

Name of Student: Pushkar Sane		
Roll Number: 45	Lab Assignment Number: 10	
Title of Lab Assignment: To learn Data Visualization and dashboard creation in Power BI		
DOP: 20-10-2024	DOS: 21-10-2024	
CO Mapped: CO6	PO Mapped: PO1, PO2, PO3, PO4, PO5, PO6, PO7, PO8, PSO1, PSO2	Signature:

AIM: TO LEARN DATA VISUALIZATION AND DASHBOARD CREATION IN POWER BI

THEORY:

1. Understanding Data Visualization

Data visualization is the graphical representation of data to make it easier for users to understand trends, patterns, and insights. Effective visualization simplifies complex data sets into charts, graphs, and maps, allowing for better decision-making. Power BI is a powerful tool for creating such visualizations with a user-friendly interface and numerous features that enable users to analyze data interactively.

2. Power BI Overview

Power BI is a business analytics tool developed by Microsoft for creating visual reports, dashboards, and data models. It allows users to connect to various data sources, transform raw data, and visualize it in a meaningful way.

Key components of Power BI include:

- **Power BI Desktop:** The primary environment for designing reports and dashboards.
- **Power BI Service:** The cloud-based platform where reports and dashboards can be shared and accessed.
- **Power BI Mobile:** The mobile app for viewing reports on the go.

3. Data Preparation and Transformation

Before visualizing data, it must be clean and well-organized. Power BI offers tools for:

- **Data Importation:** Power BI can connect to multiple data sources like Excel, SQL databases, Azure, APIs, and more.
- **Data Transformation (Power Query):** Users can clean, filter, reshape, and aggregate data using Power Query, which is built into Power BI. Operations include removing duplicates, merging tables, changing data types, and more.

4. Basic Visualization Techniques

Power BI provides several visual elements to choose from, such as:

- **Bar and Column Charts:** Used to compare values across categories.
- **Line Charts:** Display trends over time.
- **Pie and Donut Charts:** Show parts of a whole.
- **Scatter Plots:** Represent correlations between variables.
- **Maps (Choropleth and Filled Maps):** Geographical data representation.
- **Tables and Matrices:** Present data in a structured format.

Each chart has its own use cases based on the nature of the data and the insights you want to derive.

5. Creating Dashboards

Dashboards are a collection of visualizations combined into a single view to provide insights at a glance. In Power BI, you can pin visualizations from different reports onto a dashboard. Key steps include:

- **Report Design:** Start by creating multiple visualizations in Power BI Desktop.
- **Interactivity:** Power BI allows users to filter and interact with reports, such as applying slicers or drill-downs. These features make the reports dynamic and allow for better data exploration.
- **Pinning to Dashboard:** Once the report is built, specific visuals can be pinned to dashboards in the Power BI service. Dashboards give a high-level view and allow users to monitor key metrics in real-time.

6. Key Features for Dashboard Creation

- **Slicers and Filters:** Allow users to filter data directly on the dashboard, improving interactivity.
- **Drill-through and Drill-down:** These enable users to explore data at different levels of granularity.
- **Bookmarks and Selections:** Help in navigating between different views within a report.
- **Custom Visuals:** Power BI allows the import of third-party visuals to meet specific needs.

7. DAX (Data Analysis Expressions)

DAX is the formula language in Power BI used to perform calculations and create custom measures. Understanding DAX is crucial for advanced data modeling and creating meaningful KPIs. It enables operations like:

- **Creating calculated columns and measures.**
- **Aggregating data (SUM, AVERAGE, COUNT).**
- **Time intelligence (e.g., year-over-year growth, quarter-to-date).**

8. Power BI Best Practices for Visualization

To create effective dashboards and visualizations:

- **Focus on Simplicity:** Avoid clutter and unnecessary visuals; focus on key metrics.
- **Use Color Wisely:** Use consistent and meaningful colors for categories and trends. Avoid overuse of color.
- **Ensure Readability:** Use appropriate font sizes, legends, and labels for readability.
- **Context and Annotations:** Provide context for your data with titles, annotations, or tooltips.
- **Interactivity and Exploration:** Allow users to explore data interactively via slicers, filters, and tooltips.

9. Sharing and Collaboration

Once the dashboard is built, you can share it with others via the Power BI service. Power BI also supports real-time collaboration, where users can comment on dashboards, create alerts, and schedule data refreshes.

10. Power BI Use Cases

Some common use cases for dashboards and visualizations in Power BI include:

- **Sales and Financial Analysis:** Visualizing sales trends, profit margins, and forecasting.
- **Customer Segmentation:** Understanding customer behavior and segmenting markets.
- **Supply Chain Management:** Tracking inventory levels, logistics, and supplier performance.
- **HR Analytics:** Visualizing employee performance, attrition rates, and recruitment data.

