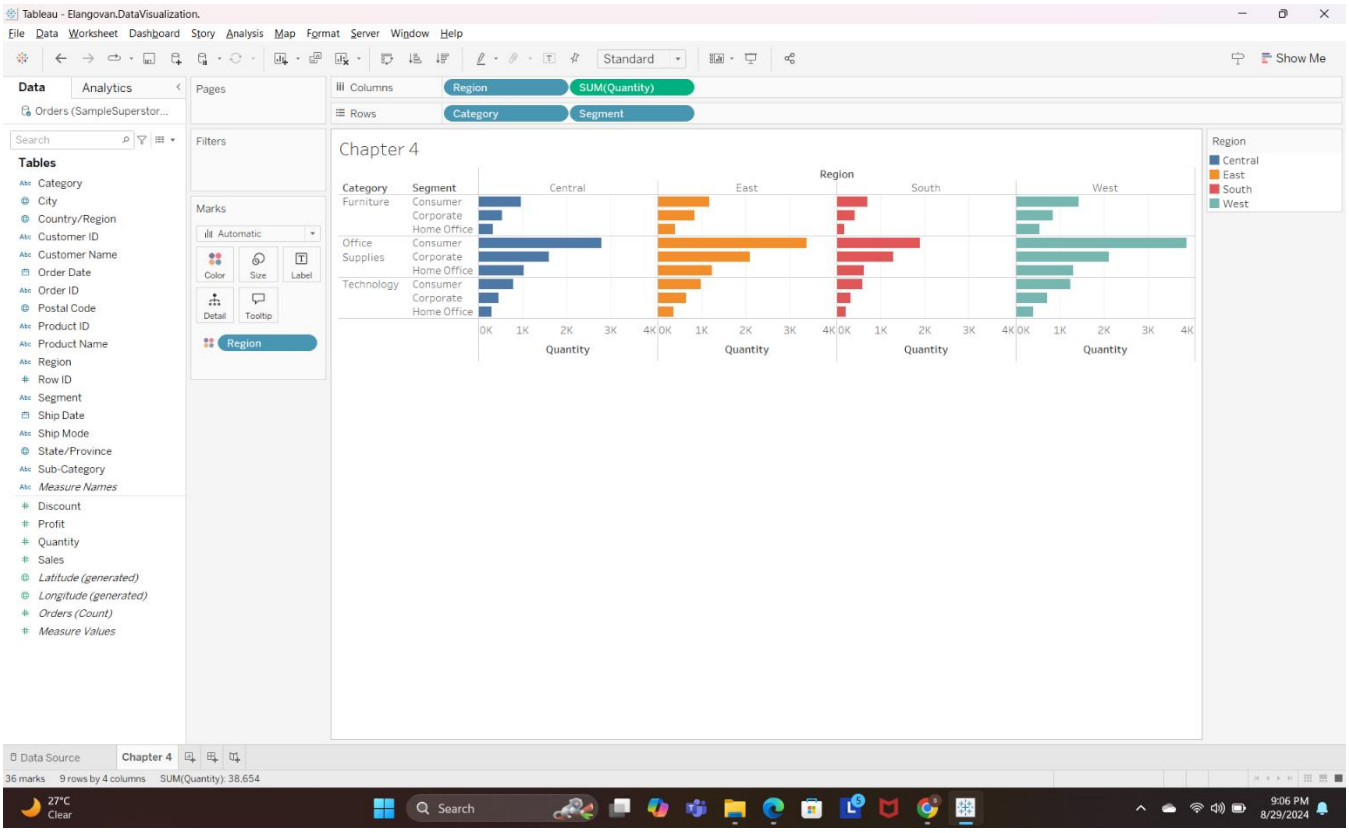
Module 2: Assignment 1 - Data Exploration & Visualization

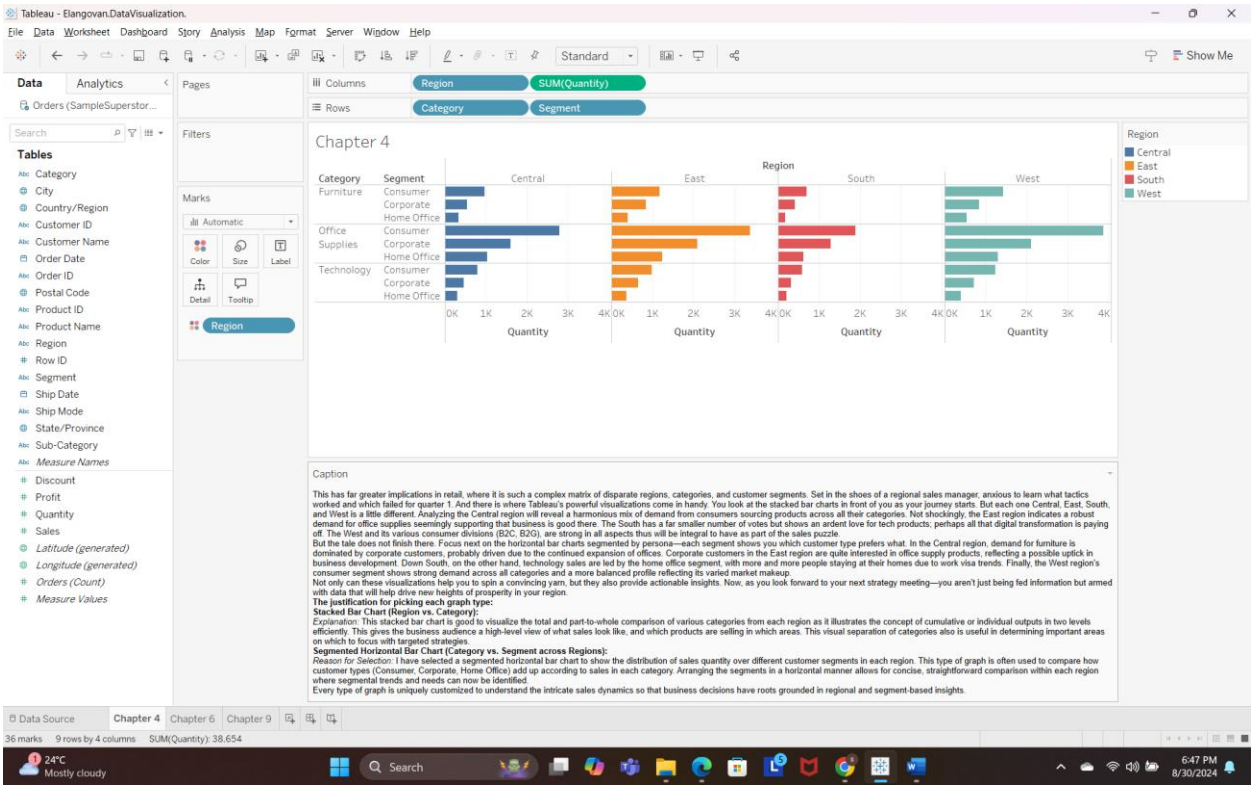
CHAPTER 4:

We can find the data in the Data Source worksheet.

Simple Visualization is Performed Step-by-step as in Chapter 4 (Khan 2016).



Caption of Chapter 4 Worksheet includes the description of graph for business audience and the rationale for choosing the specific graph and Worksheet renamed as Chapter 4.



Caption:

This has far greater implications in retail, where it is such a complex matrix of disparate regions, categories, and customer segments. Set in the shoes of a regional sales manager, anxious to learn what tactics worked and which failed for quarter 1. And there is where Tableau's powerful visualizations come in handy. You look at the stacked bar charts in front of you as your journey starts. But each one Central, East, South, and West is a little different. Analyzing the Central region will reveal a harmonious mix of demand from consumers sourcing products across all their categories. Not shockingly, the East region indicates a robust demand for office supplies seemingly supporting that business is good there. The South has a far smaller number of votes but shows an ardent love for tech products; perhaps all that digital transformation is paying off. The West and its various consumer divisions (B2C, B2G), are strong in all aspects thus will be integral to have as part of the sales puzzle.

But the tale does not finish there. Focus next on the horizontal bar charts segmented by persona—each segment shows you which customer type prefers what. In the Central region, demand for furniture is dominated by corporate customers, probably driven due to the continued expansion of offices. Corporate customers in the East region are quite interested in office supply products, reflecting a possible uptick in business development. Down South, on the other hand,

technology sales are led by the home office segment, with more and more people staying at their homes due to work visa trends. Finally, the West region's consumer segment shows strong demand across all categories and a more balanced profile reflecting its varied market makeup.

Not only can these visualizations help you to spin a convincing yarn, but they also provide actionable insights. Now, as you look forward to your next strategy meeting—you aren't just being fed information but armed with data that will help drive new heights of prosperity in your region.

The justification for picking each graph type:

Stacked Bar Chart (Region vs. Category):

Explanation: This stacked bar chart is good to visualize the total and part-to-whole comparison of various categories from each region as it illustrates the concept of cumulative or individual outputs in two levels efficiently. This gives the business audience a high-level view of what sales look like, and which products are selling in which areas. This visual separation of categories also is useful in determining important areas on which to focus with targeted strategies.

