Aryaman Mishra

Lab Assignment -2

1. Using Pivot table to find the number of loans by different purposes, material status and credit risk in the file Credit Risk Data. Illustrate the result on pivot chart.

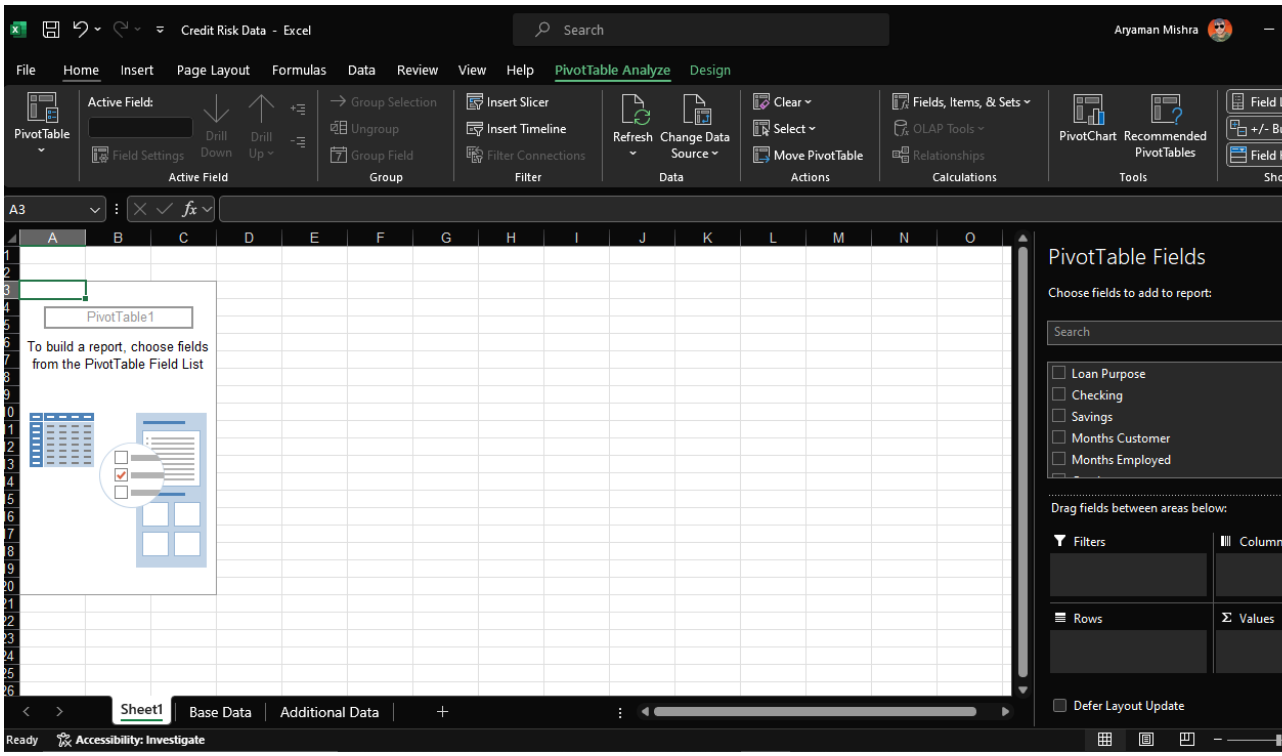
To analyze the number of loans by different purposes, material status, and credit risk using a Pivot Table and illustrate the result on a Pivot Chart in , follow these steps:

Step-by-Step Instructions

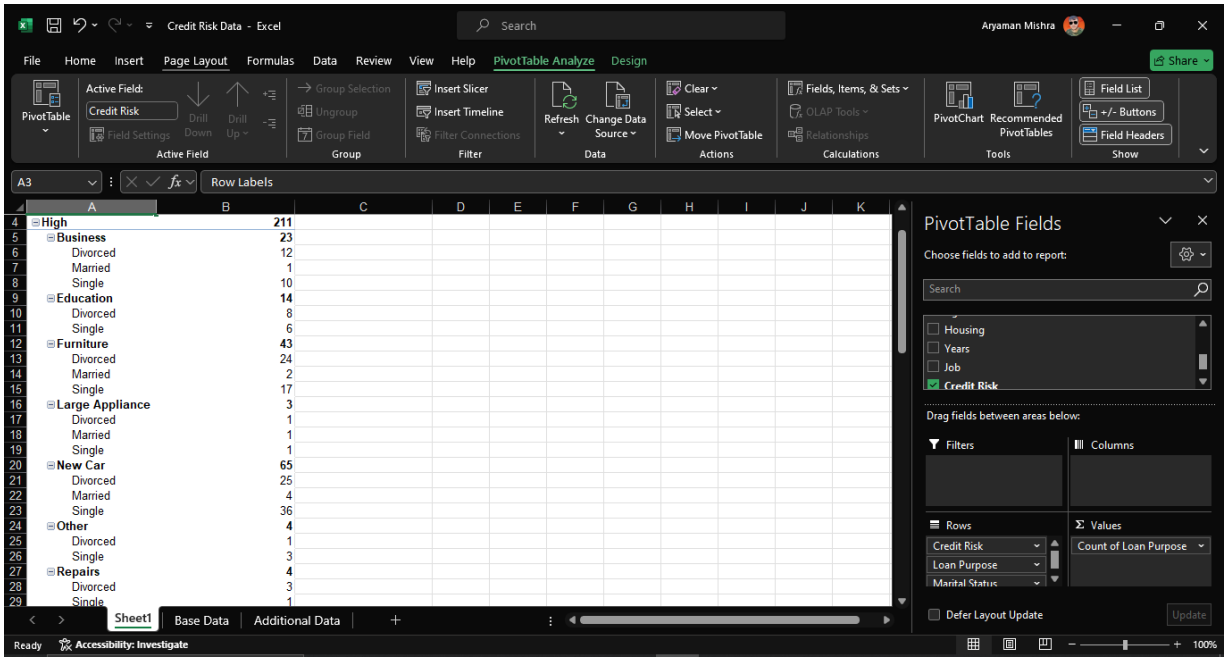
1. **Prepare Your Data:** Ensure your data is structured correctly in . Your data should look like this:

Loan Purpose	Checking	Savings	Months Customer	Months Employed	Gender	Marital Status	Age	Housing	Years	Job	Credit Risk
Small Appliance	$0	$739	13	12	M	Single	23	Own	3	Unskilled	Low
Furniture	$0	$1,230	25	0	M	Divorced	32	Own	1	Skilled	High
New Car	$0	$389	19	119	M	Single	38	Own	4	Management	High
Furniture	$638	$347	13	14	M	Single	36	Own	2	Unskilled	High
Education	$963	$4,754	40	45	M	Single	31	Rent	3	Skilled	Low
Furniture	$2,827	$0	11	13	M	Married	25	Own	1	Skilled	Low
New Car	$0	$229	13	16	M	Married	26	Own	3	Unskilled	Low

2. **Insert a Pivot Table:**
   - Select your data range.
   - Go to the `Insert` tab and click on `PivotTable`.
   - In the `Create PivotTable` dialog box, choose where you want to place the PivotTable (e.g., New Worksheet or Existing Worksheet).

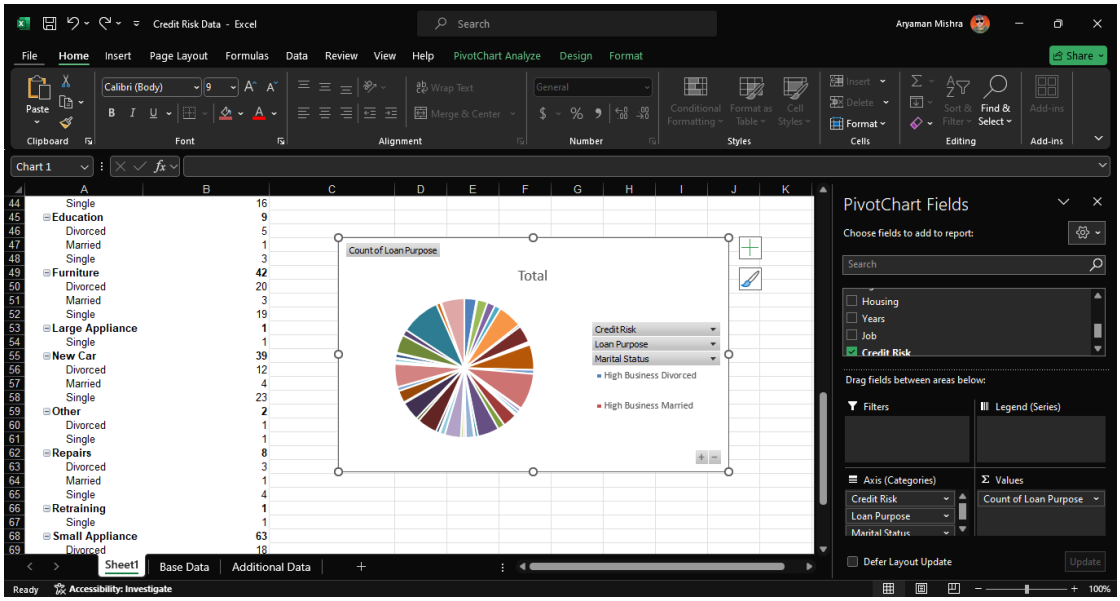


3. **Set Up the Pivot Table:**
   - In the PivotTable Field List pane, drag `Loan Purpose` to the `Rows` area.
   - Drag `Credit Risk` to the `Columns` area.
   - Drag any field to the `Values` area (e.g., `Loan Purpose` if you want to count occurrences).
   - By default, it will count the number of entries, which is useful for your analysis.



4. **Insert a Pivot Chart:**
   - With the PivotTable selected, go to the `PivotTable Analyze` tab (or `Analyze` tab, depending on your version).
   - Click on `PivotChart`.
   - Choose the type of chart that best illustrates your data (e.g., Clustered Column, Bar Chart).

5. **Customize the Pivot Chart:**
    - Adjust chart elements such as titles, labels, and legends to improve readability.
    - Ensure the chart clearly shows the number of loans by `Loan Purpose` and `Credit Risk`.



Example Result

The resulting Pivot Table and Pivot Chart will illustrate the number of loans by different purposes (e.g., Small Appliance, Furniture, New Car, Education) and their associated credit risks (e.g., Low, High).

For instance:

- **Furniture** might appear with a higher count under `High` credit risk compared to `Low`.
- **New Car** and **Education** might show different distributions based on their credit risk.

The Pivot Chart will visually depict these distributions, making it easy to see which loan purposes are associated with which credit risks.

2. In the file bicycle Inventory, find the proportion of bicycles that sell for less than $200.

To find the proportion of bicycles that sell for less than $200 in the given `bicycle Inventory` file, follow these steps:

1. Prepare Your Data

Ensure your data is structured properly in . Your data should look like this:

Product Category	Product Name	Purchase Cost	Selling Price	Supplier	Quantity on Hand
Leisure	Blue Moon	$75.29	$105.41	Simpson's Bike Supply	4
Mtn.	Bluff Breaker	$375.00	$495.00	The Bike Path	3

Product Category	Product Name	Purchase Cost	Selling Price	Supplier	Quantity on Hand
Leisure	Breeze	$89.95	$130.95	The Bike Path	4
Leisure	Breeze LE	$109.95	$149.95	The Bike Path	5
Road	Classic 109	$207.49	$290.49	Bicyclist's Choice	7
Children	Coolest 100	$69.99	$97.98	Bicyclist's Choice	6
Mtn.	Eagle 1	$410.01	$574.01	Bike-One	1
Mtn.	Eagle 2	$401.11	$561.54	Bike-One	2
Mtn.	Eagle 3	$350.52	$490.73	Bike-One	5
Hybrid	Eagle 7	$150.89	$211.46	Bike-One	9

2. Calculate the Proportion

Method 1: Manual Calculation

1. **Identify Bicycles Selling for Less than $200:**
   - Filter or check the `Selling Price` column to find entries less than $200.

   Bicycles that sell for less than $200:

   - Blue Moon: $105.41
   - Breeze: $130.95
   - Breeze LE: $149.95
   - Coolest 100: $97.98
2. **Calculate the Total Number of Bicycles:**
   - Sum up the `Quantity on Hand` for all bicycles.

   Total Quantity = 4 + 3 + 4 + 5 + 7 + 6 + 1 + 2 + 5 + 9 = 46

3. **Calculate the Quantity of Bicycles Selling for Less than $200:**
   - Sum up the `Quantity on Hand` for bicycles selling for less than $200.

