

# DS Practical

## PRACTICAL 1

### Aim: Introduction to Excel

1. Create a dataset
2. Sort Data
3. Apply Filter
4. Write Formula
5. Remove duplicates



### Steps to Perform

#### 1. Create a Dataset

- Open **Microsoft Excel**.
- In **Sheet1**, create a small dataset (example: Student data).

Name	Age	City	Marks
Ayesha	20	Delhi	85
Rahul	22	Mumbai	90
Neha	21	Delhi	78
Aman	22	Pune	85
Riya	20	Mumbai	90

- Make sure each column has a **header** (like Name, Age, etc.).

---

#### 2. Sort Data

- Select the dataset (including headers).
- Go to **Data → Sort**.
- Choose the column to sort (e.g., *Marks*).

- Choose **Ascending** (A to Z / Smallest to Largest) or **Descending** (Z to A / Largest to Smallest).
  - Click **OK**.
  - ✓ Data will now appear in sorted order.
- 

### 3. Apply Filter

- Select your dataset.
  - Go to **Data → Filter**.
  - Small dropdown arrows will appear on each header.
  - Click a dropdown → select or deselect values to display only specific data.  
(Example: Filter City = Delhi)
- 

### 4. Write a Formula

- Click an empty cell (e.g., next to Marks).
  - Try some formulas like:
    - **=AVERAGE(D2:D6)** → Finds average marks
    - **=MAX(D2:D6)** → Finds highest marks
    - **=MIN(D2:D6)** → Finds lowest marks
    - **=COUNT(D2:D6)** → Counts number of entries
- 

### 5. Remove Duplicates

- Select your entire dataset.
- Go to **Data → Remove Duplicates**.
- Select the columns you want to check (e.g., Name, City, etc.).
- Click **OK**.
- ✓ Excel will remove duplicate rows and show how many were deleted.

## PRACTICAL 2

### Aim: Data Validation in Excel

1. Create a dataset
2. Text to Column
3. Add Data using a Form tool
4. Transpose Data
5. Use VLOOKUP

### Steps to Perform

#### 1. Create a Dataset

- Open **Microsoft Excel**.
- In **Sheet1**, enter a small dataset like this:

ID	Name	Department	Salary
101	Ayesha Khan	HR	40000
102	Rahul Mehta	IT	55000
103	Neha Sharma	Finance	48000
104	Aman Verma	IT	52000
105	Riya Patel	HR	40000

#### 2. Text to Columns

- Suppose you have a column like **Full Name** (e.g., "Ayesha Khan").
  - Select that column → Go to **Data** → **Text to Columns**.
  - Choose **Delimited** → **Next** → **Space** → **Finish**.
- ✓ The full name will split into two columns: *First Name* and *Last Name*.

#### 3. Add Data Using the Form Tool

- Ensure your dataset has headers (ID, Name, Department, Salary).

- Go to **File** → **Options** → **Customize Ribbon** → **Choose Commands Not in the Ribbon**.
  - Add **Form** to your Quick Access Toolbar.
  - Select your data range → Click **Form** icon.
  - A data entry form appears — use it to easily **add, find, or delete** records.
- ✓ New entries will automatically appear in your dataset.
- 

## 4. Transpose Data

- Select the dataset → Copy it (**Ctrl + C**).
  - Go to a blank area → Right-click → **Paste Special** → **Transpose**.
- ✓ Rows become columns and columns become rows.
- 

## 5. Use VLOOKUP

- In a new cell, try a lookup formula to find salary by ID.

Example:

```
=VLOOKUP(103, A2:D6, 4, FALSE)
```

- ✓ This finds the salary of the employee with ID 103 from the dataset.

## PRACTICAL 3

### Aim: Data Validation in Excel

1. Specify a valid range of value for a cell
2. Specify a valid List of value for a cell
3. Specify a valid Birth of value for a cell

### Steps to Perform

#### 1. Specify a Valid Range of Values for a Cell

- Select the cell(s) where you want to restrict input.  
(Example: Cells in the "Marks" column.)
  - Go to **Data → Data Validation → Data Validation**.
  - In the **Settings** tab:
    - Choose **Allow:** *Whole Number* (or *Decimal*).
    - Choose **Data:** *between*.
    - Enter **Minimum:** 0 and **Maximum:** 100.
  - Click **OK**.
- ✓ Now, only numbers between 0 and 100 can be entered in those cells.
- 

## 2. Specify a Valid List of Values for a Cell

- Select the target cell(s) (e.g., under "Department").
- Go to **Data → Data Validation → Data Validation**.
- In the **Settings** tab:
  - Choose **Allow:** *List*.
  - In **Source**, type:

HR, IT, Finance, Marketing

- Click **OK**.
- ✓ The cell now shows a dropdown menu — only these listed values can be selected.
- 

## 3. Specify a Valid Birth Date Range for a Cell

- Select the cell(s) for Date of Birth entries.
- Go to **Data → Data Validation → Data Validation**.
- In **Settings**:

- Choose **Allow:** *Date*.
  - Choose **Data:** *between*.
  - Enter a valid date range, for example:
    - **Start Date:** 1/1/1990
    - **End Date:** 12/31/2010
  - Click **OK**.
- ✅ Users can only enter birth dates within that specific range.