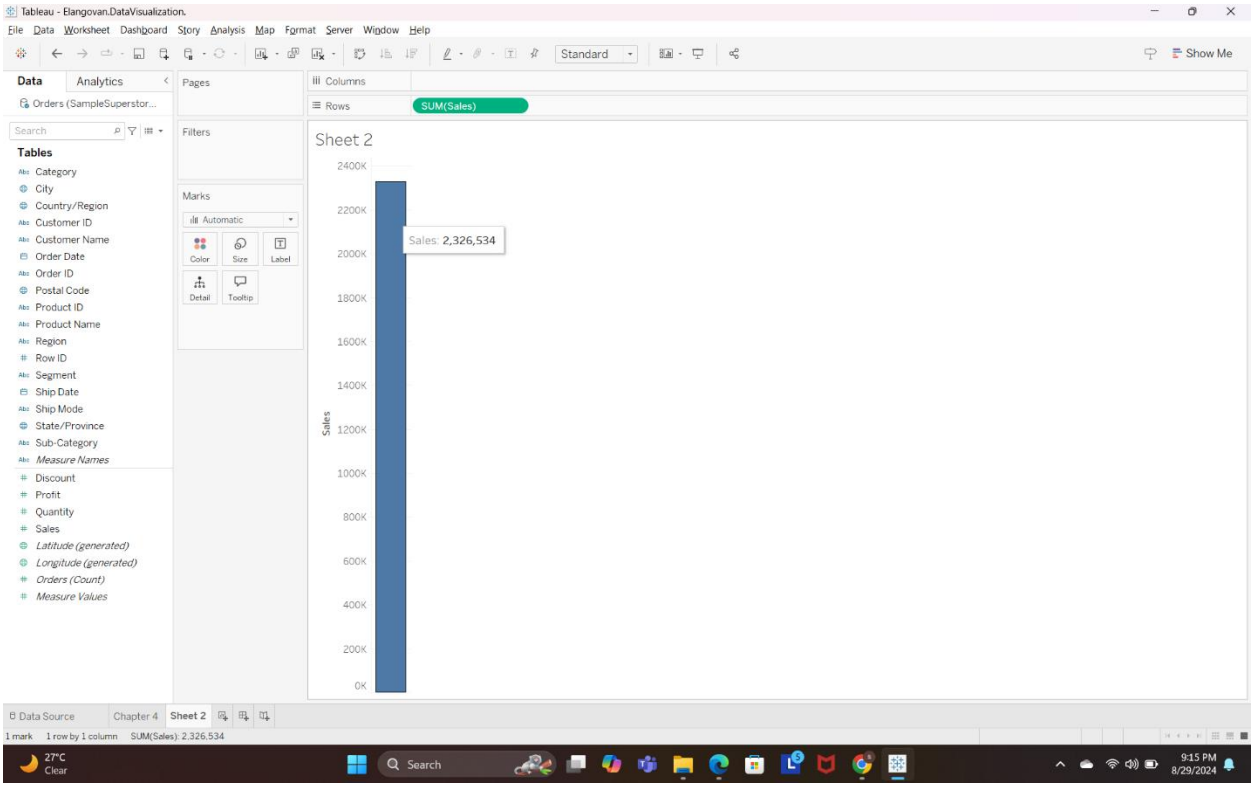
Segmented Horizontal Bar Chart (Category vs. Segment across Regions):

Reason for Selection: I have selected a segmented horizontal bar chart to show the distribution of sales quantity over different customer segments in each region. This type of graph is often used to compare how customer types (Consumer, Corporate, Home Office) add up according to sales in each category. Arranging the segments in a horizontal manner allows for concise, straightforward comparison within each region where segmental trends and needs can now be identified.

Every type of graph is uniquely customized to understand the intricate sales dynamics so that business decisions have roots grounded in regional and segment-based insights.

CHAPTER 6:

Basic Analysis of Data performed Step-by-step as in Chapter 6 (Khan 2016).



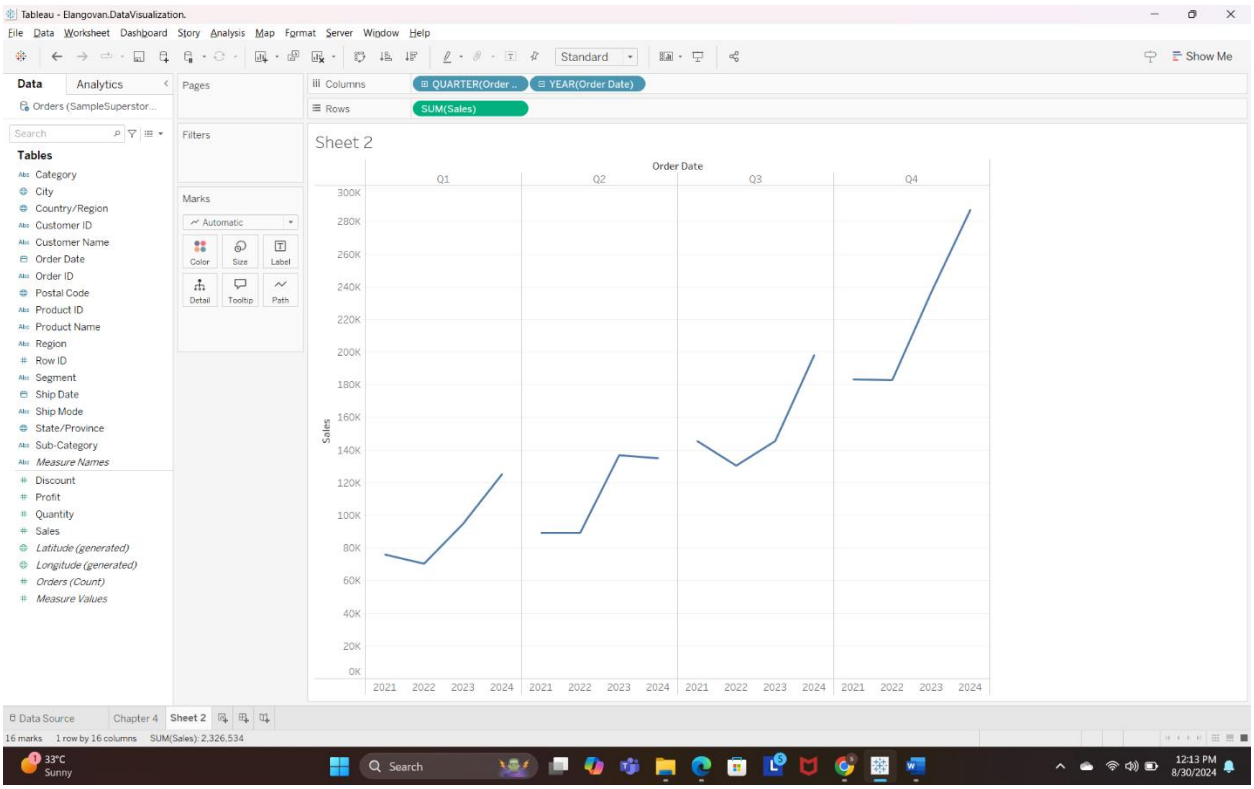


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Path
Columns
QUARTER(Order ..
YEAR(Order Date)
Rows
SUM(Sales)
Sheet 2
Order Date
Q1
Q2
Q3
Q4
Sales
300K
280K
260K
240K
220K
200K
180K
160K
140K
120K
100K
80K
60K
40K
20K
0K
2021
2022
2023
2024
Data Source
Chapter 4
Sheet 2
16 marks
1 row by 16 columns
SUM(Sales): 2,326,534
33°C
Sunny
Search
12:13 PM
8/30/2024

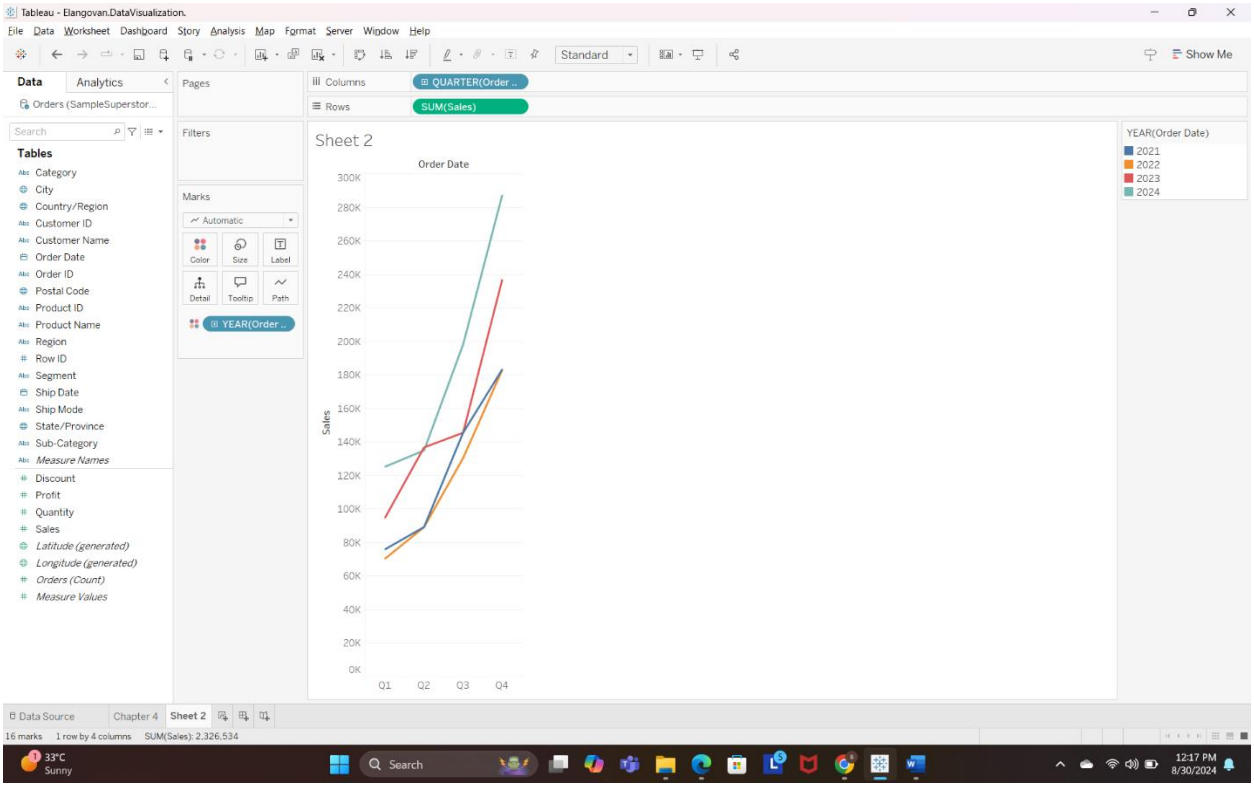


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Path
YEAR(Order ...
Columns
QUARTER(Order ..
Rows
SUM(Sales)
Sheet 2
Order Date
Q1
Q2
Q3
Q4
Sales
300K
280K
260K
240K
220K
200K
180K
160K
140K
120K
100K
80K
60K
40K
20K
0K
YEAR(Order Date)
2021
2022
2023
2024
Data Source
Chapter 4
Sheet 2
16 marks
1 row by 4 columns
SUM(Sales): 2,326,534
33°C
Sunny
Search
12:17 PM
8/30/2024