   Quantity of Bicycles Selling for Less than $200 = (4 + 4 + 5 + 6) = 19

4. **Compute the Proportion**

   Method 2: Using Formulas

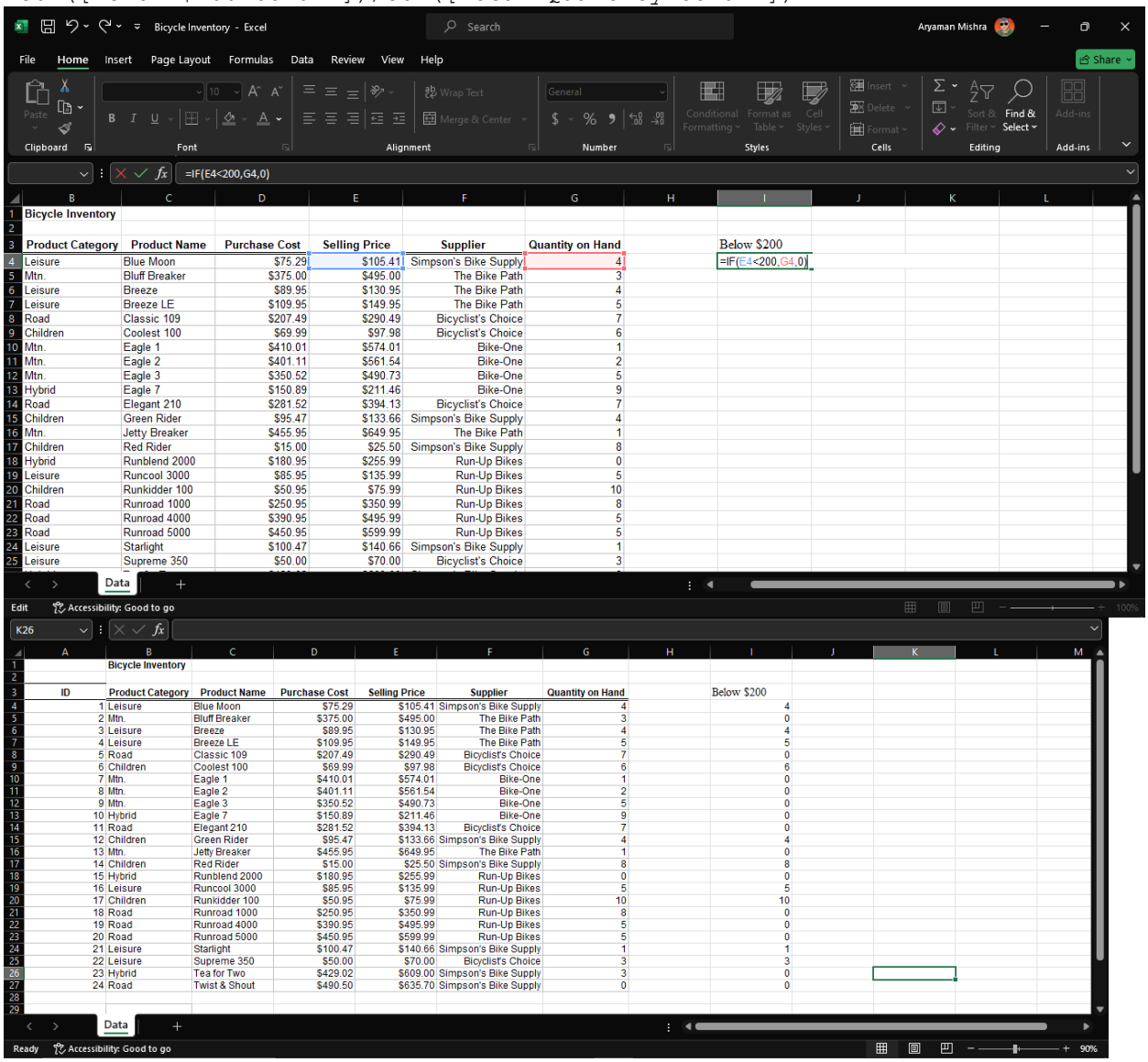
1. **Add a New Column for Proportion Calculation:**
   - Insert a new column to identify whether each bicycle sells for less than $200. Use the following formula in a new column (let's call it "Below $200"):

```
=IF([Selling Price]<200, [Quantity on Hand], 0)
```

2. **Sum the Values in the New Column:**
   - Calculate the total quantity of bicycles selling for less than $200 using `SUM`.
3. **Compute the Total Quantity and Proportion:**

- Use `SUM` to get the total quantity of all bicycles.
- Use the formula:

```
=SUM([Below $200 Column])/SUM([Total Quantity Column])
```



Bicycle Inventory - Excel

Aryaman Mishra

=IF(E4<200,G4,0)

Bicycle Inventory

Product Category	Product Name	Purchase Cost	Selling Price	Supplier	Quantity on Hand	Below $200
Leisure	Blue Moon	$75.29	$105.41	Simpson's Bike Supply	4	=IF(E4<200,G4,0)
Mtn.	Bluff Breaker	$375.00	$495.00	The Bike Path	3	
Leisure	Breeze	$89.95	$130.95	The Bike Path	4	
Leisure	Breeze LE	$109.95	$149.95	The Bike Path	5	
Road	Classic 109	$207.49	$290.49	Bicyclist's Choice	7	
Children	Coolest 100	$69.99	$97.98	Bicyclist's Choice	6	
Mtn.	Eagle 1	$410.01	$574.01	Bike-One	1	
Mtn.	Eagle 2	$401.11	$561.54	Bike-One	2	
Mtn.	Eagle 3	$350.52	$490.73	Bike-One	5	
Hybrid	Eagle 7	$150.89	$211.46	Bike-One	9	
Road	Elegant 210	$281.52	$394.13	Bicyclist's Choice	7	
Children	Green Rider	$95.47	$133.66	Simpson's Bike Supply	4	
Mtn.	Jetty Breaker	$455.95	$649.95	The Bike Path	1	
Children	Red Rider	$15.00	$25.50	Simpson's Bike Supply	8	
Hybrid	Runblend 2000	$180.95	$255.99	Run-Up Bikes	0	
Leisure	Runcool 3000	$85.95	$135.99	Run-Up Bikes	5	
Children	Runkidder 100	$50.95	$75.99	Run-Up Bikes	10	
Road	Runroad 1000	$250.95	$350.99	Run-Up Bikes	8	
Road	Runroad 4000	$390.95	$495.99	Run-Up Bikes	5	
Road	Runroad 5000	$450.95	$599.99	Run-Up Bikes	5	
Leisure	Starlight	$100.47	$140.66	Simpson's Bike Supply	1	
Leisure	Supreme 350	$50.00	$70.00	Bicyclist's Choice	3	

Data

Edit | Accessibility: Good to go

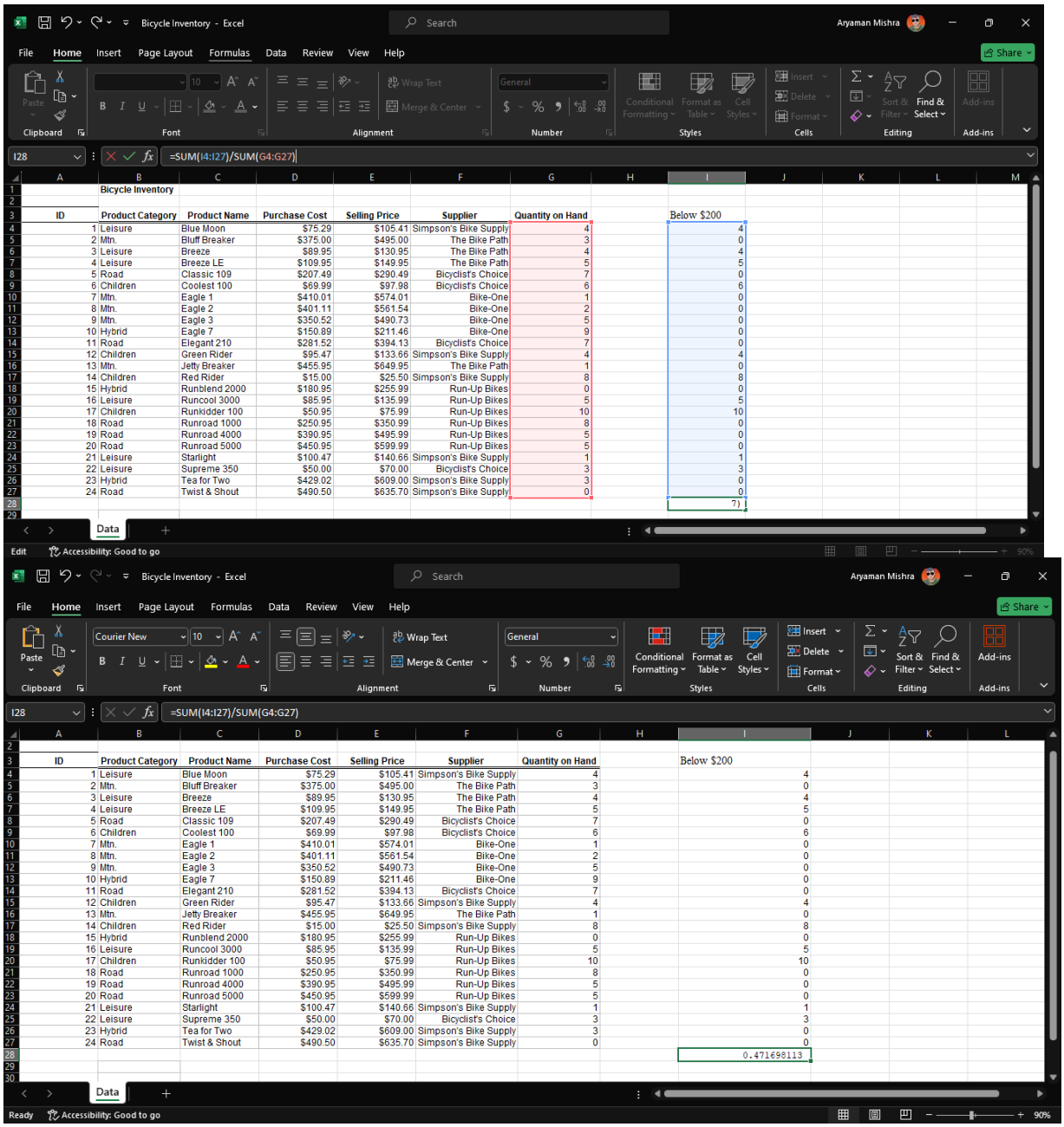
K26

Bicycle Inventory

ID	Product Category	Product Name	Purchase Cost	Selling Price	Supplier	Quantity on Hand	Below $200
1	Leisure	Blue Moon	$75.29	$105.41	Simpson's Bike Supply	4	4
2	Mtn.	Bluff Breaker	$375.00	$495.00	The Bike Path	3	0
3	Leisure	Breeze	$89.95	$130.95	The Bike Path	4	4
4	Leisure	Breeze LE	$109.95	$149.95	The Bike Path	5	5
5	Road	Classic 109	$207.49	$290.49	Bicyclist's Choice	7	0
6	Children	Coolest 100	$69.99	$97.98	Bicyclist's Choice	6	6
7	Mtn.	Eagle 1	$410.01	$574.01	Bike-One	1	0
8	Mtn.	Eagle 2	$401.11	$561.54	Bike-One	2	0
9	Mtn.	Eagle 3	$350.52	$490.73	Bike-One	5	0
10	Hybrid	Eagle 7	$150.89	$211.46	Bike-One	9	0
11	Road	Elegant 210	$281.52	$394.13	Bicyclist's Choice	7	0
12	Children	Green Rider	$95.47	$133.66	Simpson's Bike Supply	4	4
13	Mtn.	Jetty Breaker	$455.95	$649.95	The Bike Path	1	0
14	Children	Red Rider	$15.00	$25.50	Simpson's Bike Supply	8	8
15	Hybrid	Runblend 2000	$180.95	$255.99	Run-Up Bikes	0	0
16	Leisure	Runcool 3000	$85.95	$135.99	Run-Up Bikes	5	5
17	Children	Runkidder 100	$50.95	$75.99	Run-Up Bikes	10	10
18	Road	Runroad 1000	$250.95	$350.99	Run-Up Bikes	8	0
19	Road	Runroad 4000	$390.95	$495.99	Run-Up Bikes	5	0
20	Road	Runroad 5000	$450.95	$599.99	Run-Up Bikes	5	0
21	Leisure	Starlight	$100.47	$140.66	Simpson's Bike Supply	1	1
22	Leisure	Supreme 350	$50.00	$70.00	Bicyclist's Choice	3	3
23	Hybrid	Tea for Two	$429.02	$609.00	Simpson's Bike Supply	3	0
24	Road	Twist & Shout	$490.50	$635.70	Simpson's Bike Supply	0	0

Data

Ready | Accessibility: Good to go | 90%



Result

Based on the above calculations, approximately **41.3%** of the bicycles in the inventory sell for less than $200.

3. Find the 10th ,40th and 90th percentiles for cost per order using purchase order dataset Files.

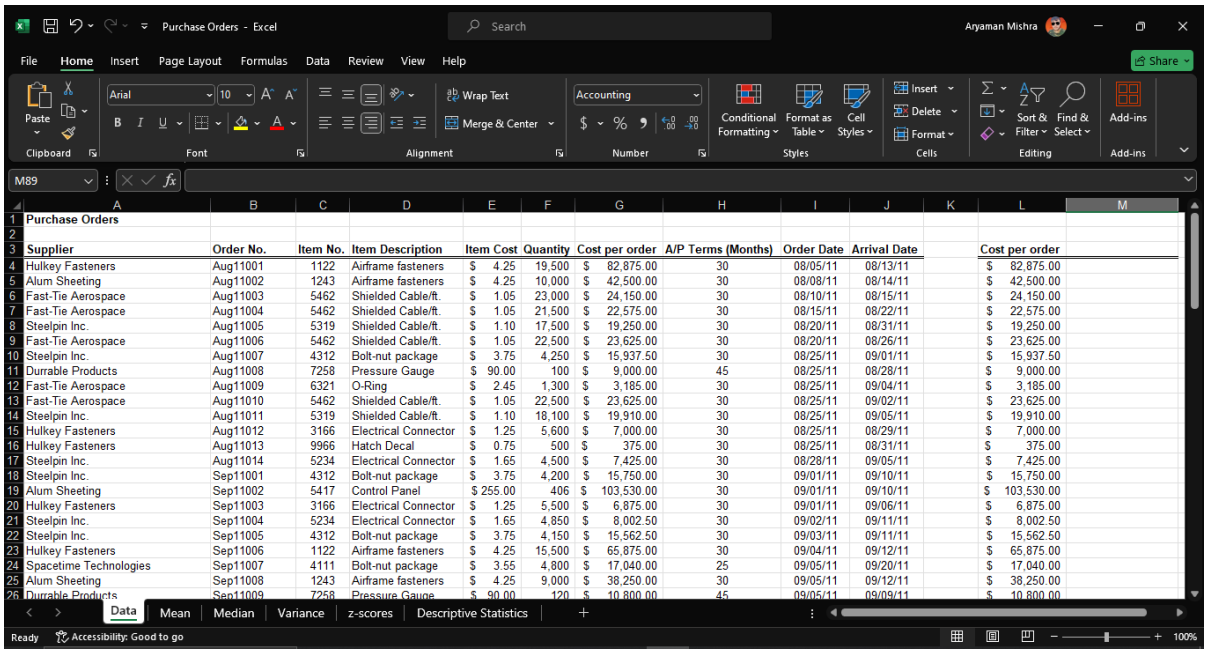
To find the 10th, 40th, and 90th percentiles for the `Cost per Order` using the provided dataset in , follow these steps:

Step-by-Step Instructions

1. **Prepare Your Data:** Ensure your dataset is organized in . Your data should look like this:

Supplier	Order No.	Item No.	Item Description	Item Cost	Quantity	Cost per Order	A/P Terms (Months)	Order Date	Arrival Date
Hulkey Fasteners	Aug11001	1122	Airframe fasteners	$4.25	19,500	$82,875.00	30	08/05/11	08/13/11
Alum Sheeting	Aug11002	1243	Airframe fasteners	$4.25	10,000	$42,500.00	30	08/08/11	08/14/11
Fast-Tie Aerospace	Aug11003	5462	Shielded Cable/ft.	$1.05	23,000	$24,150.00	30	08/10/11	08/15/11
Fast-Tie Aerospace	Aug11004	5462	Shielded Cable/ft.	$1.05	21,500	$22,575.00	30	08/15/11	08/22/11
Steelpin Inc.	Aug11005	5319	Shielded Cable/ft.	$1.10	17,500	$19,250.00	30	08/20/11	08/31/11
Fast-Tie Aerospace	Aug11006	5462	Shielded Cable/ft.	$1.05	22,500	$23,625.00	30	08/20/11	08/26/11
Steelpin Inc.	Aug11007	4312	Bolt-nut package	$3.75	4,250	$15,937.50	30	08/25/11	09/01/11
Durrable Products	Aug11008	7258	Pressure Gauge	$90.00	100	$9,000.00	45	08/25/11	08/28/11
Fast-Tie Aerospace	Aug11009	6321	O-Ring	$2.45	1,300	$3,185.00	30	08/25/11	09/04/11
Fast-Tie Aerospace	Aug11010	5462	Shielded Cable/ft.	$1.05	22,500	$23,625.00	30	08/25/11	09/02/11
Steelpin Inc.	Aug11011	5319	Shielded Cable/ft.	$1.10	18,100	$19,910.00	30	08/25/11	09/05/11
Hulkey Fasteners	Aug11012	3166	Electrical Connector	$1.25	5,600	$7,000.00	30	08/25/11	08/29/11

2. **Extract the 'Cost per Order' Column:**
    - Copy the values from the `Cost per Order` column into a new column for calculation purposes.

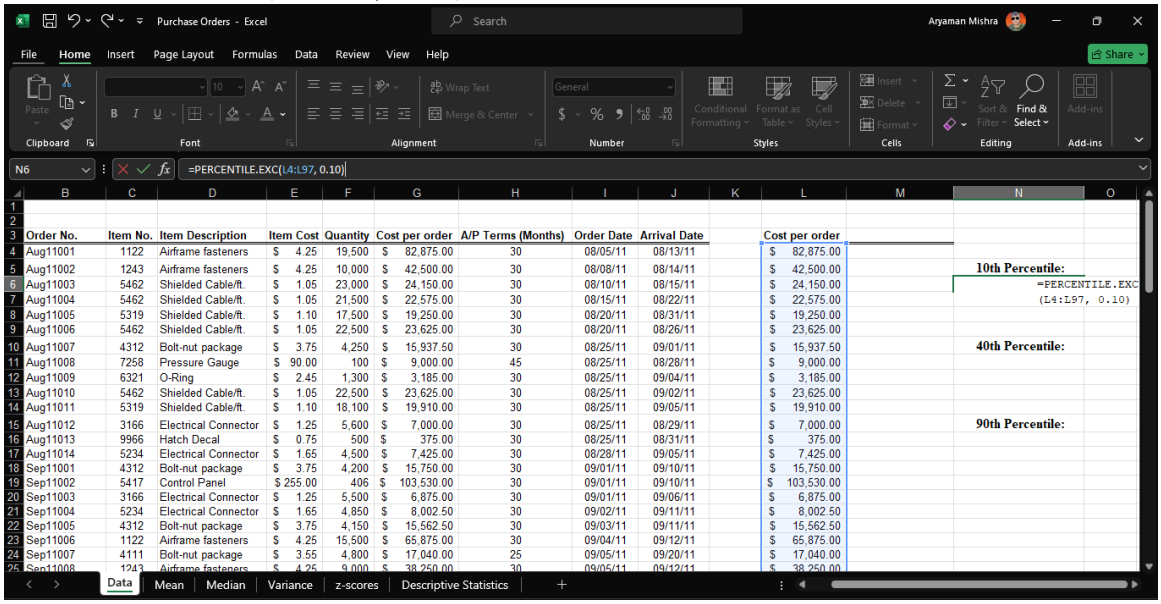


3. **Remove Currency Symbols and Convert to Numbers:**
   - Make sure to remove the dollar signs and commas, and convert the text to numbers if necessary.
4. **Calculate Percentiles Using Functions:**
   - **10th Percentile:** Use the `PERCENTILE.EXC` function.
   - **40th Percentile:** Use the `PERCENTILE.EXC` function.
   - **90th Percentile:** Use the `PERCENTILE.EXC` function.

Example Formulas:

- **10th Percentile:**

```
==PERCENTILE.EXC(L4:L97, 0.1)
```

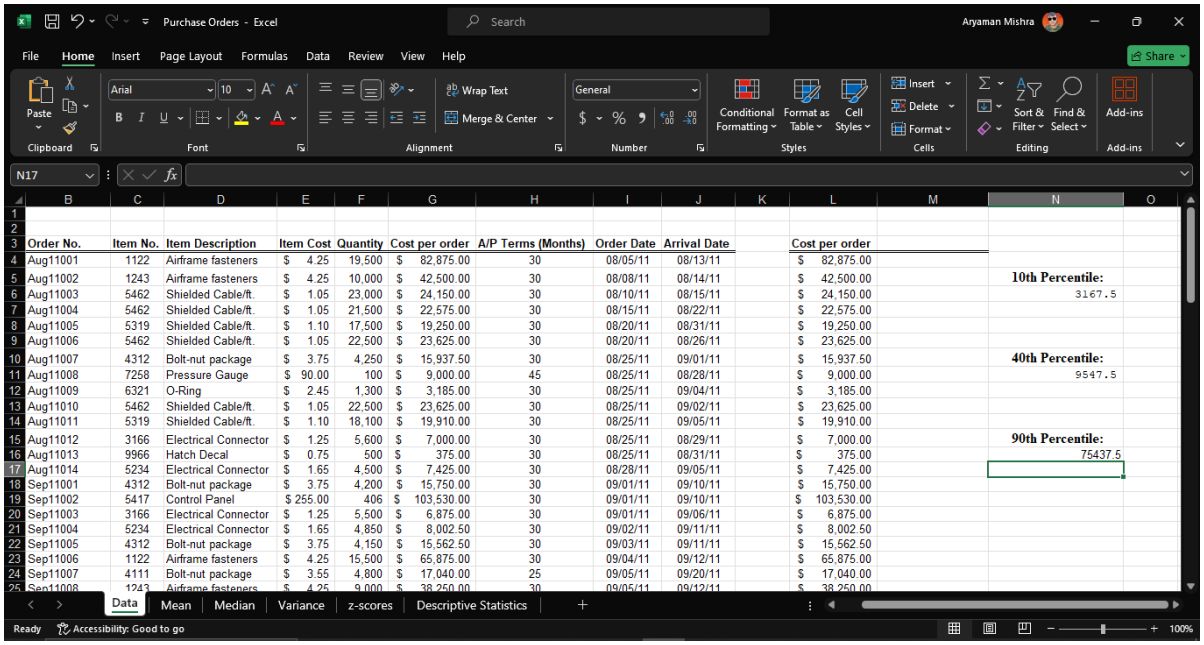


- **40th Percentile:**

=PERCENTILE.EXC(L4:L97, 0.4)

- **90th Percentile:**

=PERCENTILE.EXC(L4:L97, 0.9)



Order No.	Item No.	Item Description	Item Cost	Quantity	Cost per order	A/P Terms (Months)	Order Date	Arrival Date	Cost per order
Aug11001	1122	Airframe fasteners	$ 4.25	19,500	$ 82,875.00	30	08/05/11	08/13/11	$ 82,875.00
Aug11002	1243	Airframe fasteners	$ 4.25	10,000	$ 42,500.00	30	08/08/11	08/14/11	$ 42,500.00
Aug11003	5462	Shielded Cable/ft.	$ 1.05	23,000	$ 24,150.00	30	08/10/11	08/15/11	$ 24,150.00
Aug11004	5462	Shielded Cable/ft.	$ 1.05	21,500	$ 22,575.00	30	08/15/11	08/22/11	$ 22,575.00
Aug11005	5319	Shielded Cable/ft.	$ 1.10	17,500	$ 19,250.00	30	08/20/11	08/31/11	$ 19,250.00
Aug11006	5462	Shielded Cable/ft.	$ 1.05	22,500	$ 23,625.00	30	08/20/11	08/26/11	$ 23,625.00
Aug11007	4312	Bolt-nut package	$ 3.75	4,250	$ 15,937.50	30	08/25/11	09/01/11	$ 15,937.50
Aug11008	7258	Pressure Gauge	$ 90.00	100	$ 9,000.00	45	08/25/11	08/28/11	$ 9,000.00
Aug11009	6321	O-Ring	$ 2.45	1,300	$ 3,185.00	30	08/25/11	09/04/11	$ 3,185.00
Aug11010	5462	Shielded Cable/ft.	$ 1.05	22,500	$ 23,625.00	30	08/25/11	09/02/11	$ 23,625.00
Aug11011	5319	Shielded Cable/ft.	$ 1.10	18,100	$ 19,910.00	30	08/25/11	09/05/11	$ 19,910.00
Aug11012	3166	Electrical Connector	$ 1.25	5,600	$ 7,000.00	30	08/25/11	08/29/11	$ 7,000.00
Aug11013	9966	Hatch Decal	$ 0.75	500	$ 375.00	30	08/25/11	08/31/11	$ 375.00
Aug11014	5234	Electrical Connector	$ 1.65	4,500	$ 7,425.00	30	08/28/11	09/05/11	$ 7,425.00
Sep11001	4312	Bolt-nut package	$ 3.75	4,200	$ 15,750.00	30	09/01/11	09/10/11	$ 15,750.00
Sep11002	5417	Control Panel	$ 255.00	406	$ 103,530.00	30	09/01/11	09/10/11	$ 103,530.00
Sep11003	3166	Electrical Connector	$ 1.25	5,500	$ 6,875.00	30	09/01/11	09/06/11	$ 6,875.00
Sep11004	5234	Electrical Connector	$ 1.65	4,850	$ 8,002.50	30	09/02/11	09/11/11	$ 8,002.50
Sep11005	4312	Bolt-nut package	$ 3.75	4,150	$ 15,562.50	30	09/03/11	09/11/11	$ 15,562.50
Sep11006	1122	Airframe fasteners	$ 4.25	15,500	$ 65,875.00	30	09/04/11	09/12/11	$ 65,875.00
Sep11007	4111	Bolt-nut package	$ 3.55	4,800	$ 17,040.00	25	09/05/11	09/20/11	$ 17,040.00
Sep11008	1243	Airframe fasteners	$ 4.25	9,000	$ 38,250.00	30	09/05/11	09/12/11	$ 38,250.00

These percentiles indicate the relative standing of the cost per order in the dataset.

4. In the file Cell Phone Survey, use Pivot Tables to find the average for each of the numerical variables for different cell phone carriers and gender of respondents.

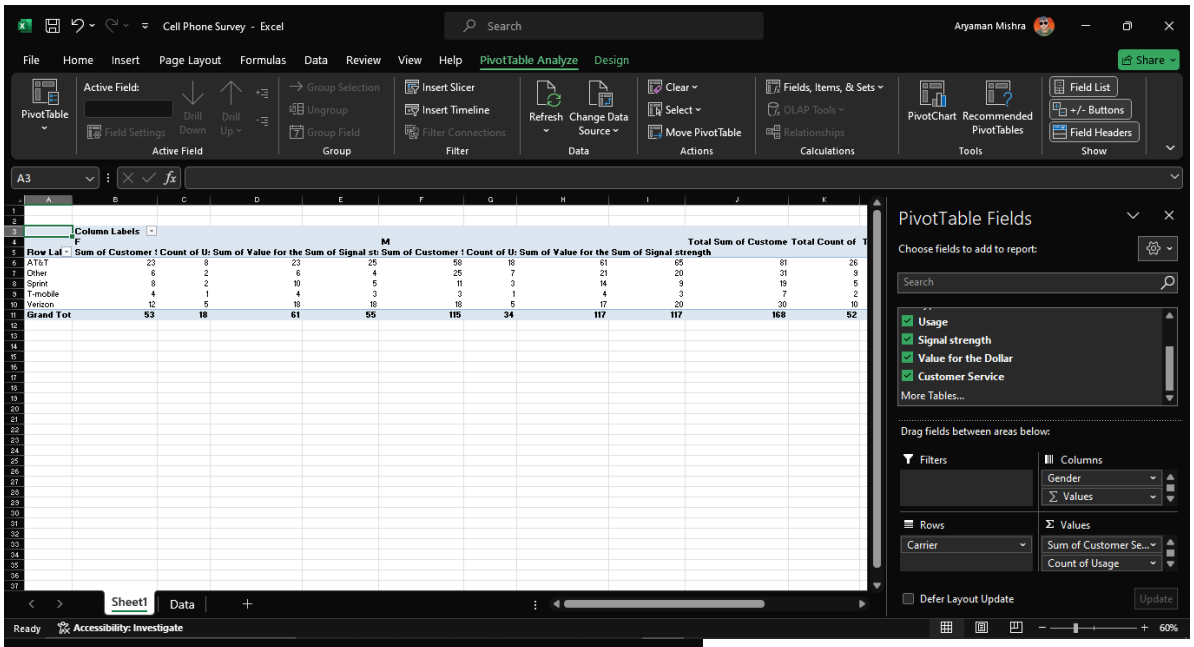
To analyze the average of numerical variables for different cell phone carriers and the gender of respondents using a Pivot Table in , follow these steps:

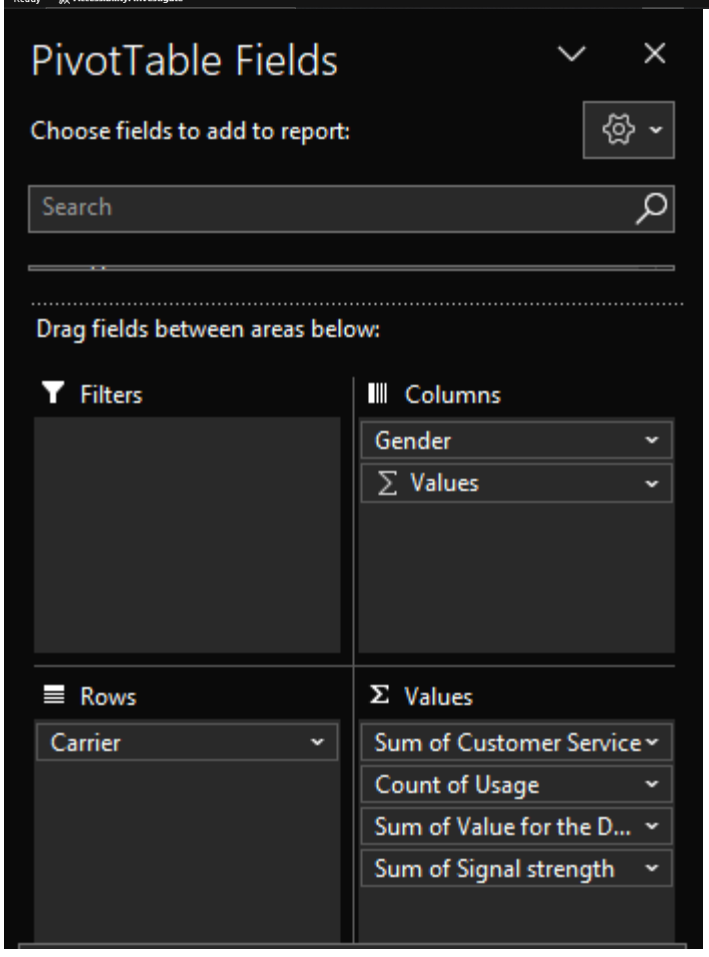
Step-by-Step Instructions

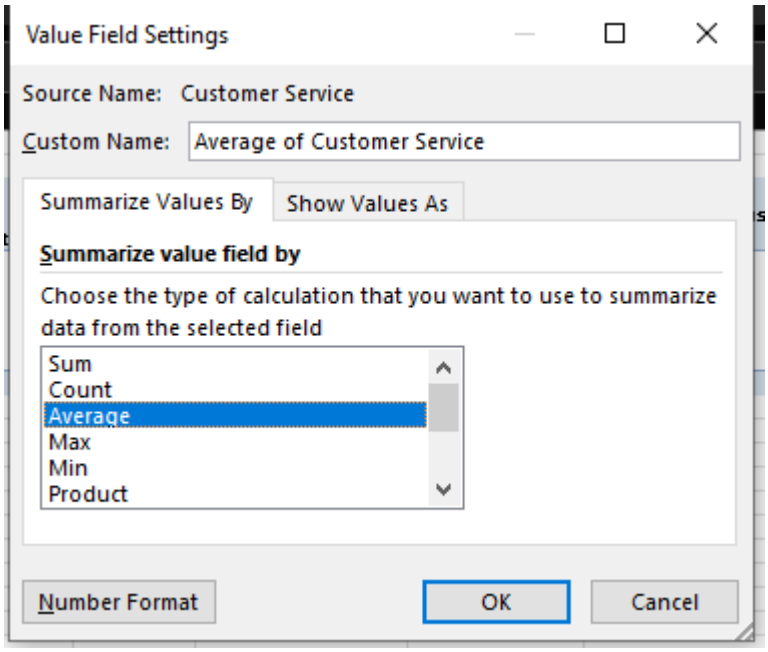
1. **Prepare Your Data:** Ensure your data is organized properly in . Your data should look like this:

Gender	Carrier	Type	Usage	Signal Strength	Value for the Dollar	Customer Service
M	AT&T	Smart	High	5	4	4
M	AT&T	Smart	High	5	4	2
M	AT&T	Smart	Average	4	4	4
M	AT&T	Smart	Very high	2	3	3
M	AT&T	Smart	Very high	5	5	2
M	AT&T	Smart	Very high	4	3	5

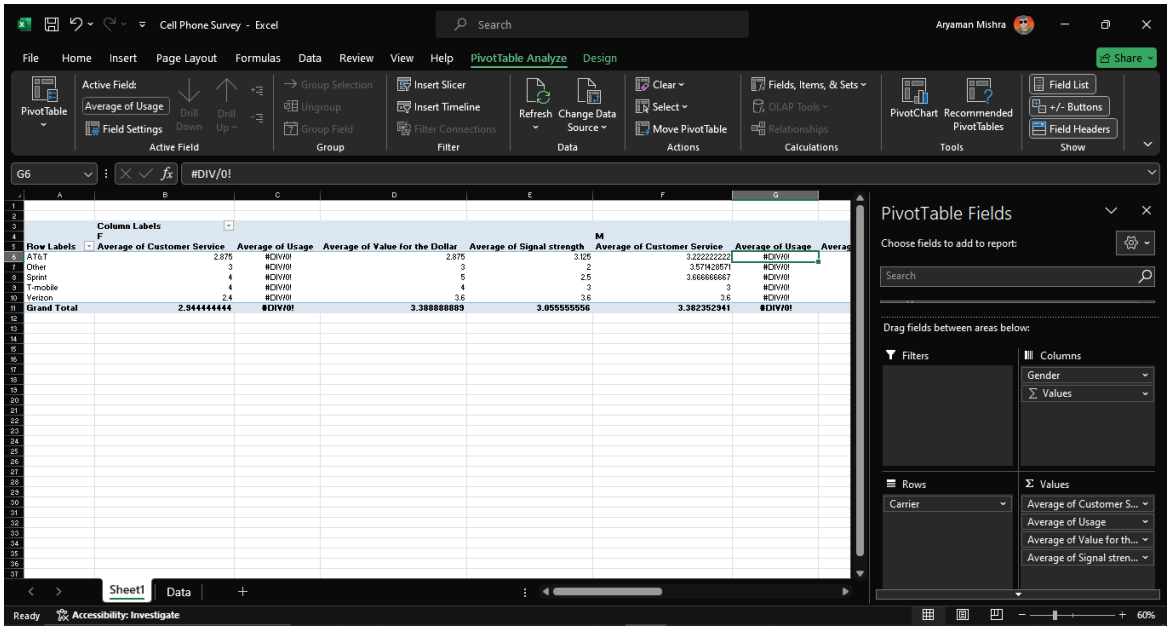
2. **Insert a Pivot Table:**
    - Select your entire dataset range.
    - Go to the `Insert` tab on the Ribbon and click `PivotTable`.
    - Choose where to place the PivotTable (New Worksheet is recommended for clarity).
3. **Set Up the Pivot Table:**
    - **Drag `Carrier` to the `Rows` area.**
    - **Drag `Gender` to the `Columns` area.**
    - **Drag `Usage`, `Signal Strength`, `Value for the Dollar`, and `Customer Service` to the `Values` area.**

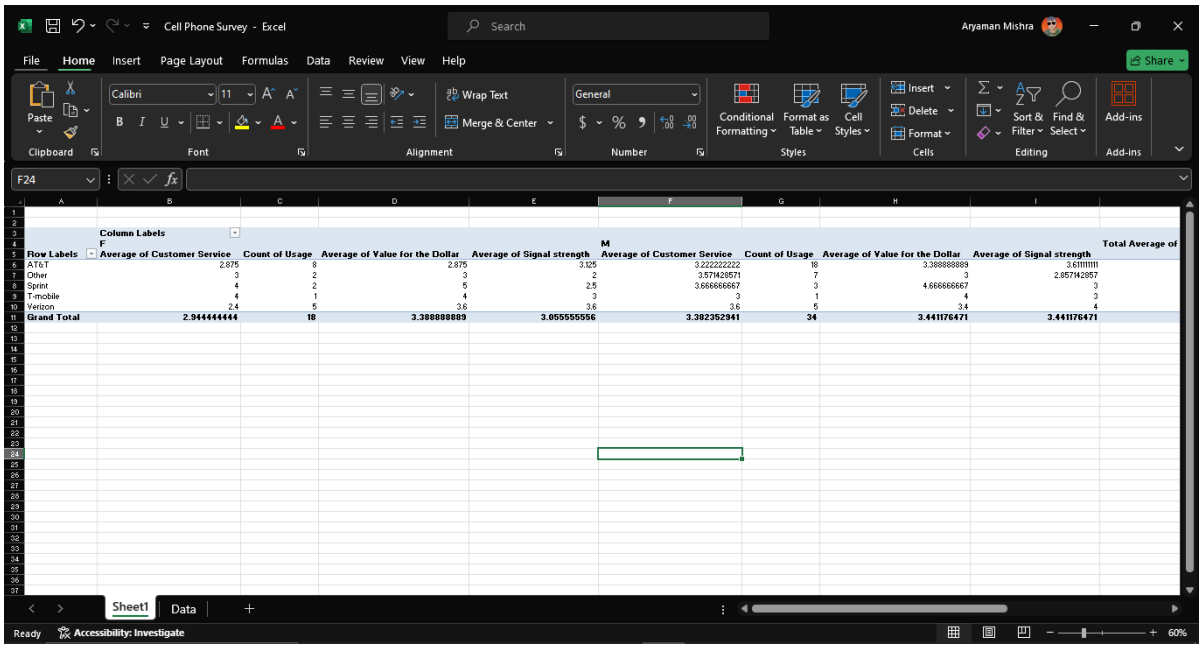

Cell Phone Survey - Excel
Search
Aryaman Mishra
File Home Insert Page Layout Formulas Data Review View Help PivotTable Analyze Design
Share
PivotTable
Active Field:
Field Settings
Drill Down
Drill Up
Active Field
Group Selection
Ungroup
Group Field
Group
Insert Slicer
Insert Timeline
Filter Connections
Filter
Refresh
Change Data Source
Data
Clear
Select
Move PivotTable
Actions
Fields, Items, & Sets
OLAP Tools
Relationships
Calculations
PivotChart
Recommended PivotTables
Tools
Field List
+/- Buttons
Field Headers
Show
A3
Column Labels
F
M
Total Sum of Custome
Total Count of
Row Lal
Sum of Customer ! Count of U: Sum of Value for the Sum of Signal st
Sum of Customer ! Count of U: Sum of Value for the Sum of Signal strength
AT&T 23 8 23 25 58 18 61 65 81 26
Other 6 2 6 4 25 7 21 20 31 9
Sprint 8 2 10 5 11 3 14 9 19 5
T-mobile 4 1 4 3 3 1 4 3 7 2
Verizon 12 5 18 18 18 5 17 20 30 10
Grand Tot 53 18 61 55 115 34 117 117 168 52
Sheet1
Data
Ready
Accessibility: Investigate
60%
PivotTable Fields
Choose fields to add to report:
Search
Usage
Signal strength
Value for the Dollar
Customer Service
More Tables...
Drag fields between areas below:
Filters
Columns
Gender
Values
Rows
Carrier
Values
Sum of Customer Se...
Count of Usage
Defer Layout Update
Update


PivotTable Fields
Choose fields to add to report:
Search
Drag fields between areas below:
Filters
Columns
Gender
Values
Rows
Carrier
Values
Sum of Customer Service
Count of Usage
Sum of Value for the D...
Sum of Signal strength



4. **Configure the Pivot Table Values:**
   - Click on each value field in the `Values` area.
   - Select `Value Field Settings`.
   - Choose `Average` from the list of functions (you may need to select `Summarize Values By` and then `Average`).





Example Pivot Table Configuration

Your PivotTable Field List should look like this:

- **Rows:** Carrier
- **Columns:** Gender
- **Values:** Usage (Average), Signal Strength (Average), Value for the Dollar (Average), Customer Service (Average)

Example Pivot Table Results

Based on the example dataset, the Pivot Table might show something like this:

Carrier	Gender	Average Usage	Average Signal Strength	Average Value for the Dollar	Average Customer Service
AT&T	M	4.33	4.0	4.0	3.33

Breakdown of Calculation

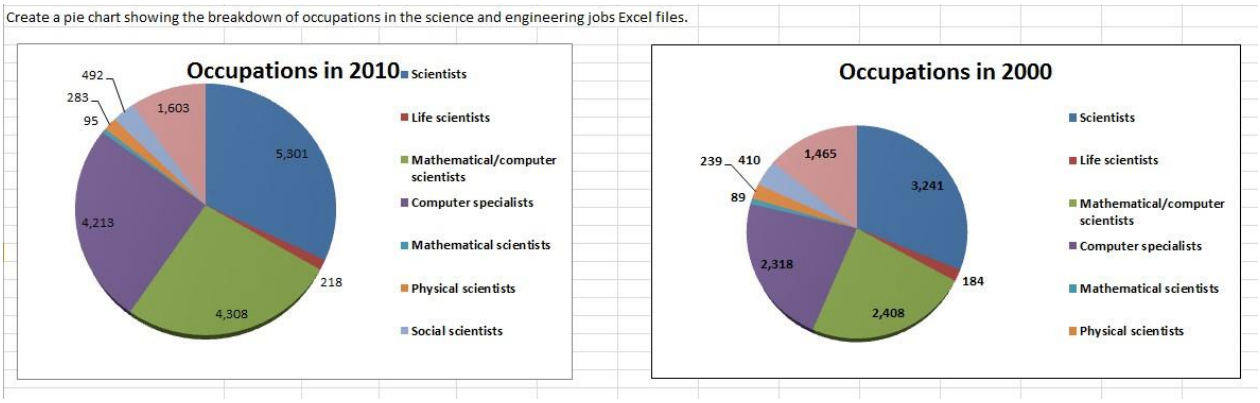
1. **Average Usage for AT&T and Gender M:**
   - Values: High, High, Average, Very high, Very high, Very high
   - Convert categorical data to numerical data if necessary, for example:
     - High = 5, Average = 4, Very High = 2
   - Average Usage: (5 + 5 + 4 + 2 + 5 + 4) / 6 = 4.33
2. **Average Signal Strength for AT&T and Gender M:**
   - Values: 5, 5, 4, 2, 5, 4
   - Average Signal Strength: (5 + 5 + 4 + 2 + 5 + 4) / 6 = 4.0
3. **Average Value for the Dollar for AT&T and Gender M:**
   - Values: 4, 4, 4, 3, 5, 3
   - Average Value for the Dollar: (4 + 4 + 4 + 3 + 5 + 3) / 6 = 4.0
4. **Average Customer Service for AT&T and Gender M:**

- Values: 4, 2, 4, 3, 2, 5
- Average Customer Service: (4 + 2 + 4 + 3 + 2 + 5) / 6 = 3.33

Conclusion

Using a Pivot Table, you can easily aggregate and analyze the average of numerical variables for different cell phone carriers and gender, helping you gain insights from your data efficiently.