## PRACTICAL 4

1.Excel Program to plot Bar Chart

2.Excel Program to plot Histogram :-

Name	Math	Science
Om	45	39
Adit	26	45
Siddhi	36	39
Manasi	37	46
Yash	29	26
Komal	45	34

Student	Marks
A	12
B	18
C	22
D	25
E	29
F	33
G	35
H	37
I	41
J	55
K	57
L	60
M	63
N	65

3.Excel Program to plot pie Chart

4.Boxplot



## 1. Excel Program to Plot a Bar Chart

**Dataset:**

Name	Math	Science
Om	45	39
Adit	26	45
Siddhi	36	39
Manasi	37	46
Yash	29	26
Komal	45	34

**Steps:**

1. Enter the above data in Excel.
2. Select the entire dataset (A1:C7).

3. Go to **Insert → Charts → Bar Chart → Clustered Bar**.
4. Excel will display a **bar chart** comparing Math and Science marks for each student.
5. Add chart elements:
  - Click **Chart Elements (+)** → add **Chart Title** → type *"Comparison of Math and Science Marks"*.
  - Add **Data Labels** if needed for clarity.

✓ **Result:** A Bar Chart showing each student's performance in Math and Science.

---



## 2. Excel Program to Plot a Histogram

**Dataset:**

Student	Marks
A	12
B	18
C	22
D	25
E	29
F	33
G	35
H	37
I	41
J	55
K	57
L	60
M	63
N	65

**Steps:**

1. Enter the above data in Excel.



2. Select the **Marks** column (B2:B15).
3. Go to **Insert → Charts → Insert Statistic Chart → Histogram**.
4. Excel automatically groups marks into “bins” (ranges like 10–20, 20–30, etc.).
5. Adjust bins if needed:
  - Right-click the X-axis → **Format Axis → Bin width** → set a custom width (e.g., 10).

✅ **Result:** A Histogram showing the frequency distribution of students' marks.



### 3. Excel Program to Plot a Pie Chart

**Dataset:**

Student	Marks
A	12
B	18
C	22
D	25
E	29
F	33
G	35
H	37
I	41
J	55
K	57
L	60
M	63
N	65

**Steps:**

1. Open Excel and enter the above dataset.

2. Select both columns (A1:B15).
3. Go to **Insert → Charts → Pie Chart → 2-D Pie**.
4. A Pie Chart will appear showing each student's contribution to the total marks.
5. Add details:
  - Click the chart → **Add Chart Element → Data Labels → Percentage**.
  - Add a **Chart Title** → e.g., *"Distribution of Marks among Students"*.

✓ **Result:** A **Pie Chart** that visualizes how each student's marks compare as a percentage of the total.

---



## 4. Excel Program to Plot a Boxplot (Box & Whisker Chart)

**Dataset (same as above):**

Student	Marks
A	12
B	18
C	22
D	25
E	29
F	33
G	35
H	37
I	41
J	55
K	57
L	60
M	63
N	65

## Steps:

1. Select only the **Marks** column (B1:B15).
2. Go to **Insert** → **Charts** → **Statistical Chart** → **Box & Whisker**.
3. Excel generates a **Boxplot** showing:
  - **Median (middle mark)**
  - **Quartiles (Q1 & Q3)**
  - **Minimum and Maximum values**
  - **Outliers**, if any
4. Add a **Chart Title** → *"Boxplot of Student Marks"*.
5. Optionally format the chart for clarity (right-click → *Format Data Series*).

✅ **Result:** A **Boxplot** representing the spread and variation of marks — showing how students are distributed across the score range.



## Final Result Summary

Chart Type	Purpose	Visualization
<b>Pie Chart</b>	To show percentage contribution of each student's marks	Circular chart divided into slices
<b>Boxplot</b>	To show the data spread, median, and outliers	Box with whiskers representing quartiles

## PRACTICAL 5

Aim: Measure of Central Tendency Calculate Mean, Median and Mode for the following data

- a. 15,16,18,19,20,24
- b. 45,48,47,43,42
- c. 100,105,115,115,106



## Aim:

To calculate the **Mean**, **Median**, and **Mode** for the given datasets using Microsoft Excel.

