**Abdullah Zahoor - 253074089**

**Analysis Report: Academic Performance Dataset**

**1. Dataset Overview**

The synthetic dataset contains information on 500 students, covering demographic, academic, and extracurricular data. The key attributes include study hours, GPA, family income, gender, extracurricular activities, and final grades.

**2. Statistical Summary**

* **Age**: Ranges from 15 to 22 years, with an average of 18.54 years.
* **Study Hours**: Average of 19.87 hours per week, ranging between 0.43 to 39.99 hours.
* **Previous GPA**: Mean GPA of 3.02, with a minimum of 2.01 and a maximum of 3.99.
* **Extracurricular Activities**: Students participate in 0 to 5 activities, with an average of 2.49.
* **Family Income**: Ranges from $20,034 to $199,474, with a mean of $108,547.
* **Final Grades**: The average final grade is 40.87, with a standard deviation of 12.78. Grades range from 0 to 78.6.

**3. Key Findings**

**Correlation Between Study Hours and Final Grades**

* A positive correlation was observed between study hours and final grades (0.47). This suggests that increased study hours are moderately associated with higher academic performance.

**Impact of Family Income on Academic Performance**

* A scatter plot (Figure 4) shows no significant linear relationship between family income and final grades. This indicates that economic background does not strongly influence grades.

**Gender Differences in Final Grades**

* The mean final grade for females is slightly higher than males (Figure 2). However, the box plot indicates substantial overlap in the distributions, suggesting that gender differences are not statistically significant.

**Relationship Between Extracurricular Activities and Final Grades**

* A box plot (Figure 5) highlights slight variations in final grades across different levels of extracurricular participation. However, no strong trend was evident, implying that moderate participation in activities does not drastically impact academic performance.

**4. Visualizations**

The visualizations provide deeper insights:

* **Correlation Matrix**: Highlights relationships among variables, with study hours and GPA showing the strongest correlation with final grades.
* **Box Plot by Gender**: Displays overlapping grade distributions for males and females.
* **Histogram of Study Hours**: Shows a fairly uniform distribution of study hours.
* **Scatter Plot of Family Income and Final Grades**: Demonstrates a lack of strong correlation.
* **Box Plot of Extracurricular Activities**: Examines grade distribution across activity levels.

**5. Conclusion**

* **Key Influencers**: Study hours and previous GPA are the most critical factors affecting final grades.
* **Minor Influences**: Gender and extracurricular activities show minimal impact on academic performance.
* **Economic Neutrality**: Family income does not significantly affect grades.

These findings emphasize the importance of academic effort (study hours) and prior performance (GPA) over demographic and socioeconomic factors. The dataset and analysis provide a foundational understanding of the drivers of academic success.