DATASET USED:**“STUDENT PERFORMANCE FACTORS”**

Link: <https://www.kaggle.com/datasets/lainguyn123/student-performance-factors>

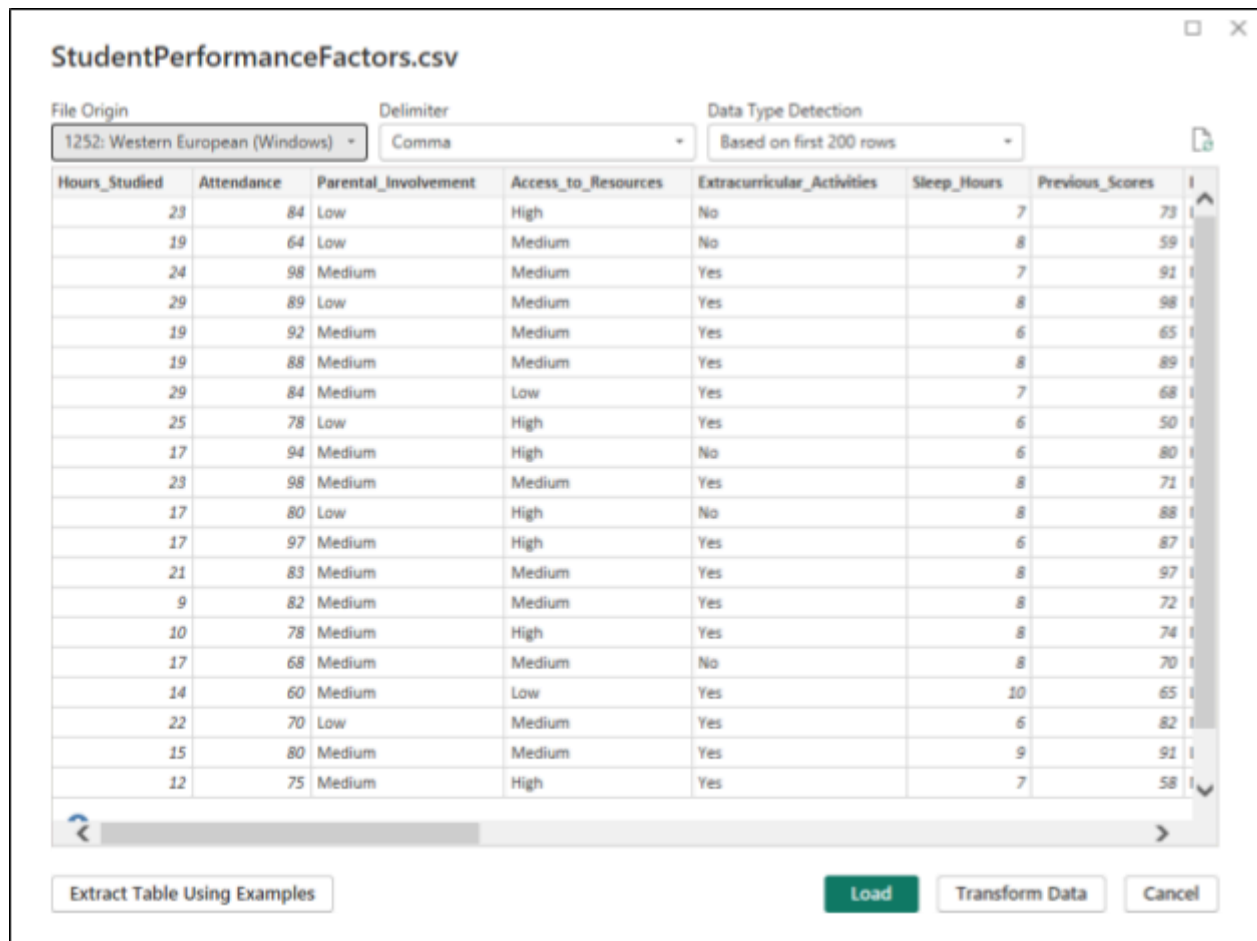
Description:

This dataset provides a comprehensive overview of various factors affecting student performance in exams. It includes information on study habits, attendance, parental involvement, and other aspects influencing academic success.

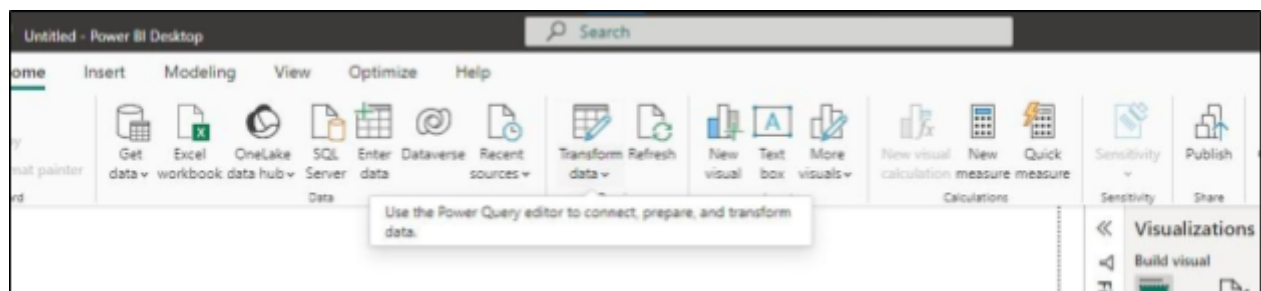
- **Hours_Studied:** Number of hours spent studying per week.
- **Attendance:** Percentage of classes attended.
- **Parental_Involvement:** Level of parental involvement in the student's education (Low, Medium, High).
- **Access_to_Resources:** Availability of educational resources (Low, Medium, High).
- **Extracurricular_Activities:** Participation in extracurricular activities (Yes, No).
- **Sleep_Hours:** Average number of hours of sleep per night.
- **Previous_Scores:** Scores from previous exams.
- **Motivation_Level:** Student's level of motivation (Low, Medium, High).
- **Internet_Access:** Availability of internet access (Yes, No).
- **Tutoring_Sessions:** Number of tutoring sessions attended per month.
- **Family_Income:** Family income level (Low, Medium, High).
- **Teacher_Quality:** Quality of the teachers (Low, Medium, High).
- **School_Type:** Type of school attended (Public, Private).
- **Peer_Influence:** Influence of peers on academic performance (Positive, Neutral, Negative).
- **Physical_Activity:** Average number of hours of physical activity per week.
- **Learning_Disabilities** : Presence of learning disabilities (Yes, No).
- **Parental_Education_Level:** Highest education level of parents (High School, College, Postgraduate).
- **Distance_from_Home:** Distance from home to school (Near, Moderate, Far).
- **Gender:** Gender of the student (Male, Female).
- **Exam_Score:** Final exam score.