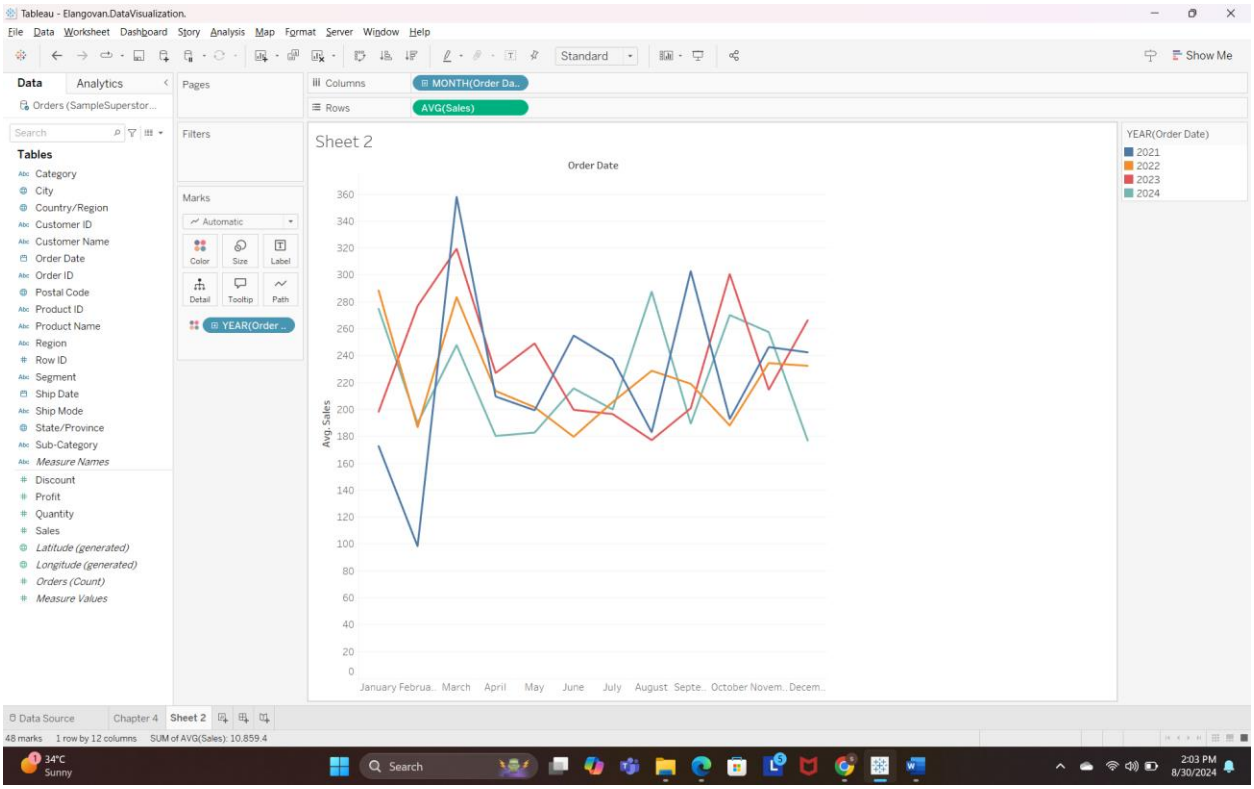


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Path
YEAR(Order ...
Columns
MONTH(Order Da..
Rows
AVG(Sales)
Sheet 2
Order Date
Avg. Sales
360
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0
January Februa.. March April May June July August Septe.. October Novem.. Decem..
YEAR(Order Date)
2021
2022
2023
2024
Data Source
Chapter 4
Sheet 2
48 marks
1 row by 12 columns
SUM of AVG(Sales): 10,859.4
34°C
Sunny
Search
2:03 PM
8/30/2024

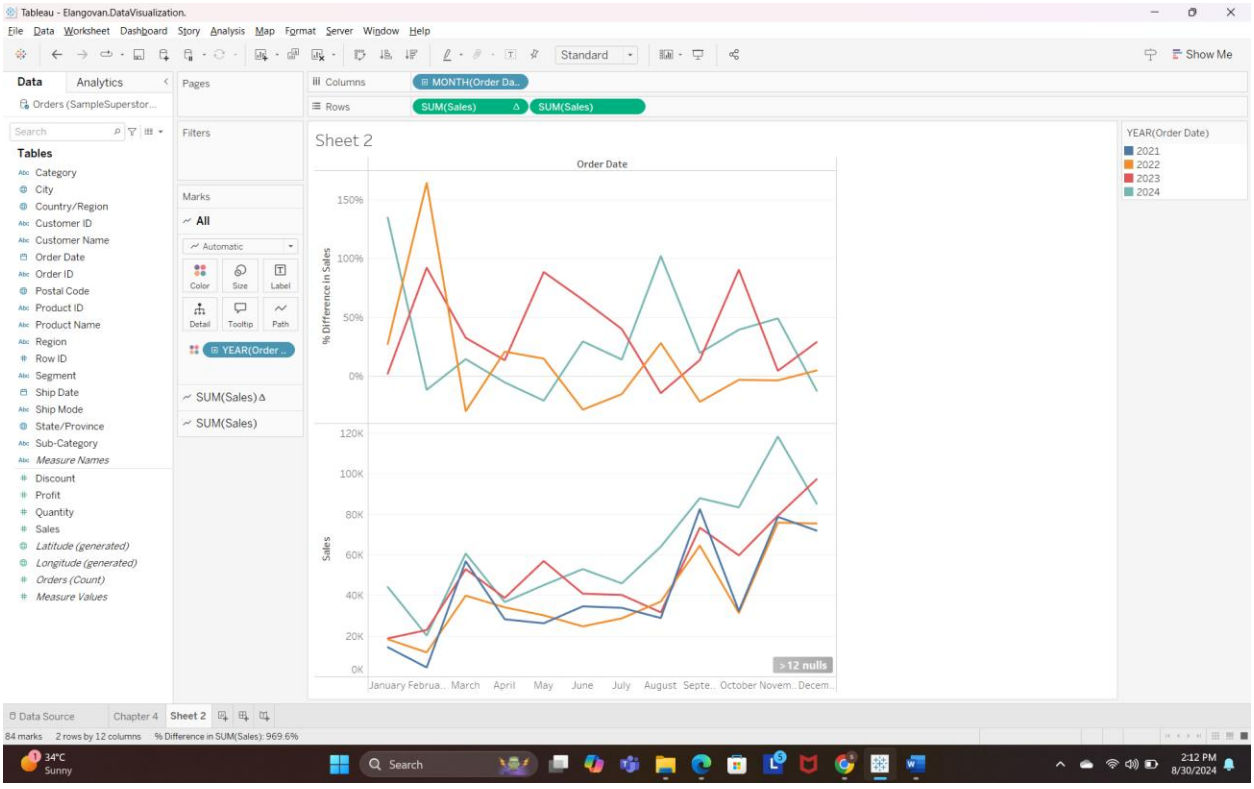
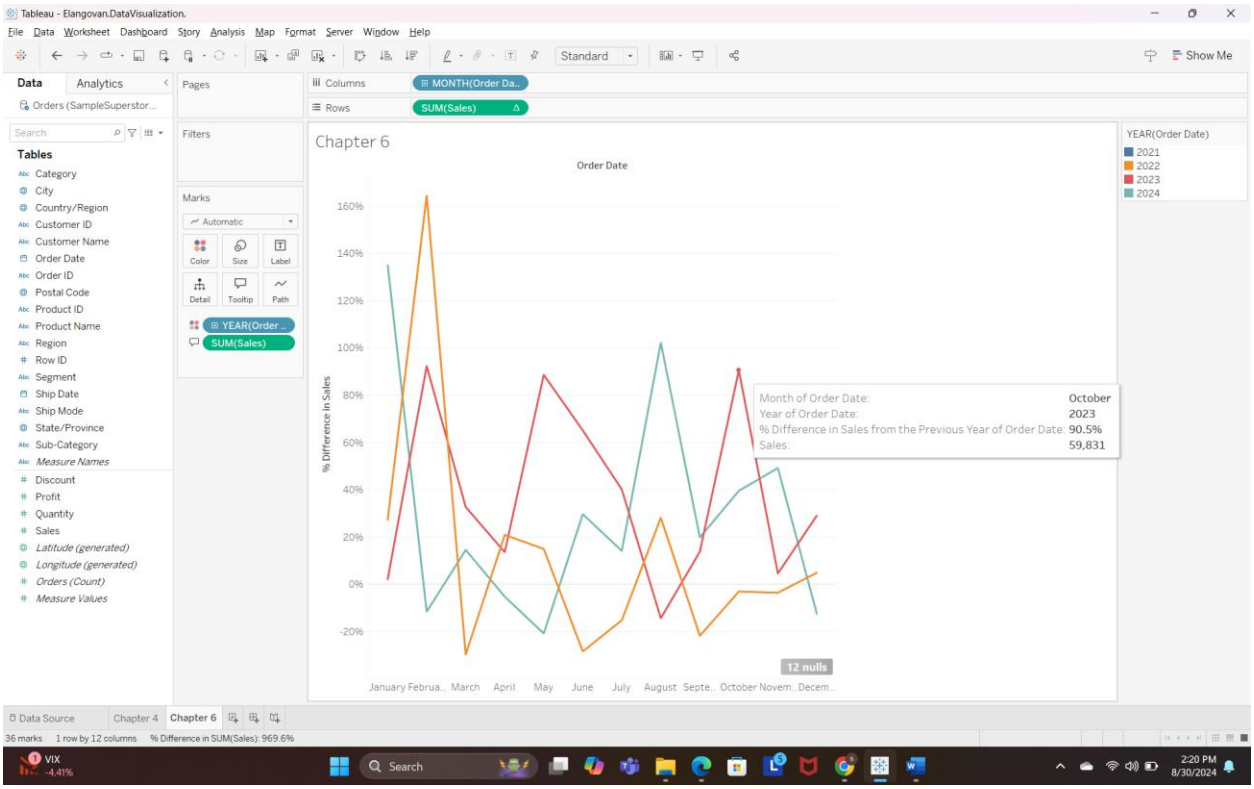
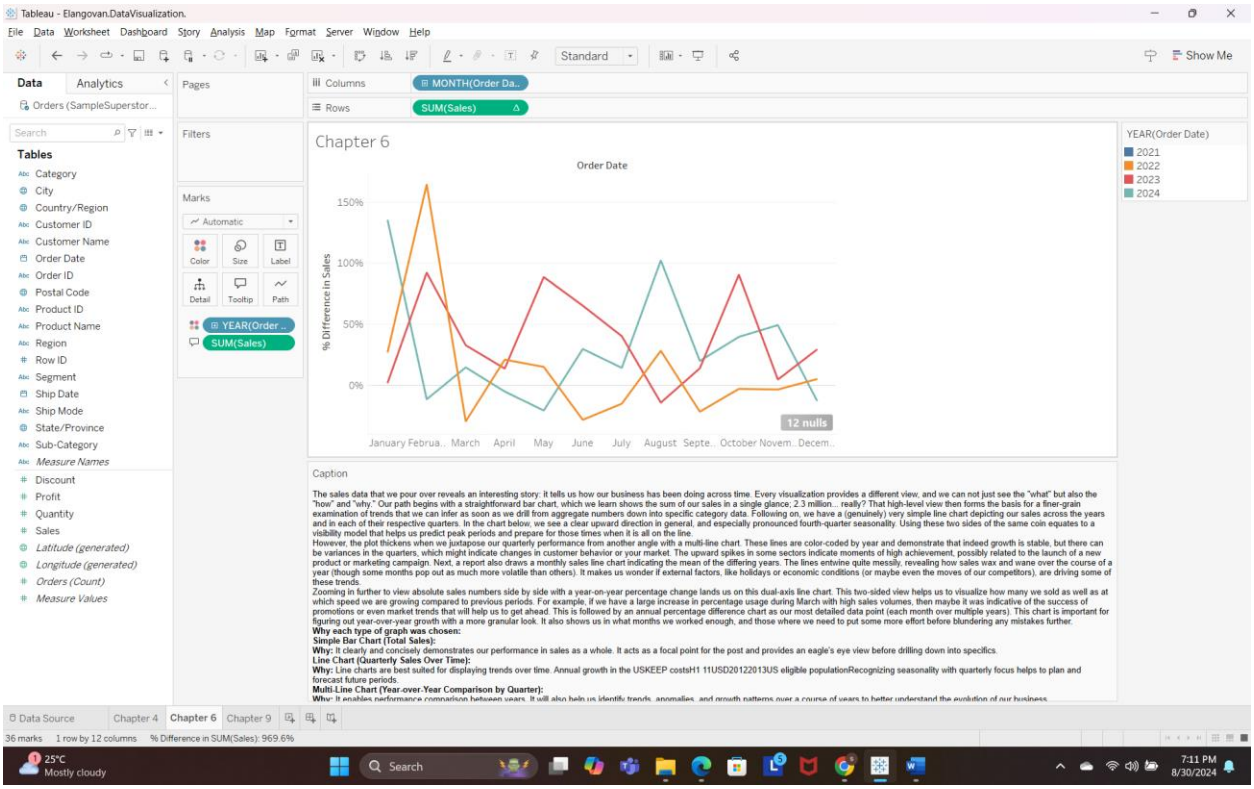


