

Week 4 by: Manisha Pal

Week 1.4:

Introduction to Pivot Tables

In videos professors was just discussing the dataset and doing analysis here I have made notes for the formulas, techniques and charts they were using

Pivot tables are a powerful feature in Excel that allows you to quickly summarize, analyze, explore, and present your data. They enable you to transform data into insightful reports and charts, making it easier to identify trends, patterns, and relationships within your dataset.

Why Use Pivot Tables?

- **Efficient Data Analysis:** Pivot tables can handle large datasets and help you perform complex calculations, such as sums, averages, counts, and percentages, without writing formulas.
- **Dynamic Reports:** Pivot tables automatically update when the underlying data changes, allowing for continuous analysis.
- **Interactive Data Exploration:** You can easily rearrange fields, apply filters, and sort data to explore different aspects of your data.

Creating a Pivot Table

1. **Select the Data Range:**
 - Highlight the range of data you want to analyze, including headers (e.g., A1)
2. **Insert the Pivot Table:**
 - Go to the Insert tab in Excel.
 - Click Pivot Table.
 - Choose whether to place the pivot table in a new worksheet or an existing one, then click OK.
3. **Understanding the Pivot Table Layout:**
 - **Rows:** Fields placed here will appear as row labels in the pivot table.
 - **Columns:** Fields placed here will appear as column labels.
 - **Values:** Fields placed here are summarized (e.g., count, sum) based on the data.
 - **Filters:** Fields placed here can filter the entire pivot table based on specific criteria.

Basic Analysis with Pivot Tables

- **Counting Data:**
 - Drag a categorical field (e.g., Department) to the Rows area.
 - Drag a numerical or categorical field (e.g., Roll Number) to the Values area.
 - By default, Excel will sum the data, but you can change this to Count by clicking on the field in the Values area, selecting Value Field Settings, and choosing Count.
- **Analyzing Data Across Multiple Dimensions:**

- To analyze data across different categories, drag another field (e.g., Hostel) into the Columns area.
- This will create a matrix, allowing you to see the interaction between rows and columns.

