DATA ANALYSIS

Using MS Excel

DECEMBER 5, 2023

- Pavan

Data analysis using Microsoft Excel

Table of Contents:

1. Introduction
    - Overview of Data Analysis
    - Importance of Excel in Data Analysis
2. Getting Started with Excel
    - Setting Up Your Excel Environment
    - Basic Excel Terminology
    - Importing Data
3. Data Cleaning and Preparation
    - Removing Duplicates
    - Handling Missing Values
    - Data Formatting and Validation
4. Exploratory Data Analysis (EDA)
    - Creating Visualizations
    - Charts and Graphs
    - Conditional Formatting
5. Data Manipulation Techniques
    - Sorting and Filtering Data
    - Using Formulas and Functions
    - PivotTables and Pivot Charts
6. Best Practices and Tips
    - Organizing Your Workbook
    - Documenting Your Analysis
    - Ensuring Accuracy

7. Conclusion
    - Summary of Key Points
    - Further Reading and Resources

8. Appendices
    - Glossary of Terms
    - Sample Data Sets

1. Introduction:

Overview of Data Analysis

Data analysis involves inspecting, cleansing, transforming, and modeling data to discover useful information, draw conclusions, and support decision-making. Microsoft Excel is a powerful tool that provides various features and functions to perform these tasks effectively.

Importance of Excel in Data Analysis

Excel is widely used for data analysis due to its accessibility, versatility, and array of analytical tools. It is especially beneficial for users who need to handle data on a smaller scale or require immediate analysis without the need for specialized software.

2. Getting Started with Excel

Setting Up Your Excel Environment

Launching Excel: Open Excel from the Start menu or desktop shortcut.

Interface Overview: Familiarize yourself with the Workbook, Worksheet, and Cells.

Basic Excel Terminology

Cell: The intersection of a row and column.

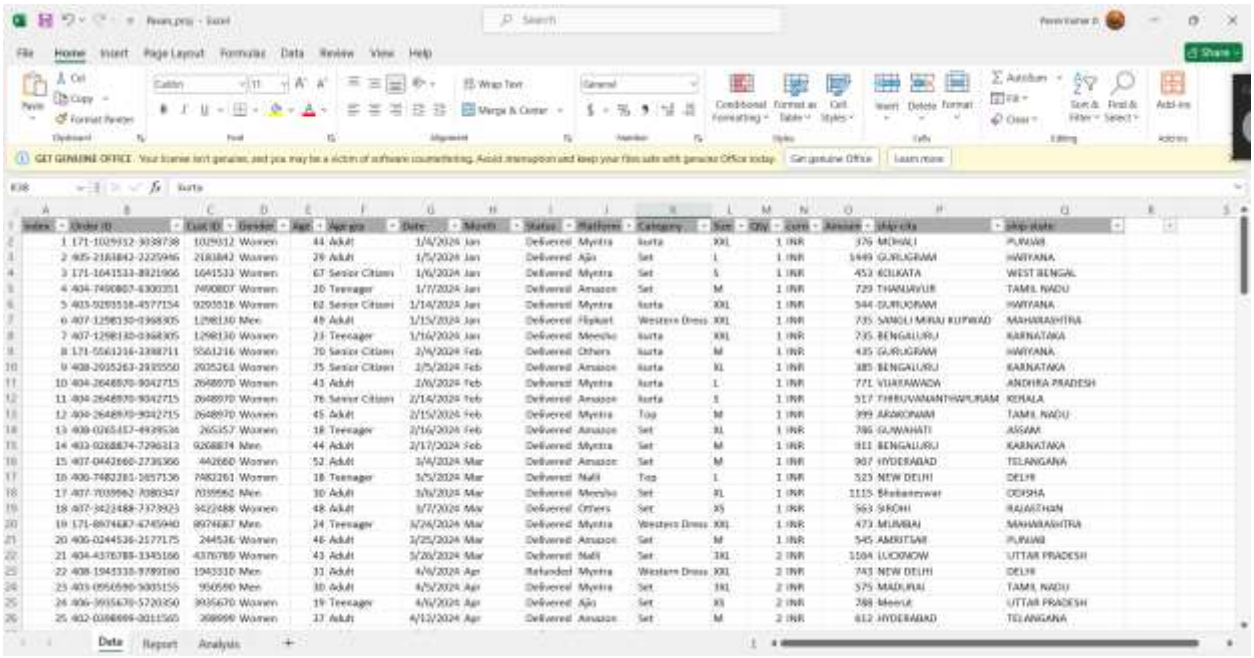
Range: A group of cells.