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
All
Automatic
Color
Size
Label
Detail
Tooltip
Path
YEAR(Order ...
SUM(Sales) Δ
SUM(Sales)
Columns
MONTH(Order Da..
Rows
SUM(Sales) Δ
SUM(Sales)
Sheet 2
Order Date
% Difference in Sales
150%
100%
50%
0%
Sales
120K
100K
80K
60K
40K
20K
0K
>12 nulls
January Februa.. March April May June July August Septe.. October Novem.. Decem..
YEAR(Order Date)
2021
2022
2023
2024
Data Source
Chapter 4
Sheet 2
84 marks
2 rows by 12 columns
% Difference in SUM(Sales): 969.6%
34°C
Sunny
Search
2:12 PM
8/30/2024



- Caption of Chapter 6 Worksheet includes the description of graph for business audience and the rationale for choosing the specific graph.
- Worksheet renamed as Chapter 6.



Caption:

Our path begins with a straightforward bar chart, which we learn shows the sum of our sales in a single glance; 2.3 million... really? That high-level view then forms the basis for a finer-grain examination of trends that we can infer as soon as we drill from aggregate numbers down into specific category data. Following on, we have a (genuinely) very simple line chart depicting our sales across the years and in each of their respective quarters. In the chart below, we see a clear upward direction in general, and especially pronounced fourth-quarter seasonality. Using these two sides of the same coin equates to a visibility model that helps us predict peak periods and prepare for those times when it is all on the line.

However, the plot thickens when we juxtapose our quarterly performance from another angle with a multi-line chart. These lines are color-coded by year and demonstrate that indeed growth is stable, but there can be variances in the quarters, which might indicate changes in customer behavior or your market. The upward spikes in some sectors indicate moments of high achievement, possibly related to the launch of a new product or marketing campaign. Next, a report also draws a monthly sales line chart indicating the mean of the differing years. The lines entwine quite messily, revealing how sales wax and wane over the course of a year (though some months pop out as much more volatile than others). It makes us wonder if external factors, like holidays or economic conditions (or maybe even the moves of our competitors), are driving some of these trends.

Zooming in further to view absolute sales numbers side by side with a year-on-year percentage change lands us on this dual-axis line chart. This two-sided view helps us to visualize how many we sold as well as at which speed we are growing compared to previous periods. For example, if we have a large increase in percentage usage during March with high sales volumes, then maybe it was indicative of the success of promotions or even market trends that will help us to get ahead. This is followed by an annual percentage difference chart as our most detailed data point (each month over multiple years). This chart is important for figuring out year-over-year growth with a more granular look. It also shows us in what months we worked enough, and those where we need to put some more effort before blundering any mistakes further.

Why each type of graph was chosen:

Simple Bar Chart (Total Sales):

Why: It clearly and concisely demonstrates our performance in sales as a whole. It acts as a focal point for the post and provides an eagle's eye view before drilling down into specifics.

Line Chart (Quarterly Sales Over Time):

Why: Line charts are best suited for displaying trends over time. It Recognizing seasonality with a quarterly focus helps to plan and forecast future periods.

Multi-Line Chart (Year-over-Year Comparison by Quarter):

Why: It enables performance comparison between years. It will also help us identify trends, anomalies, and growth patterns over a course of years to better understand the evolution of our business.

Monthly Sales Line Chart:

Why: In this case, a chain of line charts is supported in place, being compared vertically to locate trends over time within years (they illustrate it along with layer-based representation), which are matched by month.

Dual-Axis Line Chart (Sales and Percentage Difference):

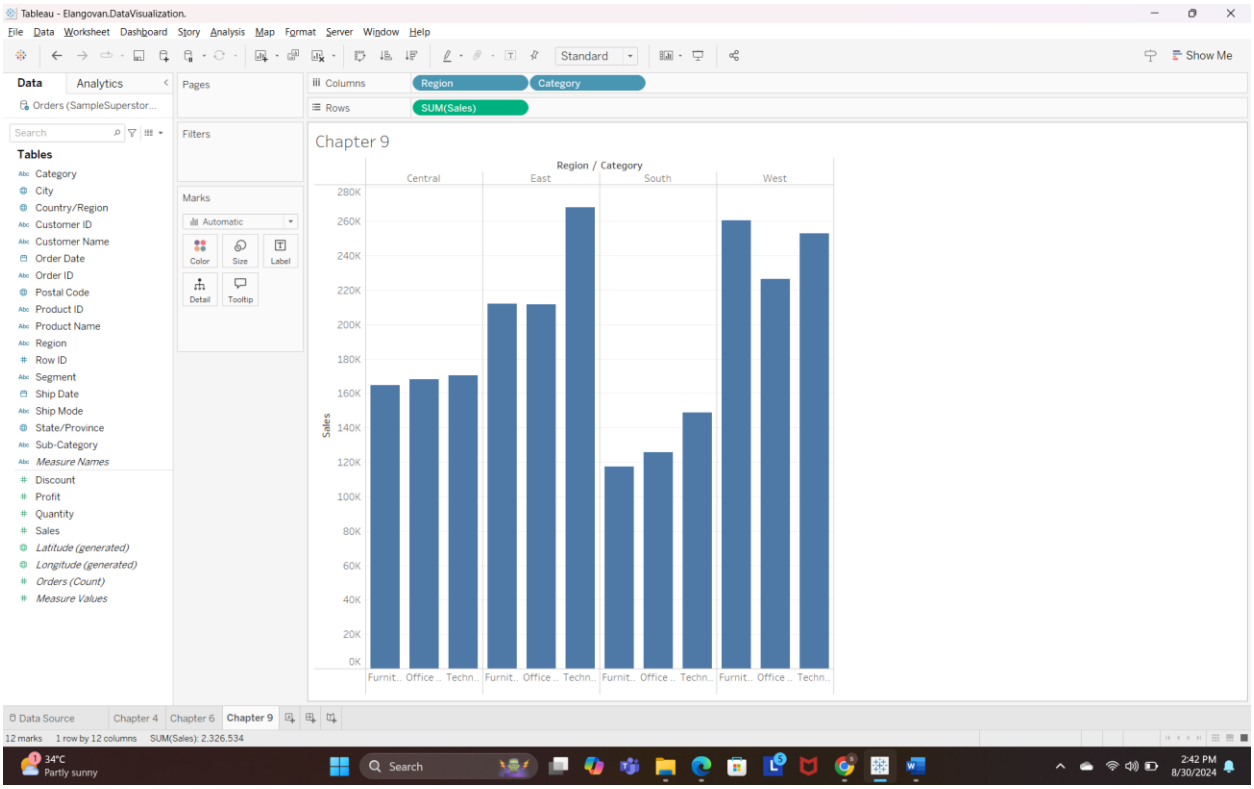
Why: The dual-axis chart is great to compare two related metrics within the same view. This gives us a view not only of the raw sales numbers, but also their rate of change, and can therefore be considered more "informative" as to how well you're doing.

