

## Project Idea: Education Performance Dataset Analysis

---

### Week 1: Data Modeling, Data Cleaning, and Preprocessing

#### Tasks:

- **Data Understanding:**

Understand the dataset structure (students, subjects, grades, attendance, demographics).

- **Data Cleaning:**

- Handle missing values
- Remove duplicates
- Standardize grades, dates, and categorical values

- **Data Modeling:**

Design a data model (Students – Subjects – Scores – Attendance).

#### Tools:

- SQL
- Python (Pandas, NumPy, Matplotlib)

#### Deliverables:

- Cleaned dataset ready for analysis
- Data preprocessing notebook
- Data model diagram (ERD or Star Schema)

---

### Week 2: Exploratory Data Analysis (EDA) & Analytical Questions

#### Tasks:

- **Define Analytical Questions**, such as:

- What is the impact of attendance on student performance?
- Which subjects have the highest failure rates?
- Does study time affect student grades?

- Is there a performance difference based on gender?
- **Exploratory Data Analysis (EDA):**
  - Grade distributions
  - Performance trends by subject and time

**Tools:**

- SQL
- Python (Pandas, Matplotlib, Seaborn)

**Deliverables:**

- List of analytical questions
  - Visualizations answering the analysis questions
  - Insights and findings summary
- 

**Week 3: Prediction / Forecasting Phase**

**Tasks:**

- **Define Prediction Questions**, such as:
  - Can we predict whether a student will pass or fail?
  - Can we predict a student's final grade based on attendance and study hours?
- **Build Predictive Models:**
  - Classification model (Pass / Fail)
  - Regression model (Grade prediction)

**Tools:**

- Python (Scikit-learn, Pandas, Matplotlib)

**Deliverables:**

- Prediction results
  - Model performance evaluation (accuracy, precision, recall, RMSE)
  - Visualizations explaining prediction outcomes
- 

**Week 4: Visualization Dashboard & Final Presentation****Tasks:**

- **Build an Interactive Dashboard:**
  - Student performance KPIs
  - Attendance impact on grades
  - Subject difficulty comparison
- **Final Report & Presentation:**
  - Problem statement
  - Data preparation process
  - Analysis results and insights
  - Recommendations for educational improvement

**Tools:**

- Power BI or Tableau
- SQL
- Python

**Deliverables:**

- Interactive visualization dashboard
  - Final project report
  - Final presentation slides
-

## **Key KPIs**

- Average Grade
  - Pass / Fail Rate
  - Attendance Rate
  - Performance by Subject
- 

## **Project Outcome**

This project aims to provide data-driven insights that help educational institutions:

- Improve student performance
- Identify at-risk students early
- Enhance teaching and learning strategies