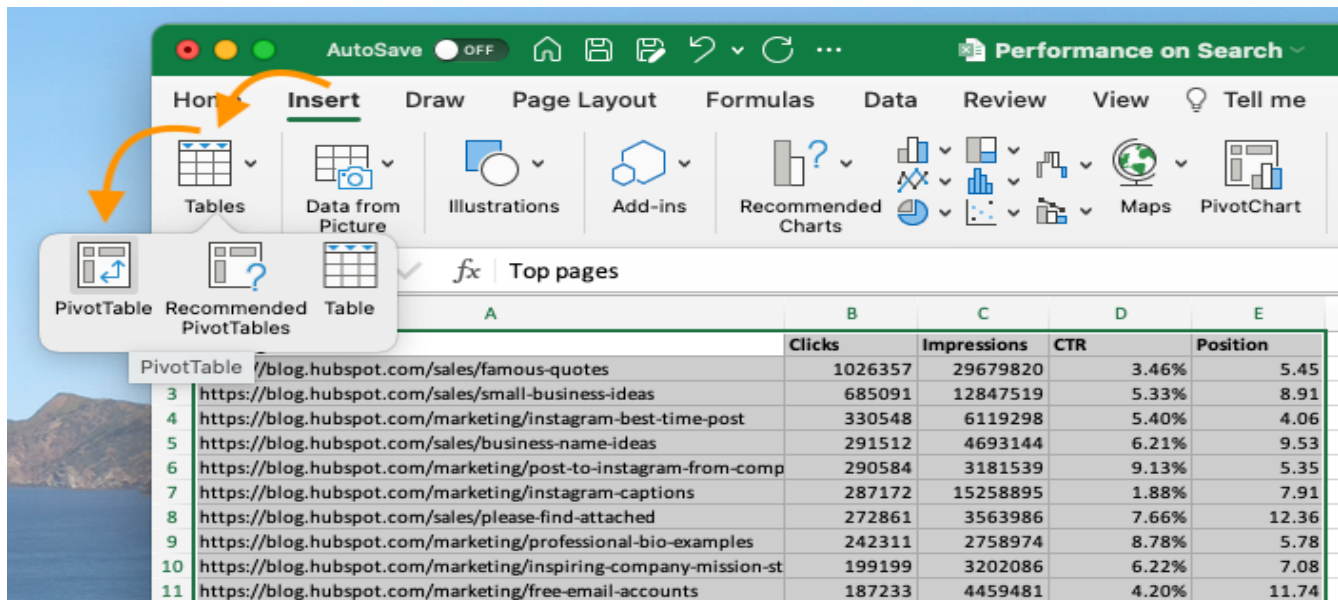
Advanced Features

- Filtering Data:
 - Drag a field into the Filters area to add a filter.
 - You can use this to focus on specific subsets of data (e.g., filtering to see only students from the CS department).
- Using Multiple Fields in Rows and Columns:
 - You can place multiple fields in both the Rows and Columns areas to create more detailed breakdowns.
 - For example, placing Hostel and Department in the Rows area will show a nested structure, breaking down the data by both criteria.

Practical Applications of Pivot Tables

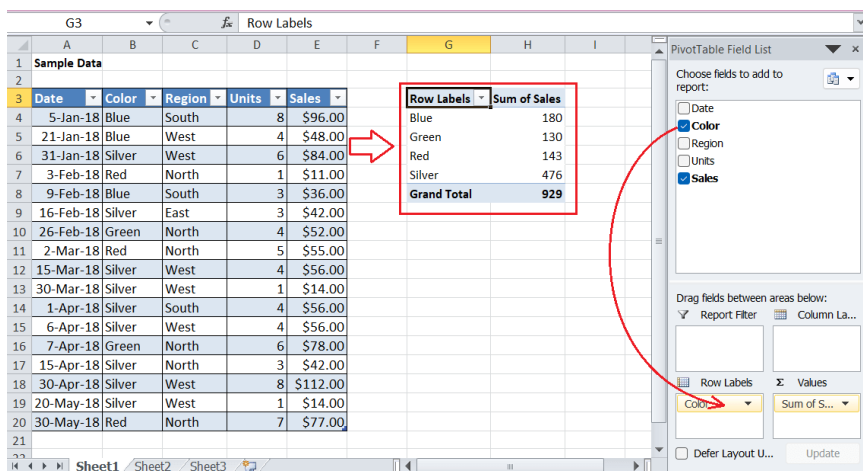
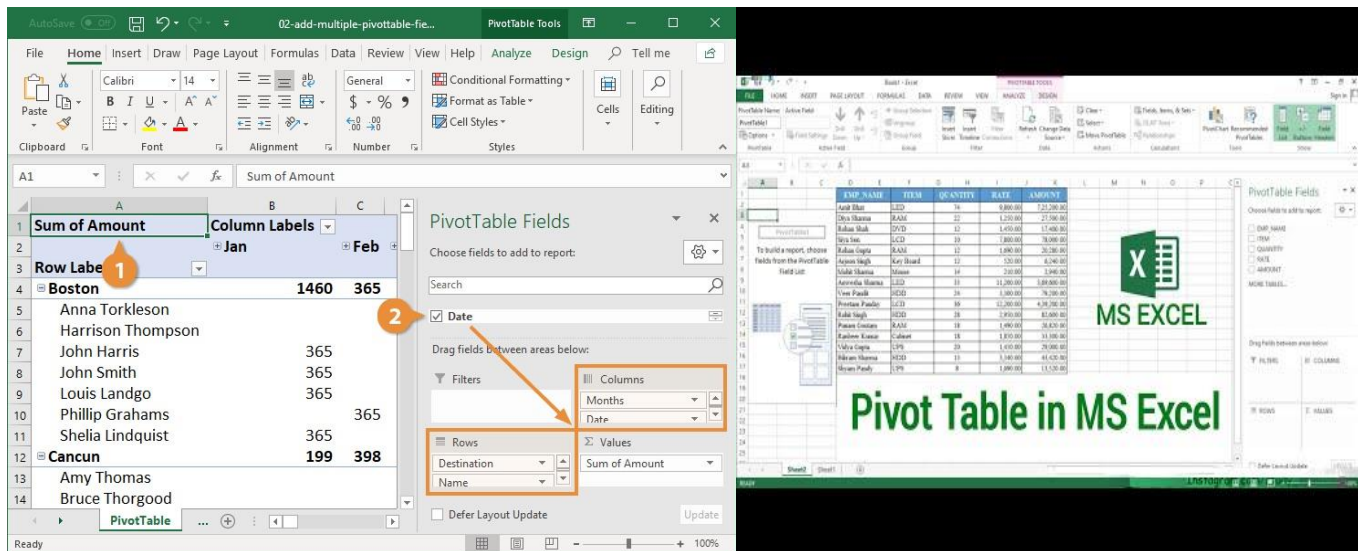
- Dynamic Data Analysis:
 - Pivot tables are dynamic, meaning they automatically update when the underlying data changes.
 - You can refresh the pivot table by right-clicking inside it and selecting Refresh.
- Use Cases:
 - Business Analysis: Summarize sales data by region, product, or sales representative.
 - Academic Analysis: Analyze student performance across different departments or hostels.
 - Event Analysis: Track participation or results in sports events.