STEPS:**DATA PREPROCESSING:****1. Data Cleaning**

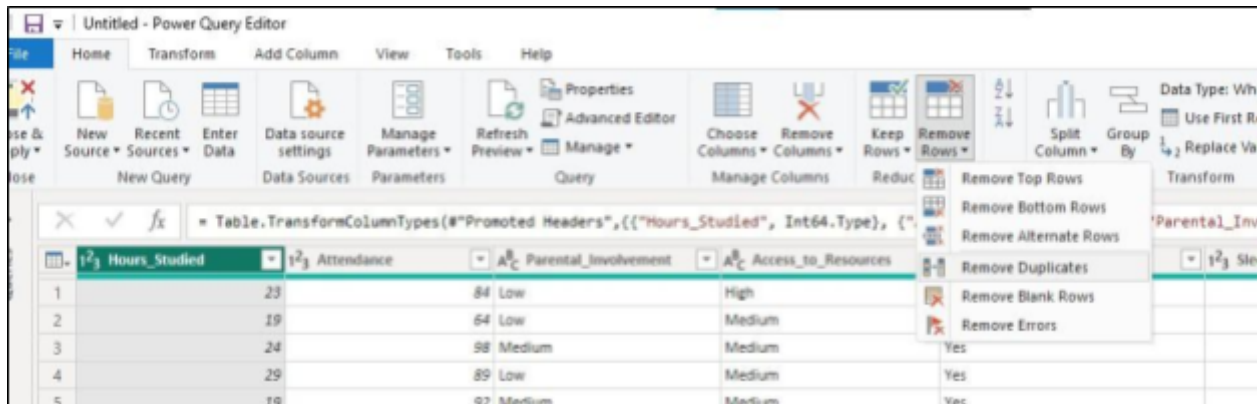
- Load your Dataset



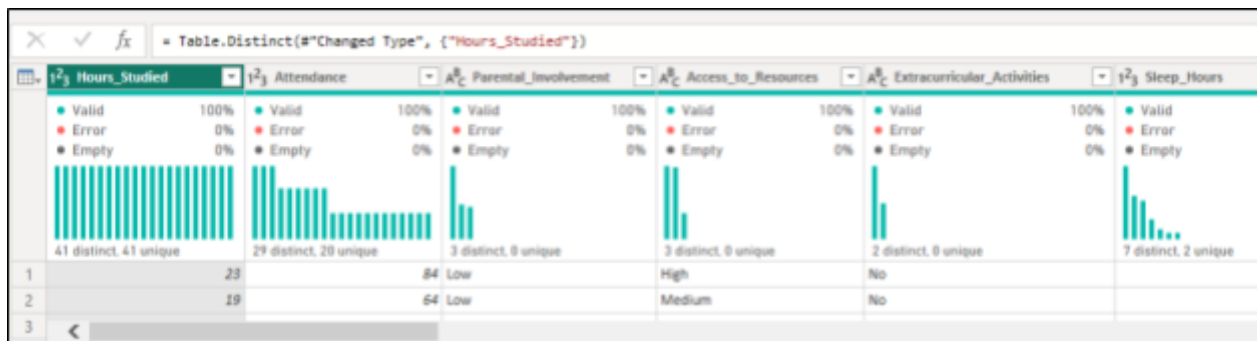
- Click on "Transform Data"



- **Remove Duplicates:** Ensure there are no duplicate rows in your dataset. In Power Query, go to *Home* → *Remove Duplicates*.



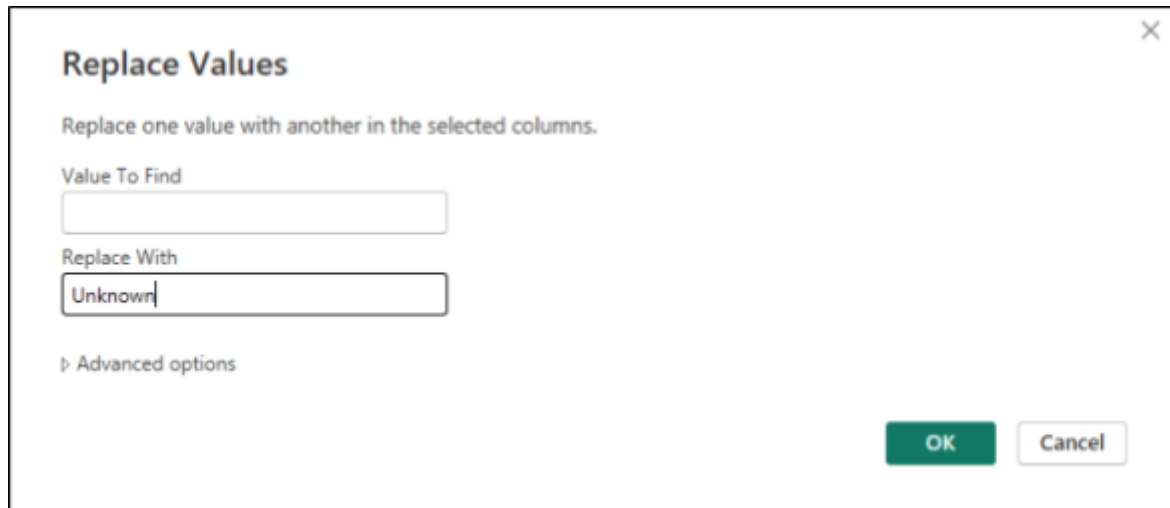
- **Handle Missing Values:** Check for missing values and decide on an approach:
 - Replace missing numeric values with a mean/median.
 - For categorical fields (e.g., Parental_Involvement), replace missing values with the mode or a default value (e.g., "Unknown").
 - In Power Query, select *Transform* → *Replace Values* to handle missing values.