---

## Given Data

Dataset	Values
(a)	15, 16, 18, 19, 20, 24
(b)	45, 48, 47, 43, 42
(c)	100, 105, 115, 115, 106

---

## Steps to Perform in Excel

### 1 Enter Data

- Open **Microsoft Excel**.
- In **Column A**, enter data as follows:

A	B
<b>a)</b>	15
	16
	18
	19
	20
	24
<b>b)</b>	45
	48
	47
	43
	42
<b>c)</b>	100
	105
	115

A	B
	115
	106

## 2 Use Excel Formulas

For each dataset, apply these formulas in separate cells:

Measure	Excel Formula	Meaning
<b>Mean</b>	<code>=AVERAGE(range)</code>	Finds the arithmetic average
<b>Median</b>	<code>=MEDIAN(range)</code>	Finds the middle value
<b>Mode</b>	<code>=MODE.SNGL(range)</code>	Finds the most frequent value



## Final Results Summary

Dataset	Mean	Median	Mode
(a)	18.67	18.5	None
(b)	45	45	None
(c)	108.2	115	115

### 2. Find Mean

## Aim:

To calculate the **Mean** for the given **Class Interval and Frequency** data using Microsoft Excel.

## Given Data

Class Interval	Frequency
1	100
2	104
3	106
4	107
5	101

Class Interval	Frequency
6	103
7	105


## Steps to Perform in Excel

### Step 1: Enter the Data

Open Excel and type the data as shown:

A (Class Interval)	B (Frequency)
1	100
2	104
3	106
4	107
5	101
6	103
7	105

### Step 2: Compute $fx$

- In **Column C**, type this header: .
- In cell **C2**, enter the formula:

```
=A2*B2
```

- Drag it down till the last row.

You'll get:

Class Interval (x)	Frequency (f)	$fx$
1	100	100
2	104	208

Class Interval (x)	Frequency (f)	fx
3	106	318
4	107	428
5	101	505
6	103	618
7	105	735

### Step 3: Find $\Sigma f$ and $\Sigma fx$

At the bottom of the Frequency and fx columns:

- In B9: `=SUM(B2:B8)` → gives  $\Sigma f = 726$
- In C9: `=SUM(C2:C8)` → gives  $\Sigma fx = 2912$

### Step 4: Calculate Mean

In a new cell (say E2), enter the formula:

`=SUM(C2:C8)/SUM(B2:B8)`

✓ Mean =  $\Sigma fx / \Sigma f = 2912 / 726 = 4.01$



### Result

Measure	Formula	Result
Mean	$\Sigma fx / \Sigma f$	4.01

### ✓ Final Answer

Mean = 4.01

3. Find Mean or Median

### Given Data

Class Interval	Frequency (f)
0 – 100	13
100 – 200	18
200 – 300	27
300 – 400	34
400 – 500	23
500 – 600	17
600 – 700	10

## 1) Sheet layout (enter these headings and values)

Put these headings in row 1 and the values in rows 2–8:

A1: Class  
 B1: Frequency  
 C1: Lower  
 D1: Upper  
 E1: Midpoint  
 F1:  $f \cdot x$   
 G1: CumFreq

Fill rows 2–8 like this:

A2: 0-100    B2: 13    C2: 0    D2: 100  
 A3: 100-200    B3: 18    C3: 100    D3: 200  
 A4: 200-300    B4: 27    C4: 200    D4: 300  
 A5: 300-400    B5: 34    C5: 300    D5: 400  
 A6: 400-500    B6: 23    C6: 400    D6: 500  
 A7: 500-600    B7: 17    C7: 500    D7: 600  
 A8: 600-700    B8: 10    C8: 600    D8: 700



## 2) Calculate Midpoint and $f \times x$ (drag formulas down row 2 → 8)

E2 (Midpoint):

```
=(C2+D2)/2
```

Drag E2 down to E8.

F2 ( $f \times x$ ):

```
=B2*E2
```

Drag F2 down to F8.

G2 (Cumulative frequency):

```
=SUM($B$2:B2)
```

Drag G2 down to G8.

## 3) Totals (put below the table, e.g. row 9)

B9 ( $\Sigma f$ ):

```
=SUM(B2:B8)
```

F9 ( $\Sigma f \times x$ ):

```
=SUM(F2:F8)
```

## 4) Mean (grouped data formula)

Put Mean in a cell, e.g. B11:

```
=F9/B9
```

This computes Mean =  $\Sigma(fx) / \Sigma f$ .

(With this dataset you should get about **339.44**.)

