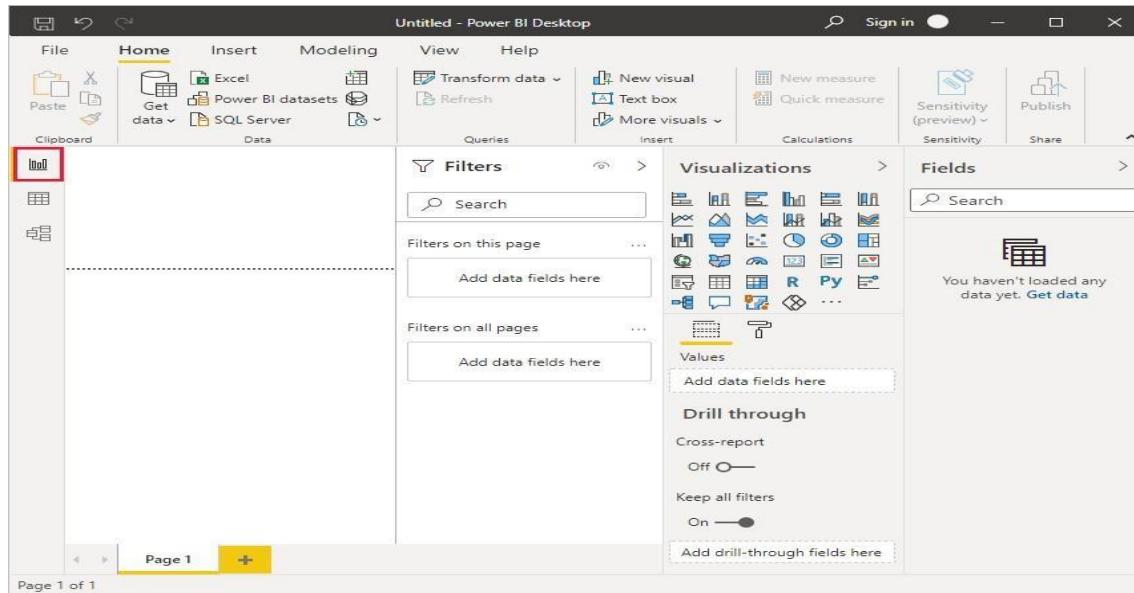


PROGRAM 6: Querying Data from CSV - Query Editor, Connecting the data from the Excel Source, Clean, Transform the data.

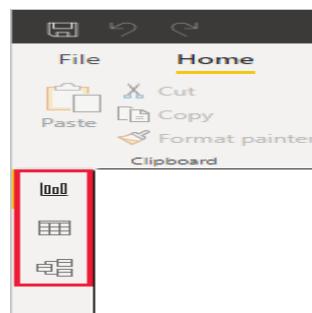
Power BI Desktop also includes the Power Query Editor, which opens in a separate window. In Power Query Editor, you can build queries and transform data, then load the refined data model into Power BI Desktop to create reports.



Along the left side of Power BI Desktop are icons for the three Power BI Desktop views:

Report, Data, and Model, from top to bottom. The current view is indicated by the yellow bar along the left, and you can change views by selecting any of the icons.

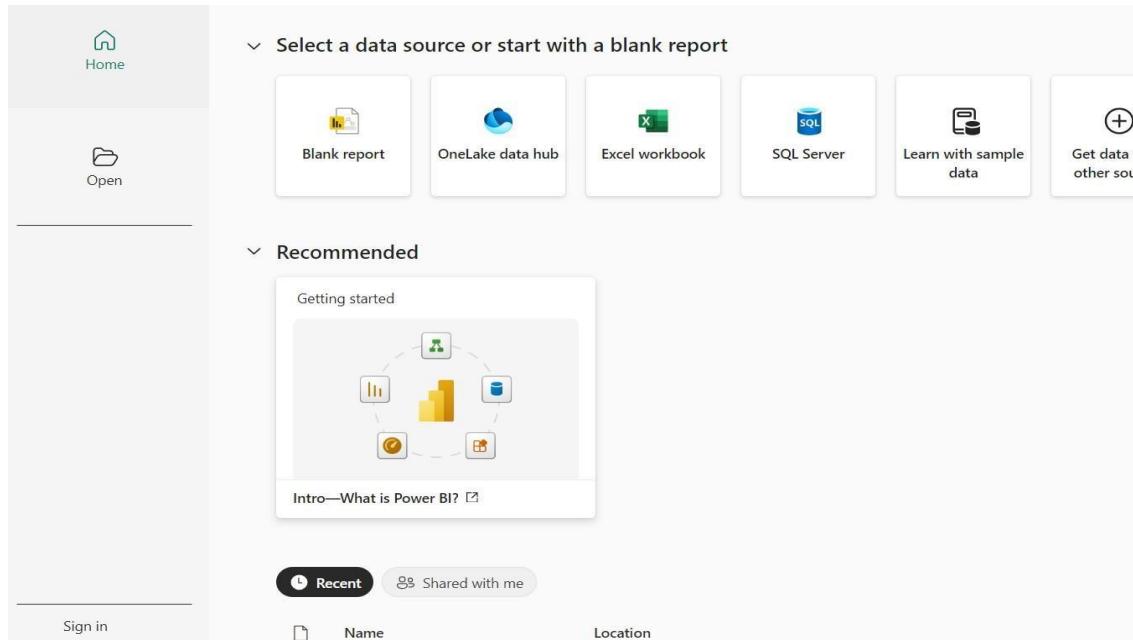
Report view is the default view.



Connect to data(Get Data from different Sources)

With Power BI Desktop installed, we can connect to the world of data. To see the many types of data sources available,

Once Power BI screen is seen click on blank Report



After clicking on blank Report the below screen appears.