Percentage Difference Line Chart (Year-over-Year Monthly Growth):

Why: This chart is meant to easily show year-over-year growth or decline. It allows us to see how well we are maintaining or improving our performance month over month from year to year.

Chapter 9 Worksheet Visualization:

Explored various Chart Types and discussed them in the Caption of Chapter 9 worksheet.



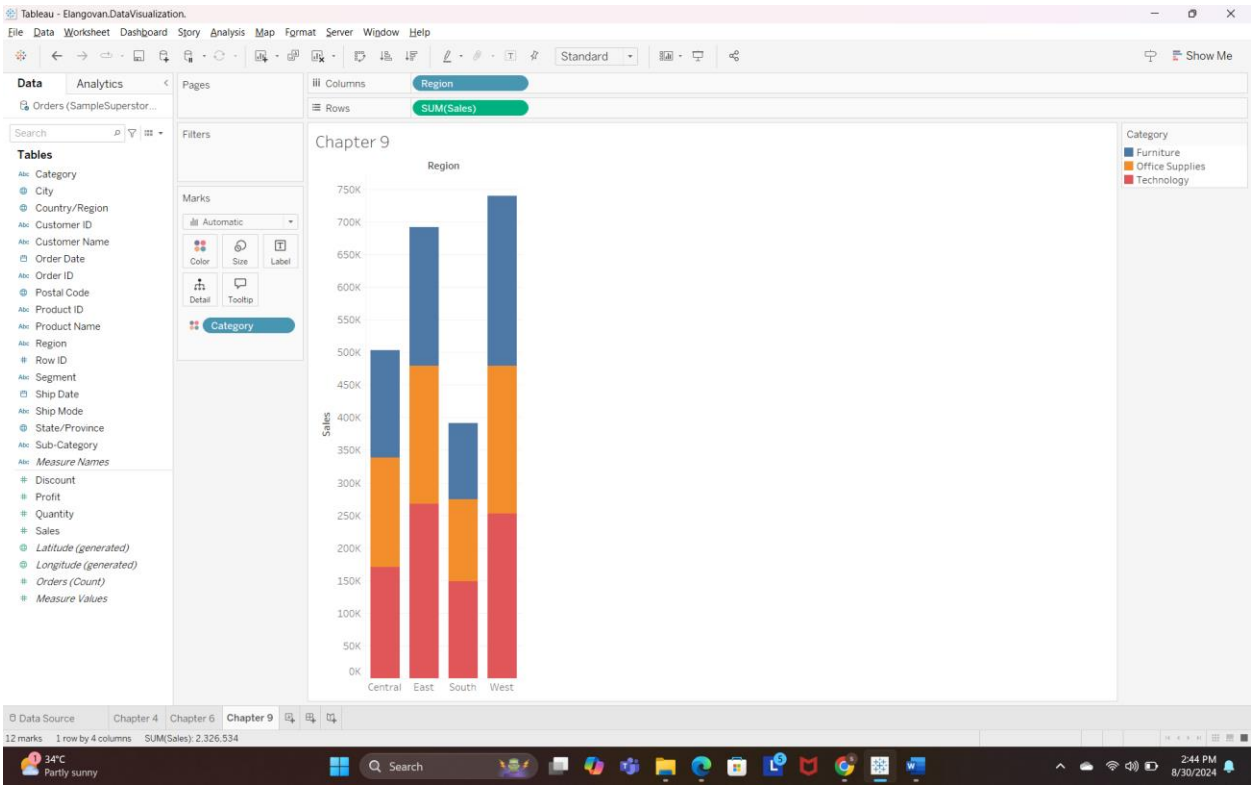


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Category
Columns
Region
Rows
SUM(Sales)
Chapter 9
Region
Sales
750K
700K
650K
600K
550K
500K
450K
400K
350K
300K
250K
200K
150K
100K
50K
0K
Central
East
South
West
Category
Furniture
Office Supplies
Technology
Data Source
Chapter 4
Chapter 6
Chapter 9
12 marks
1 row by 4 columns
SUM(Sales): 2,326,534
34°C
Partly sunny
Search
2:44 PM
8/30/2024

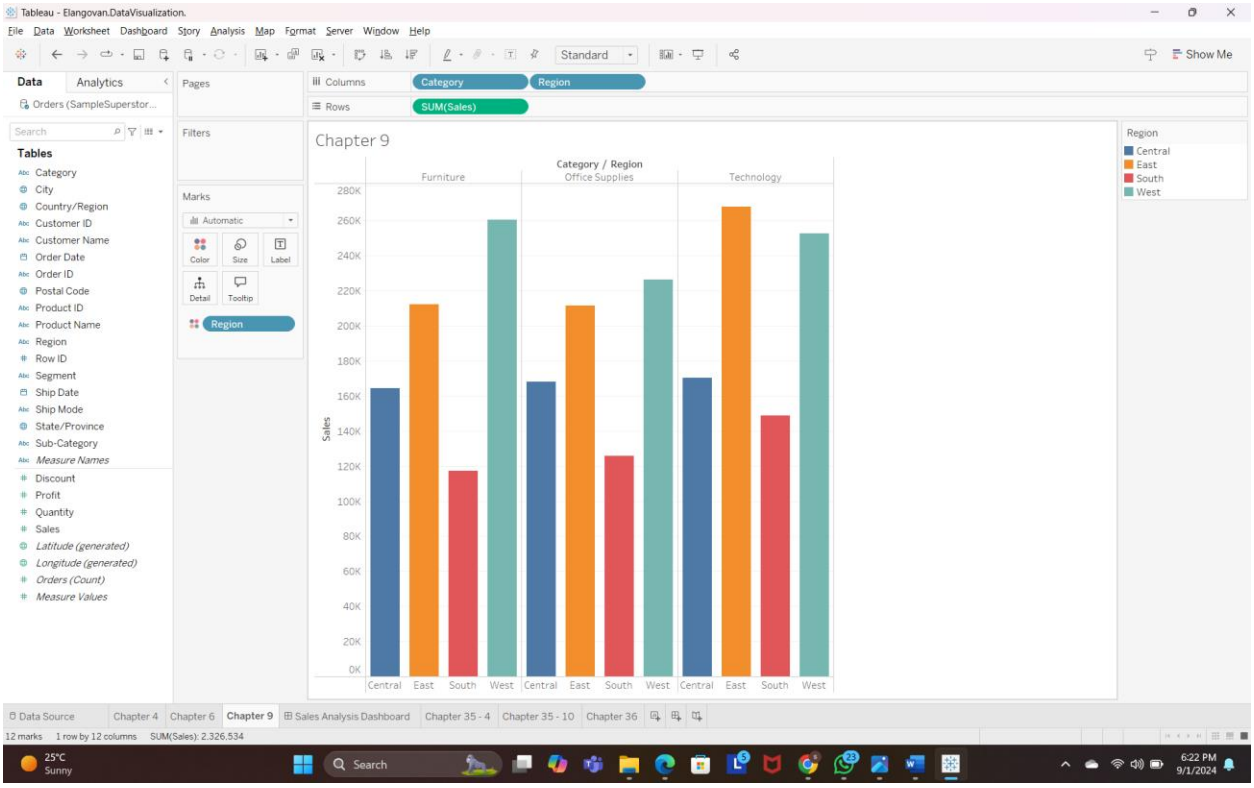


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Region
Columns
Category
Region
Rows
SUM(Sales)
Chapter 9
Category / Region
Furniture
Office Supplies
Technology
Sales
280K
260K
240K
220K
200K
180K
160K
140K
120K
100K
80K
60K
40K
20K
0K
Central
East
South
West
Region
Central
East
South
West
Data Source
Chapter 4
Chapter 6
Chapter 9
Sales Analysis Dashboard
Chapter 35 - 4
Chapter 35 - 10
Chapter 36
12 marks
1 row by 12 columns
SUM(Sales): 2,326,534
25°C
Sunny
Search
6:22 PM
9/1/2024