## 7) Copy-friendly summary (paste into Excel comments or notes)

1. Enter classes in A2:A8 and frequencies in B2:B8.
2. Put Lower/Upper limits in C2:D8.
3. E2 = `(C2+D2)/2` → drag down (midpoints).
4. F2 = `B2*E2` → drag down ( $f \times x$ ).
5. G2 = `SUM($B$2:B2)` → drag down (cum freq).
6. B9 = `SUM(B2:B8)` (N).
7. F9 = `SUM(F2:F8)` ( $\Sigma fx$ ).
8. Mean = `=F9/B9`.
9. B12 = `=B9/2`.
10. B13 = `=MATCH(TRUE,INDEX(G2:G8>=B12,0),0)` (median class index) — or  
`=MATCH(B12,G2:G8,1)+1`.
11. L = `=INDEX(C2:C8,B13)` ; cf\_prev = `=IF(B13=1,0,INDEX(G2:G8,B13-1))` ; f\_med = `=INDEX(B2:B8,B13)` ;  
h = `=INDEX(D2:D8,B13)-INDEX(C2:C8,B13)`.
12. Median = `= L + ((B12 - cf_prev) / f_med) * h`.

# PRACTICAL 6

## Aim:

To calculate **Range, Interquartile Range (IQR), Variance, Standard Deviation, Skewness, and Kurtosis** of the given datasets.

---

## Step 1: Enter Data

### Dataset 1: Student Marks

A	B
Student	Marks
A	45
B	52
C	50
D	60
E	65
F	70
G	75
H	80
I	85
J	90
K	78
L	65
M	95
N	65
O	46
P	59
Q	84
R	61
S	95
T	86

(Enter this from cell A1:B21)

## Dataset 2: List of Things

D	E
Item	Price
Book	58
Pen	65
Pencil	48
Calculator	52
Highlighter	43
Eraser	62
Scissors	45

(Enter this from cell D1:E8)

## Step 2: Calculate in Excel (Formulas)

Measure	Excel Formula	Example (Marks in B2:B21)	Example (Prices in E2:E8)
Range	<code>=MAX(range)-MIN(range)</code>	<code>=MAX(B2:B21)-MIN(B2:B21)</code>	<code>=MAX(E2:E8)-MIN(E2:E8)</code>
Q1 (First Quartile)	<code>=QUARTILE.INC(range,1)</code>	<code>=QUARTILE.INC(B2:B21,1)</code>	<code>=QUARTILE.INC(E2:E8,1)</code>
Q3 (Third Quartile)	<code>=QUARTILE.INC(range,3)</code>	<code>=QUARTILE.INC(B2:B21,3)</code>	<code>=QUARTILE.INC(E2:E8,3)</code>
IQR (Interquartile Range)	<code>=Q3-Q1</code>	<code>=QUARTILE.INC(B2:B21,3)-QUARTILE.INC(B2:B21,1)</code>	<code>=QUARTILE.INC(E2:E8,3)-QUARTILE.INC(E2:E8,1)</code>
Mean	<code>=AVERAGE(range)</code>	<code>=AVERAGE(B2:B21)</code>	<code>=AVERAGE(E2:E8)</code>
Variance	<code>=VAR.P(range)</code> (for population) or <code>=VAR.S(range)</code> (for sample)	<code>=VAR.P(B2:B21)</code>	<code>=VAR.P(E2:E8)</code>

Measure	Excel Formula	Example (Marks in B2:B21)	Example (Prices in E2:E8)
<b>Standard Deviation</b>	<code>=STDEV.P(range)</code> (for population) or <code>=STDEV.S(range)</code> (for sample)	<code>=STDEV.P(B2:B21)</code>	<code>=STDEV.P(E2:E8)</code>
<b>Skewness</b>	<code>=SKEW(range)</code>	<code>=SKEW(B2:B21)</code>	<code>=SKEW(E2:E8)</code>
<b>Kurtosis</b>	<code>=KURT(range)</code>	<code>=KURT(B2:B21)</code>	<code>=KURT(E2:E8)</code>

## Step 3: Format Your Output Table

Make a small result table below your data, like this:

Measure	Formula Used	Result
Range	<code>=MAX(B2:B21)-MIN(B2:B21)</code>	...
IQR	<code>=QUARTILE.INC(B2:B21,3)-QUARTILE.INC(B2:B21,1)</code>	...
Mean	<code>=AVERAGE(B2:B21)</code>	...
Variance	<code>=VAR.P(B2:B21)</code>	...
SD	<code>=STDEV.P(B2:B21)</code>	...
Skewness	<code>=SKEW(B2:B21)</code>	...
Kurtosis	<code>=KURT(B2:B21)</code>	...

Then repeat for the second dataset (Prices).

## ✓ Step 4: Interpret Results

After typing each formula:

- Press **Enter** to see the calculated value.
- You can adjust decimal points using **Home** → **Increase/Decrease Decimal**.
- Add titles and borders for a neat practical report.

## Optional: Visual Representation

To visualize dispersion:

1. Select your data.
2. Go to **Insert** → **Chart** → **Box & Whisker**.
  - It automatically shows **median**, **quartiles**, and **outliers**.