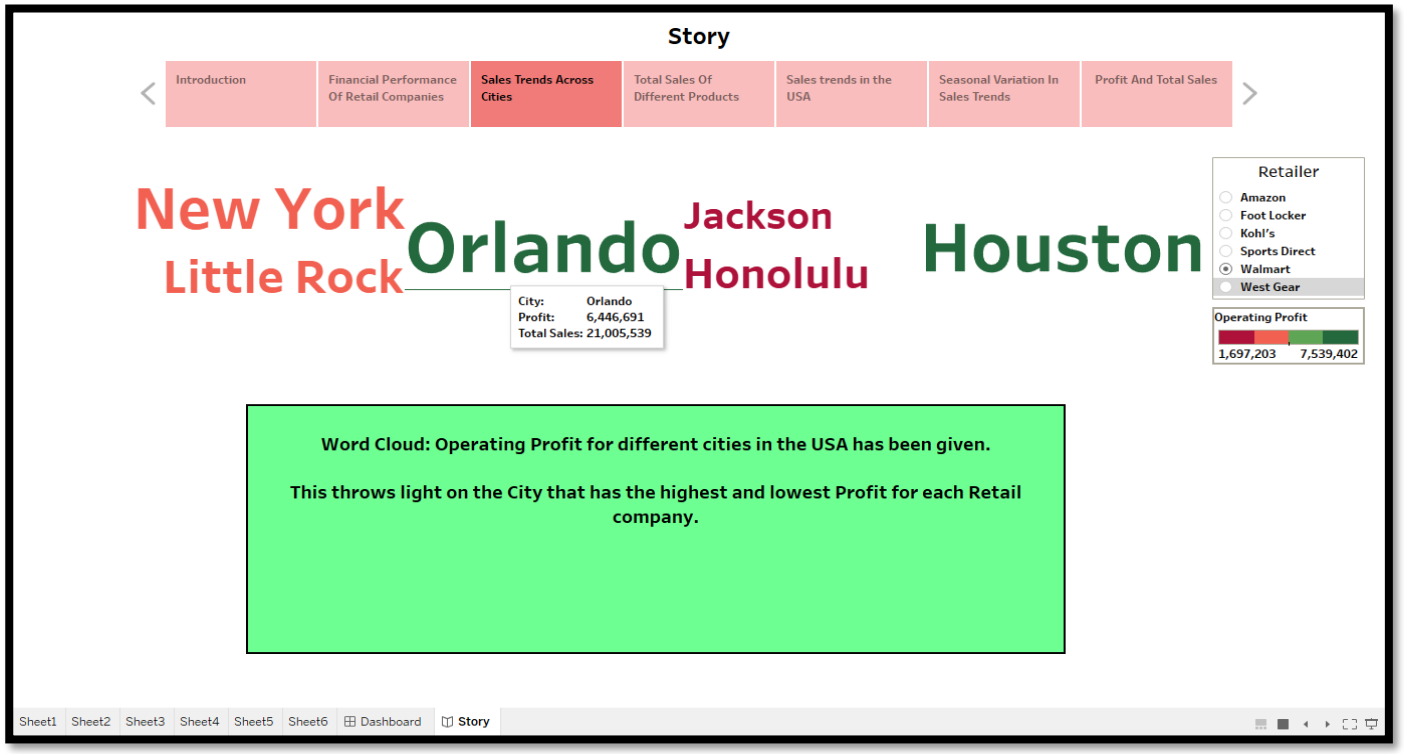


b) ***Sheet 2:*** Text Table: Financial Performance of Retail Companies

It can be seen from the sheet below that the Sports Retailer company was ranked as number 1 among the six other retailer companies for a period of 2 years in the USA. It gives an overview of the sales trends of the retailers by showing their rank and profit ratio. It also helps to answer Q2 as it is for the period of 2 years.

c) ***Sheet 3:*** Word Cloud visual - The Sales Trends Across Cities

It helps to understand the profits of retailers in the different cities of the USA. The size feature and color encoding helps to determine the profit levels in different cities. The city with the smallest size has the smallest profit. In the given visual, Orlando has the highest profit ie 6446691 whereas Jackson City seems to have the lowest profit. This helps to answer Q1 ie sales trends among retailers across different cities in the USA. In all the sheets, Retailers has been applied as a global filter.