Parental_Education_Level	
Valid	98%
Error	0%
Empty	2%
High School	
College	
Postgraduate	
High School	
College	
Postgraduate	

Replace blank values with “Unknown”

Parental_Education_Level		Distance from Home		Geography	
Valid		Copy		Valid	
Error		Remove		Error	
Empty		Remove Other Columns		Empty	
High School		Duplicate Column		High School	
College		Add Column From Examples...		College	
Postgraduate		Remove Duplicates		Postgraduate	
High School		Remove Errors		High School	
College		Change Type		College	
Postgraduate		Transform		Postgraduate	
High School		Replace Values...		High School	
College		Replace Errors...		College	
High School		Split Column		High School	
College		Group By...		College	
High School		Fill		High School	
Postgraduate		Unpivot Columns		Postgraduate	
Postgraduate		Unpivot Other Columns		Postgraduate	



Replace Values

Replace one value with another in the selected columns.

Value To Find

Replace With

Unknown

Advanced options

OK Cancel

2. Data Transformation

- **Convert Data Types:** Ensure each field has the correct data type:
 - Numeric fields (e.g., Hours_Studied, Exam_Score) should be set to *Whole Number* or *Decimal Number*.
 - Categorical fields (e.g., Parental_Involvement, Motivation_Level) should be set as *Text* or *Categorical*.
- **Create Calculated Columns:**
 - If necessary, derive new columns. For example, you might create a Performance_Level column based on the Exam_Score (e.g., "High", "Medium", "Low").

Add Conditional Column

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

New column name
Performance_Level

Column Name	Operator	Value	Output
if Exam_Score	is less than	65	Low
Else if Exam_Score	is greater than	70	High

Add Clause

Else
Medium

OK Cancel

Query Settings

PROPERTIES

Name
StudentPerformanceFactors

APPLIED STEPS

Source
Promoted Headers
Changed Type
Added Conditional Column

Gender	Exam_Score	Performance_Level
Male	67	Medium
Female	61	Low
Male	74	High
Male	71	High
Female	70	Medium
Male	71	High
Male	67	Medium
Male	66	Medium
Male	69	Medium
Male	72	High
Male	68	Medium
Male	71	High
Male	70	Medium
Male	66	Medium
Male	65	Medium
Female	64	Low
Male	60	Low

Click on Close and Apply

Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Close & Apply New Source Recent Sources Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Choose Columns Remove Columns Manage Columns

Queries [1]
StudentPerformanceFact...

fx = Table.AddColumn(#"Changed Type", "Performance_Level"

Hours_Studied	Attendance	Parental_Involvement
23	84	Low
19	64	Low
24	98	Medium

1. Introduction to Visuals

- Open Power BI Desktop and load your dataset.

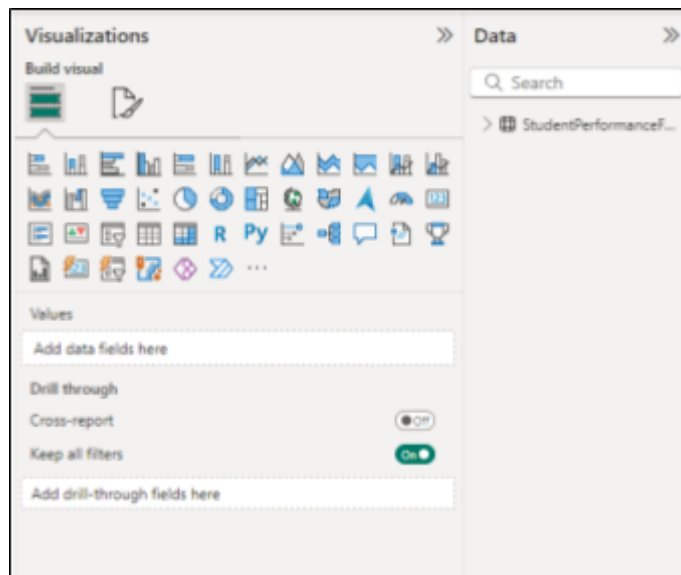
StudentPerformanceFactors.csv

File Origin: 1252: Western European (Windows) | Delimiter: Comma | Data Type Detection: Based on first 200 rows

Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores
23	84	Low	High	No	7	73
19	64	Low	Medium	No	8	59
24	98	Medium	Medium	Yes	7	91
29	89	Low	Medium	Yes	8	98
19	92	Medium	Medium	Yes	6	65
19	88	Medium	Medium	Yes	8	89
29	84	Medium	Low	Yes	7	68
25	78	Low	High	Yes	6	50
17	94	Medium	High	No	6	80
23	98	Medium	Medium	Yes	8	71
17	80	Low	High	No	8	88
17	97	Medium	High	Yes	6	87
21	83	Medium	Medium	Yes	8	97
9	82	Medium	Medium	Yes	8	72
10	78	Medium	High	Yes	8	74
17	68	Medium	Medium	No	8	70
14	60	Medium	Low	Yes	10	65
22	70	Low	Medium	Yes	6	82
15	80	Medium	Medium	Yes	9	91
12	75	Medium	High	Yes	7	58

Extract Table Using Examples | Load | Transform Data | Cancel

- From the "Visualizations" pane, select a visual (e.g., Bar Chart, Pie Chart).
- Drag dataset fields into the "Values" and "Axis" sections.