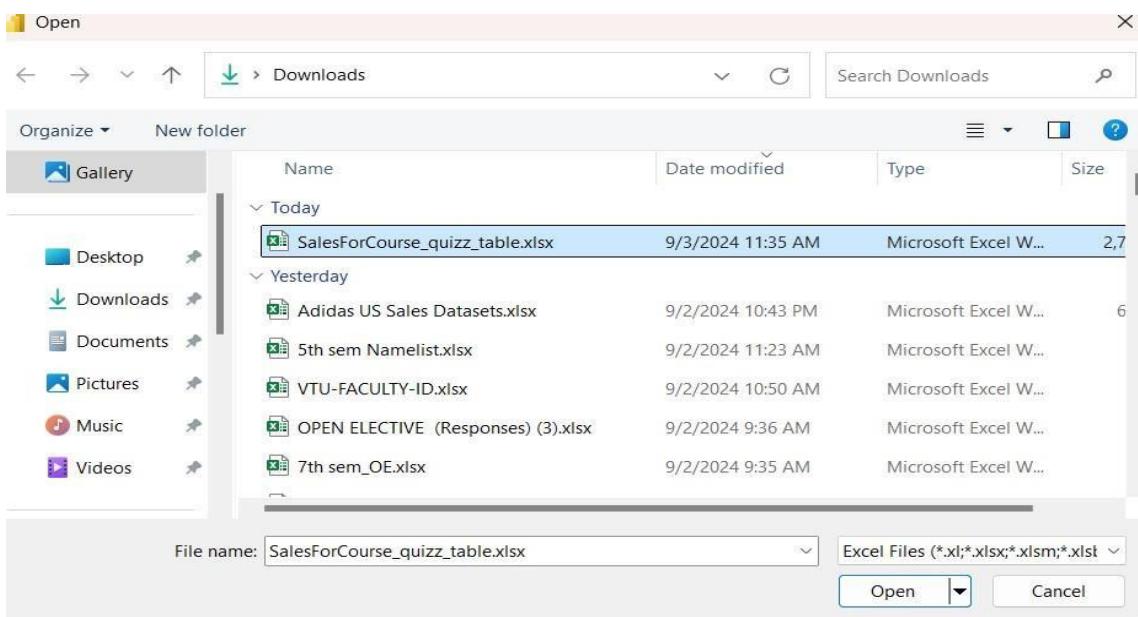
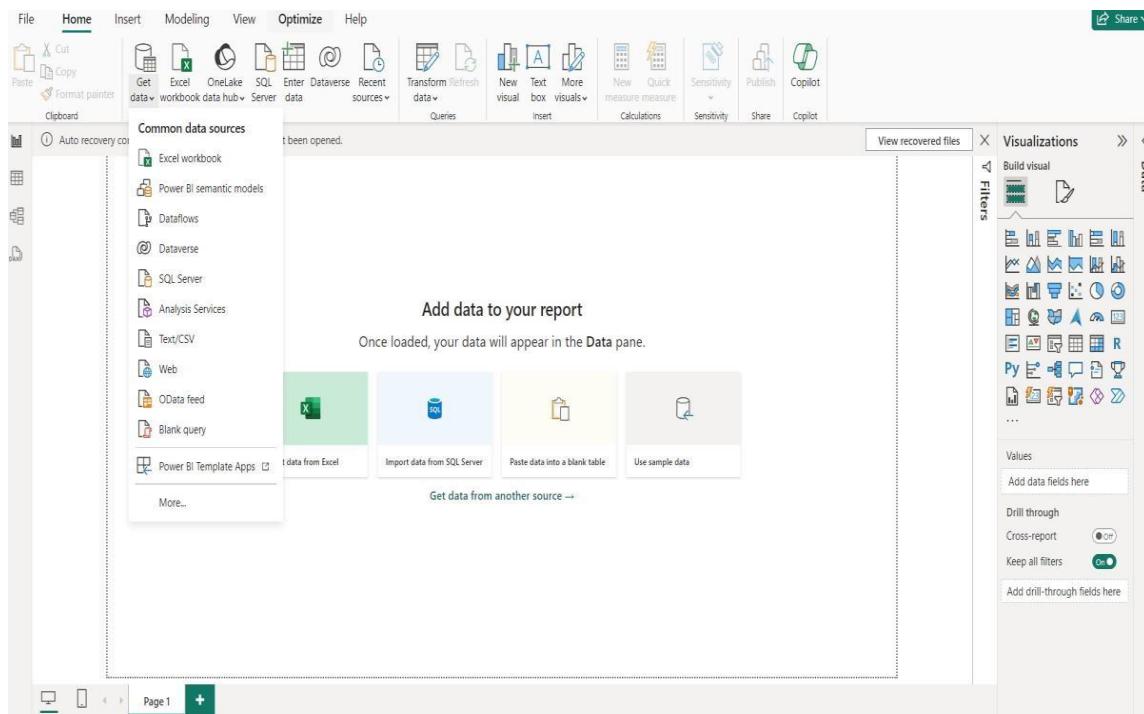
NOW TO GET DATA FROM DIFFERENT SOURCES

The steps

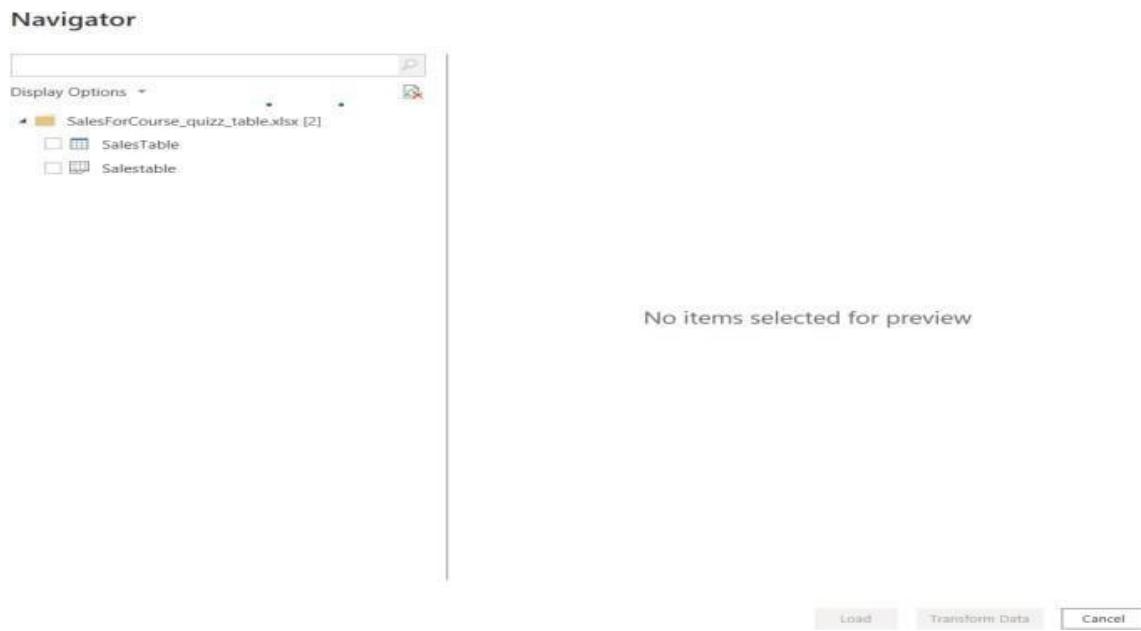
Step 1 : Select **Get Data** in the Power BI Desktop Home tab, and in the Get Data window, scroll through the list of All data sources.(like Excel,CSV,Oracle....)