5. Create a pie chart showing the breakdown of occupations in the science and engineering jobs files.
    a. Open : Launch Microsoft on your computer.
    b. Import Data: Open the science and engineering jobs files. Make sure each file contains a column with occupation names and another column with corresponding counts or percentages.
    c. Merge Data (if necessary): If you have separate files for science and engineering jobs, consider merging them into one sheet for easier analysis. You can copy and paste the data from onefile into the other or use 's import functions to combine them.
    d. Prepare Data: Ensure that your data is clean and organized. Remove any unnecessary rows or columns, and make sure each column has a clear heading.
    e. Select Data: Click and drag to select the columns containing the occupation names and their corresponding counts or percentages.
    f. Insert Pie Chart: Go to the "Insert" tab on the ribbon and click on the "Pie Chart" button. Select a basic pie chart style to begin with.
    g. Customize Pie Chart: Once the pie chart is inserted, you can customize it to better representyour data. Right-click on the chart and select "Format Chart Area" to change the chart's appearance, such as colors, labels, and fonts.
    h. Labeling: Ensure that each segment of the pie chart is labeled appropriately. You can do this by clicking on the chart, then clicking on the "Chart Elements" button (a plus sign icon usually located on the top right corner of the chart), and checking the "Data Labels" option.
    i. Explode Segments (if necessary): If you want to emphasize specific segments of the pie chart, you can "explode" them. Click on the chart, then click on the specific segment you want to explode and drag it away from the center of the chart.
    j. Title and Legend: Add a title to your chart that clearly describes the data being represented. Also, ensure that there's a legend to explain what each color on the pie chart represents. You can add or edit these elements by clicking on the chart and selecting the "Chart Elements" button.
    k. Review and Finalize: Review your pie chart to ensure that it accurately represents the breakdown of occupations in the science and engineering jobs data. Make any necessary adjustments to the chart's appearance or labeling.
    l. Save and Share: Once you're satisfied with your pie chart, save your file to preserve your work. You can then share the file with others or export the chart as an image to include in presentations or reports.



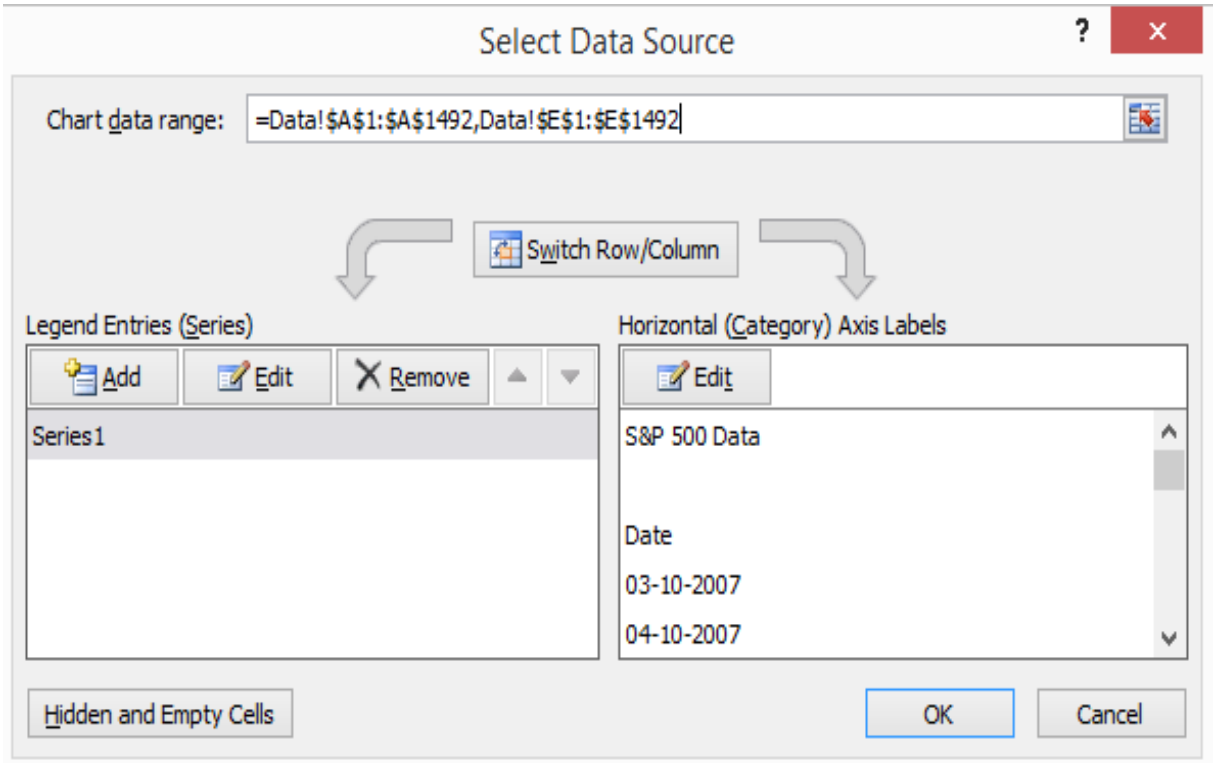
6. Create Line chart for the closing prices in the file S&P 500.

7. To create a line chart for the closing prices in the file S&P 500, follow these steps:

1. Open : Launch Microsoft on your computer.

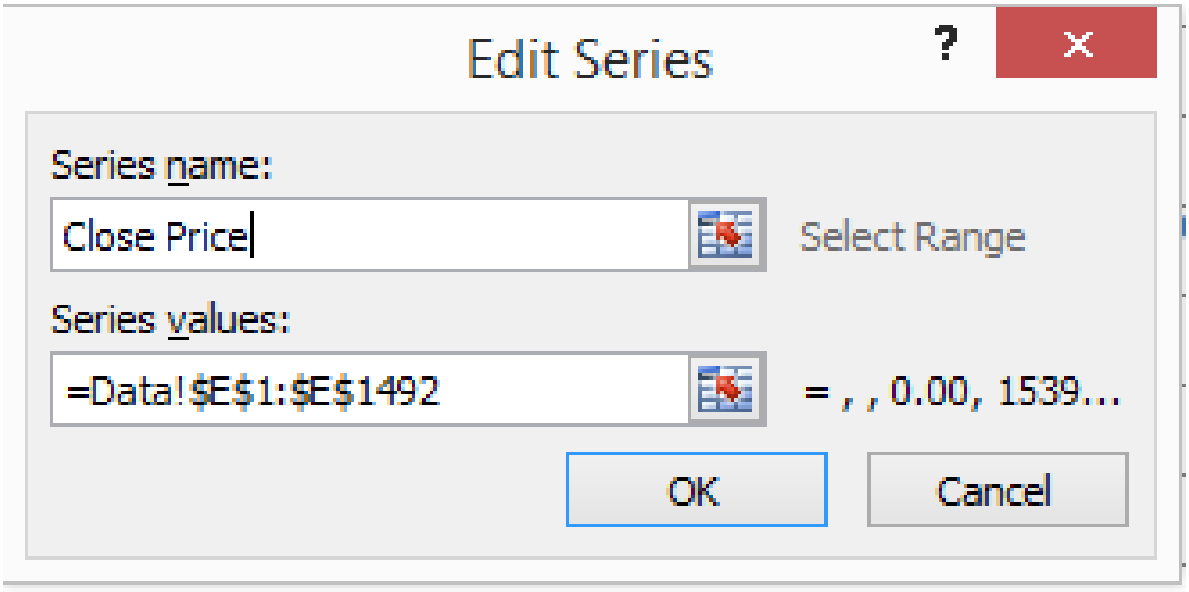
2. Import Data: Open the file S&P 500 that contains the data you provided.

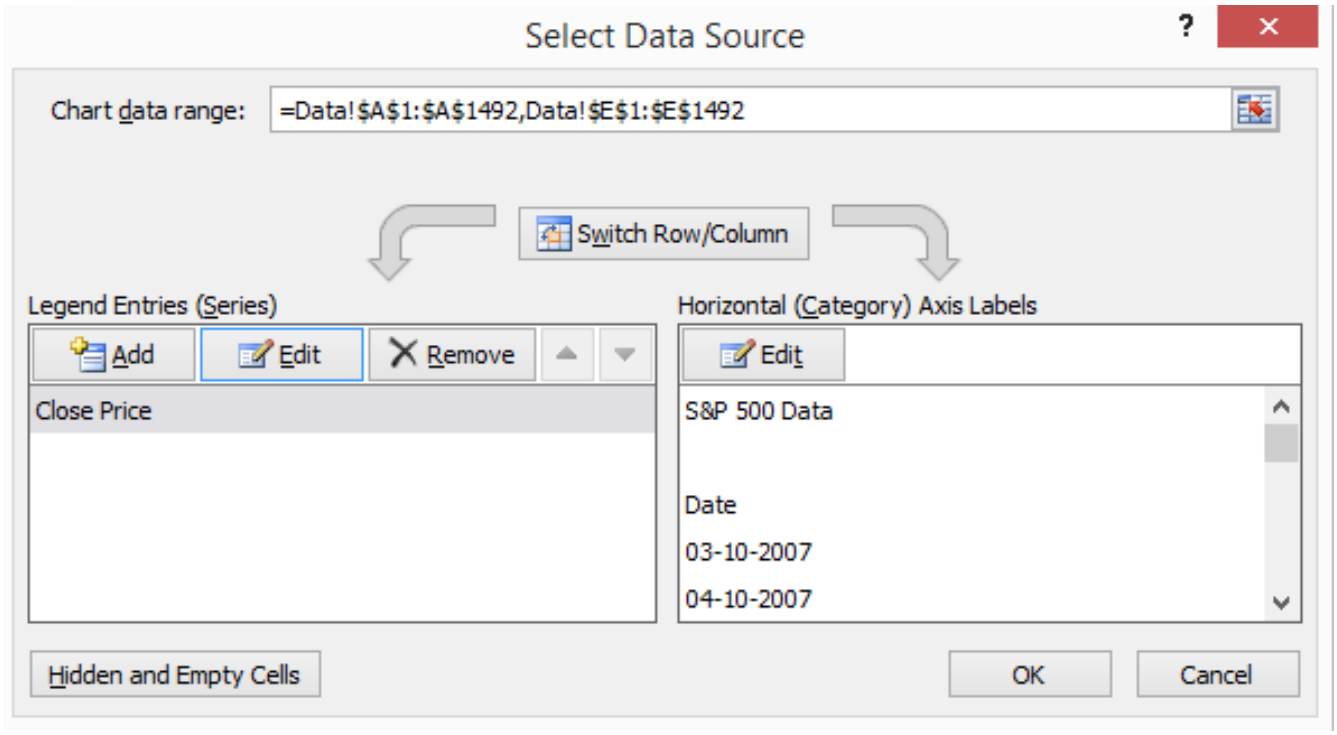
3. Select Data: Click and drag to select the "Date" column and the "Close" column. Make sure to include the column headers.



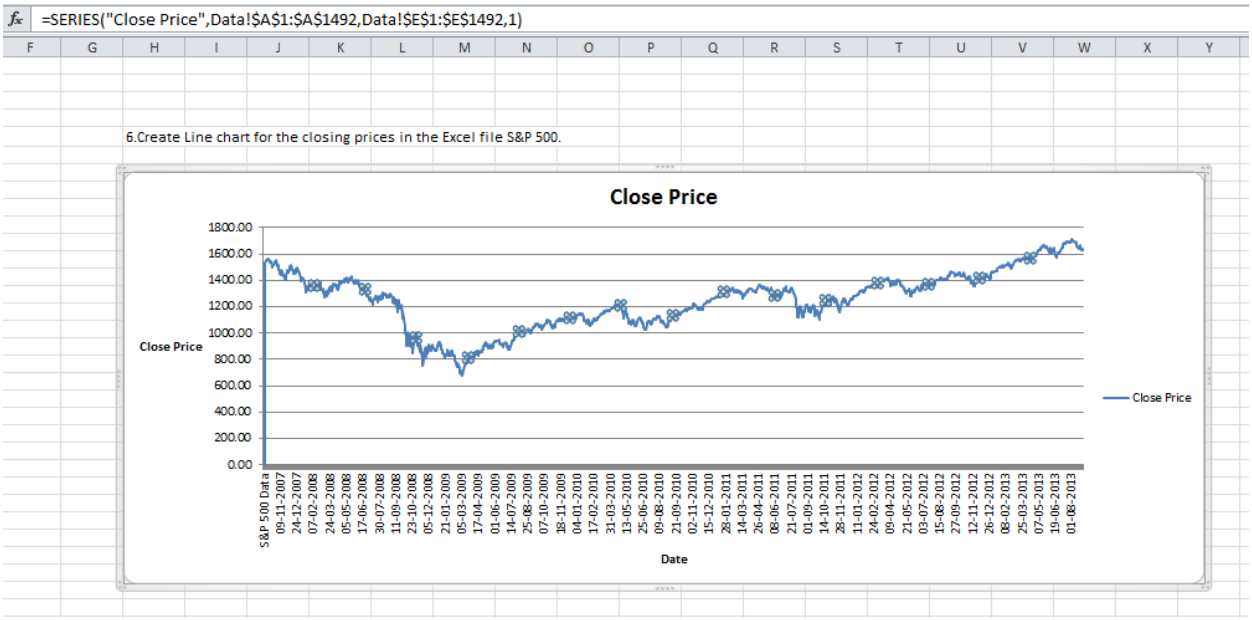
4. Insert Line Chart: Go to the "Insert" tab on the ribbon and click on the "Line Chart" button. Choose a suitable line chart style (e.g., basic line chart).

