

Project - Phase III: Decision Making

Mobile and Laptop Sales Data Analysis

IFT 533: Data Visualization and Reporting for IT
Project Group 29

Team Members:

Apurva Dange
Nidhi Thribhuvan Devadas
Trinadh Chowdary Rayapudi

Professor: Asmaa Elbadrawy

May 2, 2025

1 Final Dashboard

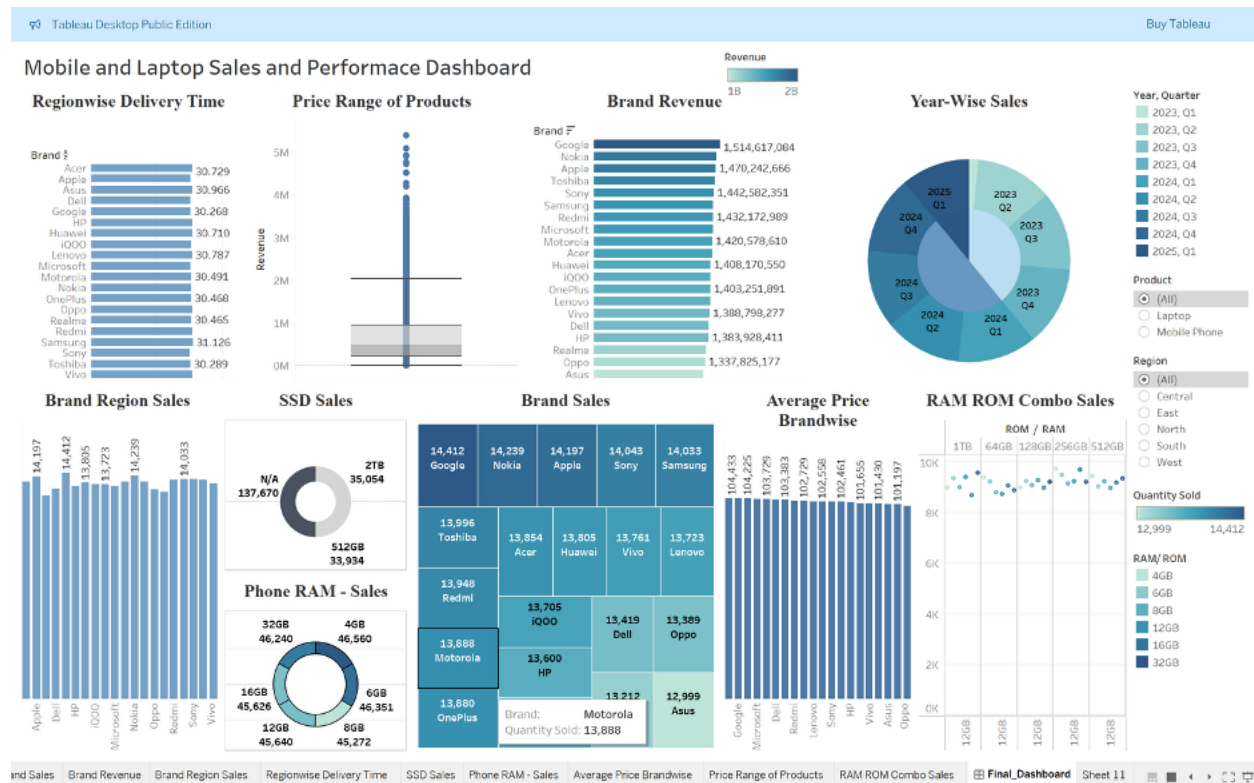


Figure 1: Mobile and Laptop Sales and Performance Dashboard

This dashboard provides a comprehensive overview of sales performance, product characteristics, and operational metrics (like delivery time) for electronic products, such as mobile phones and laptops. It allows users to visualize key performance indicators (KPIs) such as revenue, sales volume, pricing, and regional performance, facilitating data-driven decision-making.

2 The Dataset

Overview

This dataset simulates 50,000 sales records for mobile phones and laptops. Each entry represents a sales transaction and contains product details, pricing, customer data, and regional information. The data is structured tabular data with 16 columns, suitable for analyzing trends, forecasting, customer segmentation, and business strategy.

