Student Performance Factors

Milestone 1 – Infosys Internship

TOPIC: Student Performance Factors Analysis using Power BI

Introduction:

The dataset contains information on 6,607 students, focusing on various factors influencing their academic performance. It includes 20 columns with data spanning numerical and categorical features. Here's a summary:

Columns Overview:

1. **Hours_Studied**: Number of hours spent studying.
2. **Attendance**: Attendance percentage.
3. **Parental_Involvement**: Level of parental involvement in the student's education (categorical).
4. **Access_to_Resources**: Access level to educational resources (categorical).
5. **Extracurricular_Activities**: Participation in extracurricular activities (Yes/No).
6. **Sleep_Hours**: Average hours of sleep per day.
7. **Previous_Scores**: Academic performance in previous assessments.
8. **Motivation_Level**: Student's motivation level (categorical).
9. **Internet_Access**: Whether the student has internet access (Yes/No).
10. **Tutoring_Sessions**: Number of tutoring sessions attended.
11. **Family_Income**: Family income level (categorical).
12. **Teacher_Quality**: Perceived teacher quality (categorical, includes some missing values).
13. **School_Type**: Type of school (Public/Private).
14. **Peer_Influence**: Type of influence from peers (categorical).
15. **Physical_Activity**: Frequency of physical activity per week.
16. **Learning_Disabilities**: Presence of learning disabilities (Yes/No).
17. **Parental_Education_Level**: Highest education level of parents (categorical, includes some missing values).
18. **Distance_from_Home**: Distance of school from home (categorical, includes some missing values).
19. **Gender**: Gender of the student.
20. **Exam_Score**: Final exam score.

Data preprocessing Steps:

1. **Loading Data:** Import from sources like Excel, SQL Server, or CSV using **Get Data**.

2. **Data Cleaning:**

   - Remove duplicates.
   - Handle missing values (replace or remove).
   - Trim spaces, standardize text, and correct errors.

3. **Data Transformation:**

   - Rename columns/tables and change data types.
   - Split/merge columns and reshape data with pivot/unpivot.

4. **Data Filtering:** Include/exclude rows or columns based on conditions.

