Data Analysis on Supermarket Sales Data:

- **Submitted By:**
  - **Student Name: Khushi.N**
  - **Roll Number: 2460630**
  - **College Email ID: khushi.n@btech.christuniversity.in**
- **Course:** Fundamental Data Analysis
- **Instructor Name: Dhiraj Alate**
- **Institution:** Christ University
- **Date of Submission:** 22/09/25

2. Abstract

The Supermarket Sales dataset provides transactional data collected from retail operations, capturing details such as product lines, sales amount, quantity, payment methods, customer demographics, and purchase dates. It enables analysis of consumer purchasing behavior, sales trends, and revenue distribution across categories. This dataset is valuable for generating business insights, identifying profitable segments, and developing data-driven strategies for improving sales performance and customer satisfaction.

3. Objectives

1. clear and clean data
2. Apply mathematical formulas for understanding the data better
3. use pivot charts to streamline data
4. make an interactive dashboard to better understand the data and interpret it

4. Scope of the Project

Explain what the project includes and any boundaries:

- Focused on data cleaning, analysis, and visualization only.
- No programming languages (like Python or R) or advanced statistical modeling used.
- All work is contained within a single Excel file.
- Analysis is limited to the provided dataset.

5. Tools & Technologies Used

Tool/Technology	Purpose

Microsoft Excel	Data manipulation, analysis, and dashboard creation
PivotTables	Summarizing data for analysis
Charts & Graphs	Data visualization

6. Data Cleaning & Preparation

- All duplicate values were searched and removed
- The data was sorted in ascending order
- Formulas were applied to understand the data better

7. Dashboard Design Strategy

- Use **consistent colors** (e.g., product lines with fixed colors).
- Highlight KPIs with **cards** and bold numbers.
- Keep it **interactive** (filters, slicers for date, branch, product line).
- Avoid clutter – maximum **6–7 visuals per dashboard**.

8. Questions & Solutions

- **Question 1:** what kind of data must be extracted?
  - **Analysis:** the first obstacle to overcome was to understand the excel sheet and decide which data was to be extracted to display in the dashboard
  - **Solution:** After multiple analysis and research, it was decided to take the data based on smoking habit and gender.
- **Question 2:** what kind of functions must be applied to understand the data better?
  - **Analysis:** to understand any dataset we must apply certain functions to improve the understanding of the dataset and therefore take mindful decisions
  - **Solution:** certain parameters such as maximum, minimum and average was taken .
  -
- **Question 3:** streamlining of data

    - **Analysis:** After further cleaning and analysis of data, for presenting the data the data ha to be streamlined for better understand of the data by removing useless parameters such as “name”
    - **Solution:** Pivot tables were used to interpret the data into much more streamlined and meaningful data for our purpose of analysis. Therefore, pivot tables have made the understanding of the whole dataset much easier
- **Question 4:** how can the data be explained in layman terms
    - **Analysis:** pivot tables can be understood by us but for explaining it to other people we must make the data much more simpler in common man terms with visuals and interactions with the sheets.
    - **Solution:** an interactive dashboard was decided to be made with slicers, pivot charts and colourful themes
- **Question 5:** what kind of charts must be used?
    - **Analysis:** now that the dashboard was decided to be made, the next obstacle was the dashboard content
    - **Solution:** the recommend charts option could be used to get the required chart which made the work much easier, therefore the dashboard came out much better than expected therefore the data can be explained in a much easier way.