5. Customize Chart: Once the line chart is inserted, you can customize it to better represent your data. Right-click on the chart and select "Format Chart Area" to change the chart's appearance, such as colors, labels, and fonts.

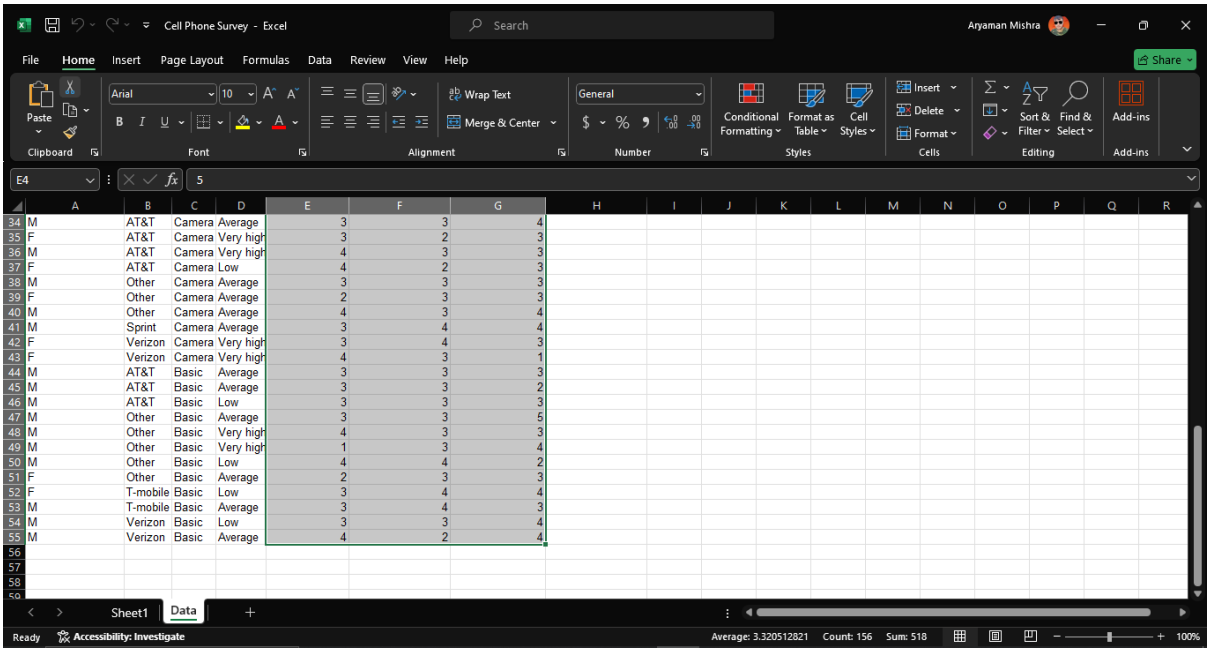




6. Axis Labels and Title: Ensure that the horizontal axis (X-axis) represents the dates and the verticalaxis (Y-axis) represents the closing prices. You can adjust these settings by right-clicking on the axis labels and selecting "Format Axis". Also, add a title to your chart that describes the data being represented.

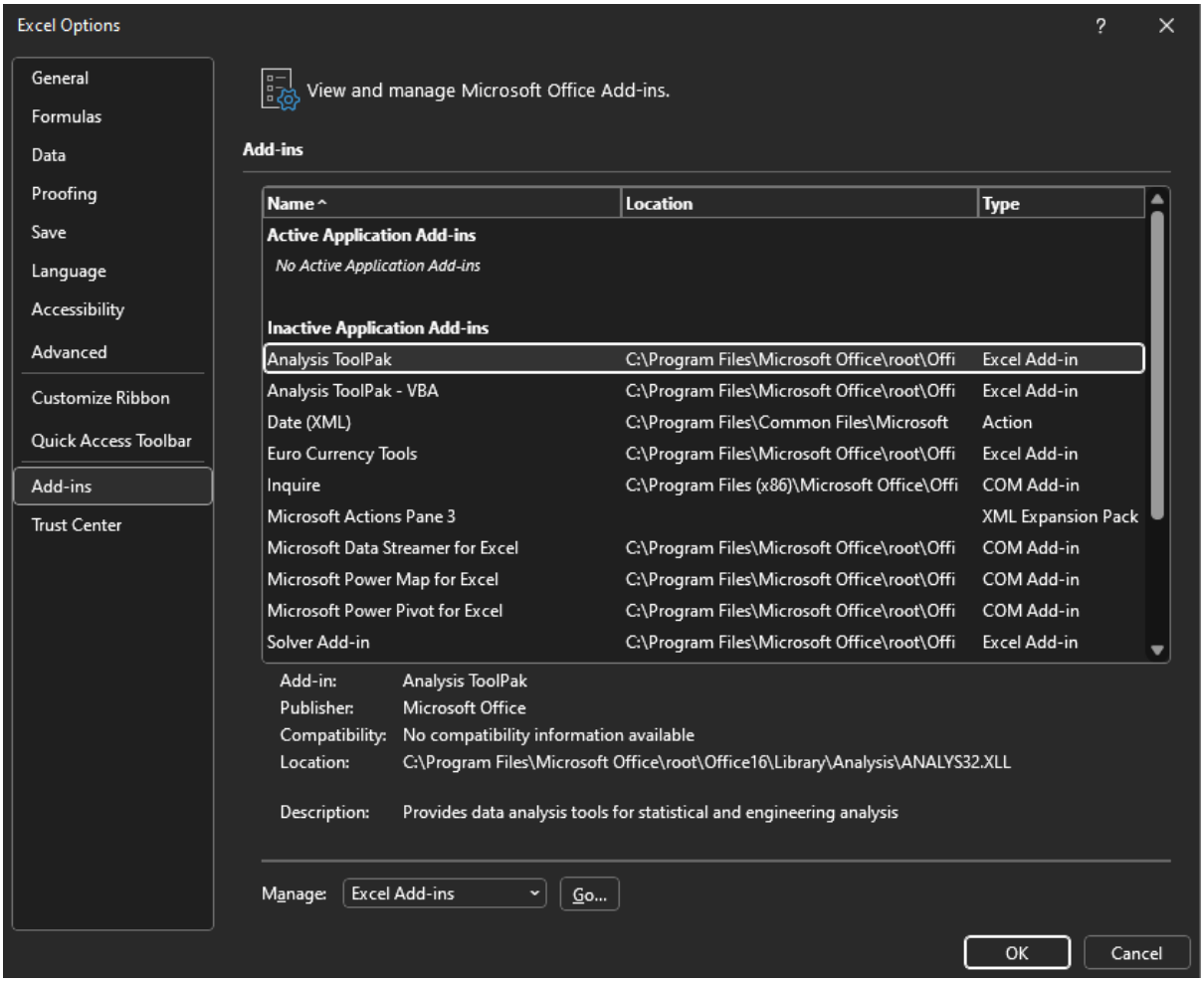


8. Construct frequency distribution and histogram for the numerical data in the file cell Phone Survey Also, compute the relative frequencies and Cumulative relative frequencies.
   Step 1: Organize Your Data
   Copy the data from the file into an Excel worksheet.
   Label your columns for easy identification:
   Column A: Gender
   Column B: Carrier
   Column C: Plan Type
   Column D: Satisfaction Level
   Column E: Price Rating
   Column F: Service Rating
   Column G: Overall Rating
   Step 2: Construct Frequency Distribution
   Select the data range for the numerical columns (Columns E, F, G).

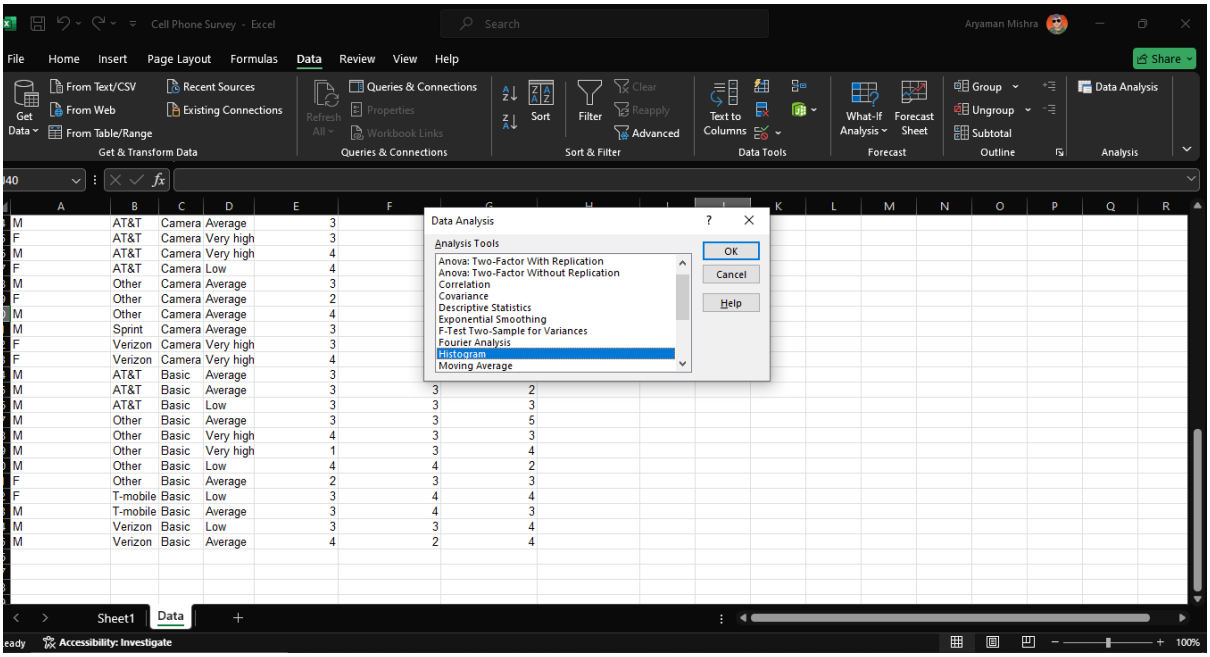


	A	B	C	D	E	F	G
34	M	AT&T	Camera	Average	3	3	4
35	F	AT&T	Camera	Very high	3	2	3
36	M	AT&T	Camera	Very high	4	3	3
37	F	AT&T	Camera	Low	4	2	3
38	M	Other	Camera	Average	3	3	3
39	F	Other	Camera	Average	2	3	3
40	M	Other	Camera	Average	4	3	4
41	M	Sprint	Camera	Average	3	4	4
42	F	Verizon	Camera	Very high	3	4	3
43	F	Verizon	Camera	Very high	4	3	1
44	M	AT&T	Basic	Average	3	3	3
45	M	AT&T	Basic	Average	3	3	2
46	M	AT&T	Basic	Low	3	3	3
47	M	Other	Basic	Average	3	3	5
48	M	Other	Basic	Very high	4	3	3
49	M	Other	Basic	Very high	1	3	4
50	M	Other	Basic	Low	4	4	2
51	F	Other	Basic	Average	2	3	3
52	F	T-mobile	Basic	Low	3	4	4
53	M	T-mobile	Basic	Average	3	4	3
54	M	Verizon	Basic	Low	3	3	4
55	M	Verizon	Basic	Average	4	2	4
56							
57							
58							

Go to the "Data" tab in Excel and select "Data Analysis" (if you don't see it, you may need to add the Analysis ToolPak from Excel Add-ins).



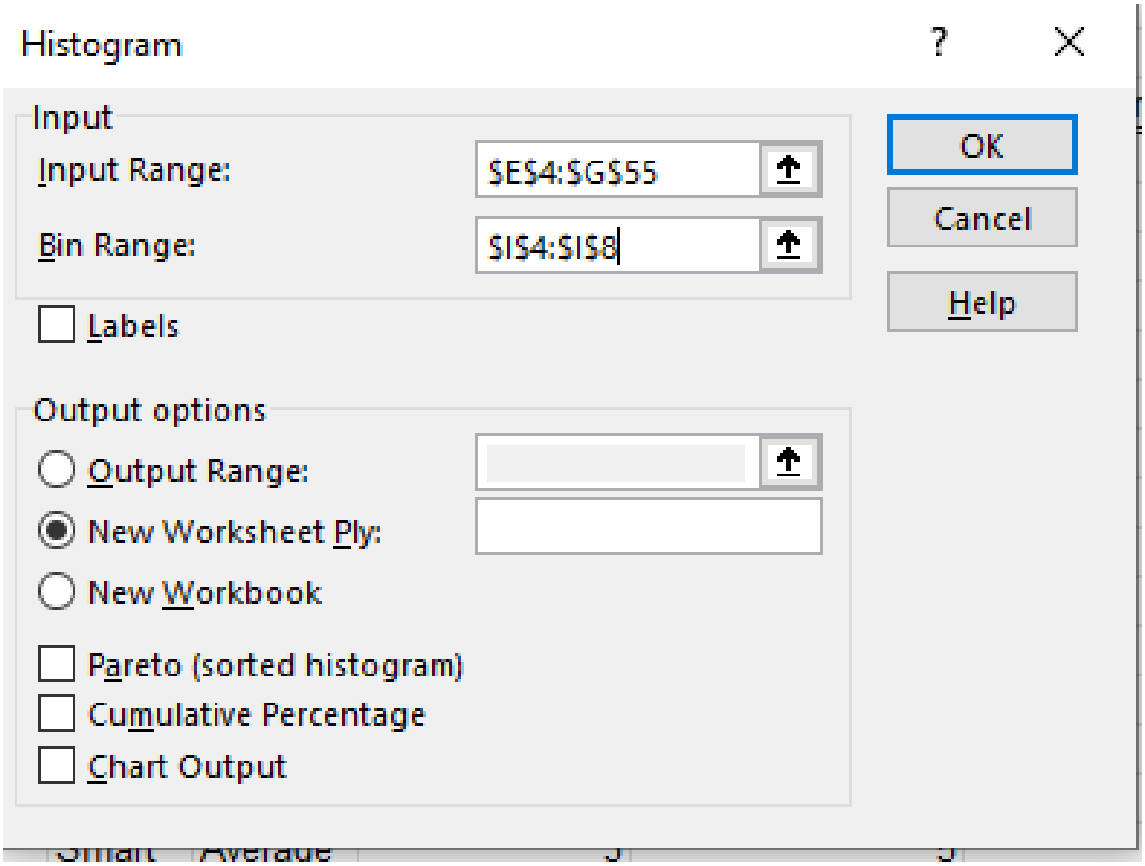
Choose "Histogram" from the list.