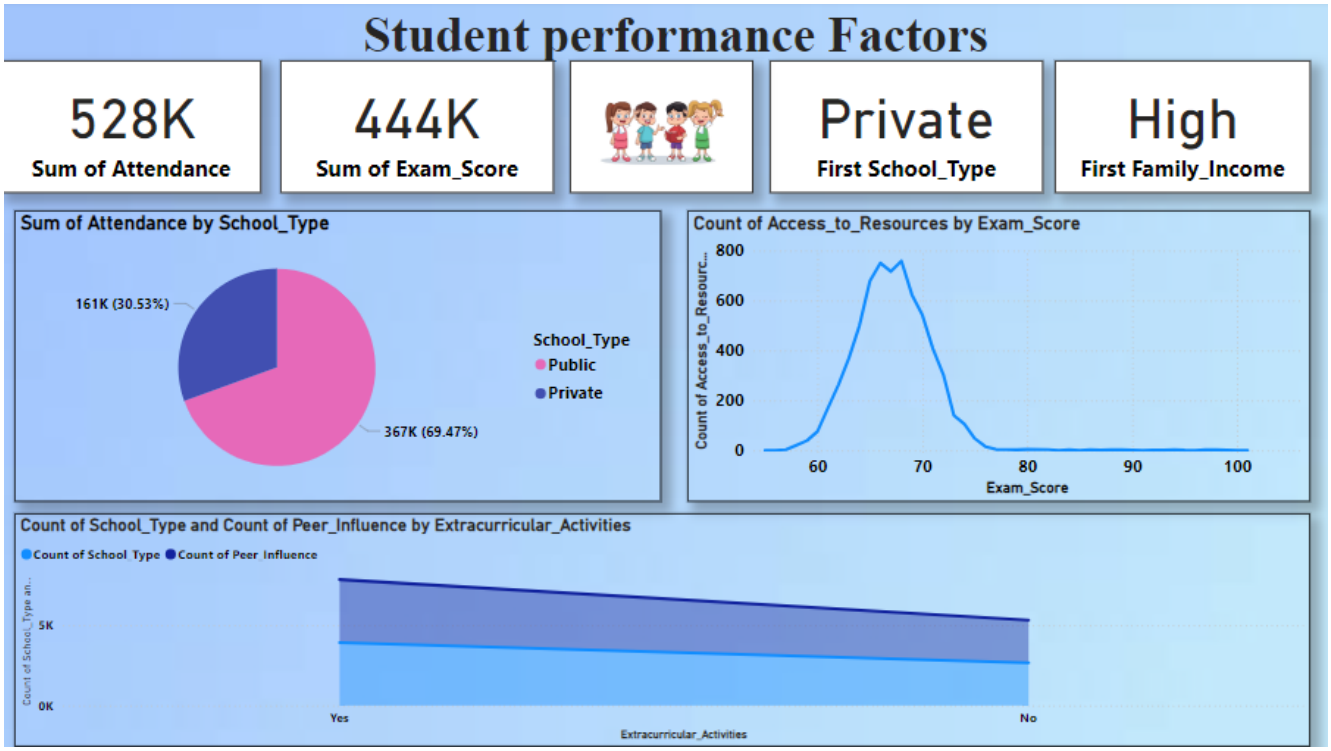
5. **Data Enrichment:**

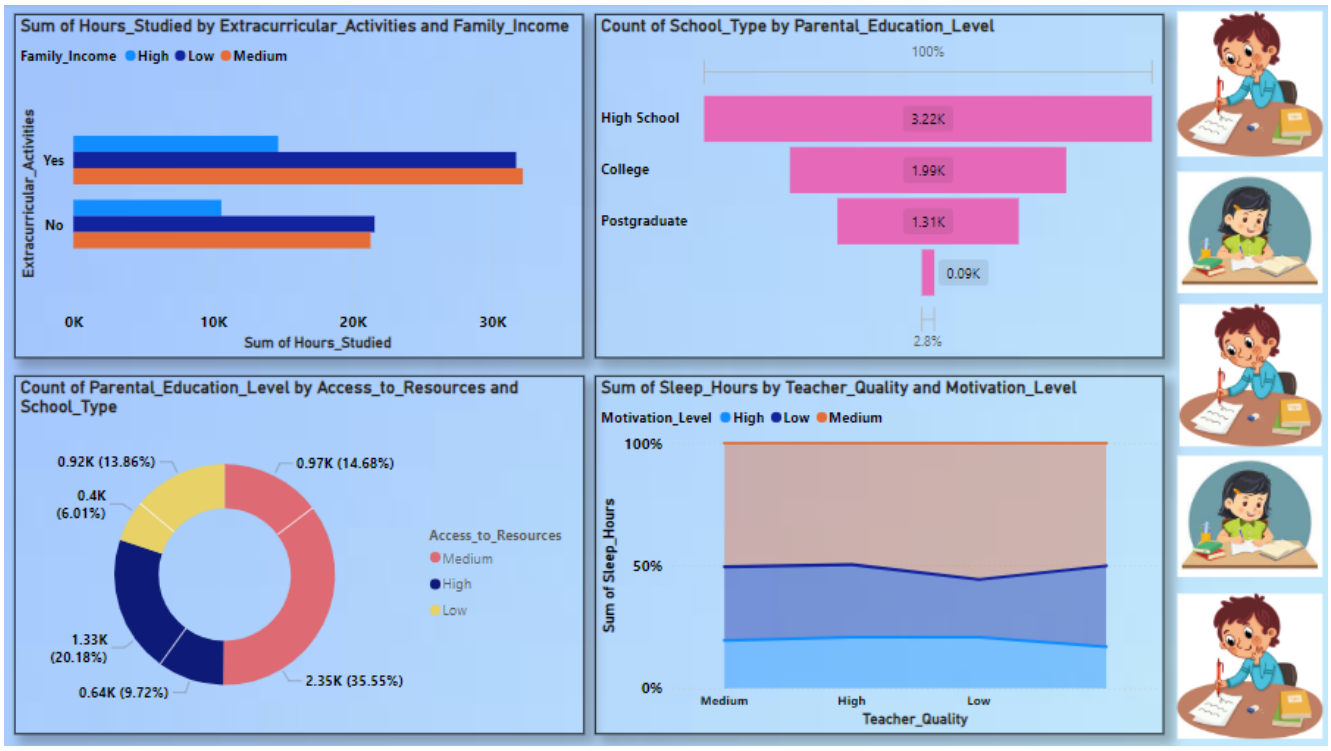
   - Add calculated columns.
   - Combine datasets using joins or append.

6. **Grouping & Aggregation:** Summarize data using sums, averages, or counts.

7. **Remove Unnecessary Data:** Delete irrelevant rows/columns for efficiency.

Report Views:





This Power BI dashboards provides a comprehensive **analysis**, on Student Performance Factors.showcasing key metrics like Sum of attendance, exam Score, School Type and Income of a Family. It uses visualizations such as

pie charts, bar charts, and line graphs, donut Chart , Funnel ,Clustered line chart, flow graph to highlight **Sum of attendance by school type**, **count of access to resources by exam score**, and **count of peer influence by extra curricular activities.** The data is further segmented by **sum of hours studied by extracurricular activities, count of school type by parental education level.** offering insights into students performance, study hours, and operational efficiency.

Conclusion:

The data suggests that factors like family income, parental education level, and motivation significantly influence students' study habits, resource access, and overall well-being. Addressing disparities in resources and promoting extracurricular engagement might enhance academic outcomes and holistic development.Students with better access to resources tend to achieve higher exam scores, particularly in the mid-range (60-80 marks). This highlights the critical role of resource availability in enhancing academic outcomes.