Workbook: An Excel file containing one or more worksheets.

Importing Data

From Text Files: Use the "Import Text Wizard" for CSV and TXT files.

From Other Applications: Use options like "Get External Data" to import data from Access, SQL, or other sources.



3. Data Cleaning and Preparation

Removing Duplicates

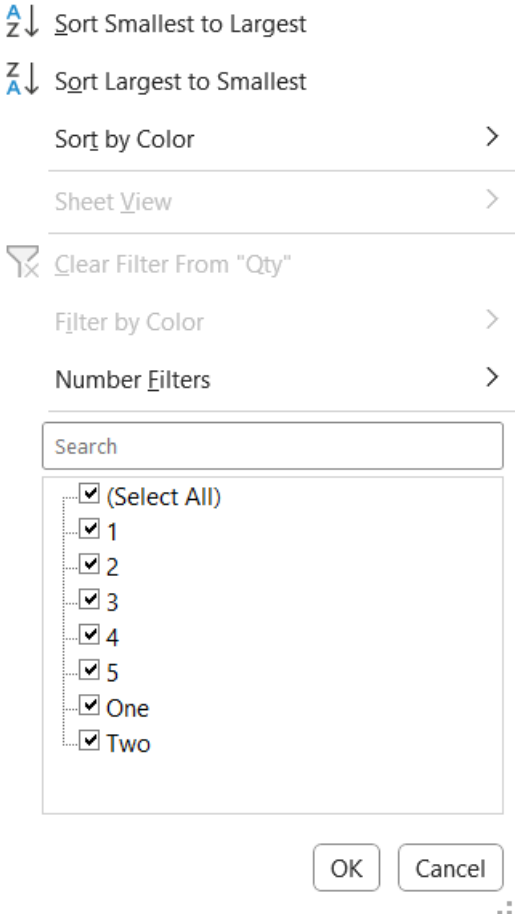
- Go to the "Data" tab, select "Remove Duplicates," and choose the columns to check.

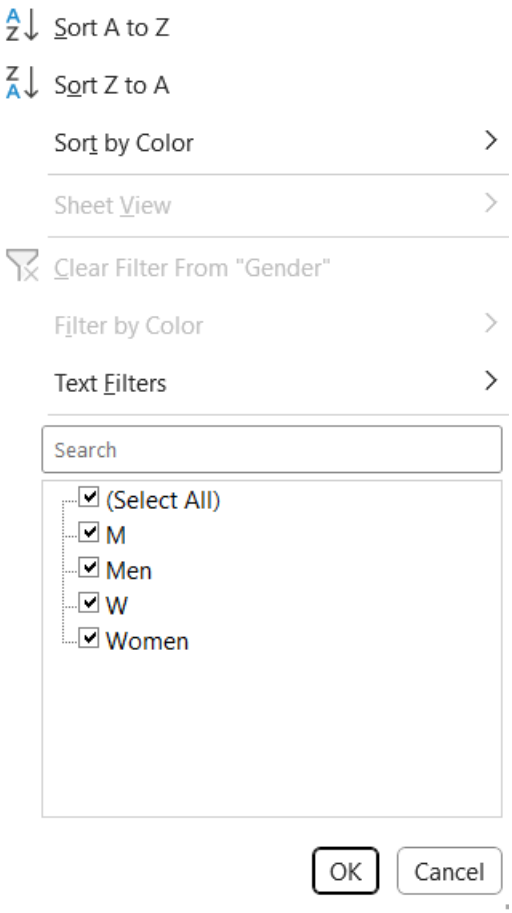
Handling Missing Values

- Use functions like `IFERROR`, `ISBLANK`, and fill missing data with appropriate values or use data imputation techniques.

Data Formatting and Validation

- Format cells (e.g., dates, currency) and use Data Validation rules to ensure data integrity.





4. Exploratory Data Analysis (EDA)

Creating Visualizations

Conditional Formatting:

- Apply formatting rules to highlight important data points.
- Also compare, apply some conditions for Analyzing
  - Comparing sales and orders using a Chart
  - Month got the Highest sales and orders
  - Highest purchase (Men/Women)
  - Different order status
  - Top 5 Selling states
  - Relation between age and gender based on sales
  - Which platform has Maximum Sales
  - Highest selling Category

Top 5 Sales States

Row Labels	Sum of Amount
MAHARASHTRA	9591
KARNATAKA	8279
TAMIL NADU	6389
HARYANA	6171
UTTAR PRADESH	6074

