Author:

Muhammad Subhan
Data Analyst | Enthusiastic Learner

🔍 Problem Statement

Title: *Student Performance Analysis Based on Study Hours*

The objective of this project is to analyze the relationship between students' study hours and their academic performance. By using exploratory data analysis (EDA) and visualization techniques, the goal is to identify patterns, clean data, handle null values, and draw meaningful insights that can inform educational strategies and student self-assessment.

📘 Project Report

1. Project Title:

Student Performance Analysis with Python

2. Objective:

To analyze and visualize student performance data to find trends, clean the dataset, and derive insights from the correlation between study hours and scores.

3. Tools & Technologies Used:

- **Python**
- **Pandas**
- **Matplotlib**
- **Seaborn**
- **NumPy**
- **Jupyter Notebook**

---

4. Dataset:

Filename: studentscores.csv.csv
Description: This dataset contains columns representing students' weekly study hours and their corresponding scores.

---

5. Data Preprocessing:

✅ **Steps Performed:**

- Dropped unnamed columns (garbage values during CSV export).
- Cleaned and renamed column headers for clarity.
- Filled or dropped null/missing values if any.
- Verified data types and integrity.

---

6. Exploratory Data Analysis (EDA):

📊 **Visualizations & Insights:**

- **Scatter Plot:** Visualized relationship between Hours and Scores.
- **Correlation Heatmap:** Checked how closely hours studied affect scores.
- **Regression Line:** Fit linear regression to model prediction capability.

---

7. Key Findings:

- There is a strong **positive correlation** between study hours and student scores.
- The dataset supports the **linear relationship** hypothesis.
- Students who studied more generally scored higher.

---

8. Conclusion:

This analysis helps in understanding how consistent study habits can significantly improve academic results. The findings can be beneficial for:

- Students to self-monitor study hours.

- Educators to set study schedules.
- Institutions to build prediction models based on learning time.

---

09. How to Use:

1. Clone the repository.
2. Open Analysis.ipynb using Jupyter Notebook.
3. Run all cells step by step to explore insights.
4. Dataset file must be placed in the same directory.