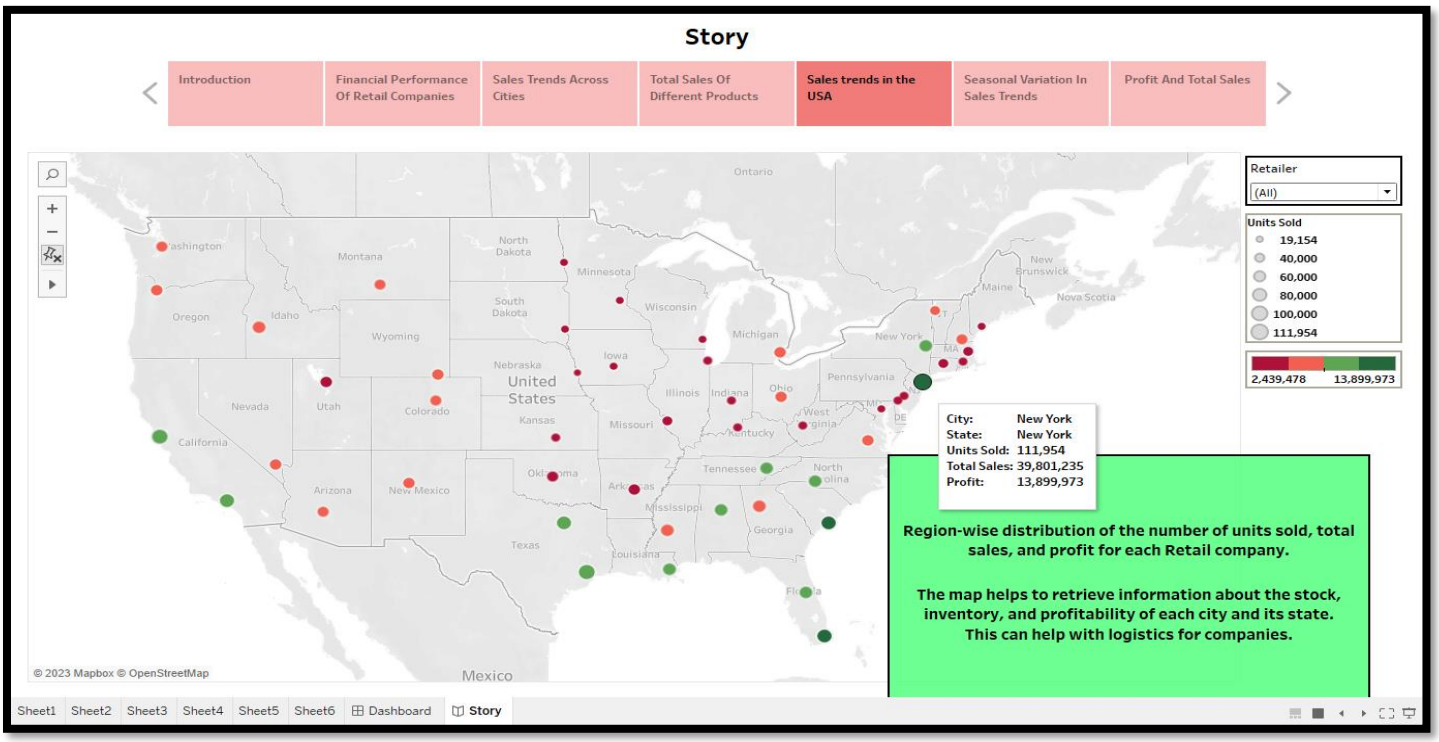


d) ***Sheet 4:*** Dumbbell Graph- Total sales of different products. performance of different retailers.