On the Power BI Desktop Home tab, select Get Data > Excel workbook

Data Visualization Laboratory (BAIL504)



1. Click on the file you need and open the file ,once you open the file below window with navigator appears ,select the file (2nd option to see the contents of the file)



2. At this point you can select Load to load the table, or Transform data to make changes in the table before you load it.
3. When you select Transform data, Power Query Editor launches, with a representative view of the table. The Query Settings pane is on the right, or you can always show it by selecting Query Settings on the View tab of Power Query Editor.

The screenshot shows the Microsoft Power Query Editor interface. The main area displays a table titled "Ranking of best and worst..." with columns: State, Overall rank, Affordability, Crime, and Culture. The table contains 15 rows of data. The Query Settings pane on the right shows the query name "Ranking of best and worst states for retire" and the applied step "Changed Type".

	Column1	Column2	Column3	Column4	Column5
1	State	Overall rank	Affordability	Crime	Culture
2	Source: Bankrate's 2019 "Bes...				
3	Nebraska	1	14	19	21
4	Iowa	2	8	15	20
5	Missouri	3	1	42	33
6	South Dakota	4	17	23	12
7	Florida	5	25	29	13
8	Kentucky	6	9	9	46
9	Kansas	7	7	39	37
10	North Carolina	7	13	28	28
11	Montana	9	16	31	2
12	Hawaii	10	45	24	9
13	Arkansas	11	4	46	39
14	Wisconsin	12	20	15	17
15	North Dakota	13	22	17	26
16					

7 COLUMNS, 52 ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 12:22 PM

Transforming the data

Once connected to a data source, you can adjust the data to meet your needs.

To transform the data, you provide Power Query Editor with step-by-step instructions for adjusting the data while loading and presenting it. Transforming doesn't affect the original data source, only this particular view of the data.

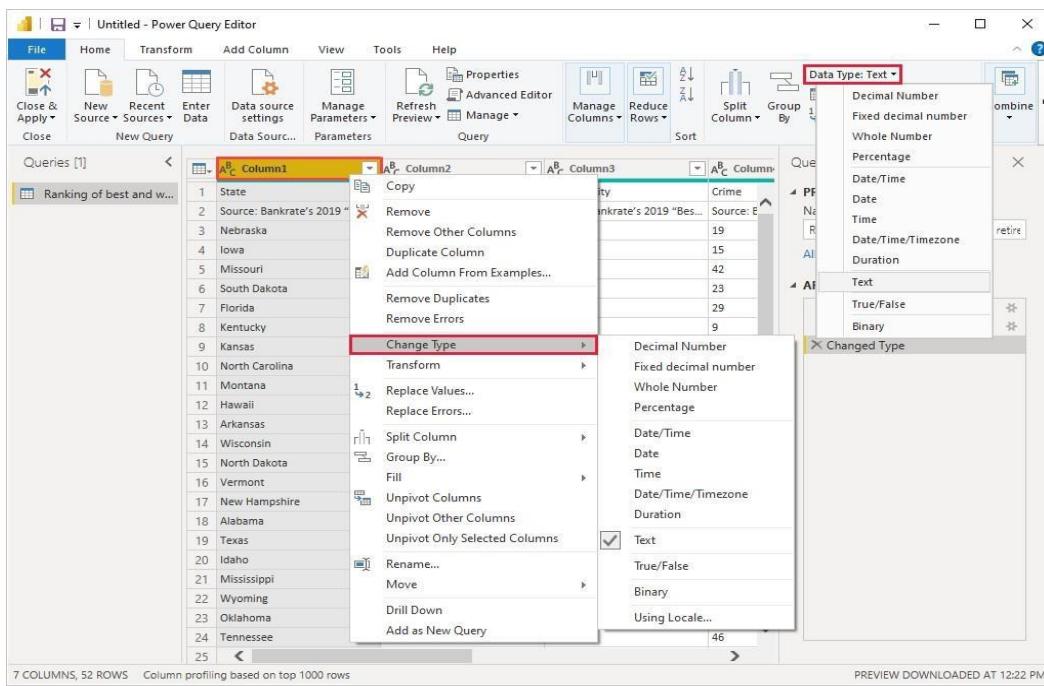
Transforming the data, includes **renaming columns or tables, removing rows or columns, or changing data types.**

Power Query Editor captures these steps sequentially under Applied Steps in the Query Settings pane.

Notice that the Applied Steps in Query Settings already contain a few steps. You can select each step to see its effect in the Power Query Editor

To Change a data type

- Select the column or columns to change.
- Hold down the Shift key to select several adjacent columns, or Ctrl to select non-adjacent columns.
- Either right-click a column header, select Change Type,
- choose a new data type from the menu, or drop down the list next to Data Type in the Transform group of the Home tab,
- select a new data type.



To Reduce/Delete the Rows

- From the Home tab select
- Reduce Rows > Remove Rows > Remove Bottom Rows.
- In the Remove Bottom Rows dialog box, enter *10*, and then select OK.