Pivot tables are essential for anyone working with large datasets in Excel. By mastering pivot tables, you'll be able to analyze and present your data more effectively, making it easier to derive meaningful insights and make informed decisions.



	A	B	C	D	E
		Clicks	Impressions	CTR	Position
3	https://blog.hubspot.com/sales/famous-quotes	1026357	29679820	3.46%	5.45
4	https://blog.hubspot.com/sales/small-business-ideas	685091	12847519	5.33%	8.91
5	https://blog.hubspot.com/marketing/instagram-best-time-post	330548	6119298	5.40%	4.06
6	https://blog.hubspot.com/sales/business-name-ideas	291512	4693144	6.21%	9.53
7	https://blog.hubspot.com/marketing/post-to-instagram-from-comp	290584	3181539	9.13%	5.35
8	https://blog.hubspot.com/marketing/instagram-captions	287172	15258895	1.88%	7.91
9	https://blog.hubspot.com/sales/please-find-attached	272861	3563986	7.66%	12.36
10	https://blog.hubspot.com/marketing/professional-bio-examples	242311	2758974	8.78%	5.78
11	https://blog.hubspot.com/marketing/inspiring-company-mission-st	199199	3202086	6.22%	7.08
12	https://blog.hubspot.com/marketing/free-email-accounts	187233	4459481	4.20%	11.74

Still facing problem in pivot tableNo problem!! Have a look at this 5 min video below:



Week 4.2

Line Chart:

1. What is a Line Chart?

A line chart is a type of graph used to display information as a series of data points connected by straight line segments. It is commonly used to visualize data trends over intervals of time, such as days, months, quarters, or years.

2. When to Use a Line Chart?

Line charts are ideal when you want to:

- Show trends over time (e.g., sales over the last year).
- Compare changes in different groups over the same period (e.g., the grades of multiple students over several semesters).
- Highlight the rate of change between data points.

3. Components of a Line Chart

- **X-axis (Horizontal Axis):** Represents the time period or category.
- **Y-axis (Vertical Axis):** Represents the values being measured, like grades, sales, or temperature.

- **Data Points:** Represent the value at each time period.
- **Line Segments:** Connect the data points to show the trend.
- **Legends:** Identify which data series corresponds to which line (especially useful when comparing multiple series).

4. Creating a Line Chart in Excel

Let's go through the steps of creating a line chart in Excel, using an example of tracking a student's grades over 6 semesters.

Step-by-Step Guide:

1. Prepare Your Data:

- Structure your data in two columns: one for the time period (e.g., Semesters) and one for the values you want to track (e.g., Grades).

Semester	Grade
----------	-------

Semester 1	78
------------	----

Semester 2	82
------------	----

Semester 3	74
------------	----

Semester 4	88
------------	----

Semester 5	85
------------	----

Semester 6	90
------------	----

2. Select the Data:

- Highlight the data you want to include in the chart. In this case, you would select the range A1:B7.

3. Insert the Line Chart:

- Go to the 'Insert' tab in the Excel ribbon.
- Click on the 'Line Chart' option in the 'Charts' group.
- Choose the style of line chart you prefer (e.g., standard line, line with markers).

5. Advanced Features of Line Charts

- **Multiple Lines:** If you want to compare multiple sets of data (e.g., grades of multiple students), you can plot them on the same line chart. Ensure each dataset is in a separate column.