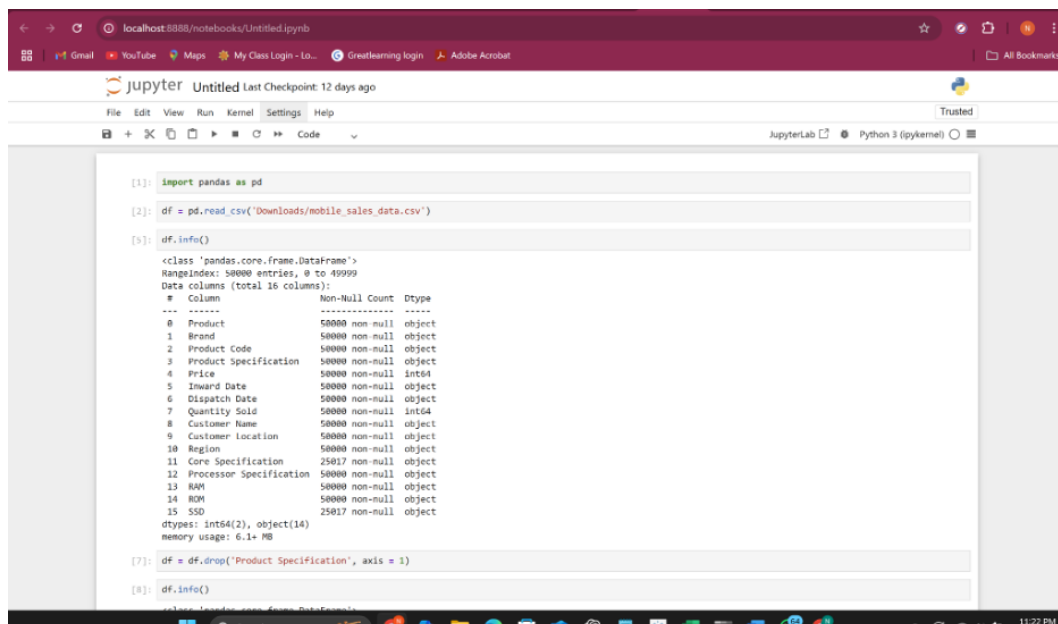
Column Data Types & Domains

The following table details the schema of the dataset:

Column Name	Data Type	Domain (Example or Range)
Product	Categorical	Mobile Phone, Laptop
Brand	Categorical	Apple, Samsung, Dell, OnePlus, etc.
Product Code	Categorical	Unique alphanumeric (e.g., "73D2A7CC")
Price	Ratio	5,008 – 199,999
Inward Date	Interval	e.g., 2023-01-01 to 2025-03-31
Dispatch Date	Interval	e.g., 2023-01-02 to 2025-04-01
Quantity Sold	Ratio	1 – 10
Customer Name	Categorical	Random names (e.g., "Leah Copeland")
Customer Location	Categorical	Example: "South Todd", "North Lisa"
Region	Categorical	North, South, East, West, Central
Core Spec	Categorical	i3, i5, i7, Ryzen 5, Ryzen 7, etc.
Processor Spec	Categorical	Snapdragon, Exynos, Apple A15
RAM	Ordinal	4GB, 8GB, 16GB, 32GB
ROM	Ordinal	64GB, 128GB, 256GB, 512GB, 1TB
SSD	Ordinal	256GB, 512GB, 1TB, 2TB, or "N/A"

Preprocessing

The column "Product Specification" was dropped during the data cleaning phase because the text-based descriptions contained in that column were deemed irrelevant to the quantitative analysis.



```

[1]: import pandas as pd

[2]: df = pd.read_csv('Downloads/mobile_sales_data.csv')

[5]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50000 entries, 0 to 49999
Data columns (total 16 columns):
 #   Column                Non-Null Count  Dtype  
---  --
 0   Product                50000 non-null object
 1   Brand                  50000 non-null object
 2   Product Code           50000 non-null object
 3   Product Specification  50000 non-null object
 4   Price                  50000 non-null int64
 5   Inward Date            50000 non-null object
 6   Dispatch Date          50000 non-null object
 7   Quantity Sold          50000 non-null int64
 8   Customer Name          50000 non-null object
 9   Customer Location      50000 non-null object
10   Region                 50000 non-null object
11   Core Specification     25017 non-null object
12   Processor Specification 50000 non-null object
13   RAM                    50000 non-null object
14   ROM                    50000 non-null object
15   SSD                    25017 non-null object
dtypes: int64(2), object(14)
memory usage: 6.1+ MB

[7]: df = df.drop("Product Specification", axis = 1)

[8]: df.info()