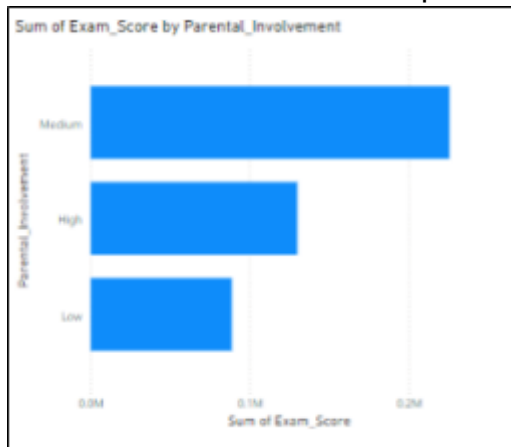


2. Visualization Charts

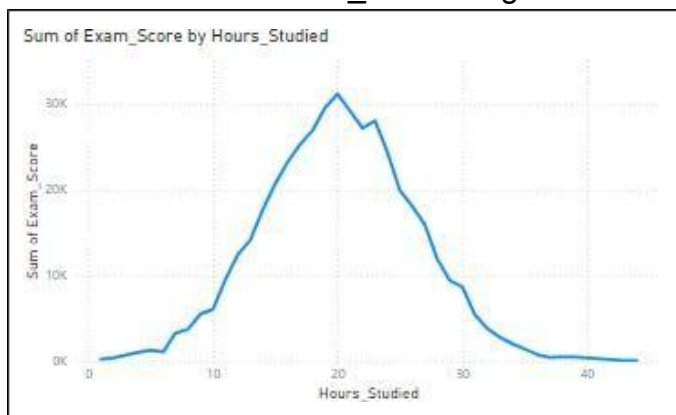
- Select your chart type from the "Visualizations" pane (e.g., bar, line, pie).

- Drag relevant fields into the chart.

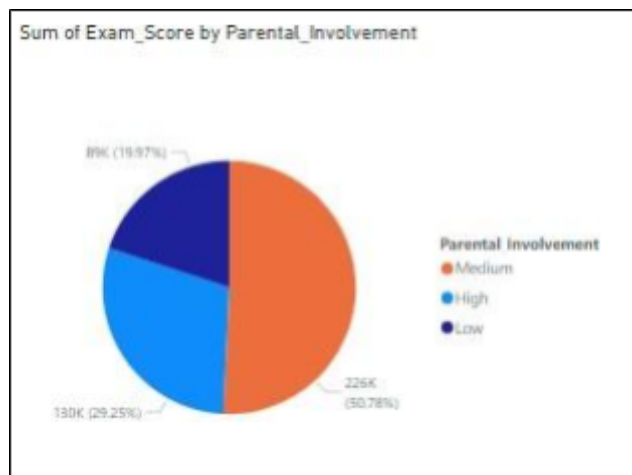
Bar Charts: Show relationships between Parental_Involvement and Exam_Score.



Line Charts: Plot Hours_Studied against Exam_Score over time (if applicable).



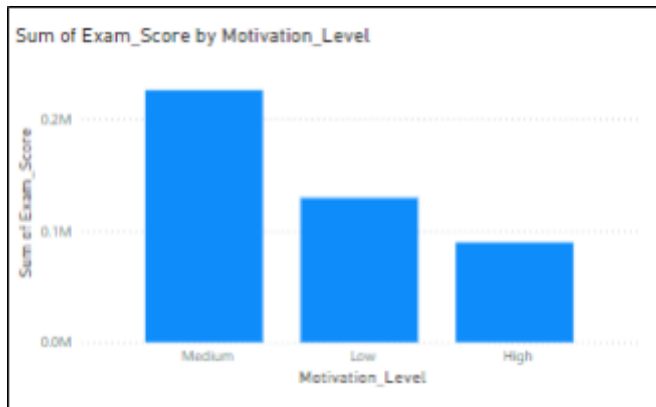
Pie Charts: Represent the proportion of students with High, Medium, and Low Parental_Involvement.



3. Filtering Options

- Use the "Filters" pane on the right.

- Drag fields (e.g., Gender, School_Type) into the filter and apply them.



Filtering by Gender= Female and School_Type= Private

Filters on this visual ...

Gender
is Female

Filter type ⓘ
Basic filtering

Search

☒ Select all

☒ Female 2793

☐ Male 3814

School_Type ^ x 🔒

is Private 🔍 🔍

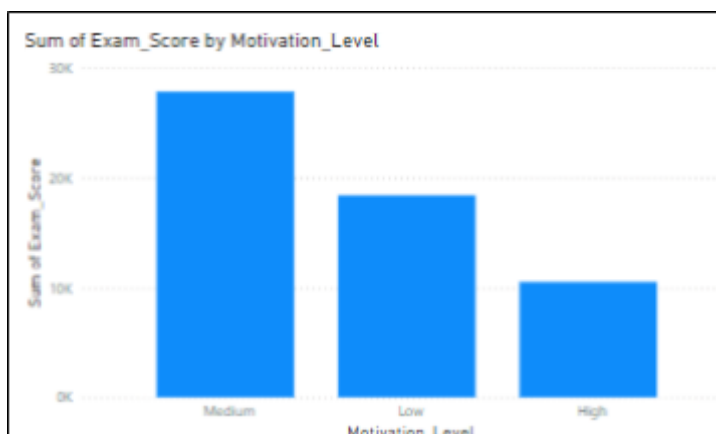
Filter type ⓘ
Basic filtering

Search

☒ Select all

☒ Private 843

☐ Public 1950



4. Exploring Matrix Visuals

- Add a "Matrix" visual from the "Visualizations" pane.

- Drag fields like Exam_Score to "Values", Gender and Parent_Education_Level to "Rows" or "Columns."

Gender	College	High School	Postgraduate	Unknown	Total
Female	57605	90620	37274	2316	187815
Male	76286	124978	51428	3719	256411
Total	133891	215598	88702	6035	444226

Rows

Gender

Columns

Parental_Education_Level

Values

Sum of Exam_Score

Drill through

Cross-report

Keep all filters

5. Filtering Data with Slicers

- Add a "Slicer" visual.
- Drag a field like Gender, Attendance, or Parental_Involvement into the slicer.