Semester	Student A	Student B
----------	-----------	-----------

Semester 1	78	85
------------	----	----

Semester 2	82	80
------------	----	----

Semester 3	74	88
------------	----	----

Semester Student A Student B

Semester 4 88 90

Semester 5 85 87

Semester 6 90 92

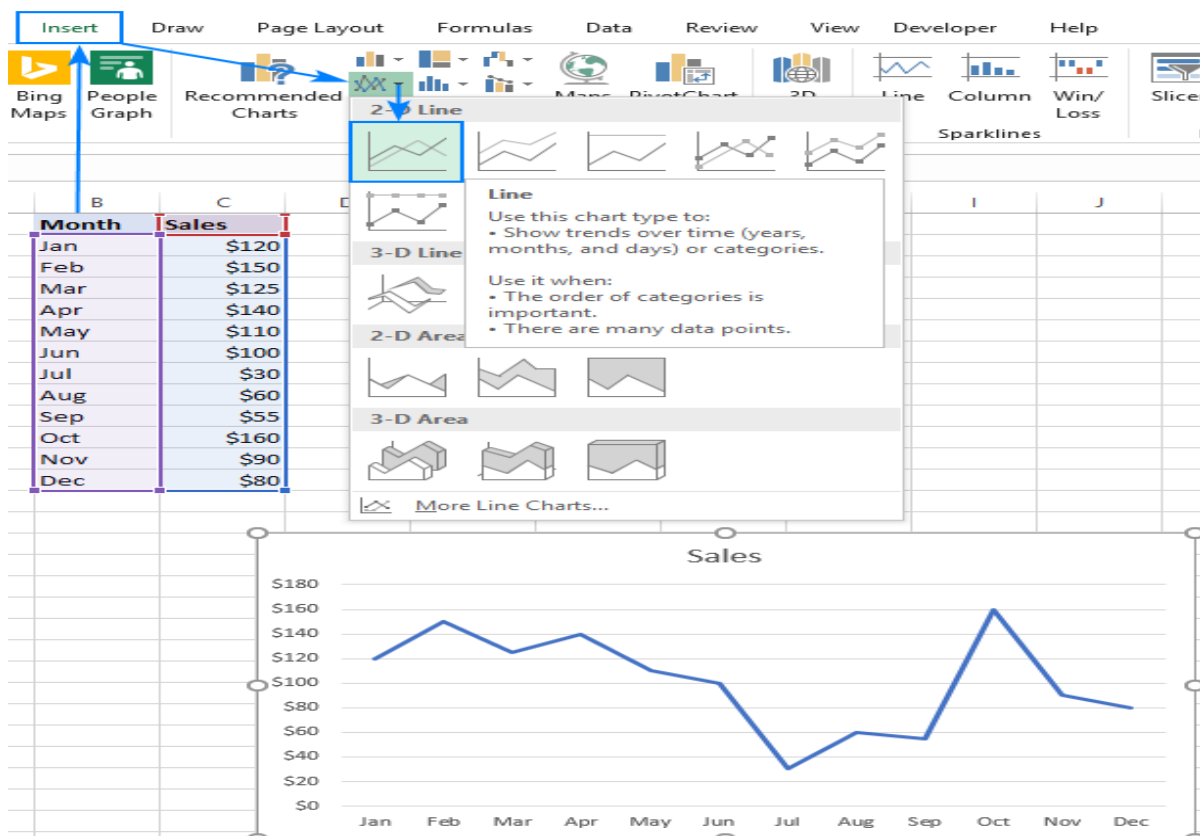
- After selecting the entire data range and inserting a line chart, Excel will automatically create lines for both students.
- Trendlines: Add a trendline to see the general direction of the data over time. This can be done by clicking on a data series in the chart, then selecting 'Add Trendline.'