```

Figure 2: Data Loading and Inspection

The screenshot shows a JupyterLab notebook with a DataFrame containing 16 columns. The 'Product Specification' column is highlighted, showing a list of product descriptions. The interface includes a top bar with navigation icons and a bottom status bar showing the time as 11:30 PM on 4/17/2025.

```

# Column      Non-Null Count  Dtype
---
0 Product      50000 non-null  object
1 Brand        50000 non-null  object
2 Product Code 50000 non-null  object
3 Price        50000 non-null  int64
4 Inward Date  50000 non-null  object
5 Dispatch Date 50000 non-null  object
6 Quantity Sold 50000 non-null  int64
7 Customer Name 50000 non-null  object
8 Customer Location 50000 non-null  object
9 Region       50000 non-null  object
10 Core Specification 25017 non-null  object
11 Processor Specification 50000 non-null  object
12 RAM         50000 non-null  object
13 ROM         50000 non-null  object
14 SSD         25017 non-null  object
dtypes: int64(2), object(14)
memory usage: 6.1+ MB

[12]: df["Product Specification"]

[12]: 0      Site candidate activity company there bit insi...
      1      Beat put care fight affect address his.
      2      Energy special low seven place audience.
      3      Friend record hard contain minute we role sea ...
      4      Program recently feeling save tree hotel people.
      ...
      49995      Pick build kind teach mean each camera.
      49996      Outside policy treat he opportunity pattern.
      49997      Test floor tax together out adult discover own...
      49998      Away also magazine receive such lay.
      49999      Role ahead common listen idea few blood half.
      Name: Product Specification, Length: 50000, dtype: object

```

Figure 3: Viewing the Irrelevant Product Specification Column

The screenshot shows the same JupyterLab notebook after the 'Product Specification' column has been dropped. The DataFrame now has 15 columns. The 'df.info()' output is shown, confirming the removal of the column. The interface includes a top bar with navigation icons and a bottom status bar showing the time as 11:32 PM.

```

14 ROM      50000 non-null  object
15 SSD      25017 non-null  object
dtypes: int64(2), object(14)
memory usage: 6.1+ MB

[7]: df = df.drop('Product Specification', axis = 1)

[8]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50000 entries, 0 to 49999
Data columns (total 15 columns):
#   Column      Non-Null Count  Dtype
---
0 Product      50000 non-null  object
1 Brand        50000 non-null  object
2 Product Code 50000 non-null  object
3 Price        50000 non-null  int64
4 Inward Date  50000 non-null  object
5 Dispatch Date 50000 non-null  object
6 Quantity Sold 50000 non-null  int64
7 Customer Name 50000 non-null  object
8 Customer Location 50000 non-null  object
9 Region       50000 non-null  object
10 Core Specification 25017 non-null  object
11 Processor Specification 50000 non-null  object
12 RAM         50000 non-null  object
13 ROM         50000 non-null  object
14 SSD         25017 non-null  object
dtypes: int64(2), object(13)
memory usage: 5.7+ MB

```

Figure 4: Dropping the Column

3 Dashboard Users

This dashboard is likely intended for various stakeholders within an electronics retail or manufacturing company, including:

- **Sales Managers:** To track sales performance by brand, region, and product; identify top-performing products and regions; and monitor sales trends over time.
- **Marketing Teams:** To understand product popularity, pricing strategies, and customer preferences (e.g., popular RAM/ROM combinations).
- **Operations/Logistics Managers:** To monitor regional delivery times and identify potential bottlenecks or areas for improvement.
- **Product Managers:** To analyze sales data for different product specifications (RAM, ROM, SSD) and inform product development decisions.
- **Business Analysts & Executives:** To gain a high-level overview of business performance, identify trends, and support strategic planning.

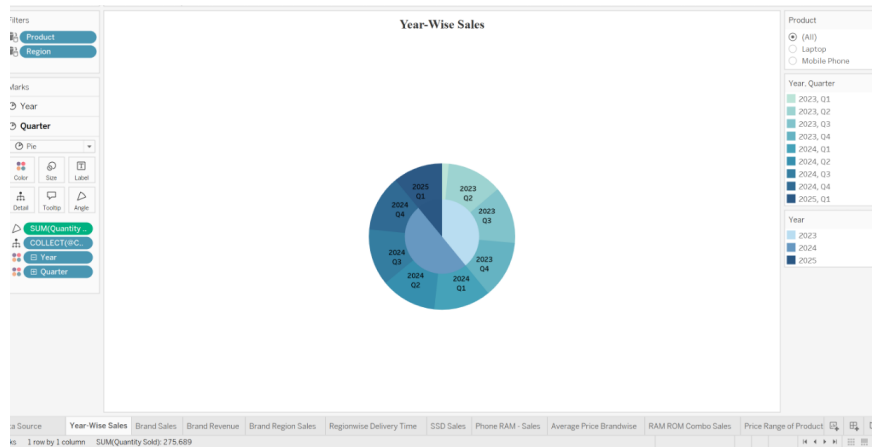
4 Key Questions Addressed

The dashboard is designed to answer questions such as:

1. Which brands generate the most revenue?
2. Which brands have the highest sales volume?
3. How do sales trends evolve year over year and quarter over quarter?
4. What is the distribution of product prices?
5. How does average delivery time vary across different brands?
6. How do sales vary across different regions for each brand?
7. What are the sales figures for products with different SSD sizes?
8. What are the sales figures for mobile phones with different RAM sizes?
9. What is the average price point for each brand?
10. Which RAM/ROM combinations are most popular or generate the most sales?
11. How do sales and revenue change when filtering by specific product types, regions, or time periods?