The screenshot shows the Power Query Editor interface. In the top ribbon, the 'Home' tab is selected. The 'Manage Columns' group is highlighted with a red box. A context menu is open over the 'Crime' column, also highlighted with a red box. The 'Remove Rows' option is selected, and its submenu is visible, with 'Remove Bottom Rows' highlighted with a red box. The main table view shows columns for State, Overall rank, and Affordability.

Remove Bottom Rows

Specify how many rows to remove from the bottom.

Number of rows: 10

OK Cancel

Query Settings

PROPERTIES

- Remove Top Rows
- Remove Bottom Rows
- Remove Alternate Rows
- Remove Duplicates
- Remove Blank Rows
- Remove Errors

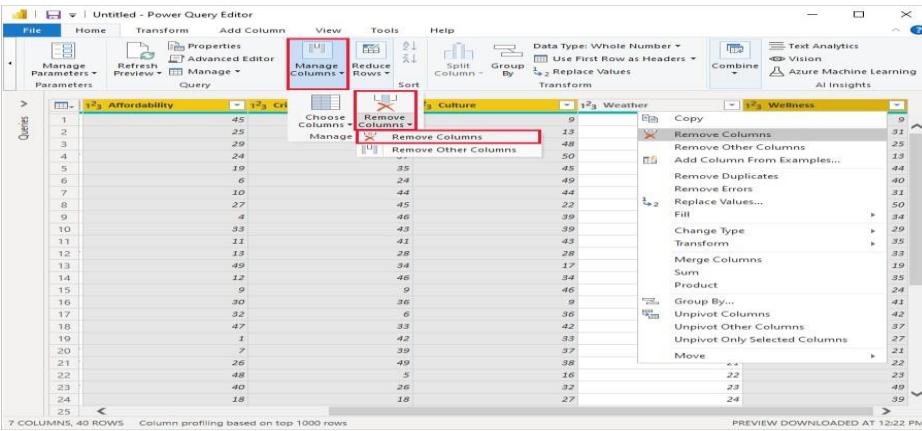
Applied Steps

- Changed Type
- Promoted Headers
- Changed Type1
- Removed Top Rows
- Changed Type2
- Sorted Rows

The bottom 10 worst rows are removed from the table, and the step Removed Bottom Rows appears in Applied Steps.

To Remove columns

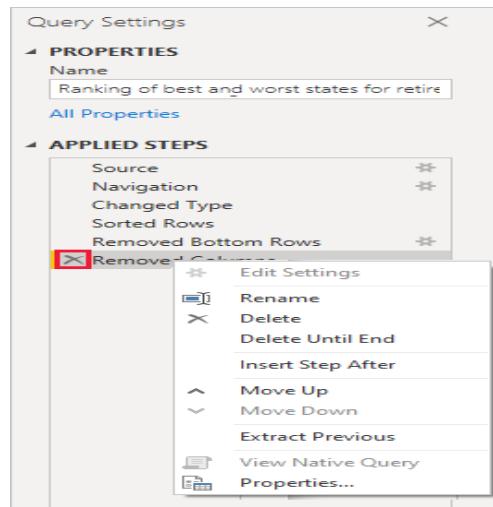
- From Home Tab Select Manage Columns group □ select Remove Columns.
- You can also right-click one of the selected column headers and select Remove Columns from the menu.
- The selected columns are removed, and the step Removed Columns appears in Applied Steps.



Applied steps in the Query setting pane

Right-click any step in the Applied Steps pane and choose to delete it, rename it, move it up or down in the sequence, or add or delete steps after it.

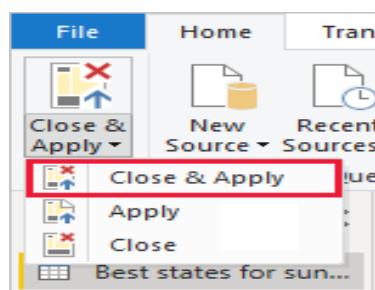
For intermediate steps, Power BI Desktop will warn you if the change could affect later steps and break your query.



Once all the required transformations are done the report should be created in the Power BI Desktop

- Apply the changes in Power Query Editor and load them into Power BI Desktop
- Selecting **Close & Apply** from the Home tab of the ribbon.

You can also select just **Apply** to keep the query open in Power Query Editor while you work in Power BI Desktop.

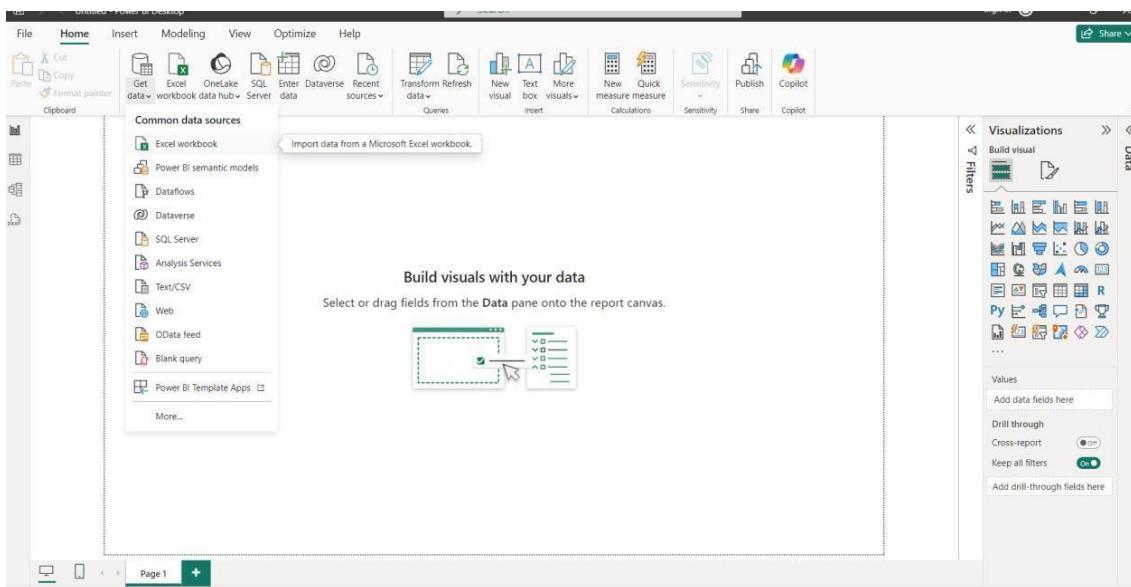


Program 7: Creating Reports & Visualizations - Different types of charts, Formatting charts with Title, Colors