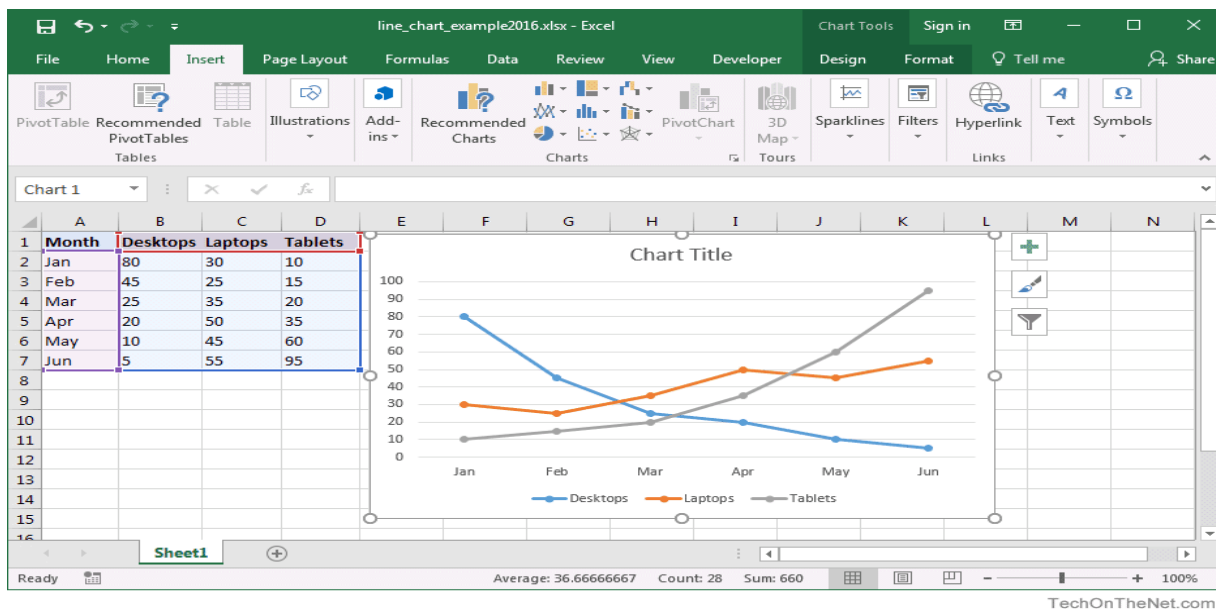
6. Interpretation of Line Charts

- Upward Slope: Indicates an increase in the data over time (e.g., improving grades).
- Downward Slope: Indicates a decrease in the data over time (e.g., declining sales).
- Flat Line: Indicates no change in the data over time (e.g., consistent performance).
- Peaks and Troughs: Indicates variability in the data, showing periods of highs and lows.

7. Practical Applications of Line Charts

- Business: Track sales, revenue, or stock prices over time.
- Education: Compare student performance across different subjects or terms.
- Finance: Monitor expenses, profits, or financial trends.
- Healthcare: Track patient vital signs over time.





Week 4.3

Stacked Bar Charts

What is a Stacked Bar Chart?

A stacked bar chart is a type of bar chart that displays multiple data series stacked on top of each other. It allows you to visualize the cumulative effect of different categories, showing both the total and the individual contributions of each category within a bar. This is useful for comparing the overall size of categories while also breaking down the components that make up each category.

When to Use Stacked Bar Charts?

- **Comparing Total Values:** When you want to compare the total values of different categories.
- **Showing Distribution:** When you need to show the distribution of parts within a whole for each category.
- **Visualizing Trends:** When comparing trends or changes in data across categories over time or between different groups.

Types of Stacked Bar Charts

1. **Standard Stacked Bar Chart:**
 - Displays the absolute values stacked on top of each other.
 - The height of each segment represents the value of that category.
2. **100% Stacked Bar Chart:**
 - Displays the relative contribution of each category as a percentage of the total.
 - Useful when comparing proportional contributions.

How to Create a Stacked Bar Chart in Excel

1. **Prepare Your Data:**
 - Organize your data in columns with each series representing a category.
2. **Select the Data:**

- Highlight the data you want to include in the chart, including the category labels and series data.

3. Insert the Stacked Bar Chart:

- Go to the Insert tab on the Ribbon.
- Click on Bar Chart in the Charts group.
- Choose Stacked Bar or 100% Stacked Bar from the drop-down menu.

4. You can customize the chart and adjust its axis

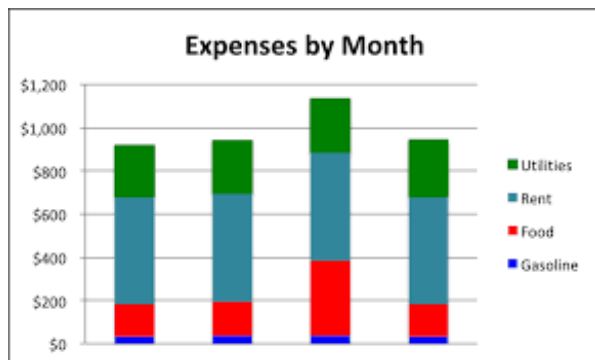
Tips for Using Stacked Bar Charts

- **Avoid Overcrowding:** Limit the number of data series to avoid cluttering the chart.
- **Use Contrasting Colors:** Ensure that the different series are easily distinguishable.
- **Label Clearly:** Make sure your axis labels and data labels are clear and concise.
- **Consider Alternatives:** If the chart becomes too complex, consider using other chart types like grouped bar charts or line charts.