5 Visualization Details

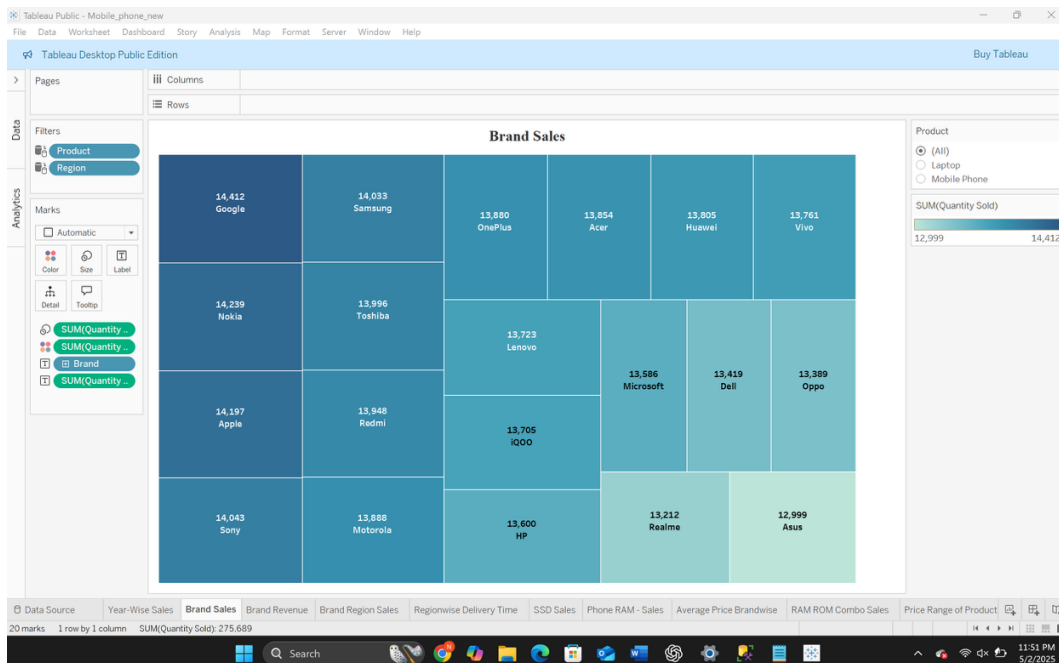
5.1 Year-Wise Sales (Donut Chart)



Question Addressed: What is the proportion of sales contributed by each year/quarter?

How it Addresses: Divides total sales into segments representing different time periods (quarters within years), showing their relative contribution.

