



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.

Index	Order ID	Customer ID	Gender	Age	Age group	Date	Month	Status	Platform	Category	Size	Qty	Price	Address	City
1	171-102932-3038738	102932	Women	44	Adult	1/4/2024	Jan	Delivered	Mytra	Kurt	XXL	1	1 INR	376 MIDHALL	PUNJAB
2	405-2181842-2225946	2181842	Women	29	Adult	1/5/2024	Jan	Delivered	Aja	Set	L	1	1 INR	5449 GURUGRAM	HARYANA
3	171-1641533-8821966	1641533	Women	67	Senior Citizen	1/6/2024	Jan	Delivered	Mytra	Set	S	1	1 INR	453 KOLKATA	WEST BENGAL
4	404-7490807-4300351	7490807	Women	20	Teenager	1/7/2024	Jan	Delivered	Amazon	Set	M	1	1 INR	729 THANJAVUR	TAMIL NADU
5	403-5295518-4577154	5295518	Women	62	Senior Citizen	1/14/2024	Jan	Delivered	Mytra	Kurt	XXL	1	1 INR	344 SURUDURAM	HARYANA
6	407-1298130-0368305	1298130	Men	49	Adult	1/15/2024	Jan	Delivered	Flipkart	Western Dress	XXL	1	1 INR	735 SAMOLI MURAI KUPWAD	MAHARASHTRA
7	407-1298130-0368305	1298130	Women	23	Teenager	1/16/2024	Jan	Delivered	Meesho	Kurt	XXL	1	1 INR	735 BENGALURU	KARNATAKA
8	171-5561216-1388711	5561216	Women	70	Senior Citizen	2/4/2024	Feb	Delivered	Others	Kurt	M	1	1 INR	435 GURUGRAM	HARYANA
9	408-2955263-2935550	2955263	Women	35	Senior Citizen	2/5/2024	Feb	Delivered	Amazon	Kurt	XL	1	1 INR	385 BENGALURU	KARNATAKA
10	404-2648870-9042715	2648870	Women	43	Adult	2/6/2024	Feb	Delivered	Mytra	Kurt	L	1	1 INR	771 VIKRAMAWADA	ANDHRA PRADESH
11	404-2648870-9042715	2648870	Women	76	Senior Citizen	2/14/2024	Feb	Delivered	Amazon	Kurt	S	1	1 INR	317 THIRUVANANTHAPURAM	KERALA
12	404-2648870-9042715	2648870	Women	45	Adult	2/15/2024	Feb	Delivered	Mytra	Top	M	1	1 INR	399 ARMOONAM	TAMIL NADU
13	409-0263157-4938536	265157	Women	18	Teenager	2/16/2024	Feb	Delivered	Amazon	Set	XL	1	1 INR	786 GUWAHATI	ASSAM
14	403-0268874-7296312	0268874	Men	44	Adult	2/17/2024	Feb	Delivered	Amazon	Set	M	1	1 INR	811 BENGALURU	KARNATAKA
15	407-0442660-2736360	442660	Women	52	Adult	2/4/2024	Mar	Delivered	Amazon	Set	M	1	1 INR	907 HYDERABAD	TELANGANA
16	406-7482381-2657136	7482381	Women	18	Teenager	3/5/2024	Mar	Delivered	Nalli	Top	L	1	1 INR	323 NEW DELHI	DELHI
17	407-7039962-7080347	7039962	Men	30	Adult	3/6/2024	Mar	Delivered	Meesho	Set	XL	1	1 INR	1115 Bhakaneswar	ODISHA
18	407-3421488-7373923	3421488	Women	48	Adult	3/7/2024	Mar	Delivered	Others	Set	XS	1	1 INR	563 SARONI	RAJASTHAN
19	171-8874687-4745940	8874687	Men	24	Teenager	3/24/2024	Mar	Delivered	Mytra	Western Dress	XXL	1	1 INR	473 MURBAJ	MAHARASHTRA
20	406-0244526-2577175	244526	Women	46	Adult	3/25/2024	Mar	Delivered	Amazon	Set	M	1	1 INR	545 AMRITGARH	PUNJAB
21	408-1945335-9799180	1945335	Women	43	Adult	5/26/2024	Mar	Delivered	Nalli	Set	3XL	2	2 INR	1104 LUCKNOW	UTTAR PRADESH
22	408-1945335-9799180	1945335	Men	31	Adult	4/6/2024	Apr	Delivered	Meesho	Western Dress	XXL	2	2 INR	743 NEW DELHI	DELHI
23	403-0956590-5035155	505690	Men	30	Adult	4/5/2024	Apr	Delivered	Mytra	Set	XXL	2	2 INR	575 MADURAI	TAMIL NADU
24	406-3935470-5720350	3935470	Women	19	Teenager	4/6/2024	Apr	Delivered	Aja	Set	XS	2	2 INR	288 Meerut	UTTAR PRADESH
25	402-0386894-3011565	386999	Women	27	Adult	4/12/2024	Apr	Delivered	Amazon	Set	M	2	2 INR	812 HYDERABAD	TELANGANA