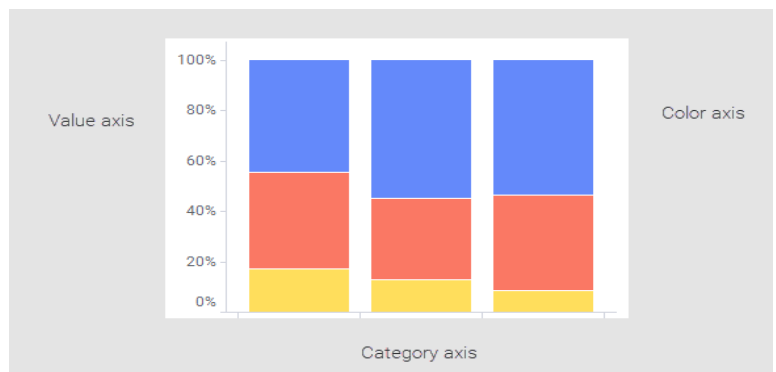
Example Use Case

Suppose you are analyzing sales data across different regions. You can use a stacked bar chart to show the total sales for each region and break down the sales by different product categories within each region.

This method of visualization gives a clear picture of both the total sales and how each product category contributes to the total, making it easier to analyze patterns and trends across regions.



Simple stack bar chart



100 percent stack bar chart

Conditional Formatting in Excel

1. What is Conditional Formatting? Conditional Formatting in Excel allows you to apply specific formatting to cells that meet certain criteria. This feature helps in visualizing data, identifying trends, and spotting important information quickly.

2. Key Features of Conditional Formatting:

- **Highlighting Cells:**
 - **Purpose:** Used to highlight cells that meet a specific condition, such as values above a certain threshold, dates before a certain time, or duplicate entries.

- **How to Use:**

1. Select the range of cells you want to apply conditional formatting to.
2. Go to the “Home” tab and click on “Conditional Formatting.”
3. Choose “Highlight Cells Rules” and then select the condition you want to apply, such as “Greater Than,” “Less Than,” or “Equal To.”
4. Set your condition and formatting style (e.g., font color, fill color).
5. Click “OK” to apply.

- **Data Bars:**

- **Purpose:** Data bars visually represent the values in a range of cells, making it easy to compare sizes. The length of the bar correlates with the cell value, so you can quickly see which values are larger or smaller.
- **How to Use:**
 1. Select the cells you want to format.
 2. Go to “Conditional Formatting” in the “Home” tab.
 3. Choose “Data Bars” and then select a color style you prefer.
 4. The data bars will automatically appear within the cells, with longer bars representing higher values.

3. Other Useful Conditional Formatting Features:

- **Color Scales:**

- **Purpose:** Applies a gradient of colors based on the cell values, with different colors representing high, medium, and low values.
- **How to Use:**
 1. Select your data range.
 2. Go to “Conditional Formatting” and choose “Color Scales.”
 3. Pick a color scale that fits your data visualization needs.

- **Icon Sets:**

- **Purpose:** Adds icons within cells to represent values, often used for performance metrics (e.g., arrows for up/down trends, traffic lights for status).
- **How to Use:**
 1. Select the data range.
 2. Under “Conditional Formatting,” select “Icon Sets.”
 3. Choose the set that best represents your data (e.g., arrows, flags, stars).

- **Custom Rules:**

- **Purpose:** Allows you to create more complex and customized formatting rules.

- **How to Use:**

1. Go to “Conditional Formatting” and click “New Rule.”
2. Select “Use a formula to determine which cells to format.”
3. Enter your custom formula (e.g., =A1>100 to format cells where the value is greater than 100).
4. Set the formatting style and click “OK.”

4. Managing and Clearing Conditional Formatting:

- **Manage Rules:**

- Go to “Conditional Formatting” and select “Manage Rules” to view, edit, or delete existing formatting rules.