Parental_Involvement, Gender

High

Low

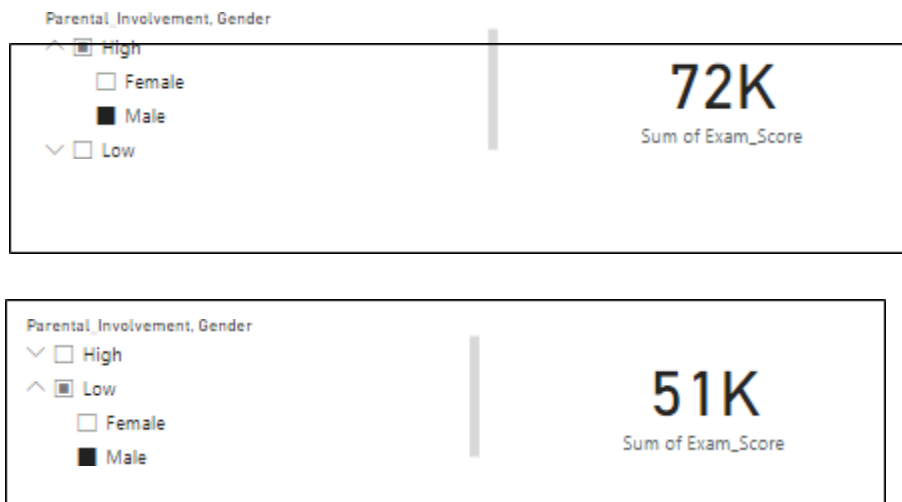
Female

Male

Gender	College	High School	Postgraduate	Unknown	Total
Female	11136	18869	6824	589	37418
Total	11136	18869	6824	589	37418

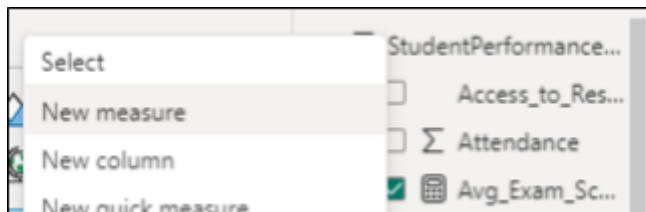
6. Number Cards and Text Cards

- Select a "Card" visual.
- Drag the field (e.g., Exam_Score) into the card for key metrics.



7. KPI Visuals

- Add a "KPI" visual.
- Set up the target as the **average of Previous Scores** or another aggregated measure:



Average of Previous Scores: You can calculate the average of all the previous scores as a reference point.

Target_Score = AVERAGE('StudentPerformanceFactors'[Previous_Scores])

Then use this measure in the **Target** field of your KPI.

Select **Average of Exam_Score** from the dropdown in the "Value" field. Use the following DAX formula to calculate the average Exam_Score:

Avg_Exam_Score = AVERAGE('StudentPerformanceFactors'[Exam_Score])

Steps to Configure the KPI:

1. **Value:** Select **Average of Exam_Score** from the dropdown in the "Value" field.
2. **Trend Axis:** selected **Previous_Scores**, which is great for showing trends based on historical data.

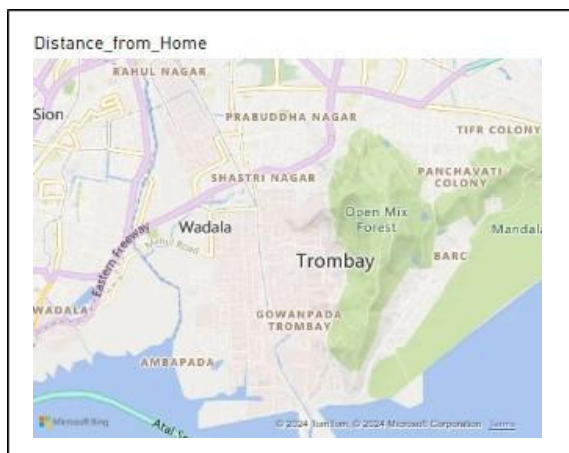
3. **Target:** create a calculated measure for the **Target** and use it here.



Value
Avg_Exam_Score
Trend axis
Previous_Scores
Target
Target_Score

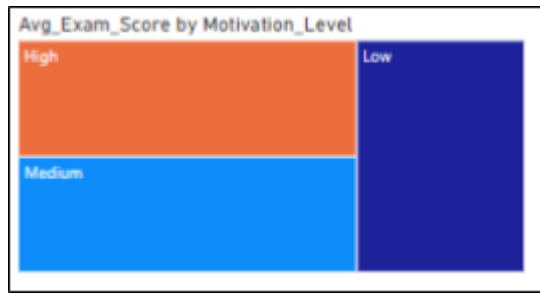
8. Visualizing Data with Maps

- Add a "Map" visual.
- Drag a geographical field (if available) or Distance_from_Home into the "Location" section.



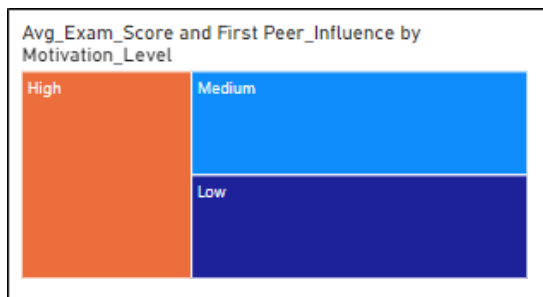
9. TreeMap

- Add a "TreeMap" visual.
- Drag fields like Motivation_Level into "Group" and Avg_Exam_Score into "Values."