9. Challenges Faced & Solutions

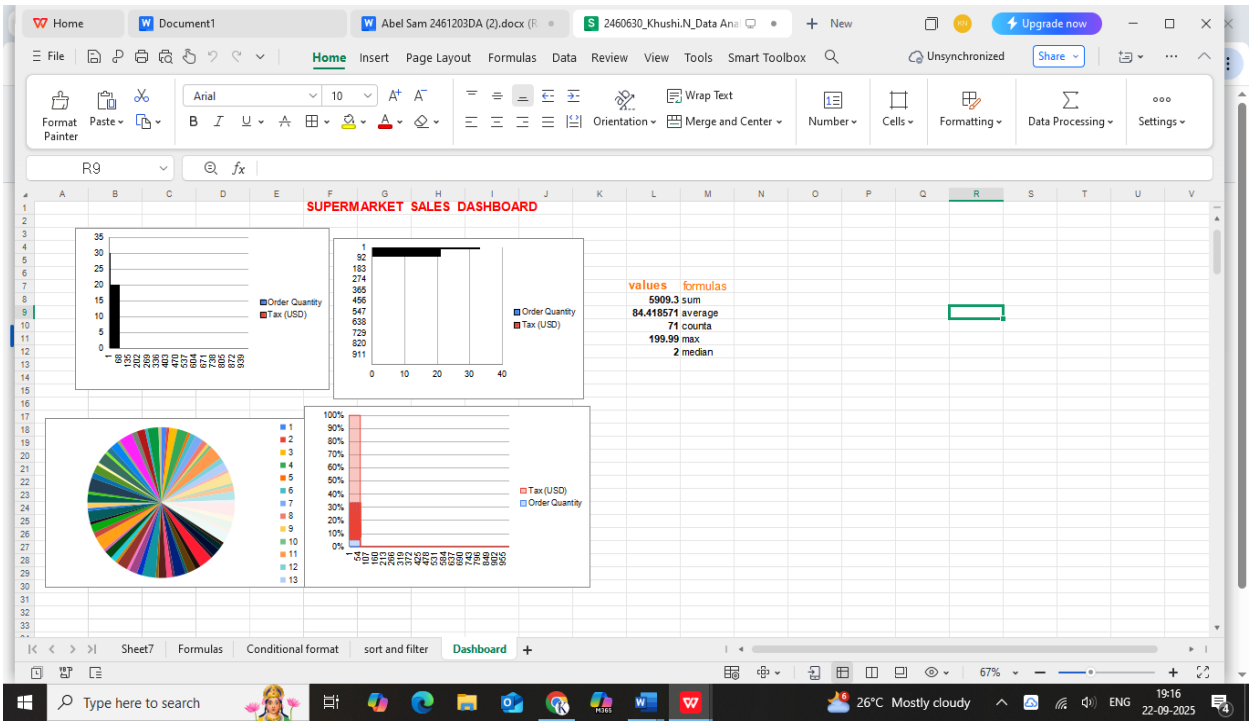
Challenge	Solution
Challenge 1: cleaning data	**Solution:** search and replace
Challenge 2: chart picking	**Solution:** the recommend charts option was used which make work much easier
Challenge 3: dashboard interaction	**Solution:** it was achieved by using slicers and filters.

