

# **Capstone Project: Student Performance Analysis**

Candidate Name: Durva Patil

Batch Code: 09

Location: Pune

Guide Name: Shivani Mam

year: 2026

## **Index**

1. Introduction
2. Literature Review / Background Study
3. Problem Statement & Objectives
4. Methodology / System Design
5. Implementation / Execution
6. Results & Analysis
7. Conclusion
8. References
9. Appendices

## **1. Introduction**

Background: This project analyses a student performance dataset to explore factors associated with differences in academic scores.

Significance of the project: Understanding patterns in student performance helps educators, Elders, and parents identify at-risk groups and target interventions.

Scope of work: Data cleaning, exploratory data analysis, visualization, simple statistical summaries, and actionable observations based on the dataset.

## **2. Literature Review / Background Study**

Review of existing solutions/studies: Prior studies often investigate the relationship between demographic factors (Subject, Grade , Attendance, Performance), test preparation and students' scores. Typical methods include descriptive statistics, regression analysis, and visualization.

Gaps identified: This analysis focuses on descriptive EDA from the provided dataset; further work could include predictive modeling and causal analysis.

### **3. Problem Statement & Objectives**

Problem statement: Which student-level factors in the dataset are associated with higher or lower academic scores?

Project objectives:

- Explore distribution of scores and identify patterns.
- Compare average scores across demographic groups .
- Visualize relationships between different score types and provide interpretations.

### **4. Methodology / System Design**

Tools & Technologies Used: Python (pandas, matplotlib), Google docs for document generation.

Phases of project development:

- Data ingestion and cleaning
- Exploratory data analysis
- Visualization and interpretation
- Document preparation

### **5. Implementation / Execution**

Step-by-step details: Data read from the provided Excel file, basic cleaning applied (handled missing values where necessary), summary statistics computed, and visualizations created.

Screenshots: See plots below

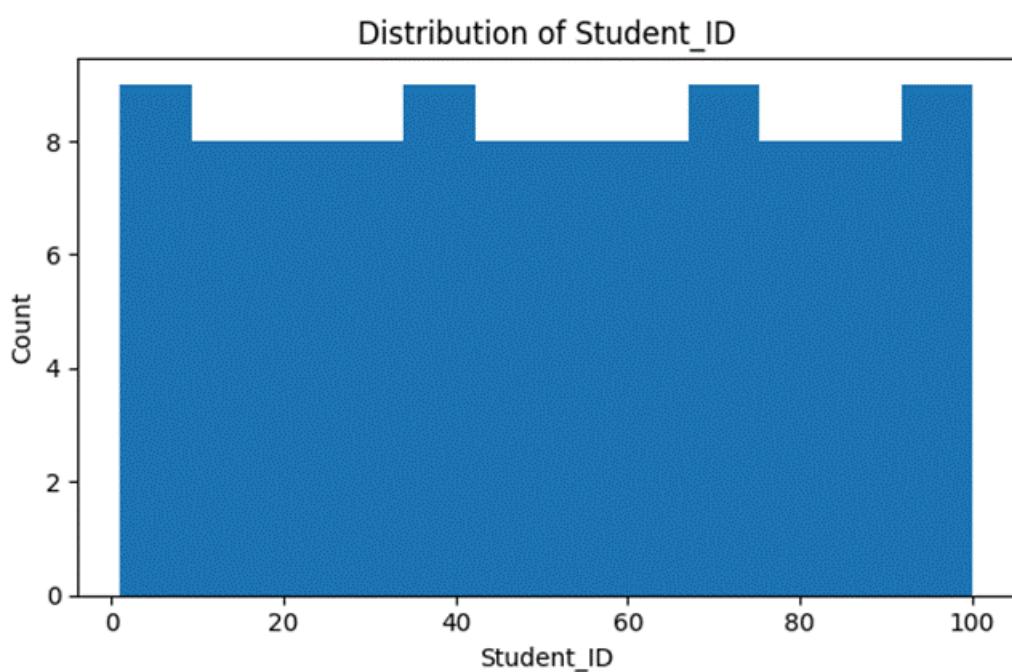
### **6. Results & Analysis**

Summary statistics for score-related columns:

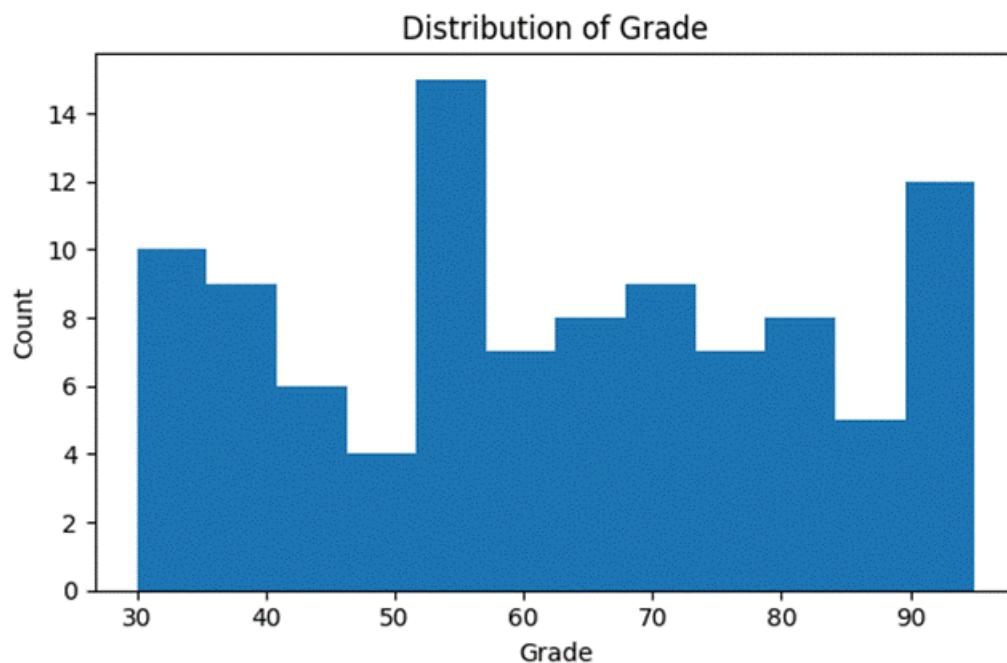
Score Column	count	mean	std	min	25%	50%	75%	max
Student_ID	100	50.5	29.01	1.0	25.75	50.5	75.25	100.0
Grade	100	62.31	19.61	30.0	46.5	62.0	78.25	95.0
Attendance_%	100	71.8	16.72	45.0	57.0	71.0	87.0	98.0
Subject wise avg of students	100	62.3	1.84	59.74	59.74	63.19	63.95	63.95

Charts and Visualizations:

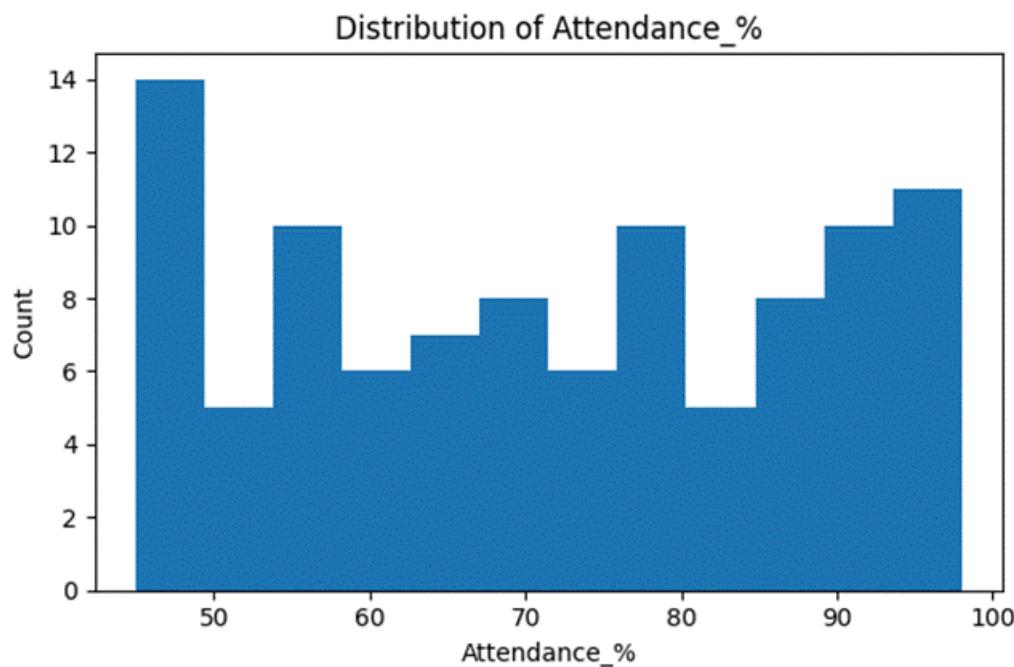
hist StudentID.png



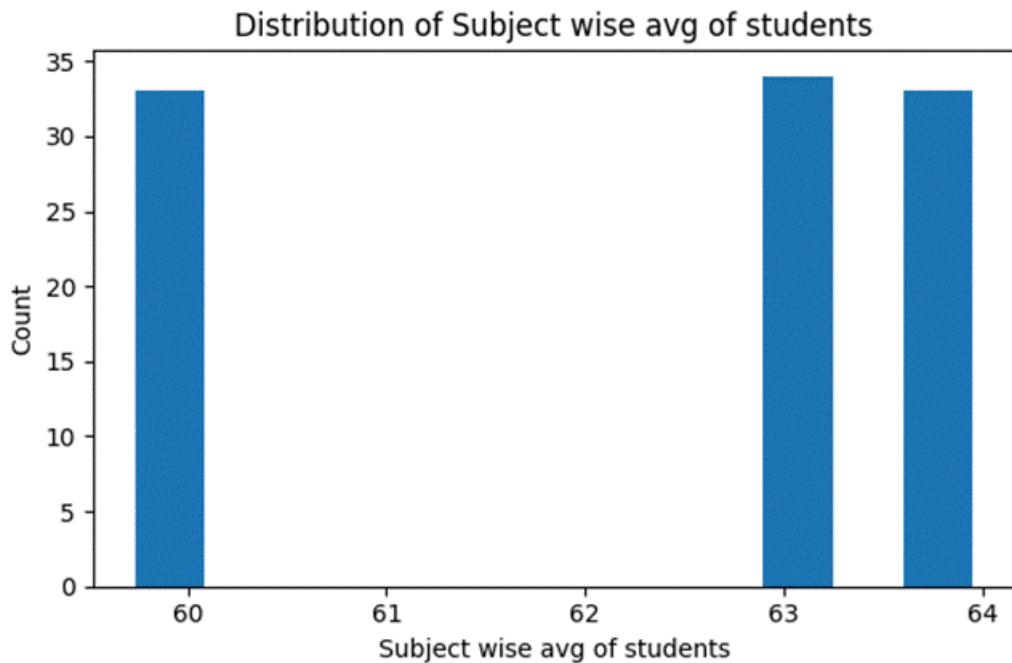
hist Grade.png



hist Attendance%.png



hist Subject wise avg of students.png



## 7. Conclusion

Summary of work done: Performed exploratory analysis on the provided student performance dataset, generated descriptive statistics and visualizations, and documented findings . Further work could include predictive modeling and deeper statistical testing.

## 8. References

Dataset: Student\_Performance

## 10. Appendices

### Appendix A – Business Requirements Document (BRD)

Includes detailed functional and non-functional requirements, stakeholder analysis, project scope, assumptions, and constraints defined during the requirement gathering phase.

## **Appendix B – Process Flow Diagrams**

Contains workflow diagrams and process maps illustrating the current and proposed system processes.

## **Appendix C – Dashboard Screenshots**

Screenshots of the developed dashboards highlighting key performance indicators (KPIs), data visualizations, and insights generated from the analysis.

## **Appendix D – Tools and Technologies Used**

- Python (Data Analysis & EDA)
- SQL (Data Extraction & Querying)
- Power BI / Tableau (Dashboard Development)
- MS Excel (Data Cleaning & Preparation)
- GitHub (Version Control)