17 Most Common Charts available in Power BI:

- Bar Chart
- Line Chart
- Scatterplot
- Sparkline
- Pie Chart
- Gauge
- Waterfall Chart
- Funnel Chart
- Heat Map / Matrix
- Histogram
- Box Plot
- Maps
- Tables
- Indicators
- Area Chart
- Radar or Spider Chart
- Tree Map
- Open **Power BI Desktop**
- Click on **Get data** in ribbon pane
- Click on **Excel worksheet** option

Data Visualization Laboratory (BAIL504)



- Choose specific dataset and open it. Example: HR Data.csv
- Click on **Transform Data** button

Power Query Editor window will open.

Data Visualization Laboratory (BAIL504)

The screenshot shows the Power Query Editor interface with a table containing 28 rows of data. The columns are labeled Column1 through Column7. The first row is highlighted as the header. The 'Use First Row as Headers' button is selected in the ribbon. The 'APPLIED STEPS' pane shows a single step: 'Changed Type'.

- We have to perform some transformation on this table

- Select row 1 and click on **Use first row as header**

The screenshot shows the Power Query Editor interface with the same table as before, but now the first row is the header and the data rows start from row 2. The 'APPLIED STEPS' pane shows the 'Use First Row as Headers' step.

- Then, we have to create new column for **attrition count**. For this, select attrition column click on **Add Column** new window will open then add details as follows. Once you are done with this, attrition count column will be added as a last row of the table

- Change the datatype of this column to **whole number**

The screenshot shows the Microsoft Power Query Editor interface. A modal dialog box titled "Add Conditional Column" is open, prompting the user to define a new column based on existing values. The rule defined is:

```

if Attrition equals Yes then 1
else 0
  
```

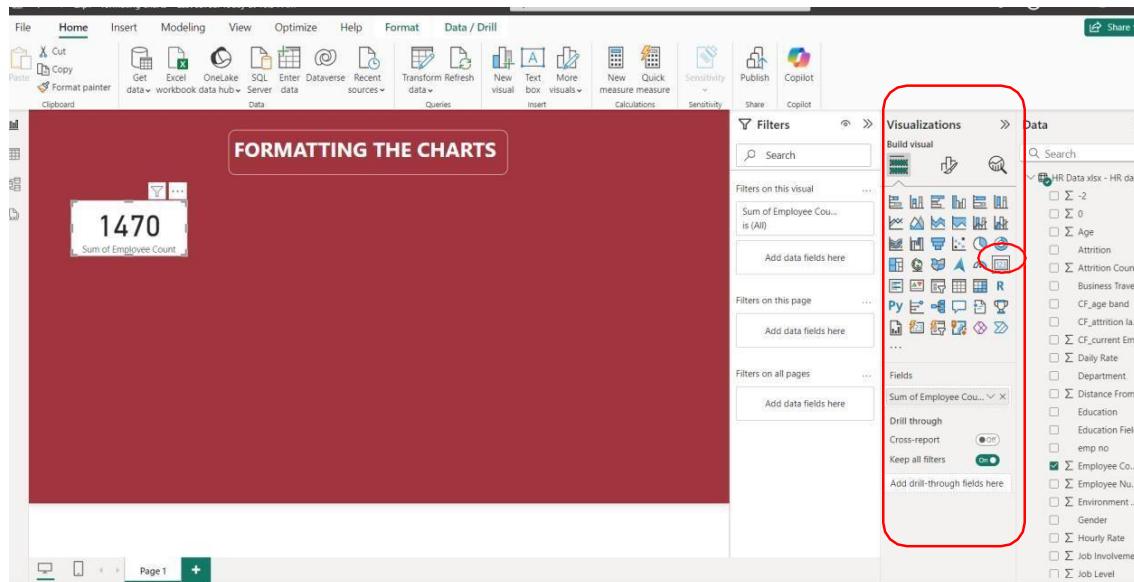
The main area of the editor shows a preview of the data, which includes columns such as Attrition, Business Travel, CF_age_band, CF_attrition_label, Department, Education Field, and emp no. The "APPLIED STEPS" pane on the right lists the transformation step "Changed Type1".

- Click on **Close & Apply**.

The screenshot shows the Microsoft Power BI Desktop interface. The ribbon menu at the top includes File, Home, Insert, Modeling, View, Optimize, and Help. The "Data" tab is selected. On the right side of the screen, there is a "Data" pane containing a list of columns from the "HR Data.xlsx - HR data" table. A red box highlights this pane. The central workspace is labeled "Build visuals with your data" and shows a small visual representation of a table.