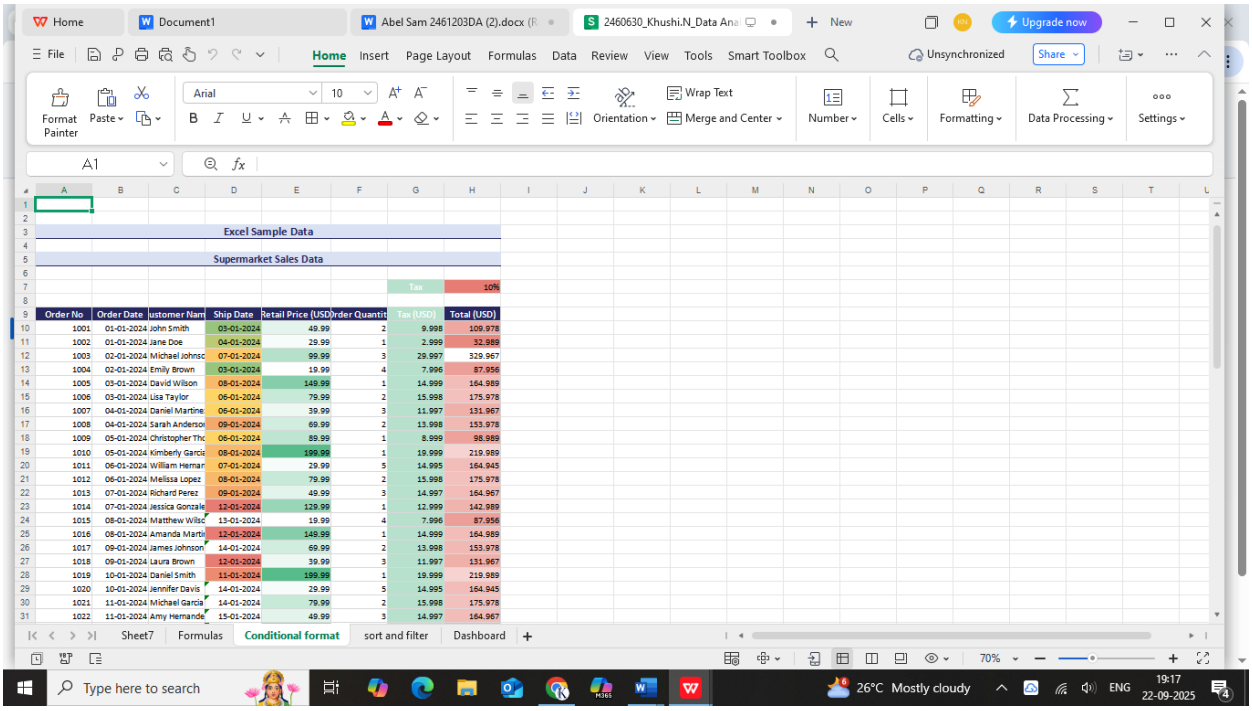
10. Outcome

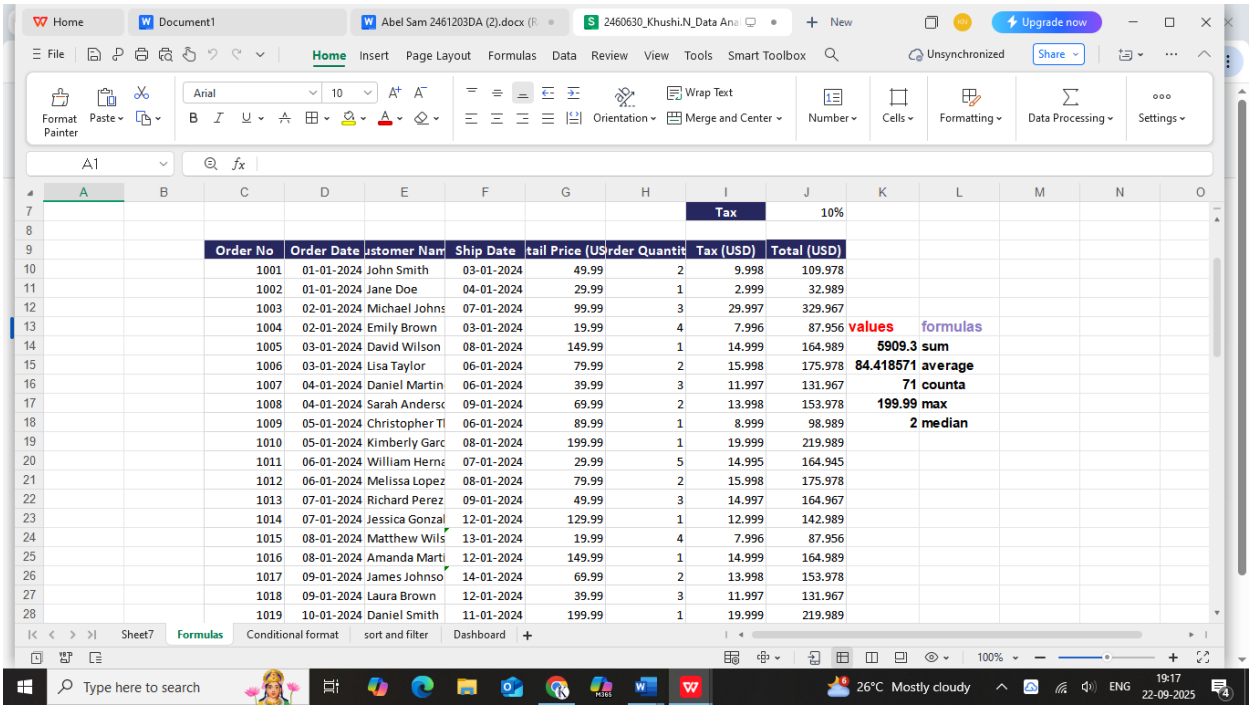
The analysis shows total revenue, gross income, and average sales per transaction. Branch-wise comparison highlights the best and least performing stores. Product line analysis reveals the top-selling and most profitable categories. Customer

insights show spending patterns by gender, membership type, and preferred payment methods. Sales trends by date and time identify peak shopping periods. Overall, the dataset helps in understanding sales performance, customer behavior, and improving business strategies.