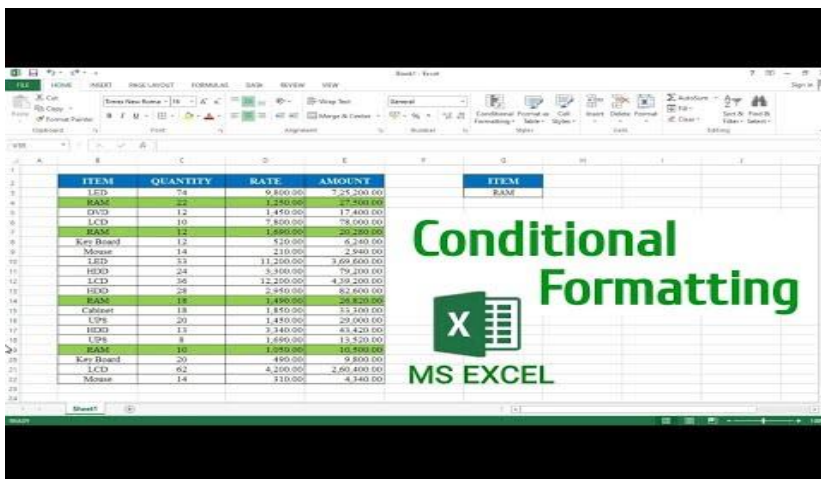
- **Clear Rules:**

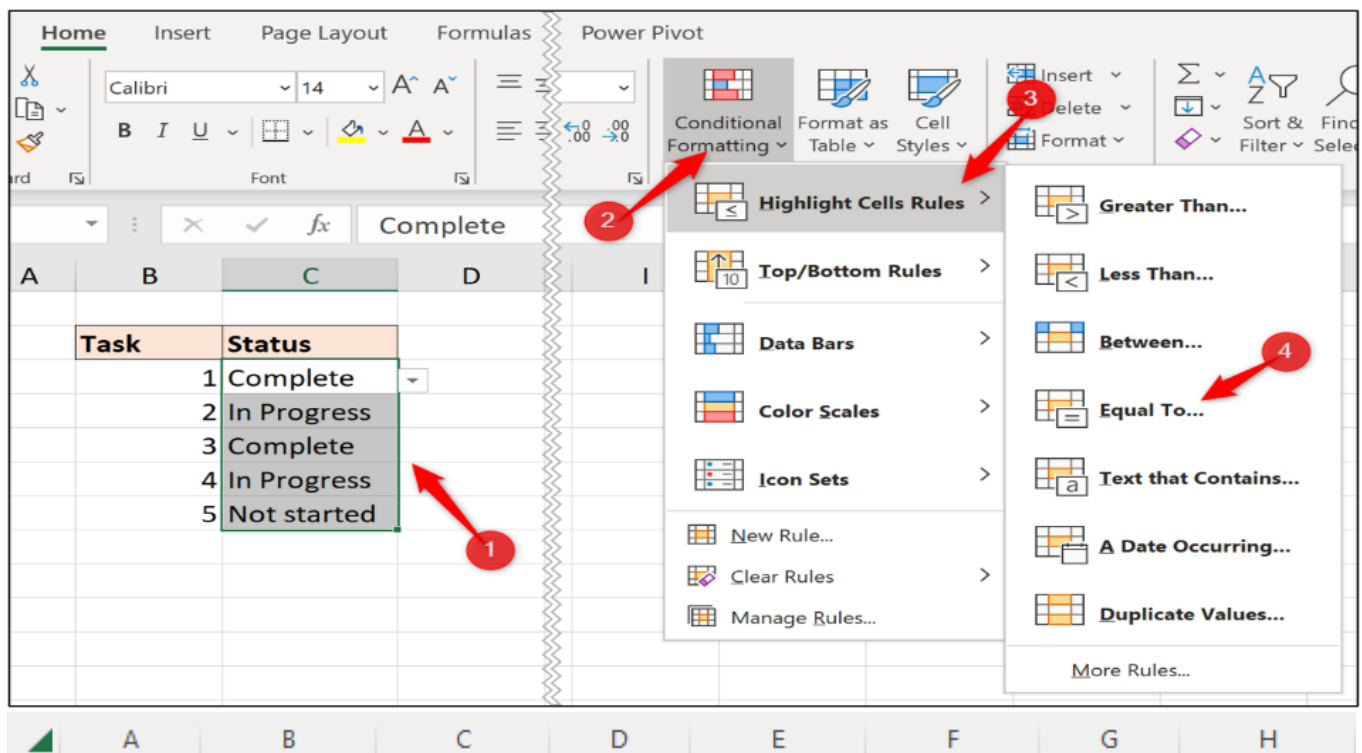
- To remove conditional formatting, select the cells, go to “Conditional Formatting,” and choose “Clear Rules.”

5. Practical Applications:

- Highlighting overdue payments by formatting dates older than today’s date.
- Using data bars to compare loan amounts at a glance.
- Applying color scales to differentiate between low, medium, and high-risk loans.