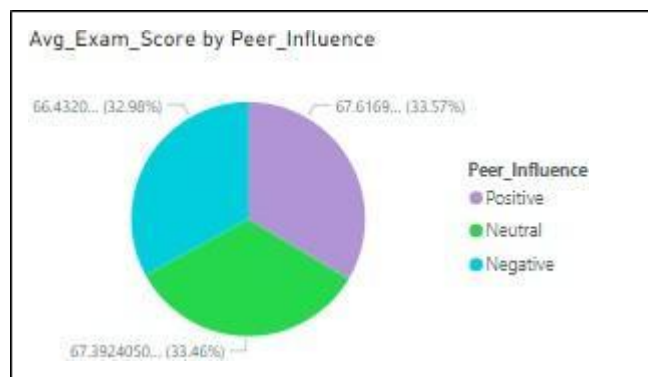
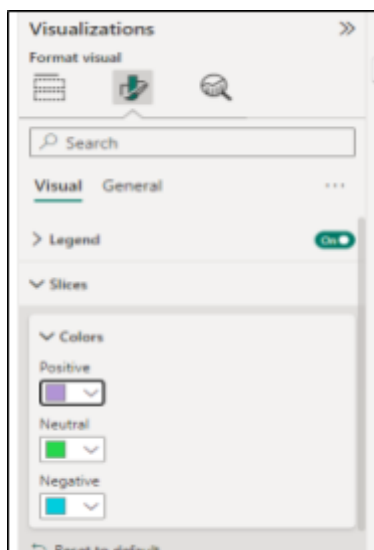
10. Tool Tips in Power BI

- Select a visual.
- In the "Fields" pane, drag additional data fields (e.g., Peer_Influence) to the "Tooltips" area.



11. Modifying Colors in Charts and Visuals

- Select a visual.
- In the "Format" pane, under "Data colors," modify the color for each value category.

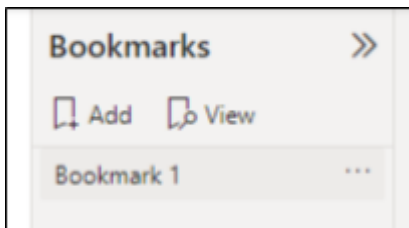
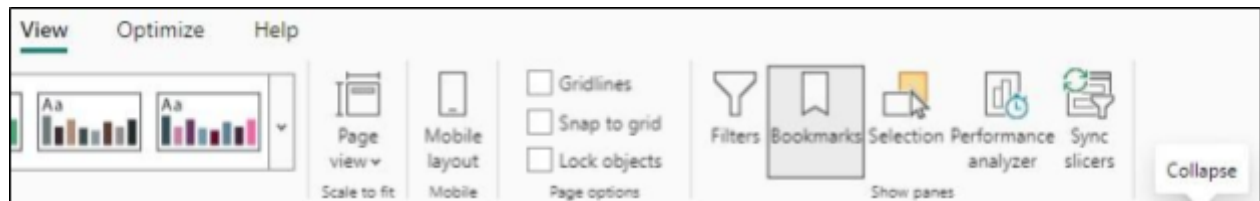


12.

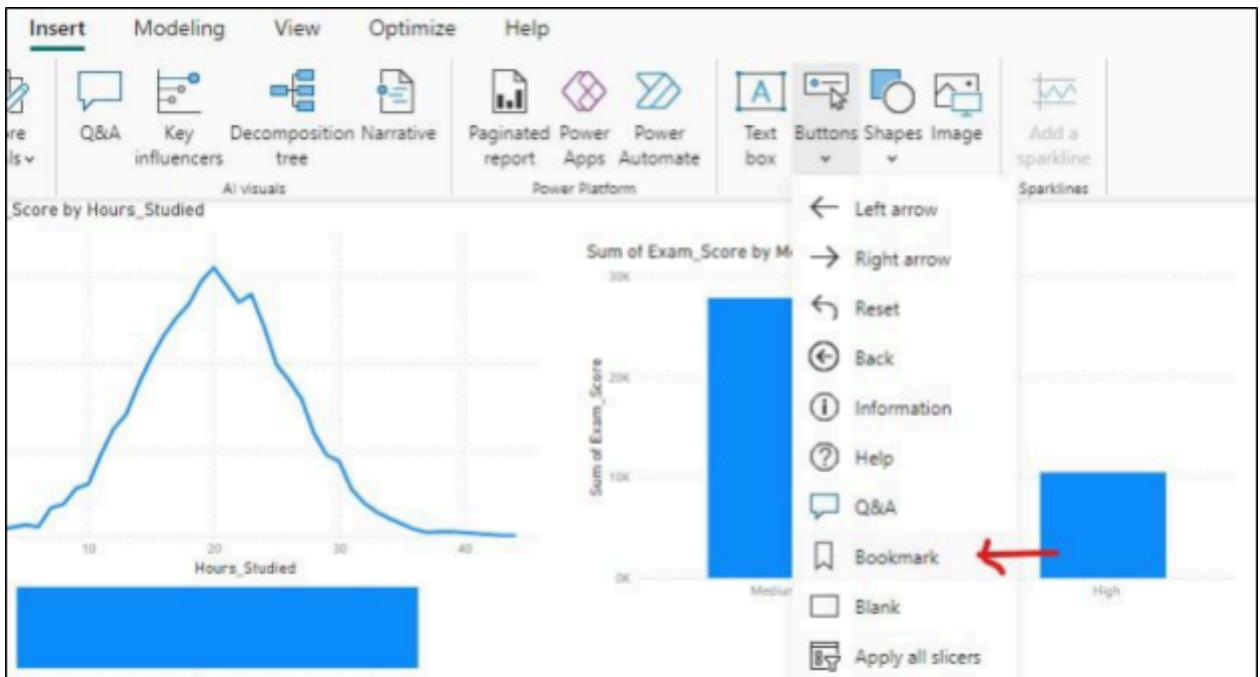
Bookmarks and Buttons

"Bookmarks" → "Add."