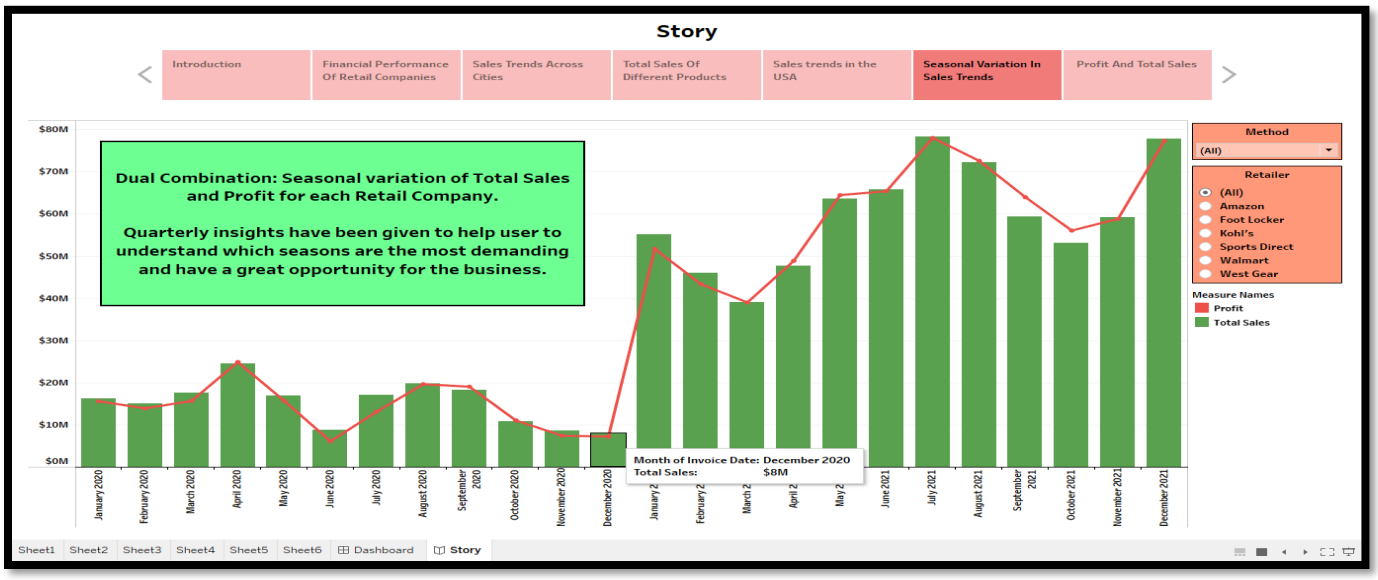
This visual provides information about products sold by Adidas Company. They are basically divided into two broader categories ie male and female. Further divided into Apparel, Athletic footwear, and Street footwear for both genders. So in total, there were a total of six products. Plotting the dumbbell graph answers Q2 ie the impact of seasonality on the business For example, the product category Men's Street footwear had the highest growth from 2020 to 2021.

e) ***Sheet 5:*** Symbol Map - Sales Trends in the USA

The Symbol Map well portrays the profit trends with the help of stepped colors and the quantities of the units sold with the help of the size of filled dots. Additionally, the URL was added as an interactive tool for helping the user to receive moe information about the states in the USA. The Wikipedia link has been shared for the same. This map helps to answer Q1 for eg, the only darkest green dot in the entire map is NewYork indicating that New York contributes to the maximum sales and profits for Adidas company. Also, the size of the fitted dot bubble is the largest for NewYork implying it has the largest inventory of Adidas goods. This insight into the units of quantities sold helps to maintain the logistics of the company



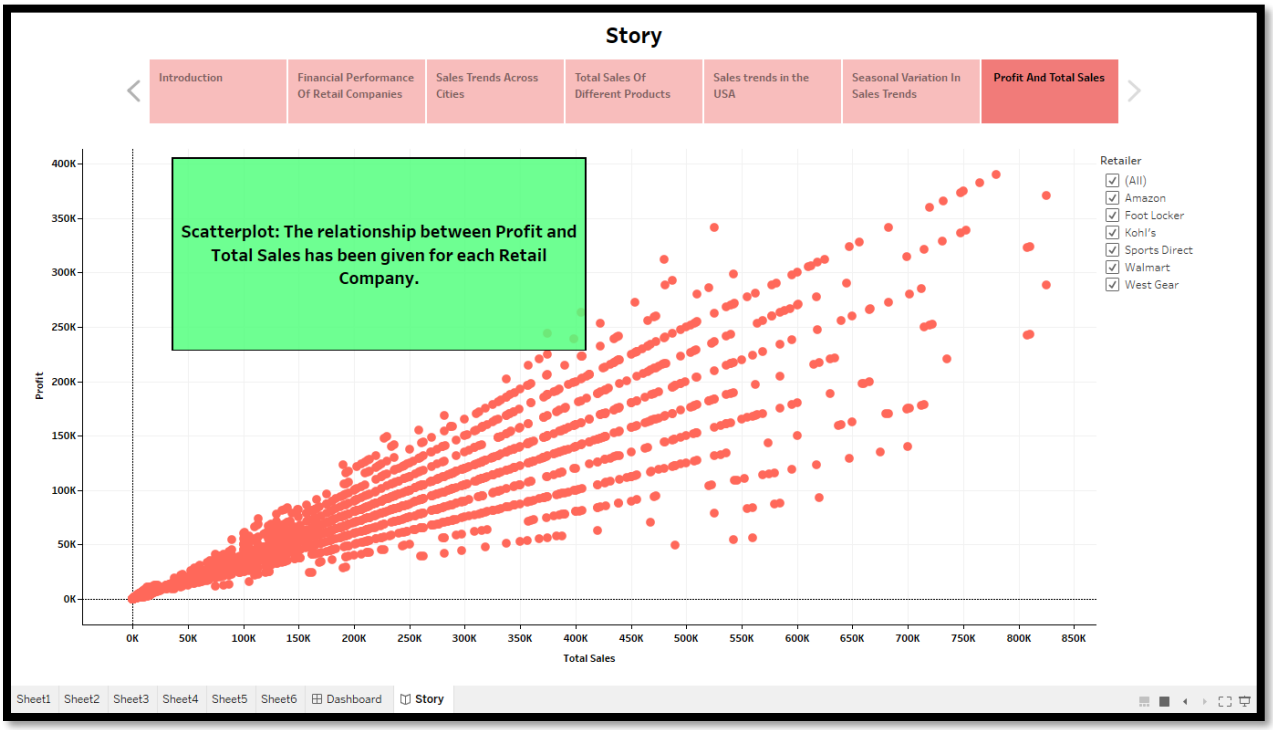
f) ***Sheet 6*:** Dual combination – Seasonal variation of Total Sales and Profit for each retail companyThe dual axis bar and line plot was used to analyze the seasonal i.e. quarterly sales for retail companies. For all the retail companies, the minimum sales