Still facing problem in understanding Conditional formatting... No problem!! Have a look at this 5 min video:





	A	B	C	D	E	F	G	H
	Excel Conditional Formatting							
1								
2	Month	Price		Month	Price		Month	Price
3	Jan	\$120.00		Jan	\$120.00		Jan	★ \$120.00
4	Feb	\$60.00		Feb	\$60.00		Feb	☆ \$60.00
5	Mar	\$80.00		Mar	\$80.00		Mar	★ \$80.00
6	Apr	\$102.49		Apr	\$102.49		Apr	★ \$102.49
7	May	\$70.00		May	\$70.00		May	★ \$70.00
8	Jun	\$99.12		Jun	\$99.12		Jun	★ \$99.12
9	Jul	\$54.00		Jul	\$54.00		Jul	☆ \$54.00
10	Aug	\$105.79		Aug	\$105.79		Aug	★ \$105.79
11	Sep	\$45.00		Sep	\$45.00		Sep	☆ \$45.00
12	Oct	\$70.00		Oct	\$70.00		Oct	★ \$70.00
13	Nov	\$97.52		Nov	\$97.52		Nov	★ \$97.52
14	Dec	\$105.94		Dec	\$105.94		Dec	★ \$105.94

Week 4.4

IF Statements

Purpose:

The **IF** function is used to perform logical tests and return one value if the condition is true and another value if the condition is false. This is helpful for segmenting and analyzing data based on specific criteria.

Syntax: =IF(logical_test, value_if_true, value_if_false)

Components:

- **logical_test:** The condition you want to test (e.g., A1 > 10).
- **value_if_true:** The value to return if the condition is true.
- **value_if_false:** The value to return if the condition is false.

Example: Suppose you have a column with sales amounts and you want to label sales above \$500 as "High" and others as "Low".

In cell B2, you could use: `=IF(A2 > 500, "High", "Low")`

This formula checks if the value in cell A2 is greater than 500. If true, it returns "High"; otherwise, it returns "Low".

Week 4.5

Subtotals

Subtotals help summarize data within a group by performing calculations such as sum, average, or count.

Using Subtotals:

1. **Sort Data:** Ensure your data is sorted by the column you want to subtotal.
2. **Subtotal Function:** Go to the Data tab, click Subtotal, and choose the column to subtotal and the function (e.g., Sum, Average)

	A	B	C	D
1	Region	Item	Sales	
2	East	Apples	\$285	
3	East	Grapes	\$340	
4	East	Apples	\$280	
5	East	Oranges	\$335	
6	North	Apples	\$260	
7	North	Grapes	\$250	
8	North	Lemons	\$255	
9	Total		L(

Example:

If you have a list of sales transactions sorted by Product, you can subtotal the Sales Amount for each product.

MIN and MAX Functions:

- **MIN:** Returns the smallest value in a range.
 - **Formula:** `=MIN(range)`
 - **Example:** `=MIN(B2:B10)` returns the smallest value in cells B2 to B10.
- **MAX:** Returns the largest value in a range.
 - **Formula:** `=MAX(range)`
 - **Example:** `=MAX(B2:B10)` returns the largest value in cells B2 to B10.

AVERAGE Function:

- **AVERAGE:** Calculates the mean of a range of values.
 - **Formula:** =AVERAGE(range)
 - **Example:** =AVERAGE(B2:B10) returns the average value in cells B2 to B10.

Standard Deviation:

- **STDEV:** Measures the amount of variation or dispersion of a set of values.
 - **Formula:** =STDEV(range)
 - **Example:** =STDEV(B2:B10) calculates the standard deviation of values in cells B2 to B10

Week 4.6

Using the Remove Duplicates Tool:

- Steps:
 1. Select the range of cells from which you want to remove duplicates.
 2. Go to the Data tab on the Ribbon.
 3. Click on Remove Duplicates in the Data Tools group.
 4. Choose the column(s) to check for duplicates and click OK.
 5. Excel will remove duplicate values and keep only the unique ones.

FILTER Formula: The FILTER formula filters a range based on a condition. It helps you create subsets of data that meet specific criteria.

Syntax: =FILTER(array, include, [if_empty])

Example: To filter people who have bought an AC, if column B contains "Yes" or "No" for AC purchase:

=FILTER(A2:B100, B2:B100="Yes")

This will return rows from A2:B100 where column B is "Yes".

SUMIF Formula: The SUMIF formula sums the values in a range based on a condition.

Syntax: =SUMIF(range, criteria, [sum_range])

Example: To sum total income of people who have bought ACs: =SUMIF(B2:B100, "Yes", A2:A100)

This sums values in A2:A100 where corresponding entries in B2:B100 are "Yes".



Best of luck