Men vs Women

Row Labels	Sum of Amount
Men	23246
Women	43884

Order Status

Row Labels	Count of Order ID
Cancelled	2
Delivered	95
Refunded	2
Returned	1

Orders: Age and Gender

Count of Order	Column Labels	
Row Labels	Men	Women
Adult	21.00%	45.00%
Senior Citizen	3.00%	12.00%
Teenager	6.00%	13.00%

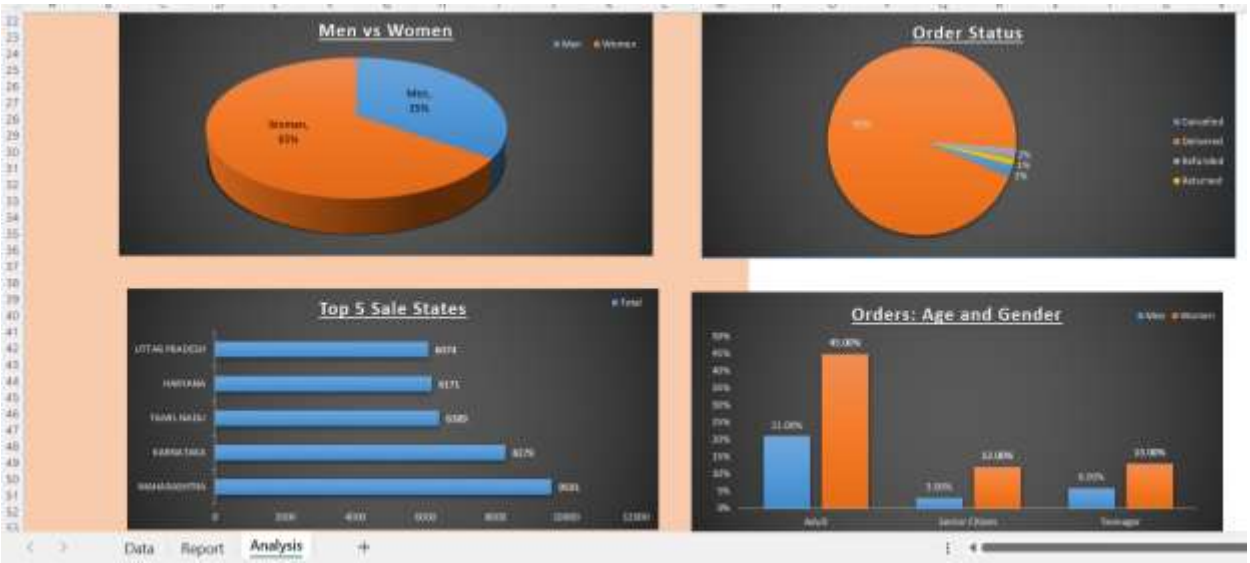
Sales vs Orders

Row Labels	Sum of Qty	Sum of Amount
Jan	7	5021
Feb	7	4204
Mar	8	5350
Apr	12	3735
May	12	3958
Jun	14	5528
Jul	33	8415
Aug	27	7033
Sep	24	3176
Oct	72	11067
Nov	35	4406
Dec	19	5237

Max used Platform

Row Labels	Sum of Qty
Ajio	2.96%
Amazon	32.22%
Flipkart	18.89%
Meesho	4.44%
Myntra	28.89%
Nalli	3.70%
Others	8.89%

Charts and Graphs: Insert charts (e.g., bar, line, pie) from the "Insert" tab.



5. Data Manipulation Techniques

Sorting and Filtering Data

- Use the "Sort & Filter" options in the "Data" tab to organize and view specific subsets of data.
- Using Formulas and Functions
- PivotTables and Pivot Charts
  - Create PivotTables to summarize data and Pivot Charts to visualize these summaries.

9. Best Practices and Tips

Organizing Your Workbook

Use meaningful names for sheets, columns, and ranges. Keep data and calculations separate.

Documenting Your Analysis

Use comments and text boxes to explain your analysis and make your workbook more understandable.

Ensuring Accuracy

Double-check formulas and data sources. Use Excel's auditing tools to track formula dependencies.

10.Sample Insights

- Women are more likely to buy compared to men (65%)
- Maharashtra, Karnataka and Tamil Nadu are top 3
- Adult age Group (25-60) is max contributing (66%)
- Amazon, Myntra and Flipkart are Max Contributors

11. Conclusion

Target **Women** Customers of age group (25-60) living in Maharashtra, Karnataka and Tamil Nadu by showing ads/offers/coupons available on Amazon, Myntra and Flipkart for More Sales and Profits.