Student Performance Analysis Using Demographic and Behavioral Factors

In this project, I analyzed a dataset containing student scores in math, reading, and writing, exploring how various demographic and behavioral factors influence academic performance. Key tasks included:

1. **Data Preparation**:
    - Loaded the dataset and performed initial exploratory data analysis (EDA) to understand its structure, including summary statistics and identifying missing values.
    - Cleaned the dataset by removing unnecessary columns to focus on the most relevant features.
2. **Gender Distribution Analysis**:
    - Visualized the gender distribution within the dataset, ensuring a balanced representation that allows for meaningful insights.
3. **Performance Analysis by Demographics**:
    - Grouped the data by parental education level, marital status, weekly study hours, and the number of siblings.
    - Calculated the average scores in math, reading, and writing for each group to identify trends and correlations.
4. **Heatmap Visualizations**:
    - Created heatmaps to visualize the relationship between these demographic factors and student performance across the three subjects. This approach highlighted key areas where certain factors, like higher parental education or more study hours, correlate with improved student outcomes.

This analysis provides actionable insights into how different demographic and behavioral factors influence academic performance, making it a valuable tool for educators and policymakers. The project demonstrates my ability to clean, analyze, and visualize complex datasets, and to derive meaningful conclusions from the data.