- Create a bookmark by going to "View" →



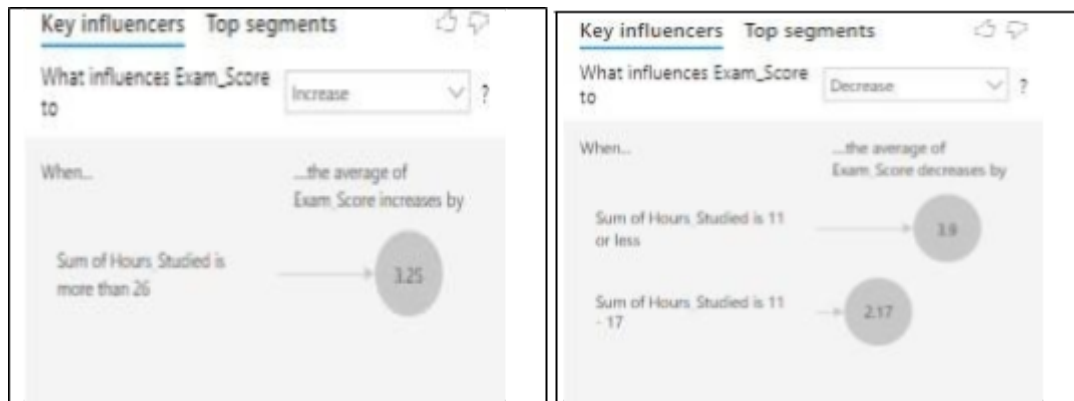
- Add buttons via "Insert" → "Buttons" to navigate between bookmarks.



13. AI Visuals

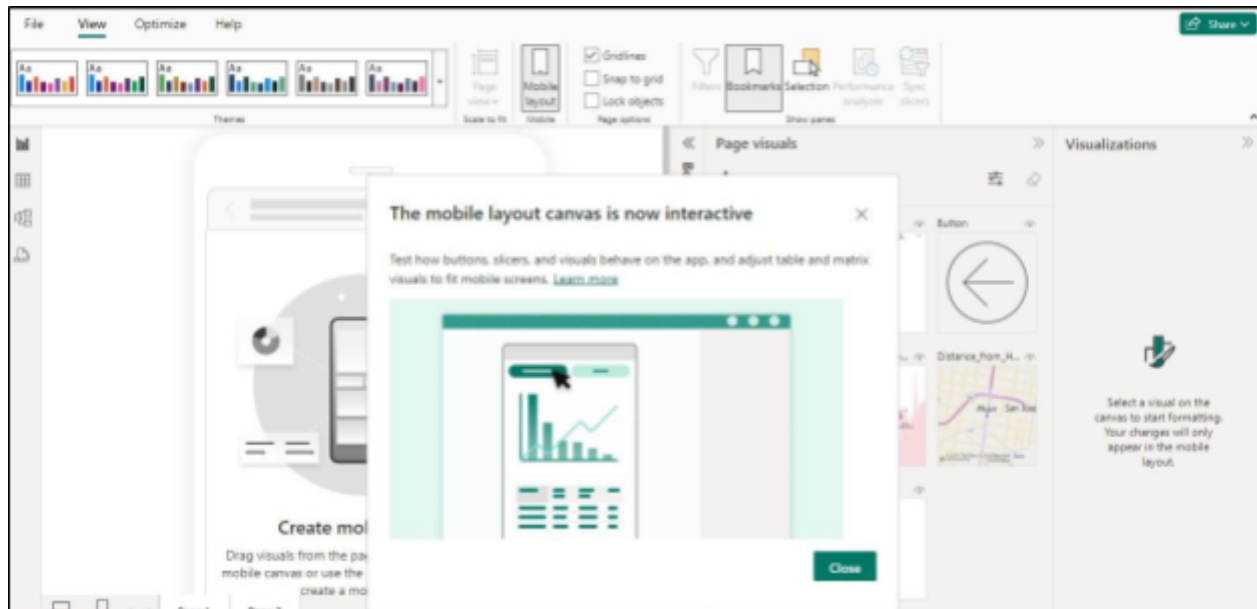
- Add the "Key Influencers" visual.

- Drag Exam_Score into the "Analyze" section and influencing factors (e.g., Hours_Studied, Motivation_Level) into "Explain By."



14. Designing for Phone vs Desktop Report Viewers

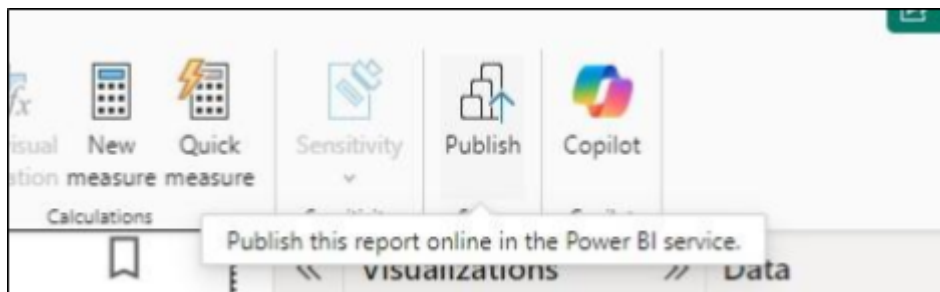
- In Power BI Desktop, go to "View" → "Phone Layout."
- Adjust your visuals for mobile view by dragging and resizing elements.





15. Publishing Reports to Power BI Services

- In Power BI Desktop, click "Publish" in the toolbar.
- Sign into Power BI, select your workspace, and publish the report.



Dashboard:

**CONCLUSION:**

In this practical, we focused on data visualization techniques using Power BI to analyze the student performance dataset. Various visuals, including KPI, were used to track Exam_Score trends and compare them against targets like the average of previous scores. The use of charts, slicers, and filters provided insights into student performance patterns, making it easier to interpret data and identify areas for improvement.