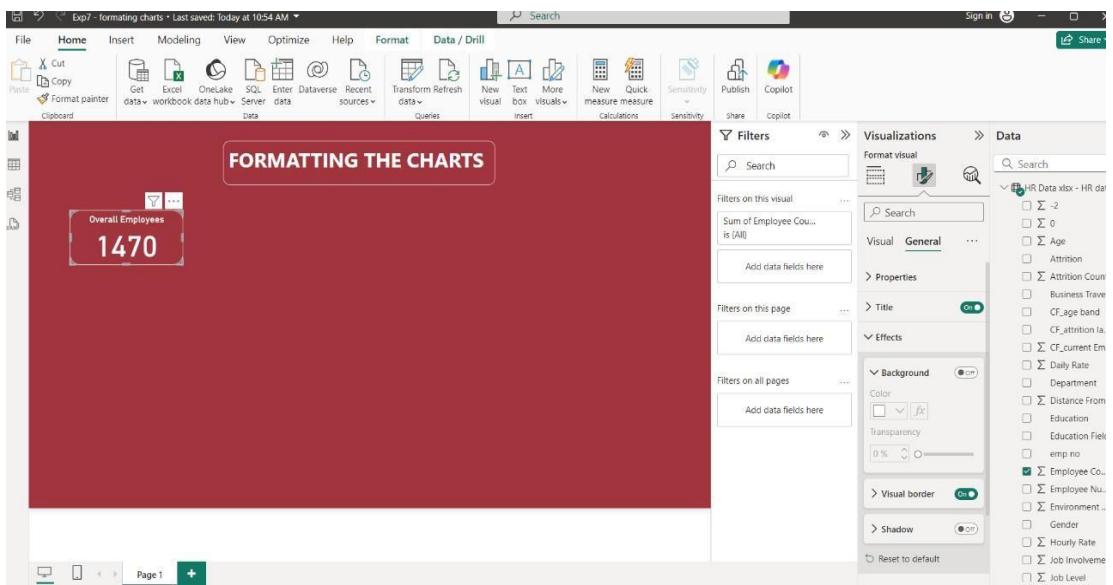
- You will be back on canvas area with table loaded in **Data Pane** (in right side).

- We will start with **KPI Chart**
- A Key Performance Indicator (KPI) is a visual cue that communicates the amount of progress made toward a measurable goal



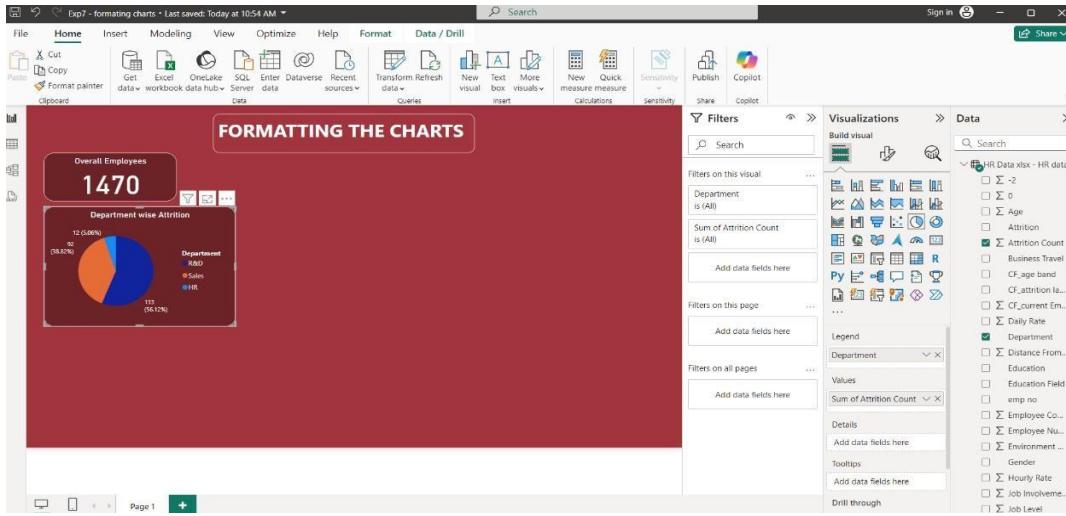
Now format this particular visual with title, size, colour.

1. Click on “Format your visual” in **Visualization Pane**
2. Go to General tab
 - a. click on Title type “Overall Employees” in Text box, Horizontal alignment and colour of your choice
 - b. expand effects **OFF** the background of KPI chart
 - c. Effects **ON** visual border change the color and 20 rounded corners
3. Now, go to Visual tab **OFF** the category label
4. In visual tab, callout value change the font color



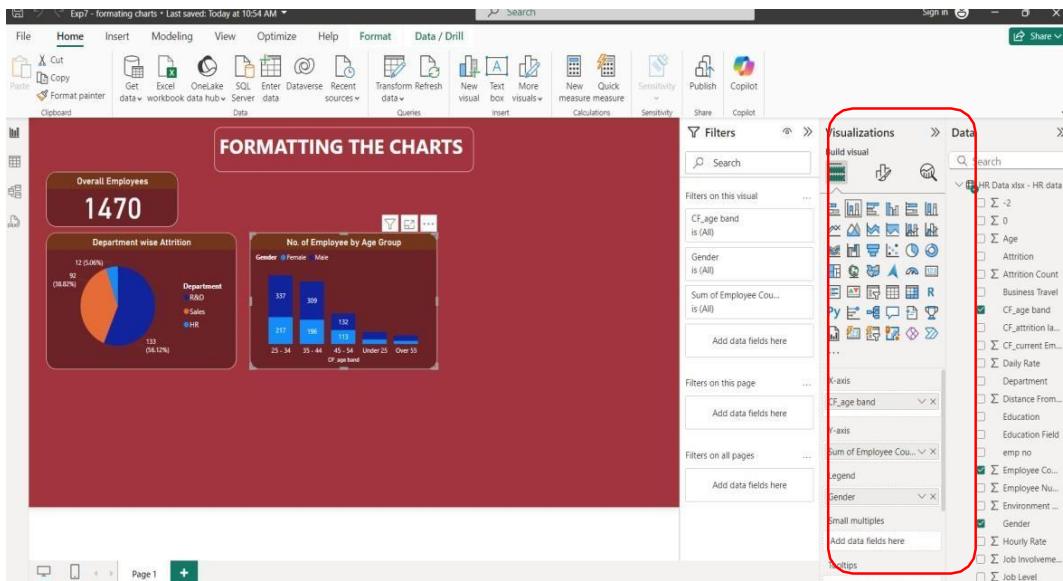
Kindly Note: If you want same format for all visuals, complete the formatting with one of the visual, click on format painter and click on the visual for which you want the formatting. Little bit formatting will be required as properties for each visual will be different

Select Pie Chart.



- Now apply same steps for creating **STACKED COLUMN CHART**.

A column chart, commonly referred to as a vertical bar graph, is a visual tool utilized to display and compare numerical data across different categories. Each column within the chart corresponds to a specific category, with the height of the column proportionally representing the associated value.



Optional: As you can see, age is not sorted correctly, so we have to create additional column.