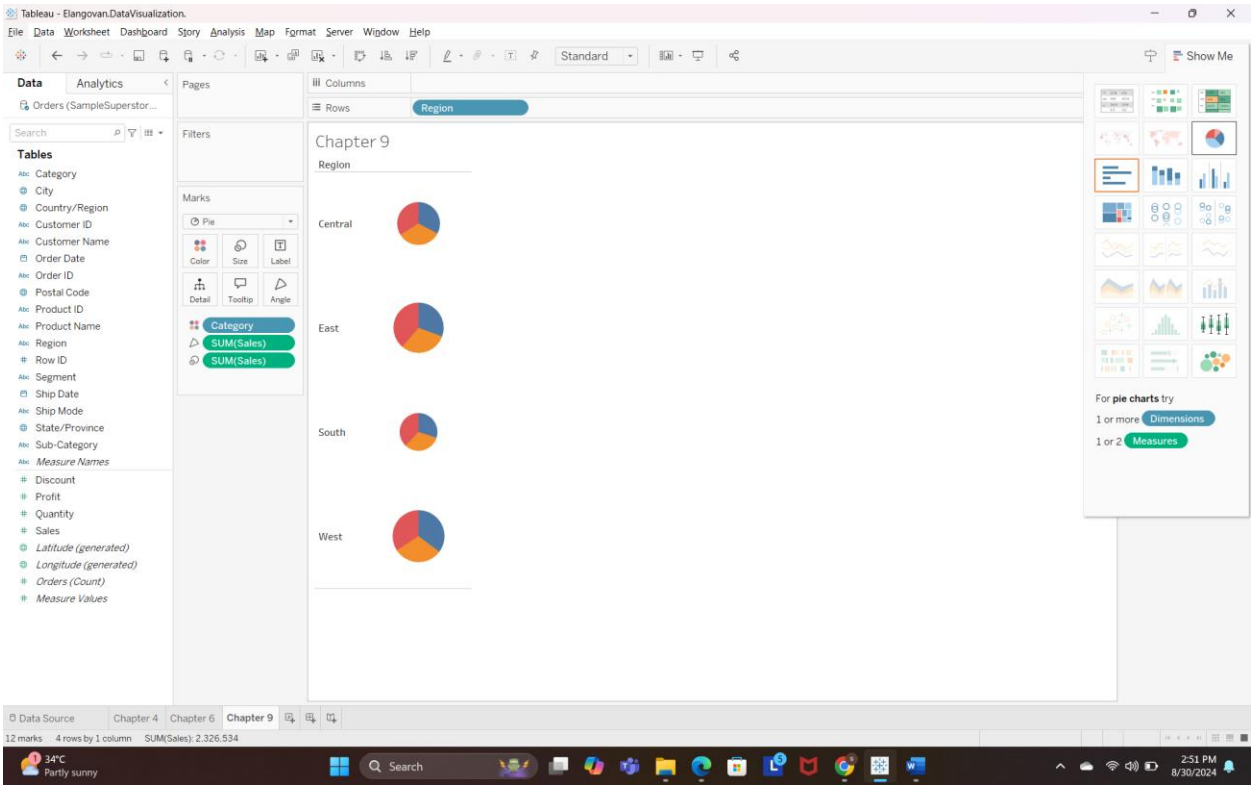


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
Pie
Color
Size
Label
Detail
Tooltip
Angle
Category
SUM(Sales)
SUM(Sales)
Columns
Rows
Region
Chapter 9
Region
Central
East
South
West
For pie charts try
1 or more Dimensions
1 or 2 Measures
Data Source
Chapter 4
Chapter 6
Chapter 9
12 marks
4 rows by 1 column
SUM(Sales): 2,326,534
34°C
Partly sunny
Search
2:51 PM
8/30/2024

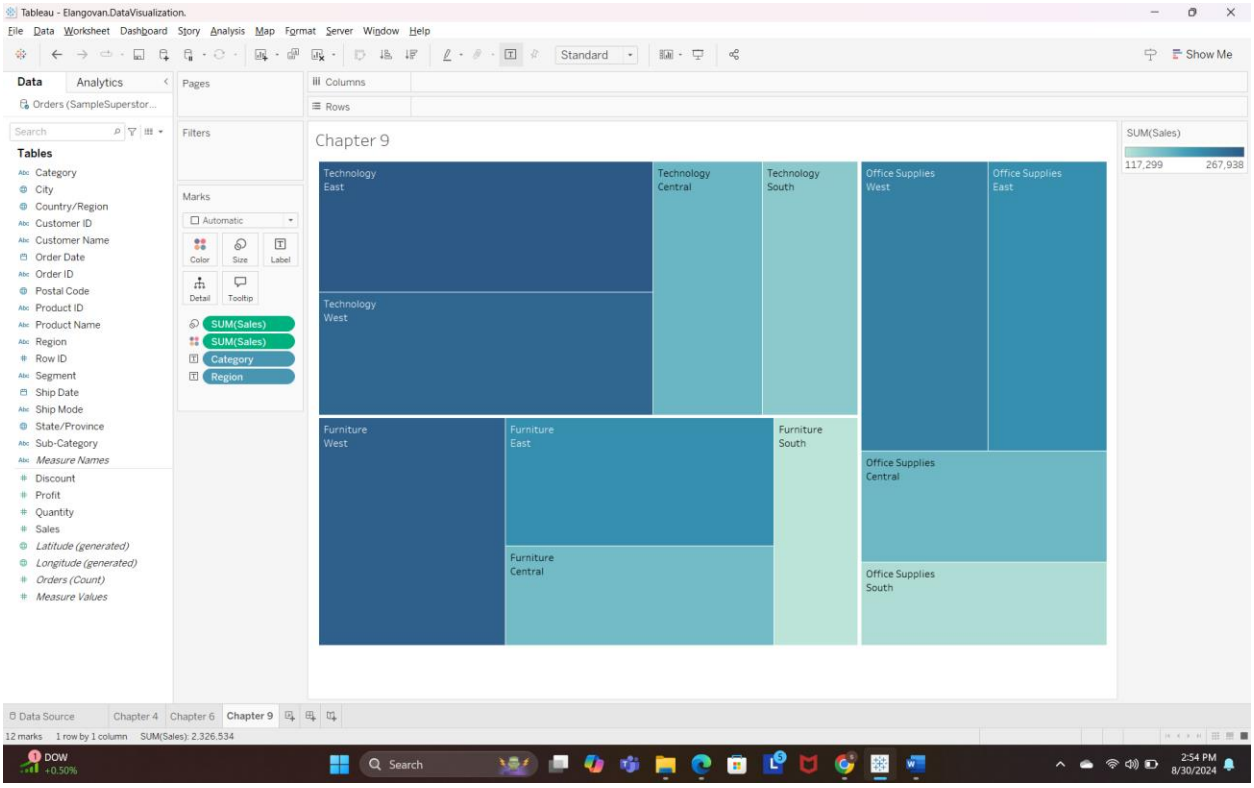
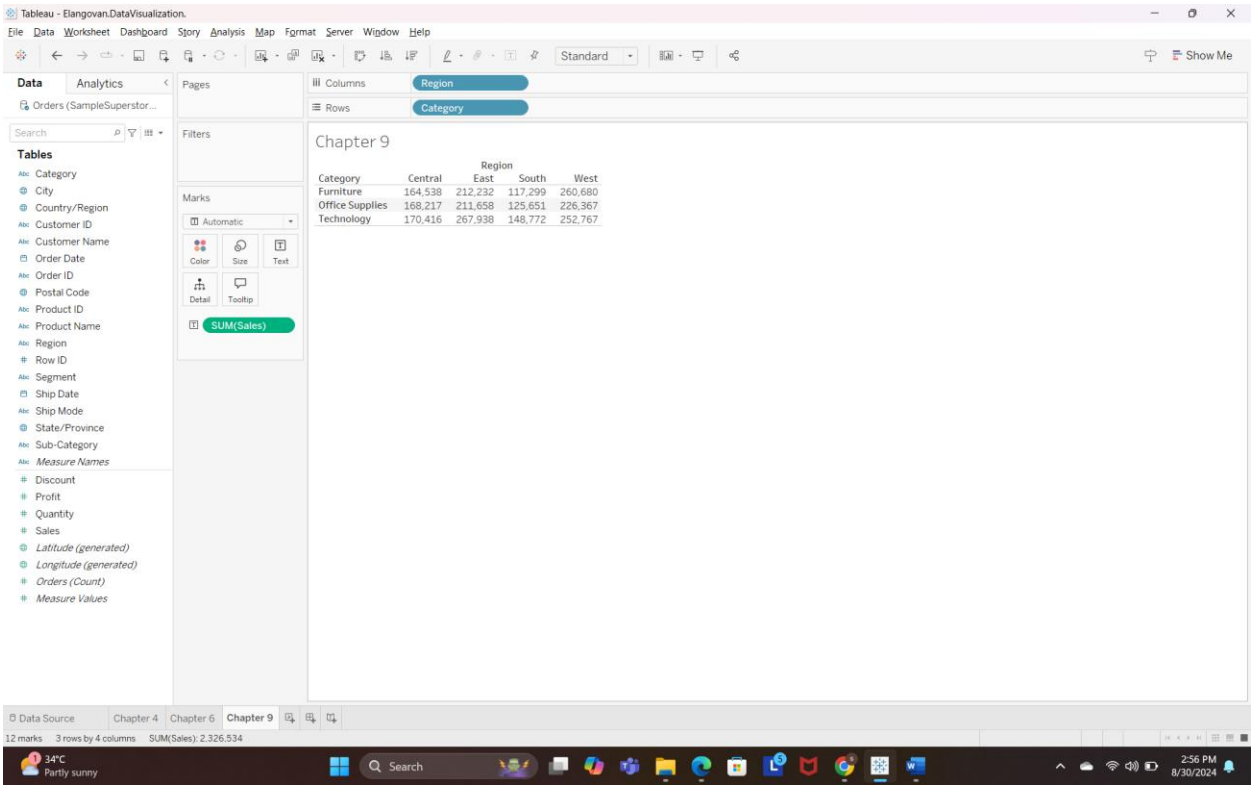
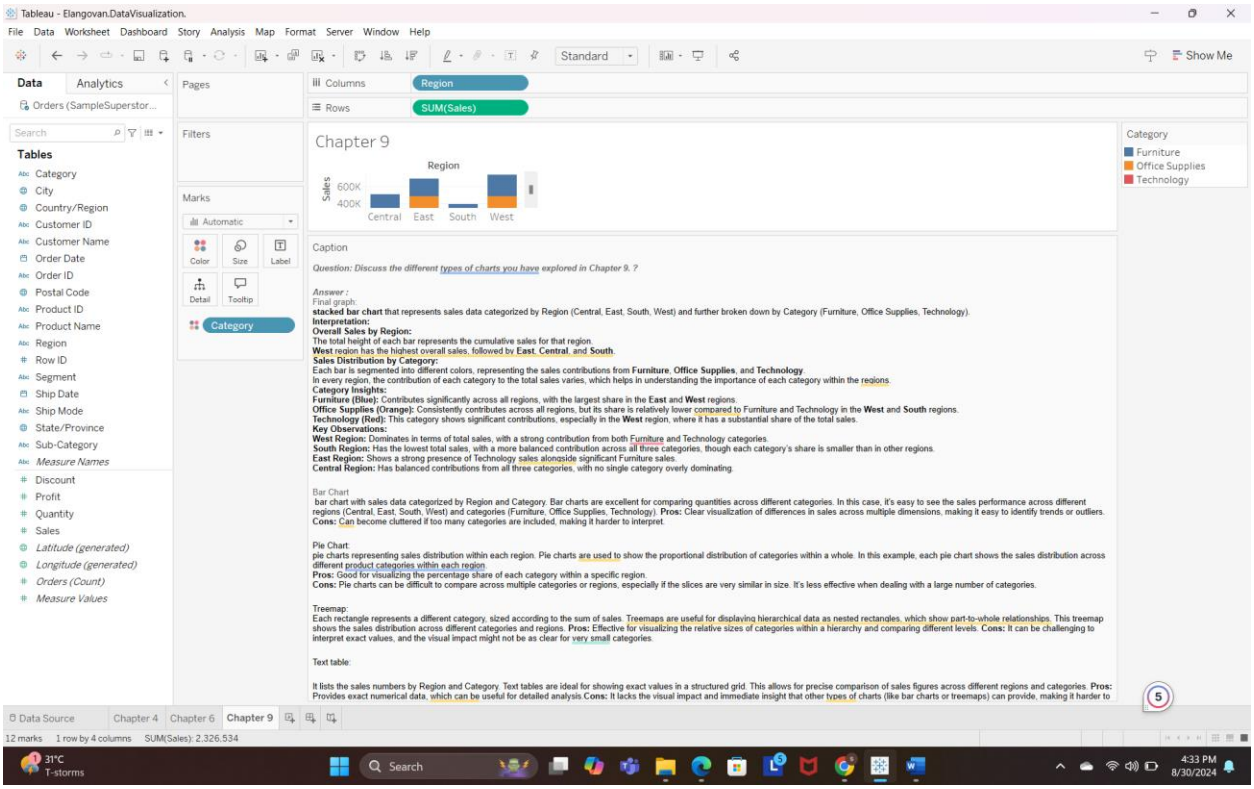