5.2 Brand Revenue (TreeMap)

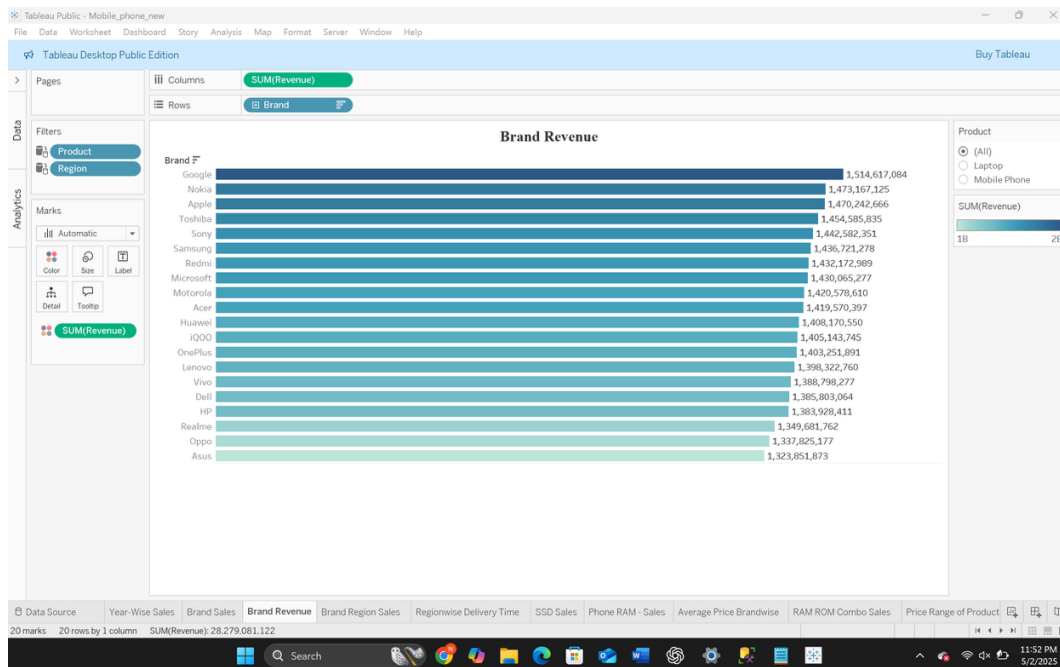


Question Addressed: How do sales for each brand vary across different regions? Which product types or regions contribute most to brand sales?

How it Addresses:

- **Size:** Represents quantity sold (SUM).
- **Color:** Gradient reinforces quantity sold (darker = higher).
- **Labels:** Show actual sales figures for precision.

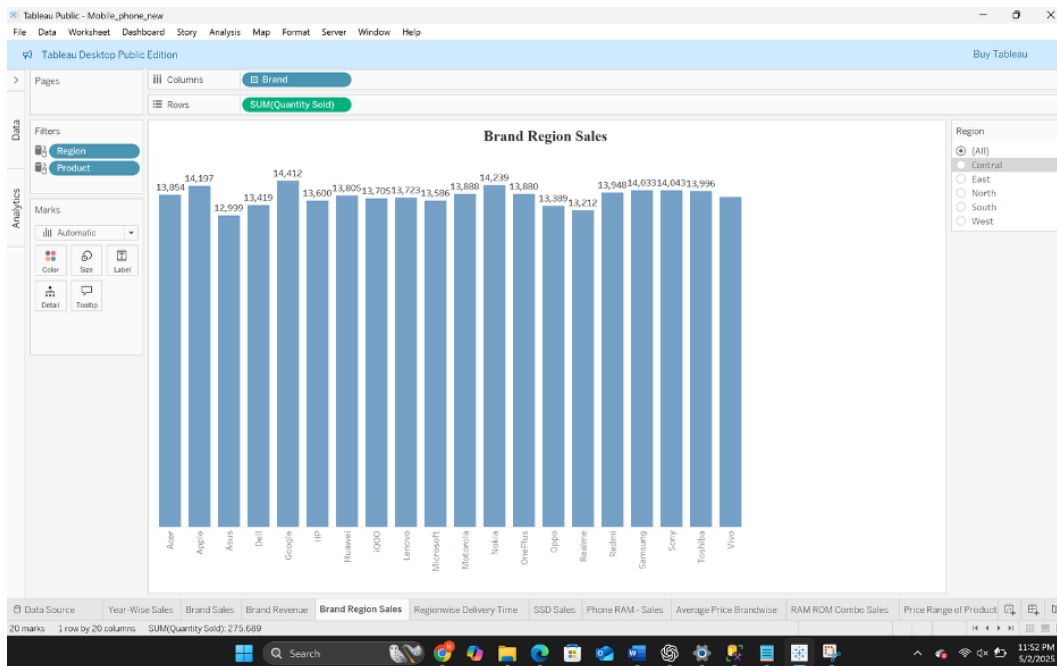
5.3 Brand Revenue (Bar Chart)



Question Addressed: Which brands generate the most revenue?

How it Addresses: Directly compares total revenue by brand in descending order. The length of the bar indicates total revenue, and the color gradient provides a visual cue for magnitude.