Once **sort age** column is created change the datatype of column if its not in whole number.

Click on “Close & Apply”

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name: sort age

Column Name	Operator	Value	Output	
If	equals	ABC	Under 25	Then 1
Else If	equals	ABC	25 - 34	Then 2
Else If	equals	ABC	35 - 44	Then 3
Else If	equals	ABC	45 - 54	Then 4
Else		ABC		5

OK Cancel

Now, on canvas, in data pane

- select CP_age_band
- click on sort by column
- select newly created column sort age and now

click on visual and follow the steps, Finally, output will be as follows:

FORMATTING THE CHARTS

Overall Employees: 1470

Department wise Attrition:

- Department: R&D (56.12%)
- Department: Sales (34.88%)
- Department: HR (8.00%)

No. of Employee by Age Group:

Age Group	Count
Under 25	132
Over 55	132
45 - 54	39
35 - 44	196
25 - 34	277

Sort by column: CF_age band, Sort descending

Filters on this page:

- CF_age band
- Sum of Employee Cou...
- Gender
- Small multiples
- Tooltips

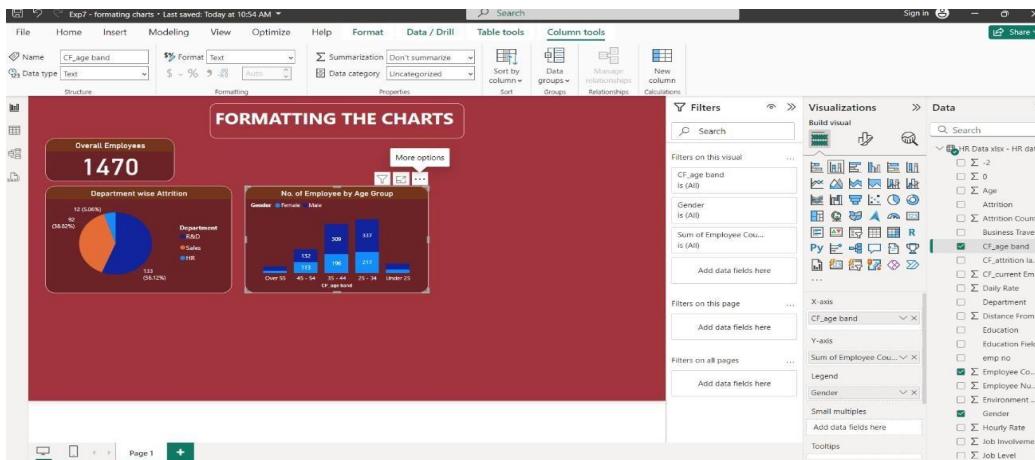
Filters on all pages:

- Sum of Employee Cou...
- Gender

Legend:

- Gender

Finally, output will be as follows



:

- Now apply same steps for creating **MATRIX**.

The matrix visual is a type of table visual that supports a stepped layout. A table supports two dimensions, but a matrix makes it easier to display data meaningfully across multiple dimensions. Often, report designers include matrixes in reports and dashboards to allow users to select one or more element (rows, columns, cells) in the matrix to cross-highlight other visuals on a report page.

Format the **row header & column header** □ **text color & background color**

Also, format the **Row grand total & column grand total**

The screenshot shows the Power BI desktop interface with three visualizations on the canvas:

- Overall Employees:** A large number 1470.
- Department wise Attrition:** A pie chart showing the distribution of employees by department: R&D (56.42%), Sales (38.09%), and HR (5.49%).
- No. of Employee by Age Group:** A stacked bar chart showing the count of employees across five age groups: Over 55 (132), 45-54 (309), 35-44 (277), 25-34 (337), and Under 25 (25).

The ribbon at the top includes Home, Insert, Modeling, View, Optimize, Help, and various data and visualization tools. The Data pane on the right lists data fields from an HR Data.xlsx source, including columns like Department, Age, Attrition, and Job Satisfaction Rating.

Now apply same steps for creating **Stacked Bar Chart**.

The screenshot shows the Power BI desktop interface with the same three visualizations. A red box highlights the 'Education' filter in the 'Visualizations' pane, which is set to 'is (All)'. This indicates that the stacked bar chart's Y-axis is currently set to show data for all education levels.

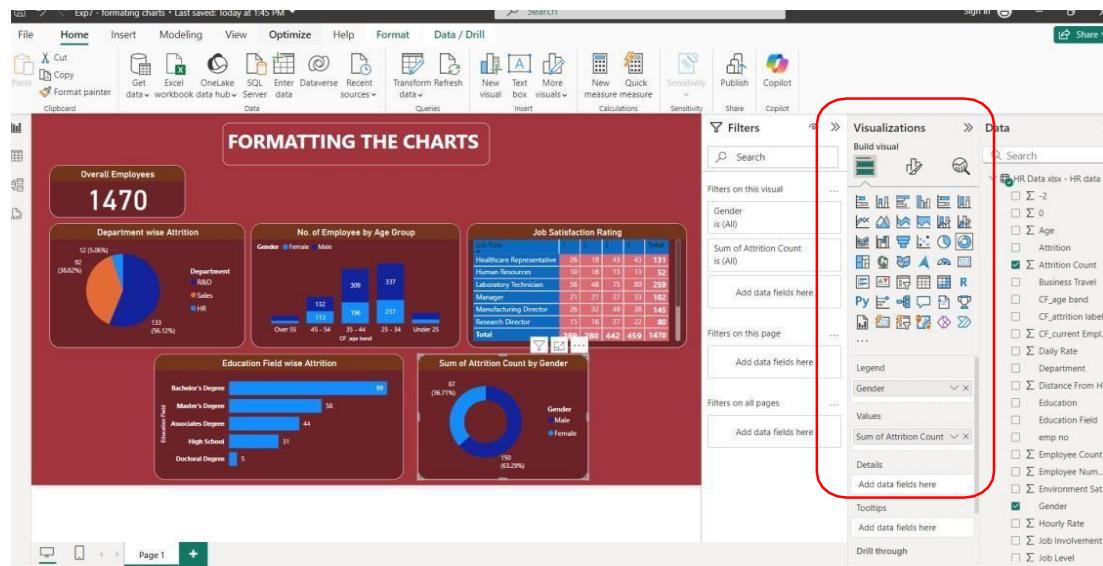
The ribbon at the top includes Home, Insert, Modeling, View, Optimize, Help, Format, and Data / Drill. The Data pane on the right lists data fields from an HR Data.xlsx source, including columns like Department, Age, Attrition, and Job Satisfaction Rating.