Tableau - Elangovan.DataVisualization.
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders (SampleSuperstor...
Search
Tables
Category
City
Country/Region
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State/Province
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Orders (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
SUM(Sales)
SUM(Sales)
Category
Region
Columns
Rows
Chapter 9
Technology East
Technology Central
Technology South
Technology West
Office Supplies West
Office Supplies East
Furniture West
Furniture East
Furniture South
Office Supplies Central
Furniture Central
Office Supplies South
SUM(Sales)
117,299
267,938
Data Source
Chapter 4
Chapter 6
Chapter 9
12 marks
1 row by 1 column
SUM(Sales): 2,326,534
DOW
+0.50%
Search
2:54 PM
8/30/2024



- Provided reasoning for the final chart and interpreted the information represented by final chart in the caption of Chapter 9 worksheet. Also, caption include the description of the visualization as an interesting story for a business audience.

- Worksheet renamed as Chapter 9.



Caption:

Question: Discuss the different types of charts you have explored in Chapter 9. ?

Answer :

Final graph:

stacked bar chart that represents sales data categorized by Region (Central, East, South, West) and further broken down by Category (Furniture, Office Supplies, Technology).

Interpretation:

Overall Sales by Region:

The total height of each bar represents the cumulative sales for that region.

West region has the highest overall sales, followed by **East**, **Central**, and **South**.

Sales Distribution by Category:

Each bar is segmented into different colors, representing the sales contributions from **Furniture**, **Office Supplies**, and **Technology**.

In every region, the contribution of each category to the total sales varies, which helps in understanding the importance of each category within the regions.

Category Insights:

Furniture (Blue): Contributes significantly across all regions, with the largest share in the **East** and **West** regions.

Office Supplies (Orange): Consistently contributes across all regions, but its share is relatively lower compared to Furniture and Technology in the **West** and **South** regions.

Technology (Red): This category shows significant contributions, especially in the **West** region, where it has a substantial share of the total sales.

Key Observations:

West Region: Dominates in terms of total sales, with a strong contribution from both Furniture and Technology categories.

South Region: Has the lowest total sales, with a more balanced contribution across all three categories, though each category's share is smaller than in other regions.

East Region: Shows a strong presence of Technology sales alongside significant Furniture sales.

Central Region: Has balanced contributions from all three categories, with no single category overly dominating.

chart:

The graph is a chart showing sales data by product category (Furniture, Office Supplies, Technology) and region (Central, East, South, West).

Bar Chart

bar chart with sales data categorized by Region and Category. Bar charts are excellent for comparing quantities across different categories. In this case, it's easy to see the sales performance across different regions (Central, East, South, West) and categories (Furniture, Office Supplies, Technology).

Pros: Clear visualization of differences in sales across multiple dimensions, making it easy to identify trends or outliers.

Cons: Can become cluttered if too many categories are included, making it harder to interpret.

Bar Chart:

This visualization displays sales across four regions (Central, East, South, and West) for three product categories: Furniture, Office Supplies, and Technology. **X-Axis:** Product categories,

further divided by region. **Y-Axis:** Sales totals (up to around 280,000). **Regions:** Color-coded (Central: blue, East: orange, South: red, West: green).

Key Insights: Furniture: West leads in sales, followed by East, Central, and South. **Office Supplies:** East has the highest sales, with West close behind. Central has the lowest. **Technology:** East dominates, while sales are weakest in the South.

Analysis:

The East performs well in Office Supplies and Technology, while the West is strong across all categories. The South, particularly in Technology, has lower sales.

Pie Chart:

pie charts representing sales distribution within each region. Pie charts are used to show the proportional distribution of categories within a whole. In this example, each pie chart shows the sales distribution across different product categories within each region.

Pros: Good for visualizing the percentage share of each category within a specific region.

Cons: Pie charts can be difficult to compare across multiple categories or regions, especially if the slices are very similar in size. It's less effective when dealing with a large number of categories.

Treemap:

Each rectangle represents a different category, sized according to the sum of sales. Treemaps are useful for displaying hierarchical data as nested rectangles, which show part-to-whole relationships. This treemap shows the sales distribution across different categories and regions.

Pros: Effective for visualizing the relative sizes of categories within a hierarchy and comparing different levels.

Cons: It can be challenging to interpret exact values, and the visual impact might not be as clear for very small categories.

Text table:

It lists the sales numbers by Region and Category. Text tables are ideal for showing exact values in a structured grid. This allows for precise comparison of sales figures across different regions and categories.

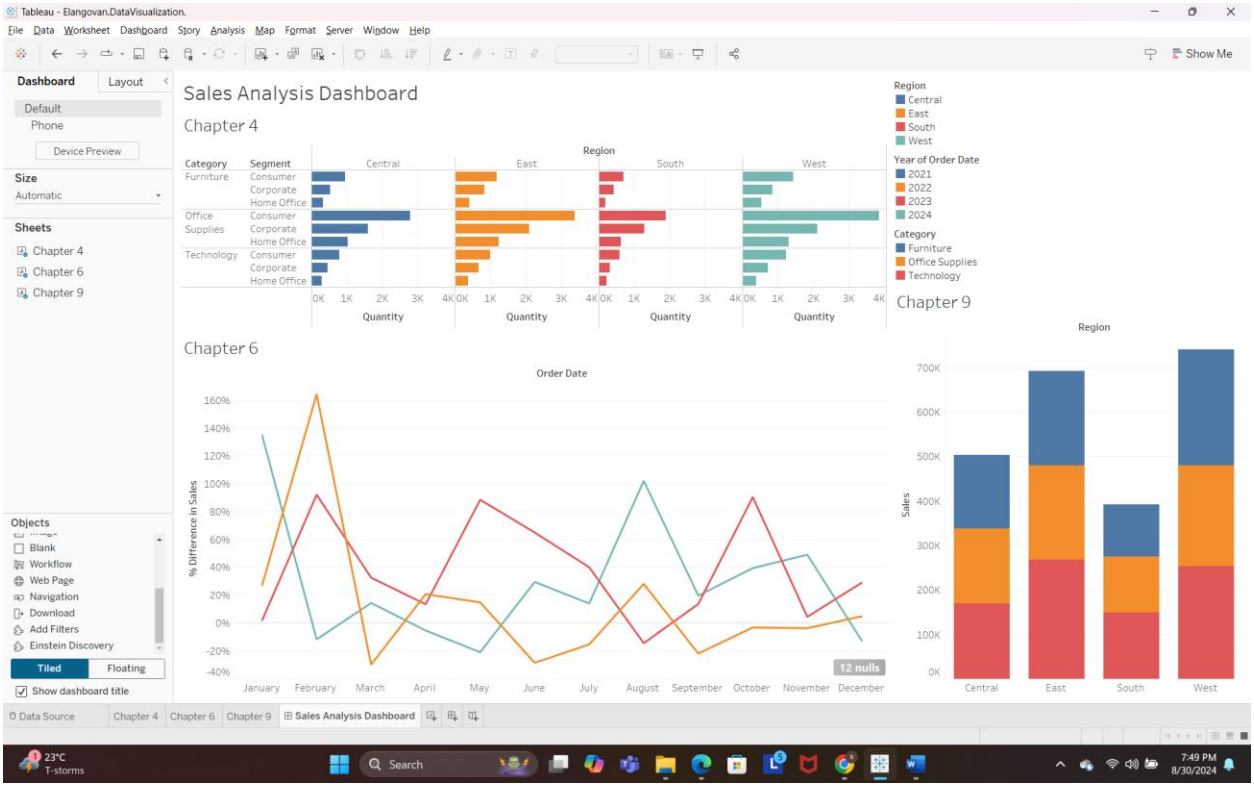
Pros: Provides exact numerical data, which can be useful for detailed analysis.

Cons: It lacks the visual impact and immediate insight that other types of charts (like bar charts or treemaps) can provide, making it harder to spot trends at a glance.

DASHBOARD:

Chapter 4, Chapter 6, Chapter 9 worksheets are presented in dashboard to reflect interesting story and caption include the description of the visualization as an interesting story for a business audience.