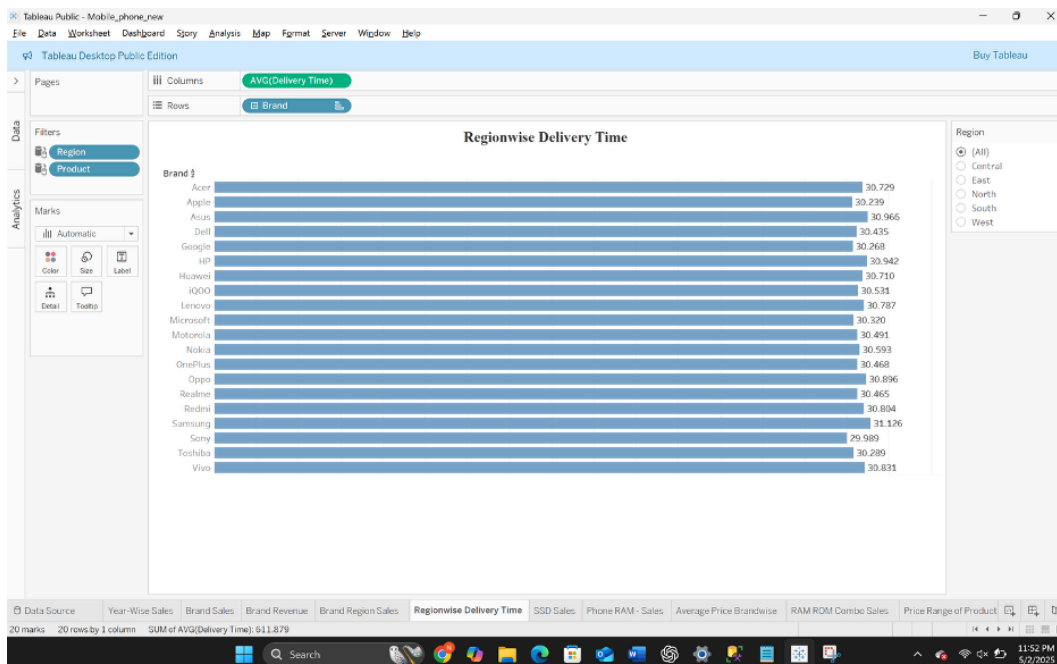
5.4 Brand Region Sales (Bar Chart)



Question Addressed: How do sales for each brand vary across different regions?

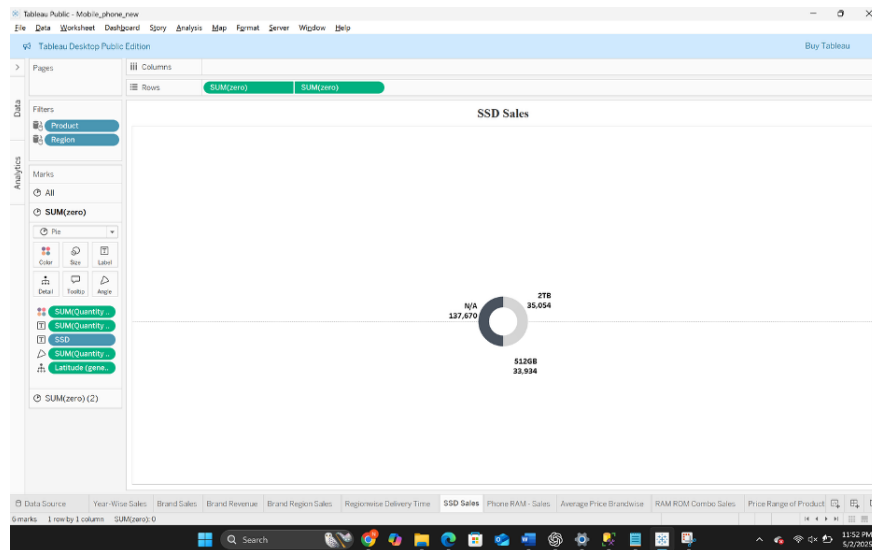
How it Addresses: Displays total quantity sold per brand. A region filter allows users to isolate data (e.g., "Central") to compare regional performance.

5.5 Regionwise Delivery Time (Horizontal Bar Chart)



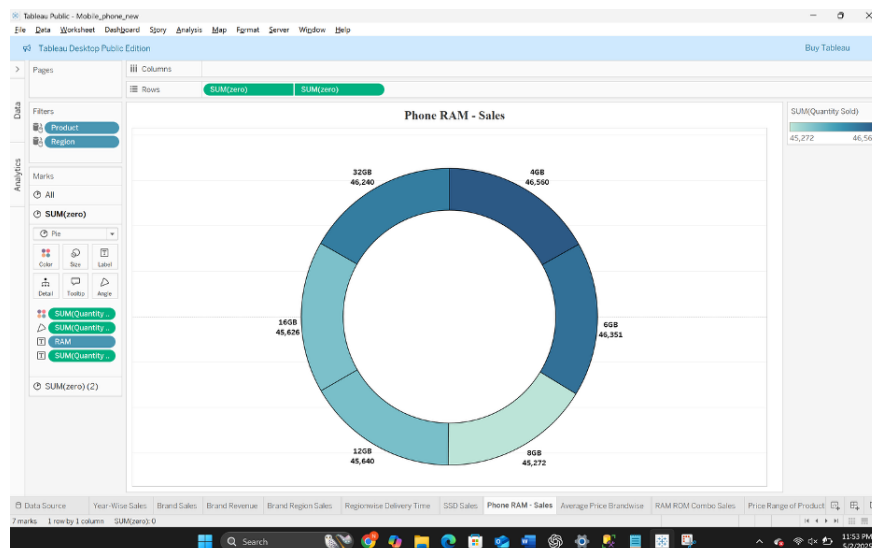
Question Addressed: What is the average delivery time for each brand across different regions?
How it Addresses: Ranks brands based on delivery speed (AVG Delivery Time). This helps identify logistics efficiency.

