# Student Performance Analysis using Exploratory Data Analysis

## Abstract

This project studies classroom sample data to understand the relationship between attendance and academic performance. The aim is to help teachers identify weak students early and improve learning outcomes.

## Introduction

Educational institutes collect marks but rarely analyze them deeply. Exploratory Data Analysis (EDA) allows us to discover patterns such as subject difficulty and attendance impact.

## Dataset Sample

| Name | Math | English | Science | Attendance |
|---|---|---|---|---|
| Aman | 78 | 65 | 70 | 88 |
| Riya | 92 | 90 | 94 | 95 |
| Kunal | 55 | 60 | 58 | 72 |
| Neha | 66 | 70 | 62 | 80 |
| Pooja | 84 | 82 | 79 | 91 |

## Methodology

1. Data Cleaning
2. Average Score Calculation
3. Attendance vs Marks Comparison
4. Weak Subject Identification
5. Performance Categorization

## Python Implementation

```
import pandas as pd
df=pd.read_csv('students.csv')
df['Average']=df[['Math','English','Science']].mean(axis=1)
weak=df[['Math','English','Science']].mean().idxmin()
print(weak)
```

## Results

Students above 85% attendance consistently scored higher averages. Mathematics had the lowest average score indicating conceptual difficulty.

## Future Work

Future versions may include machine learning prediction of final grades based on past performance.

Created by: Anamika Dubey
February 2026