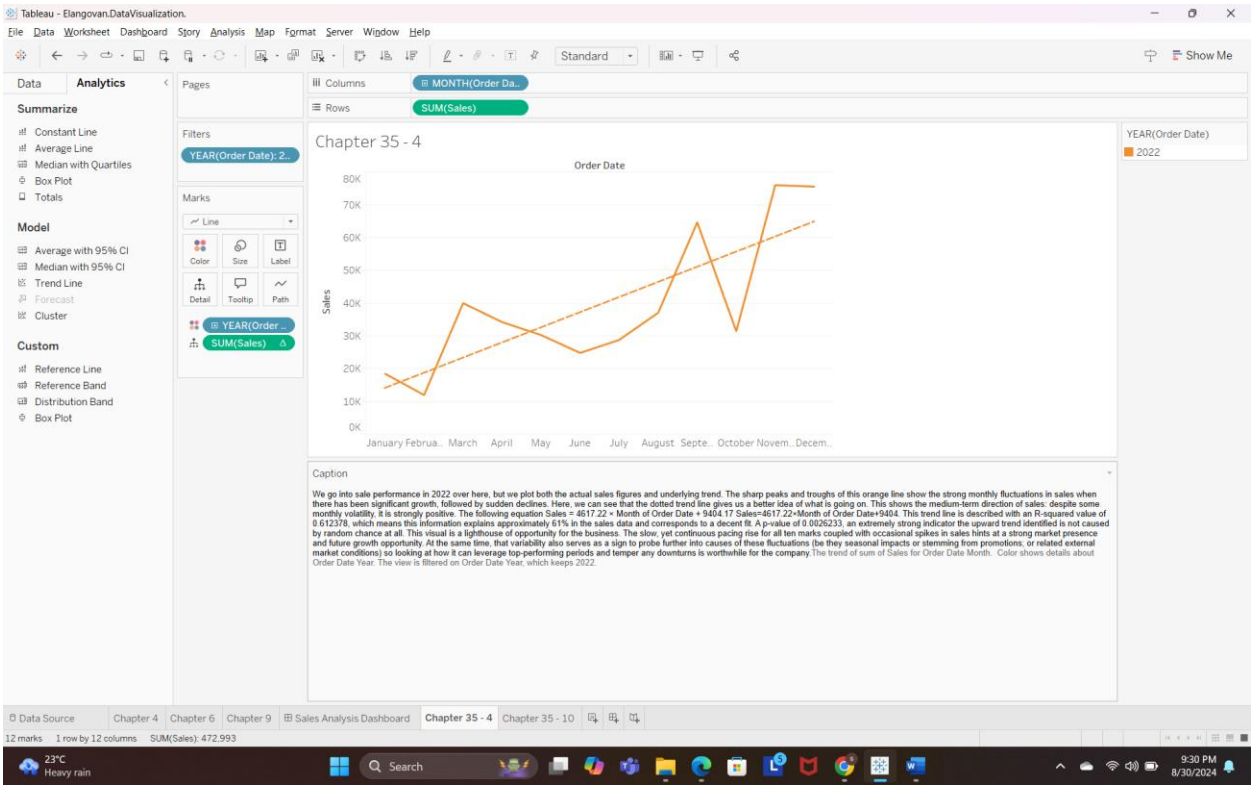
Dashboard title changed from Dashboard 1 to Sales Analysis Dashboard.



Chapter 35-4:

- Trend analysis is Performed Step-by-step as in Chapter 35 (Khan 2016) and reflects the Trend as shown in Figure 35-4.

- Caption of Chapter 35-4 Worksheet includes the description of graph for business audience and the rationale for choosing the specific graph.
- Worksheet renamed as Chapter 35-4.



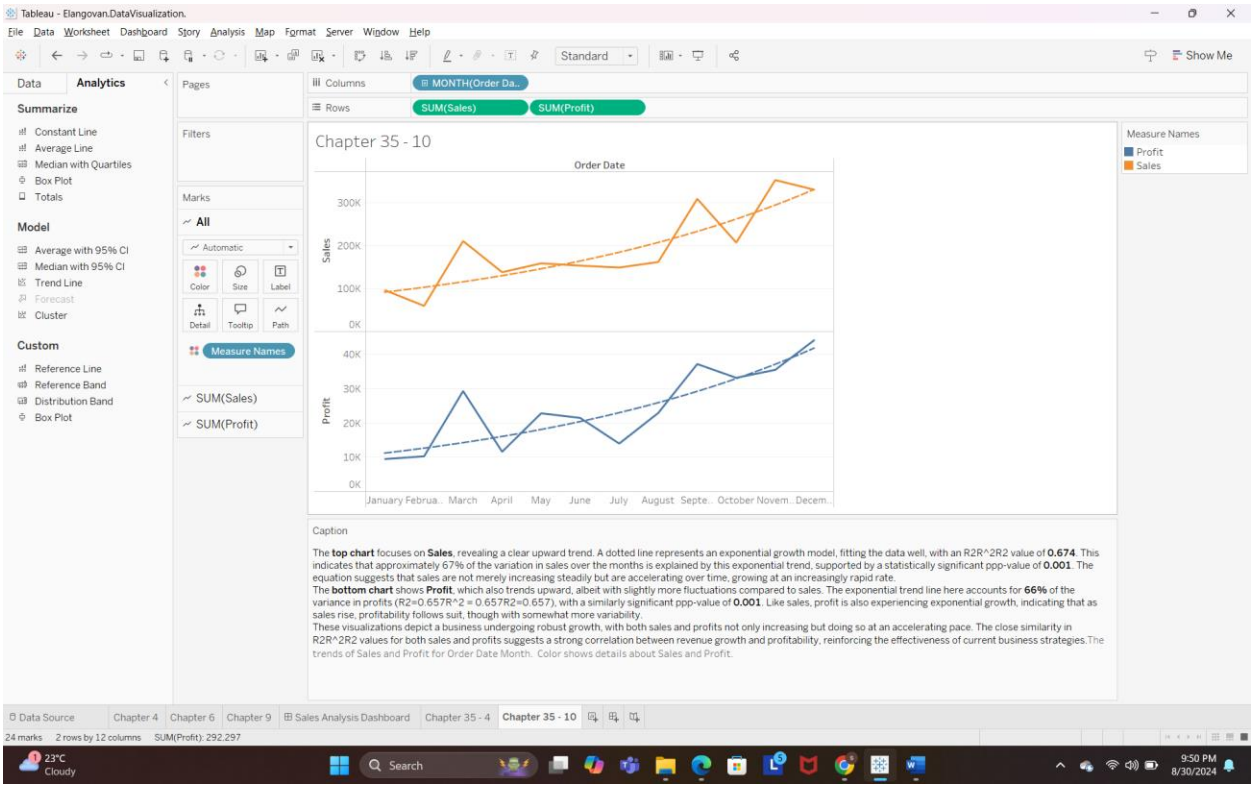
Caption:

We go into sale performance in 2022 over here, but we plot both the actual sales figures and underlying trend. The sharp peaks and troughs of this orange line show the strong monthly fluctuations in sales when there has been significant growth, followed by sudden declines. Here, we can see that the dotted trend line gives us a better idea of what is going on. This shows the medium-term direction of sales: despite some monthly volatility, it is strongly positive. The following equation Sales = 4617.22 × Month of Order Date + 9404.17 Sales=4617.22×Month of Order Date+9404. This trend line is described with an R-squared value of 0.612378, which means this information explains approximately 61% in the sales data and corresponds to a decent fit. A p-value of 0.0026233, an extremely strong indicator the upward trend identified is not caused by random chance at all. This visual is a lighthouse of opportunity for the business. The slow, yet continuous pacing rise for all ten marks coupled with occasional spikes in sales hints at a strong market presence and future growth opportunity. At the same time, that variability also serves as a sign to probe further into causes of these fluctuations (be they seasonal impacts or

stemming from promotions; or related external market conditions) so looking at how it can leverage top-performing periods and temper any downturns is worthwhile for the company.The trend of sum of Sales for Order Date Month. Color shows details about Order Date Year. The view is filtered on Order Date Year, which keeps 2022.

Chapter 35-10:

- Trend analysis is Performed Step-by-step as in Chapter 35 (Khan 2016) and reflects the Trend as shown in Figure 35-10.
- Caption of Chapter 35-10 Worksheet includes the description of graph for business audience and the rationale for choosing the specific graph.
- Worksheet renamed as Chapter 35-10.



Caption:

The **top chart** focuses on **Sales**, revealing a clear upward trend. A dotted line represents an exponential growth model, fitting the data well, with an R2R^2R2 value of **0.674**. This indicates that approximately 67% of the variation in sales over the months is explained by this exponential trend, supported by a statistically significant ppp-value of **0.001**. The equation suggests that sales

are not merely increasing steadily but are accelerating over time, growing at an increasingly rapid rate.

The **bottom chart** shows **Profit**, which also trends upward, albeit with slightly more fluctuations compared to sales. The exponential trend line here accounts for **66%** of the variance in profits (R2=0.657R^2 = 0.657R2=0.657), with a similarly significant ppp-value of **0.001**. Like sales, profit is also experiencing exponential growth, indicating that as sales rise, profitability follows suit, though with somewhat more variability.

These visualizations depict a business undergoing robust growth, with both sales and profits not only increasing but doing so at an accelerating pace. The close similarity in R2R^2R2 values for both sales and profits suggests a strong correlation between revenue growth and profitability, reinforcing the effectiveness of current business strategies.The trends of Sales and Profit for Order Date Month. Color shows details about Sales and Profit.

Chapter 36:

- Forecasting is performed Step-by-step as in Chapter 36 (Khan 2016) and reflects the Forecasting for Next two years.
- Caption of Chapter 36 Worksheet includes the description of forecasting for business audience and the rationale for choosing the specific graph.
- Worksheet renamed as Chapter 35-10.

Caption:

The visualization provides an intriguing narrative on the sales increase throughout time, with a considerable focus on 2021 and extended to estimations even through into 2026.

Wandering through the data, it is pretty evident from the graph that sales do increase gradually, and there may be an increasing trend over time. It is a roller coaster ride from 2021 for the index but has gained on a monthly closing basis. There are ups and downs in every year, likely with some seasonal forces or other periodic factors at play, but the overall trend is upward.

As a result, the forecasted sales curve should steepen by 2024 to signal that increased growth will arrive within these years. The light-blue-shaded region is the predicted range with a confidence interval (range of potential outcomes). Estimated to sell at even more elevated sales heights through 2026, the forecast is bullish in its projection of future growth.

This visualization exhibits the journey of a thriving business with increasing sales, which are forecasted to escalate in the future. The model is meant to show not just upside (+) optimism but also downside (-) probabilities and the uncertainty of these projections, for a more complete picture across the board. This is a positive trend indication, which possibly results from strategic

initiatives taken up by companies to grow their business or from an increase in customer demand, and one can now project it forward while planning future growth strategies.