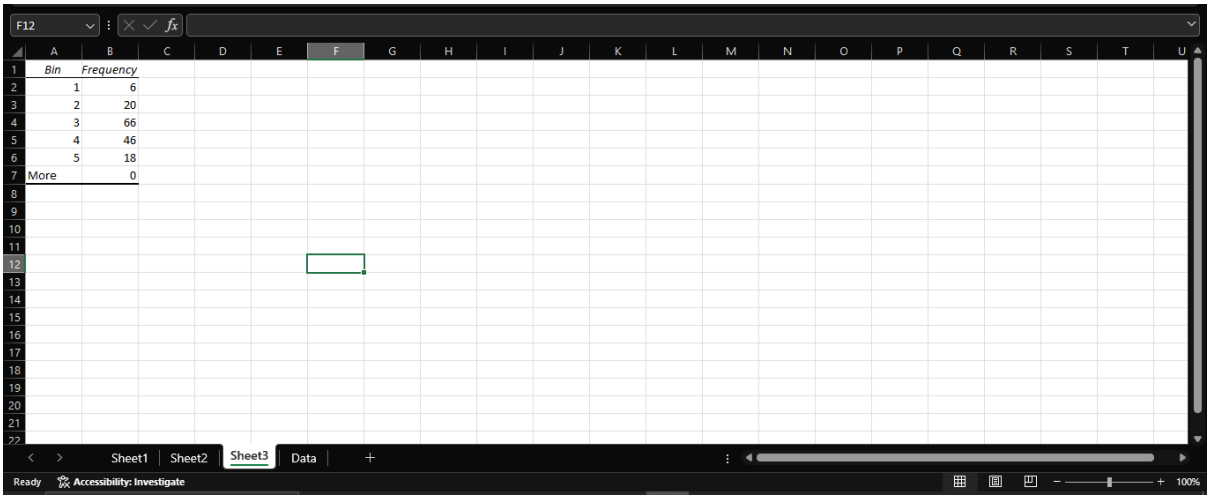


Input Range: Select the range of cells containing your numerical data (e.g., E2
).

Bin Range: Create a separate column with bin values (e.g., 1, 2, 3, 4, 5).

Output Range: Choose where you want the frequency distribution to appear.





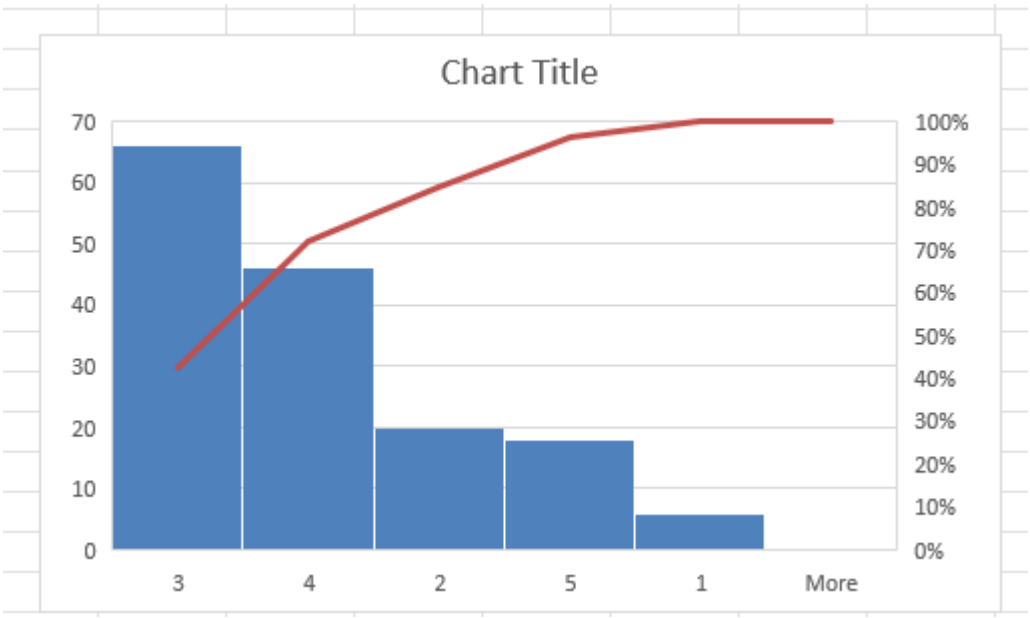
Step 3: Create Histogram

Highlight the frequency distribution results.

Go to the "Insert" tab in Excel.

Select "Histogram" from the "Charts" group.

Customize the histogram as needed (e.g., change titles, axis labels, etc.).



Step 4: Calculate Relative Frequencies

In a new column, divide each frequency by the total number of observations.

For example, if the frequency is in cell H2, the total count is in H15, use the formula: =H2/H15.

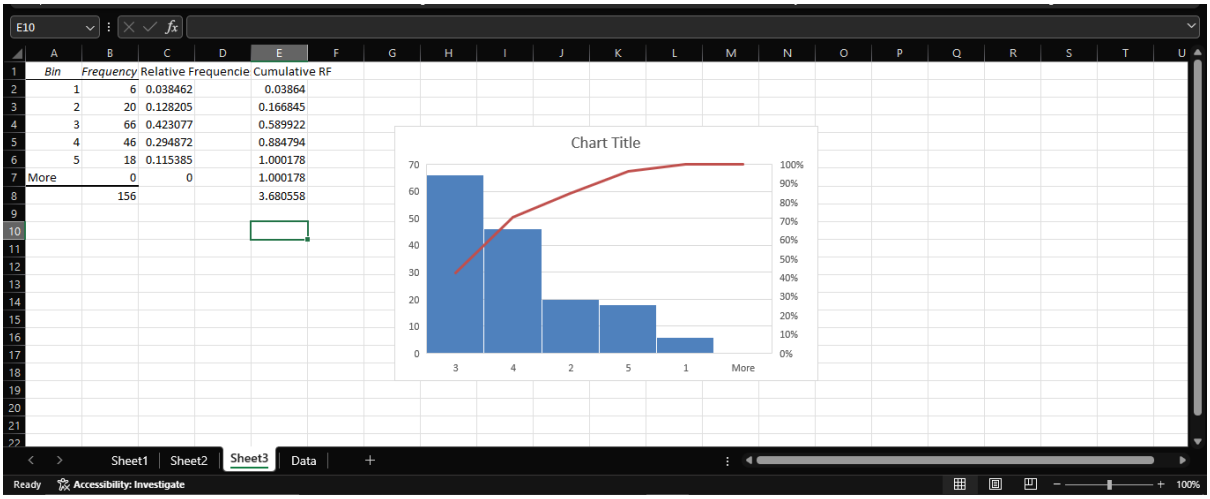
Format the new column as a percentage.

Step 5: Calculate Cumulative Relative Frequencies

In the next column, calculate cumulative relative frequencies by adding the current relative frequency to the sum of the previous ones.

For example, if the first relative frequency is in I2, the cumulative for the next row will be: =I2+I3.

Drag the formula down to fill the column.



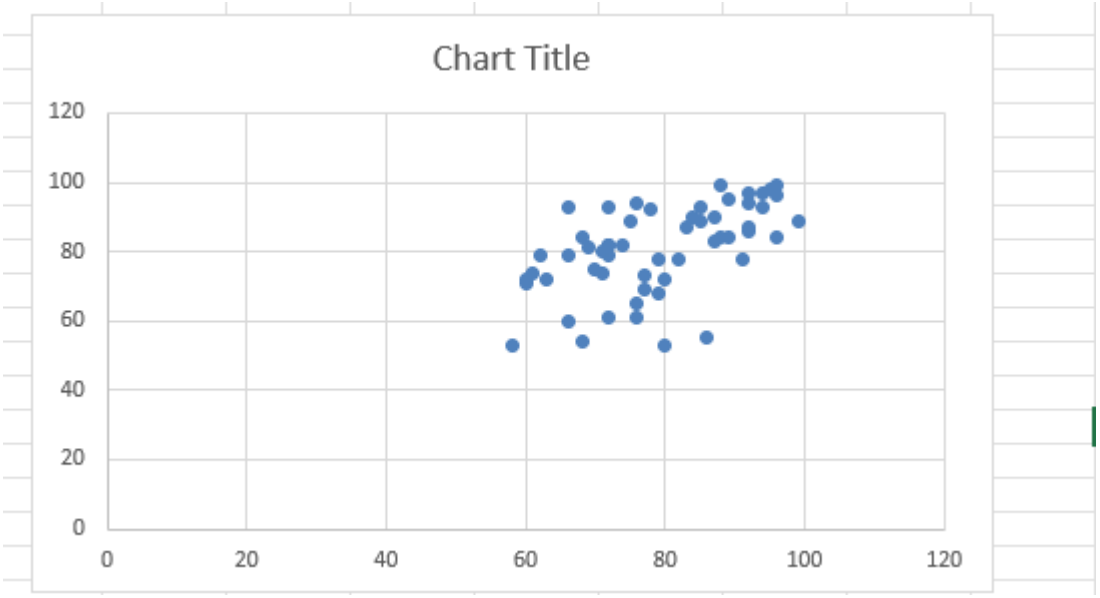
9. Using the data in the files Student Grades, construct a Scatter chart for midterm versus final exam and add a linear trend line. What is the regression model? if a student scores 85 on the midterm, what would be you predict her grade on the final exam to be?

Step 1: Enter the Data

1. **Copy the data** into an Excel worksheet.
2. **Label the columns** as follows:
   - Column A: "Student"
   - Column B: "Midterm"
   - Column C: "Final Exam"

Step 2: Create a Scatter Plot

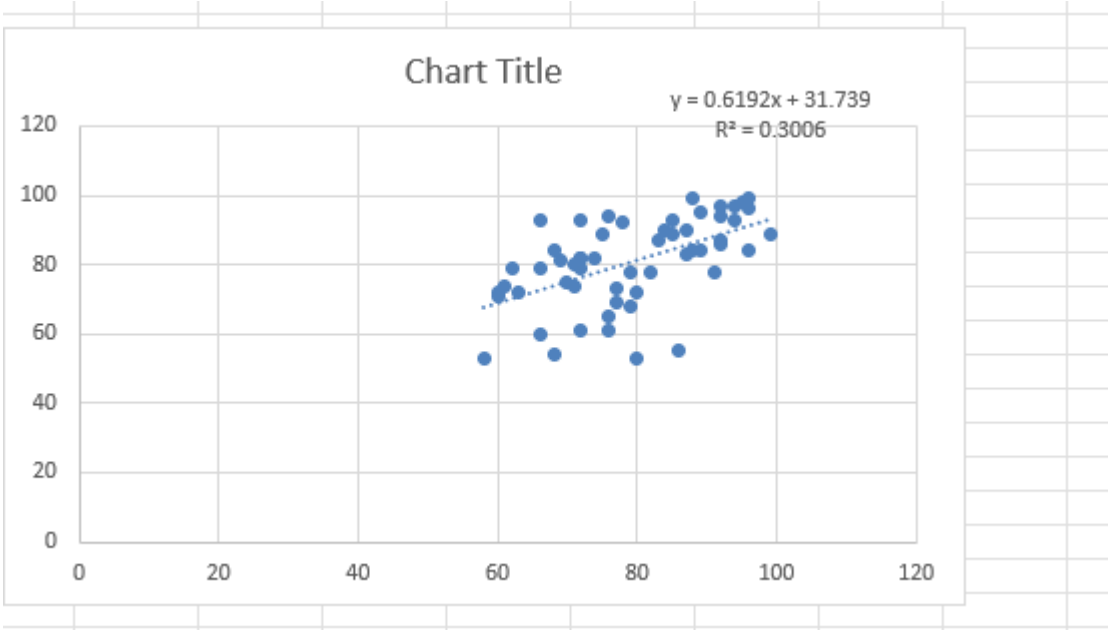
1. **Highlight the data in the "Midterm" and "Final Exam" columns** (Columns B and C).
2. **Go to the "Insert" tab** in Excel.
3. **Choose "Scatter"** from the "Charts" group and select the "Scatter with only Markers" option.



Step 3: Add a Linear Trend Line

1. **Click on the chart** to select it.
2. **Click on one of the data points** to select the data series.
3. **Right-click on the data series** and select "Add Trendline."
4. **In the Format Trendline pane**:

- Choose "Linear."
- Check the box for "Display Equation on chart."
- Check the box for "Display R-squared value on chart."



Step 4: Determine the Regression Model

- **The regression model will be displayed on the chart** in the form of an equation (e.g., y = mx + b).