5.6 SSD Sales (Donut Chart)



Question Addressed: What is the distribution of SSD storage options sold across all products?
How it Addresses: Displays quantity sold for different SSD capacities (e.g., 2TB, 512GB). Segment size represents total units sold.

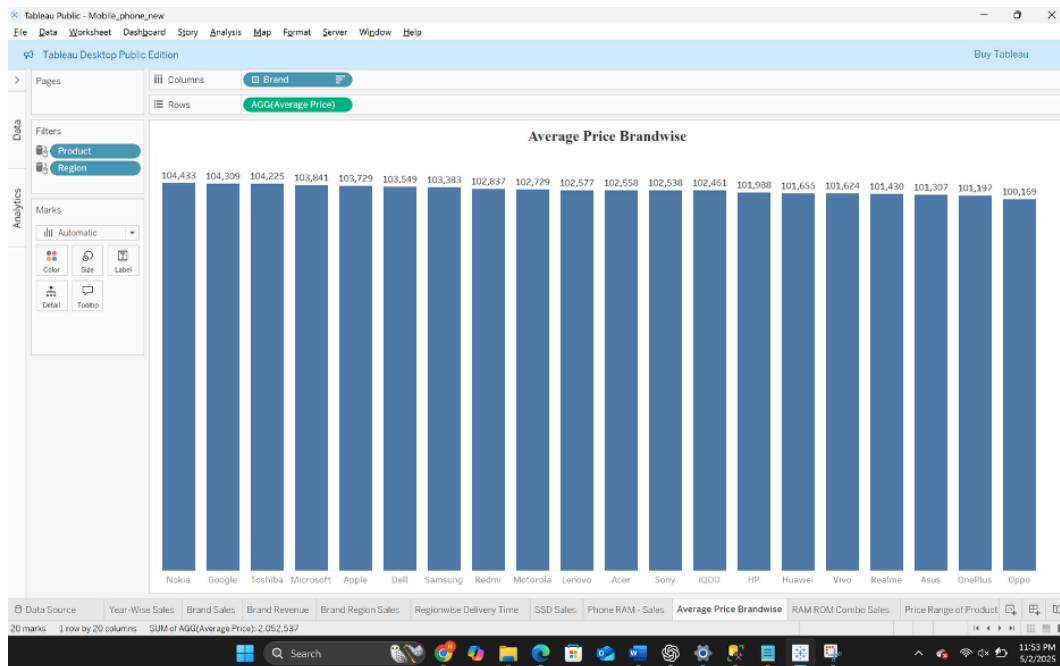
5.7 Phone RAM - Sales (Donut Chart)



Question Addressed: What is the distribution of mobile phones sold by RAM size?
How it Addresses: Shows sales quantity for RAM configurations (4GB, 6GB, etc.). Labels provide

exact unit counts.

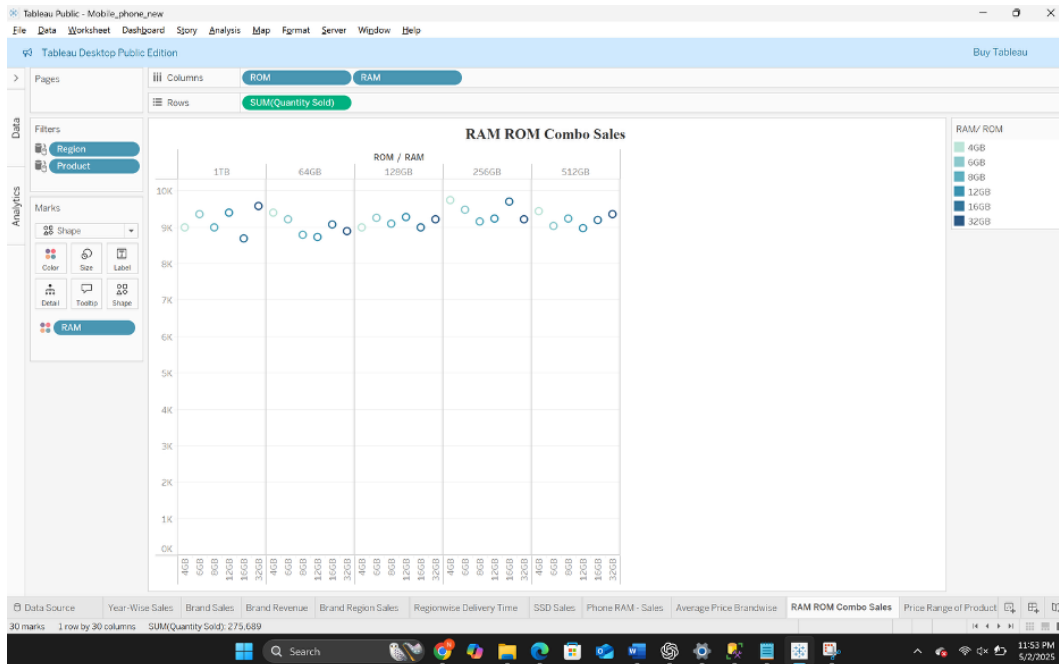
5.8 Average Price Brandwise (Bar Chart)