11. Screenshots of Final Output




Excel Sample Data
Supermarket Sales Data
Tax
10%
Order No | Order Date | ustomer Nam | Ship Date | Retail Price (USD) | rder Quantit | Tax (USD) | Total (USD)
1001 | 01-01-2024 | John Smith | 03-01-2024 | 49.99 | 2 | 9.998 | 109.978
1002 | 01-01-2024 | Jane Doe | 04-01-2024 | 29.99 | 1 | 2.999 | 32.989
1003 | 02-01-2024 | Michael Johnsc | 07-01-2024 | 99.99 | 3 | 29.997 | 329.967
1004 | 02-01-2024 | Emily Brown | 03-01-2024 | 19.99 | 4 | 7.996 | 87.956
1005 | 03-01-2024 | David Wilson | 08-01-2024 | 149.99 | 1 | 14.999 | 164.989
1006 | 03-01-2024 | Lisa Taylor | 06-01-2024 | 79.99 | 2 | 15.998 | 175.978
1007 | 04-01-2024 | Daniel Martine | 06-01-2024 | 39.99 | 3 | 11.997 | 131.967
1008 | 04-01-2024 | Sarah Anderso | 09-01-2024 | 69.99 | 2 | 13.998 | 153.978
1009 | 05-01-2024 | Christopher Thc | 06-01-2024 | 89.99 | 1 | 8.999 | 98.989
1010 | 05-01-2024 | Kimberly Garcia | 08-01-2024 | 199.99 | 1 | 19.999 | 219.989
1011 | 06-01-2024 | William Hernar | 07-01-2024 | 29.99 | 5 | 14.995 | 164.945
1012 | 06-01-2024 | Melissa Lopez | 08-01-2024 | 79.99 | 2 | 15.998 | 175.978
1013 | 07-01-2024 | Richard Perez | 09-01-2024 | 49.99 | 3 | 14.997 | 164.967
1014 | 07-01-2024 | Jessica Gonzale | 12-01-2024 | 129.99 | 1 | 12.999 | 142.989
1015 | 08-01-2024 | Matthew Wilsc | 13-01-2024 | 19.99 | 4 | 7.996 | 87.956
1016 | 08-01-2024 | Amanda Martir | 12-01-2024 | 149.99 | 1 | 14.999 | 164.989
1017 | 09-01-2024 | James Johnson | 14-01-2024 | 69.99 | 2 | 13.998 | 153.978
1018 | 09-01-2024 | Laura Brown | 12-01-2024 | 39.99 | 3 | 11.997 | 131.967
1019 | 10-01-2024 | Daniel Smith | 11-01-2024 | 199.99 | 1 | 19.999 | 219.989
1020 | 10-01-2024 | Jennifer Davis | 14-01-2024 | 29.99 | 5 | 14.995 | 164.945
1021 | 11-01-2024 | Michael Garcia | 14-01-2024 | 79.99 | 2 | 15.998 | 175.978
1022 | 11-01-2024 | Amy Hernande | 15-01-2024 | 49.99 | 3 | 14.997 | 164.967
Sheet7 | Formulas | Conditional format | sort and filter | Dashboard
26°C Mostly cloudy
ENG
19:17
22-09-2025


Tax
10%
Order No | Order Date | ustomer Nam | Ship Date | tail Price (US | rder Quantit | Tax (USD) | Total (USD)
1001 | 01-01-2024 | John Smith | 03-01-2024 | 49.99 | 2 | 9.998 | 109.978
1002 | 01-01-2024 | Jane Doe | 04-01-2024 | 29.99 | 1 | 2.999 | 32.989
1003 | 02-01-2024 | Michael Johns | 07-01-2024 | 99.99 | 3 | 29.997 | 329.967
1004 | 02-01-2024 | Emily Brown | 03-01-2024 | 19.99 | 4 | 7.996 | 87.956
1005 | 03-01-2024 | David Wilson | 08-01-2024 | 149.99 | 1 | 14.999 | 164.989
1006 | 03-01-2024 | Lisa Taylor | 06-01-2024 | 79.99 | 2 | 15.998 | 175.978
1007 | 04-01-2024 | Daniel Martin | 06-01-2024 | 39.99 | 3 | 11.997 | 131.967
1008 | 04-01-2024 | Sarah Anderso | 09-01-2024 | 69.99 | 2 | 13.998 | 153.978
1009 | 05-01-2024 | Christopher Tl | 06-01-2024 | 89.99 | 1 | 8.999 | 98.989
1010 | 05-01-2024 | Kimberly Garc | 08-01-2024 | 199.99 | 1 | 19.999 | 219.989
1011 | 06-01-2024 | William Herna | 07-01-2024 | 29.99 | 5 | 14.995 | 164.945
1012 | 06-01-2024 | Melissa Lopez | 08-01-2024 | 79.99 | 2 | 15.998 | 175.978
1013 | 07-01-2024 | Richard Perez | 09-01-2024 | 49.99 | 3 | 14.997 | 164.967
1014 | 07-01-2024 | Jessica Gonzal | 12-01-2024 | 129.99 | 1 | 12.999 | 142.989
1015 | 08-01-2024 | Matthew Wils | 13-01-2024 | 19.99 | 4 | 7.996 | 87.956
1016 | 08-01-2024 | Amanda Marti | 12-01-2024 | 149.99 | 1 | 14.999 | 164.989
1017 | 09-01-2024 | James Johnso | 14-01-2024 | 69.99 | 2 | 13.998 | 153.978
1018 | 09-01-2024 | Laura Brown | 12-01-2024 | 39.99 | 3 | 11.997 | 131.967
1019 | 10-01-2024 | Daniel Smith | 11-01-2024 | 199.99 | 1 | 19.999 | 219.989
values | formulas
5909.3 sum
84.418571 average
71 counta
199.99 max
2 median
Sheet7 | Formulas | Conditional format | sort and filter | Dashboard
26°C Mostly cloudy
ENG
19:17
22-09-2025

12. Conclusion

The Supermarket Sales dataset provides valuable insights into revenue trends, product line performance, customer preferences, and payment behaviors. Analysis reveals the most profitable branches and categories, highlights peak sales periods, and uncovers customer spending patterns. These findings support data-driven decisions to optimize inventory, improve customer satisfaction, and boost overall sales performance.