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.

A↓ Sort Smallest to Largest
 Z↓ Sort Largest to Smallest
 Sort by Color >
 Sheet View >
 Clear Filter From "Qty"
 Filter by Color >
 Number Filters >
 Search
☒ (Select All)
☒ 1
☒ 2
☒ 3
☒ 4
☒ 5
☒ One
☒ Two
 OK Cancel

A↓ Sort A to Z
 Z↓ Sort Z to A
 Sort by Color >
 Sheet View >
 Clear Filter From "Gender"
 Filter by Color >
 Text Filters >
 Search
☒ (Select All)
☒ M
☒ Men
☒ W
☒ Women
 OK Cancel

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 PRADES	6074	

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%

Men vs Women		
Row Labels	Sum of Amount	
Men	23246	
Women	43884	

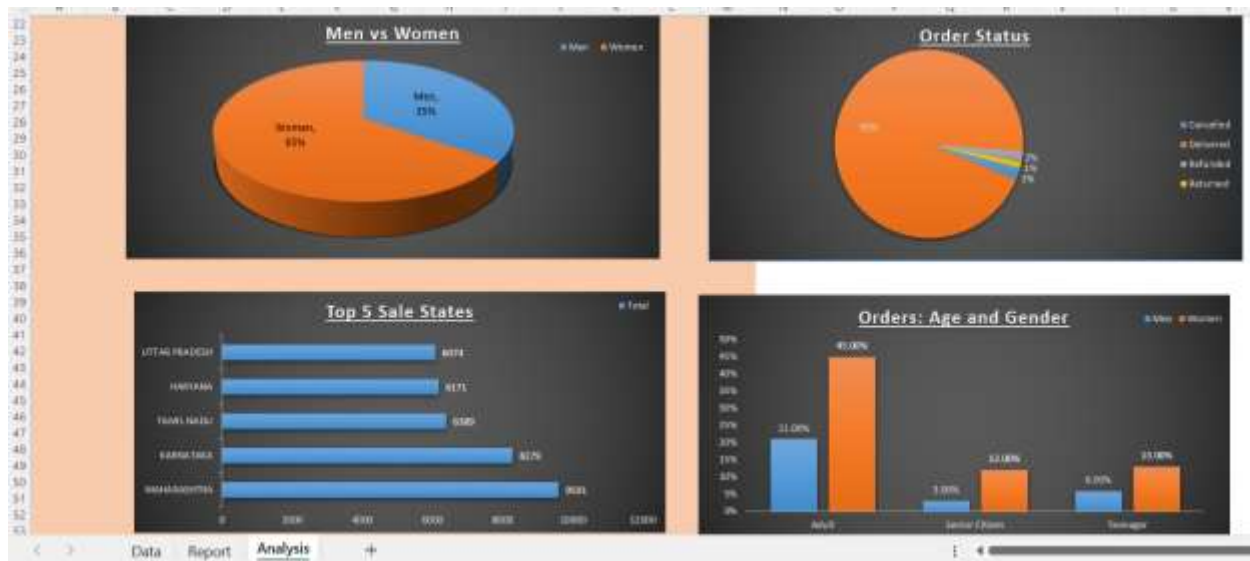
Order Status		
Row Labels	Count of Order ID	
Cancelled	2	
Delivered	95	
Refunded	2	
Returned	1	

Sales vs Orders		
Row Labels	Sum of Qty	Sum of Amount
Jan	7	9021
Feb	7	4204
Mar	8	5350
Apr	12	3735
May	12	3958
Jun	14	5528
Jul	13	8435
Aug	22	2033
Sep	24	3176
Oct	72	11067
Nov	35	4404
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.