Question Addressed: What is the average price of products for each brand?

How it Addresses: Vertical bar chart showing the aggregated average selling price, allowing for quick comparison of pricing strategies.

5.9 RAM ROM Combo Sales (Dot Plot)

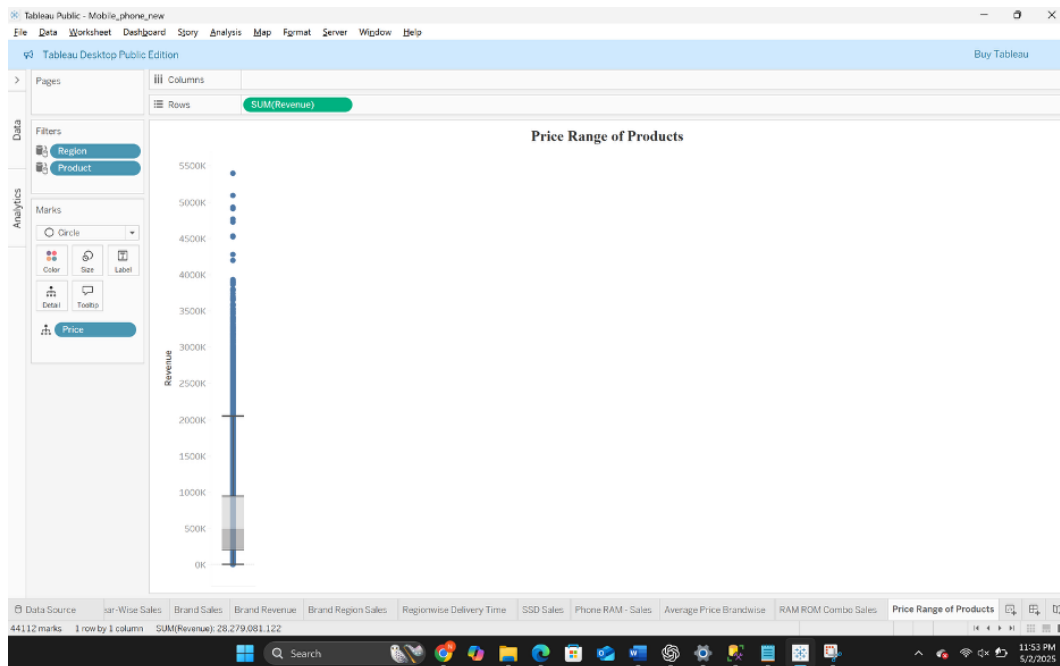


Question Addressed: Which RAM and ROM combinations are most popular?

How it Addresses:

- **X-axis:** ROM sizes (e.g., 64GB, 128GB).
- **Color:** RAM size (4GB to 32GB).
- **Y-axis:** Units sold.

5.10 Price Range of Products (Box Plot)



Question Addressed: What is the distribution and range of product prices?

How it Addresses: A box plot overlaid with data points (circles) to show the spread, median, quartiles, and outliers of product prices relative to revenue.

6 Dashboard Interactivity

The dashboard includes several interactive controls (filters) located on the right side, allowing users to refine the data displayed across most, if not all, plots:

- **Year, Quarter Filter:** Allows users to select specific years and quarters to analyze sales trends within those periods.
- **Product Filter:** Enables filtering the data by product type (e.g., Laptop, Mobile Phone). This affects all plots.
- **Region Filter:** Allows users to select one or more regions (Central, East, North, South, West). This updates plots to show data specific to the chosen geographical areas.
- **RAM / ROM Filters:** Separate filters for RAM and ROM allow users to focus on products with specific memory configurations.

7 References

- **Tableau Dashboard:**
https://public.tableau.com/app/profile/nidhi.devadas/viz/Group_29_Dashboard/Final_Dashboard?publish=yes

- **Dataset Source:**
<https://www.kaggle.com/datasets/vinothkannaece/mobiles-and-laptop-sales-data/data>
- **Mural Board:**
[Link to Project Mural Board](#)