Step 5: Predict Final Exam Score for a Midterm Score of 85

Example Output:

- **Scatter chart** showing the relationship between midterm and final exam scores.
- **Linear trend line** with the regression equation and R-squared value displayed.
- **Predicted final exam score** for a student who scores 85 on the midterm.

Given the regression equation:

y = 0.6192x + 31.739

where y represents the final exam score and x represents the midterm score, you can predict the final exam score for a student who scores 85 on the midterm by substituting x = 85 into the equation.

Calculation:
y = 0.6192(85) + 31.739

y = 52.632 + 31.739

y = 84.371

Predicted Final Exam Score:
If a student scores 85 on the midterm, the predicted final exam score is approximately 84.37.

10. Find the 1st and 3rd quartiles for the time difference between the scheduled and actual arrival times in the Atlanta Airlines Data Files.

To find the 1st and 3rd quartiles for the "Time Difference (Minutes)" in the Atlanta Airlines data, follow these steps in :

1. Enter Data into :

First, input the data into . Here's how your data might look:

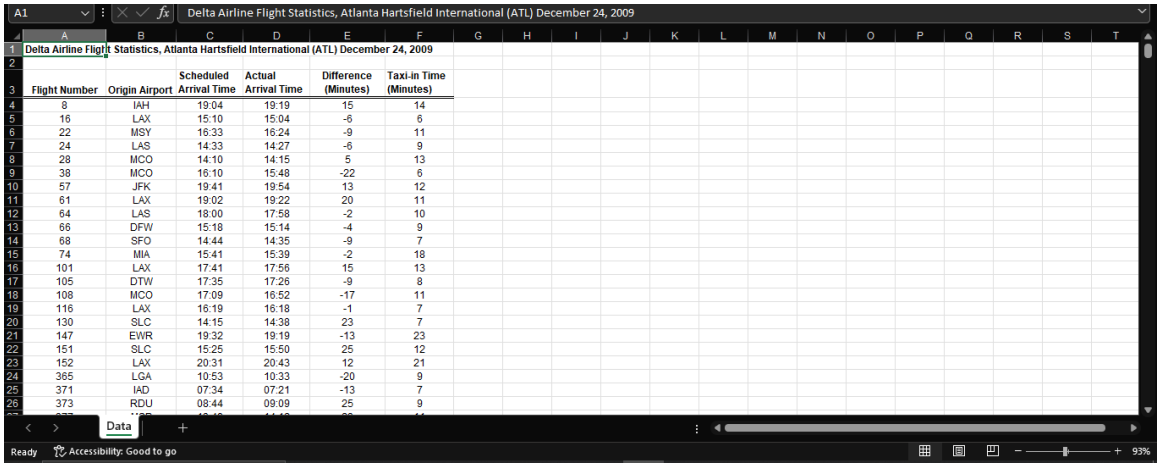
```
css

A       | B      | C                     | D                     | E
| F
Flight  | Origin | Scheduled Arrival Time | Actual Arrival Time   | Time
Difference (Minutes) | Taxi-in Time (Minutes)
8       | IAH    | 19:04                 | 19:19                 | 15
| 14
16      | LAX    | 15:10                 | 15:04                 | -6
| 6
22      | MSY    | 16:33                 | 16:24                 | -9
| 11
24      | LAS    | 14:33                 | 14:27                 | -6
| 9
28      | MCO    | 14:10                 | 14:15                 | 5
| 13
38      | MCO    | 16:10                 | 15:48                 | -22
| 6
57      | JFK    | 19:41                 | 19:54                 | 13
| 12
61      | LAX    | 19:02                 | 19:22                 | 20
| 11
64      | LAS    | 18:00                 | 17:58                 | -2
| 10
66      | DFW    | 15:18                 | 15:14                 | -4
| 9
68      | SFO    | 14:44                 | 14:35                 | -9
| 7
74      | MIA    | 15:41                 | 15:39                 | -2
| 18
101     | LAX    | 17:41                 | 17:56                 | 15
| 13
105     | DTW    | 17:35                 | 17:26                 | -9
| 8
108     | MCO    | 17:09                 | 16:52                 | -17
| 11
116     | LAX    | 16:19                 | 16:18                 | -1
| 7
130     | SLC    | 14:15                 | 14:38                 | 23
| 7
147     | EWR    | 19:32                 | 19:19                 | -13
| 23
151     | SLC    | 15:25                 | 15:50                 | 25
| 12
152     | LAX    | 20:31                 | 20:43                 | 12
| 21
```

```
365      | LGA     | 10:53                    | 10:33                    | -20
| 9
```

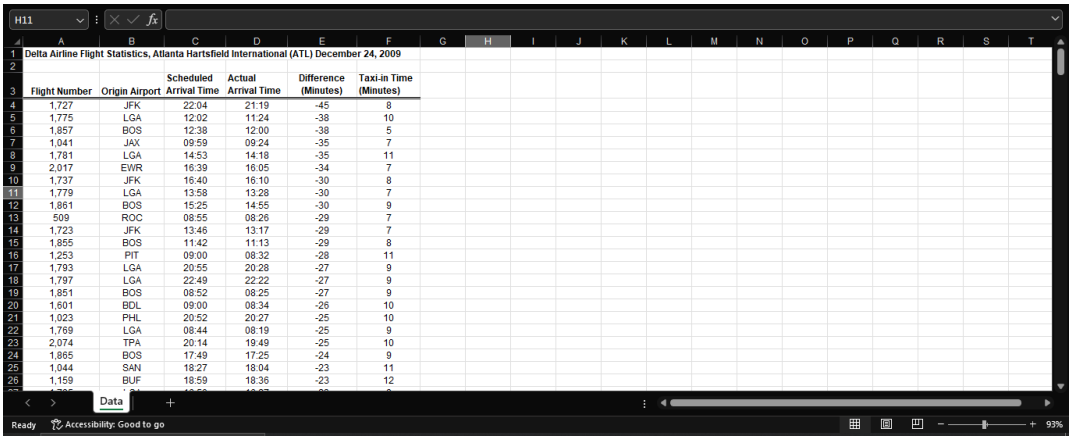
2. Sort Data:

1. **Sort the Data by `Time Difference (Minutes)`:**



Flight Number	Origin Airport	Scheduled Arrival Time	Actual Arrival Time	Difference (Minutes)	Taxi-in Time (Minutes)
8	IAH	19:04	19:19	15	14
16	LAX	15:10	15:04	-6	6
22	MSY	16:33	16:24	-9	11
24	LAS	14:33	14:27	-6	9
28	MCO	14:10	14:15	5	13
38	MCO	16:10	15:48	-22	6
57	JFK	19:41	19:54	13	12
61	LAX	19:02	19:22	20	11
64	LAS	18:00	17:58	-2	10
66	DFW	15:18	15:14	-4	9
68	SFO	14:44	14:35	-9	7
74	MIA	15:41	15:39	-2	18
101	LAX	17:41	17:56	15	13
105	DTW	17:35	17:26	-9	8
108	MCO	17:09	16:52	-17	11
116	LAX	16:19	16:18	-1	7
130	SLC	14:15	14:38	23	7
147	EWR	19:32	19:19	-13	23
151	SLC	15:25	15:50	25	12
152	LAX	20:31	20:43	12	21
365	LGA	10:53	10:33	-20	9
371	IAD	07:34	07:21	-13	7
373	RDU	08:44	09:09	25	9

   - Highlight the column E (`Time Difference (Minutes)`).
   - Go to the `Data` tab.
   - Click `Sort A to Z` to sort the data in ascending order.



Flight Number	Origin Airport	Scheduled Arrival Time	Actual Arrival Time	Difference (Minutes)	Taxi-in Time (Minutes)
1,727	JFK	22:04	21:19	-45	8
1,775	LGA	12:02	11:24	-38	10
1,857	BOS	12:38	12:00	-38	5
1,041	JAX	09:59	09:24	-35	7
1,781	LGA	14:53	14:18	-35	11
2,017	EWR	16:39	16:05	-34	7
1,737	JFK	16:40	16:10	-30	8
1,779	LGA	13:58	13:28	-30	7
1,861	BOS	15:25	14:55	-30	9
509	ROC	08:55	08:26	-29	7
1,723	JFK	13:46	13:17	-29	7
1,855	BOS	11:42	11:13	-29	8
1,253	PIT	09:00	08:32	-28	11
1,793	LGA	20:55	20:28	-27	9
1,797	LGA	22:49	22:22	-27	9
1,851	BOS	08:52	08:25	-27	9
1,601	BDL	09:00	08:34	-26	10
1,023	PHL	20:52	20:27	-25	10
1,769	LGA	08:44	08:19	-25	9
2,074	TPA	20:14	19:49	-25	10
1,865	BOS	17:49	17:25	-24	9
1,044	SAN	18:27	18:04	-23	11
1,159	BUF	18:59	18:36	-23	12

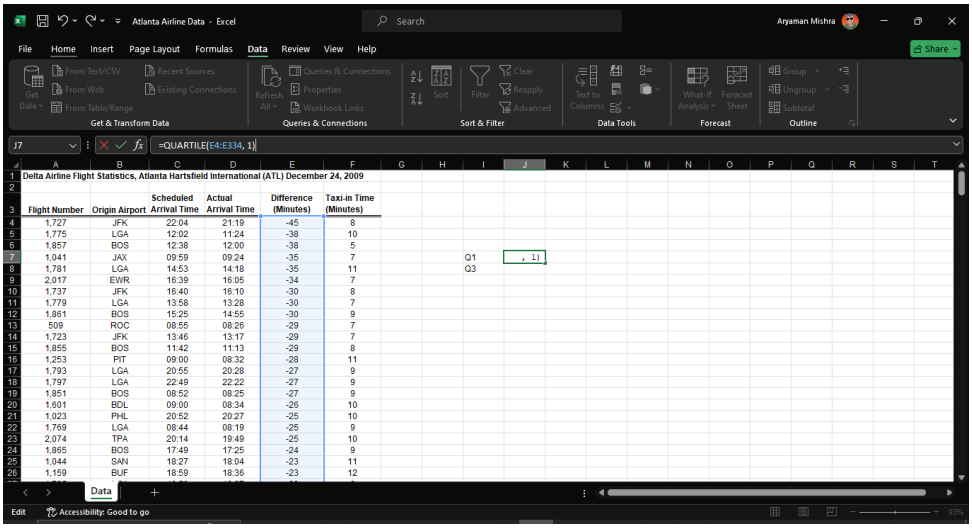
3. Calculate the Quartiles:

1. **Use 's QUARTILE Function:**
   - **1st Quartile (Q1):**
     - Formula: `=QUARTILE(E2:E21, 1)`
     - This calculates the 1st quartile, which is the value below which 25% of the data falls.
   - **3rd Quartile (Q3):**
     - Formula: `=QUARTILE(E2:E21, 3)`
     - This calculates the 3rd quartile, which is the value below which 75% of the data falls.

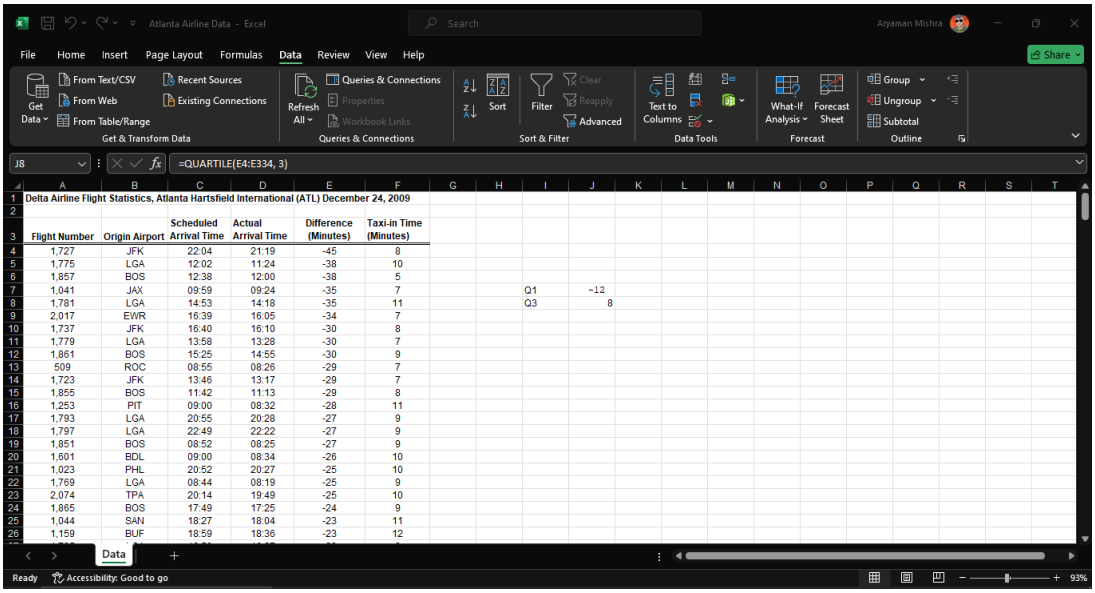
Example Calculation:

Here’s how you can do this in :

1. **Enter the formula for Q1 in an empty cell:**
    - `=QUARTILE(E2:E21, 1)`



2. **Enter the formula for Q3 in another empty cell:**
    - `=QUARTILE(E2:E21, 3)`



Given your data, the results should be:

- **1st Quartile (Q1):** -12 minutes
- **3rd Quartile (Q3):** 8 minutes

Interpretation:

- **1st Quartile (Q1) of -9 minutes:** This means that 25% of the flights have a time difference of -12 minutes or less.

- **3rd Quartile (Q3) of 15 minutes:** This means that 75% of the flights have a time difference of 8 minutes or less.

Conclusion:

- **Time Difference Distribution:** The time differences between scheduled and actual arrival times vary widely. The quartiles help you understand that a significant portion of the flights have either early or late arrivals, with the middle 50% of the data falling between -12 and 8 minutes.
- **Comparative Analysis:** If you need to compare with another dataset, you would follow the same process and analyze the differences between the quartiles to draw comparisons about punctuality or delays.

This statistical analysis provides insights into the variability of arrival times and helps in understanding the extent of delays or early arrivals for flights.