Now apply same steps for creating **Donut**.

A doughnut chart is similar to a pie chart in that it shows the relationship of parts to a whole.

The only difference is that the center is blank and allows space for a label or icon.

Doughnut charts work best when you use them to compare a particular section to the whole, rather than comparing individual sections with each other.



Slicers: A slicer is a standalone chart that can be used to filter the other visuals on the page. Slicers come in many different formats (category, range, date, etc.) and can be formatted to allow selection of only one, many, or all of the available values.

Slicers are a great choice to:

- Display commonly used or important filters on the report canvas for easier access.
- Make it easier to see the current filtered state without having to open a drop-down list.
- Filter by columns that are unneeded and hidden in the data tables.
- Create more focused reports by putting slicers next to important visuals.

The screenshot shows a Microsoft Power BI dashboard titled "FORMATTING THE CHARTS". The dashboard contains the following visualizations:

- Overall Employees:** A large number 282.
- Department wise Attrition:** A pie chart showing attrition by department: R&D (31.62%), Sales (23.22%), and HR (18.15%).
- No. of Employee by Age Group:** A bar chart showing the count of employees by age group: Over 55 (34), 45-54 (37), 35-44 (64), 25-34 (55), and Under 25 (47).
- Job Satisfaction Rating:** A table showing job satisfaction ratings across various roles.
- Education Field wise Attrition:** A horizontal bar chart showing attrition by education field: Associates Degree (44).
- Sum of Attrition Count by Gender:** A pie chart showing attrition by gender: Male (25) and Female (19).

The Power BI ribbon is visible at the top, and the Data pane on the right side is highlighted with a red box. The Data pane lists various fields and filters, including:

- Filters: Education (selected), Drill through, Cross-report, Keep all filters, Add drill-through fields here.
- Visualizations: Build visual, ...
- Data: Search, HR Data.xlsx - HR data, ...

Experiment No. 8: Dashboards - Filters in Power BI, Formatting dashboards

Filters remove all but the data you want to focus on.

Filter Pane: You can apply filters in the Filters pane, or make selections in slicers directly on the report page itself. The Filters pane shows the fields in individual visuals and any other filters the report designer adds.

There are four standard types of filters that you create in the Filters pane.

- **Visual filter** applies to a single visual on a report page. You see visual-level filters when you select a visual on the report canvas. Even if you can't edit a report, you can select a visual and filter it.
- **Page filter** applies to all the visuals on the report page.
- **Report filter** applies to all pages in the report.
- **Drill through filter** With drill through in the Power BI service and Power BI Desktop, you create a *destination* report page that focuses on a specific entity, such as a supplier. From the other report pages, users can right-click a data point for that entity and drill through to the focused page.

We will be using HR dataset (same used for Exp 7).

Extending same dashboard with using filters & let's format the final dashboard. Let's apply filter for department (Particular visual).

1. Drag Department from Data Pane to Filters □ Filters on this page textbox. □ Basic Filtering
2. Now, you can see, HR dept. is selected and now in below picture, you can see that only HR data is visible, whereas, R&D and Sales data will not be shown.

Now, HR and R&D departments are selected.

Overall Employees: 1024

Department wise Attrition:

Department	Count	Percentage
R&D	133	(91.72%)
HR	13	(8.28%)

No. of Employee by Age Group:

Age Group	Male	Female	Total
Over 55	95	71	166
45 - 54	142	142	284
35 - 44	205	205	410
25 - 34	116	116	232
Under 25	20	20	40

Job Satisfaction Rating:

Role	Male	Female	Total
Healthcare Representative	26	19	45
Human Resources	10	16	26
Laboratory Technician	56	48	104
Manager	16	10	26
Manufacturing Director	28	32	60
Research Director	15	16	31
Total	203	194	397

Education Field wise Attrition:

Field	Male	Female	Total
Bachelor's Degree	60	40	100
Master's Degree	35	25	60
Associate's Degree	26	15	41
High School	20	10	30
Doctoral Degree	4	2	6

Sum of Attrition Count by Gender:

Gender	Count	Percentage
Male	96	(66.21%)
Female	49	(33.79%)

3. Now, lets try for advanced filter

Let's find out results for **salary greater than equal to 10000**.

Drag and drop **Monthly Income** in filter Filter type: advanced filtering

Show items when the value: is greater than or equal to Value: 10000

Overall Employees: 281

Department wise Attrition:

Department	Count	Percentage
Sales	13	(52%)
R&D	11	(44%)
HR	1	(4%)

No. of Employee by Age Group:

Age Group	Male	Female	Total
Over 55	56	22	78
45 - 54	58	56	114
35 - 44	40	40	80
25 - 34	20	20	40

Job Satisfaction Rating:

Role	Male	Female	Total
Healthcare Representative	2	6	8
Human Resources	7	11	18
Laboratory Technician	21	27	48
Manager	8	12	20
Manufacturing Director	4	8	12
Research Director	15	16	31
Sales Executive	12	4	16
Total	55	56	111

Education Field wise Attrition:

Field	Male	Female	Total
Bachelor's Degree	11	9	20
Master's Degree	9	7	16
High School	3	2	5
Associate's Degree	1	1	2
Doctoral Degree	1	1	2

Sum of Attrition Count by Gender:

Gender	Count	Percentage
Male	18	(72%)
Female	7	(28%)

1. Now, try for Top N filtering. Try to display top 4 Job roles having highest job satisfaction.

- Select “Job satisfaction rating” visual
- In Filters Pane, Filters on this visual

Job Role Filter type: Top N

Show item: Top : 4

By value: Sum of Job Satisfaction