and profits for Adidas products were from October 2020 to December 2020, which probably indicates the lockdown due to COVID-19. This visual helps to answer Q2 by providing the effect of seasons on the business performance of each retailer. Retail companies are used as a global filter. On the contrary, the maximum sale and profit were on July and December 2021.

g) ***Sheet 7:*** Scatterplot – Profit and Total Sales

A scatterplot helps to predict the expected levels of profit earnings from the total sales amount. The plot shows a positive correlation. However, there is homogenous clustering of data points for the low sales amount giving fewer returns. The linear regression machine learning algorithm can be performed to derive better insights on predicting the profit based on the sales of the company. This visual helps to answer Q2 where the correlation of financial parameters given in the Adidas store dataset is analyzed. There have been instances e.g. sales at $500 K giving low-profit levels ($50K) on the other hand, sales at $550K are giving a profit of $325K.The spread of data and variance is increasing with an increase in sales.



Reflective Discussion:

This coursework gave a comprehensive view of the data analytics domain in the industry. Learning Tableau taught me to visualize the data points in a more detailed and insightful way. I realized the power of visualization on the cognitive processing of the brain. Visualizations make huge, boring data interesting and beautiful. This contributed to a better understanding of the data and hence the underlying problem statement. The data analysis study on the Adidas Sales Data helped me to understand the use of visual analytics tools for business intelligence and its efficiency in quickly administering the output. Data is everywhere and there can be infinite ways to interpret the data, detect patterns and gain meaning, and knowledge. The persona for my question was a Retailer at a Sports company. I tried exploring all the business-related KPIs which would help the retailer to get the maximum insights on the different Retail companies, their financial performance, and competitive edges. Also, working with class-mates for data collection was an amazing experience. My goal was to simplify the entire dataset for better interpretability, clarity, and visual acuity. I tried learning PowerBI as well, but I found that to be aesthetically boring from the end-user perspective. Hence, I personally prefer Tableau over Power Bi for presenting reports to clients. I would like to learn about Machine learning methods in Tableau, eg linear regression. Also, I would like to learn about connecting to various external servers and exploring different types of input files. This project was a great opportunity to work as a team, learning new methods and insights about business using visual analytics.

Conclusion:

The key objective of the project was to understand the overall process of data visualization ie from identifying the research problem to creating prototypes, being clear in the requirements, and implementing and analyzing the trends. The research question focused on finding the sales trends of retail companies seasonally and region-wise.The persona was a Retailer.This problem was answered using six visuals in the Tableau dashboard.

References:

www.datacamp.com. (n.d.). Power BI vs Tableau: Which Should You Choose in 2023? [online] Available at: https://www.datacamp.com/blog/power-bi-vs-tableau-which-one-should-you-choose.

Tableau. (n.d.). How to make dumbbell charts in Tableau. [online] Available at: https://www.tableau.com/blog/how-make-dumbbell-charts-tableau-60319#:~:text=Master%20Ryan%20Sleeper.- [Accessed 24 Apr. 2023].

CS5703 Week 23 Teaching Materials (2022). Lecture 7:Analysis and Design. Available at https://brightspace.brunel.ac.uk, accessed 16 April 2022.

CS5703 Week 19 Teaching Materials (2022). Lecture 3:Representation. Available at https://brightspace.brunel.ac.uk, accessed 14 April 2022.

maggiesMSFT (n.d.). Power BI documentation - Power BI. [online] learn.microsoft.com. Available at: https://learn.microsoft.com/en-us/power-bi/.