

Week 4: Statistical Model Selection and Hypothesis Development

Student Performance Prediction
Project



Problem Statement:



01

Objective

Understand what influences student academic performance.

02

Prediction Goal

Forecast final exam scores or categorize performance levels.

03

Analytical Approach

Use statistical models (Regression, ANOVA, Logistic Regression).

Selected Statistical Techniques

Linear Regression

Predict continuous outcomes such as final exam scores using predictors like hours studied and attendance.

ANOVA

Compare average performance across different groups, for example, students with versus without part-time jobs.

Logistic Regression

Classify students into binary categories (e.g., pass/fail or high/low performance).

Feature Selection

Hours studied per week

Participation in extracurricular activities

Parental education level

Attendance percentage

Part-time job status

Hypotheses



01

H1

Students who study more hours per week will achieve significantly higher final scores.

02

H2

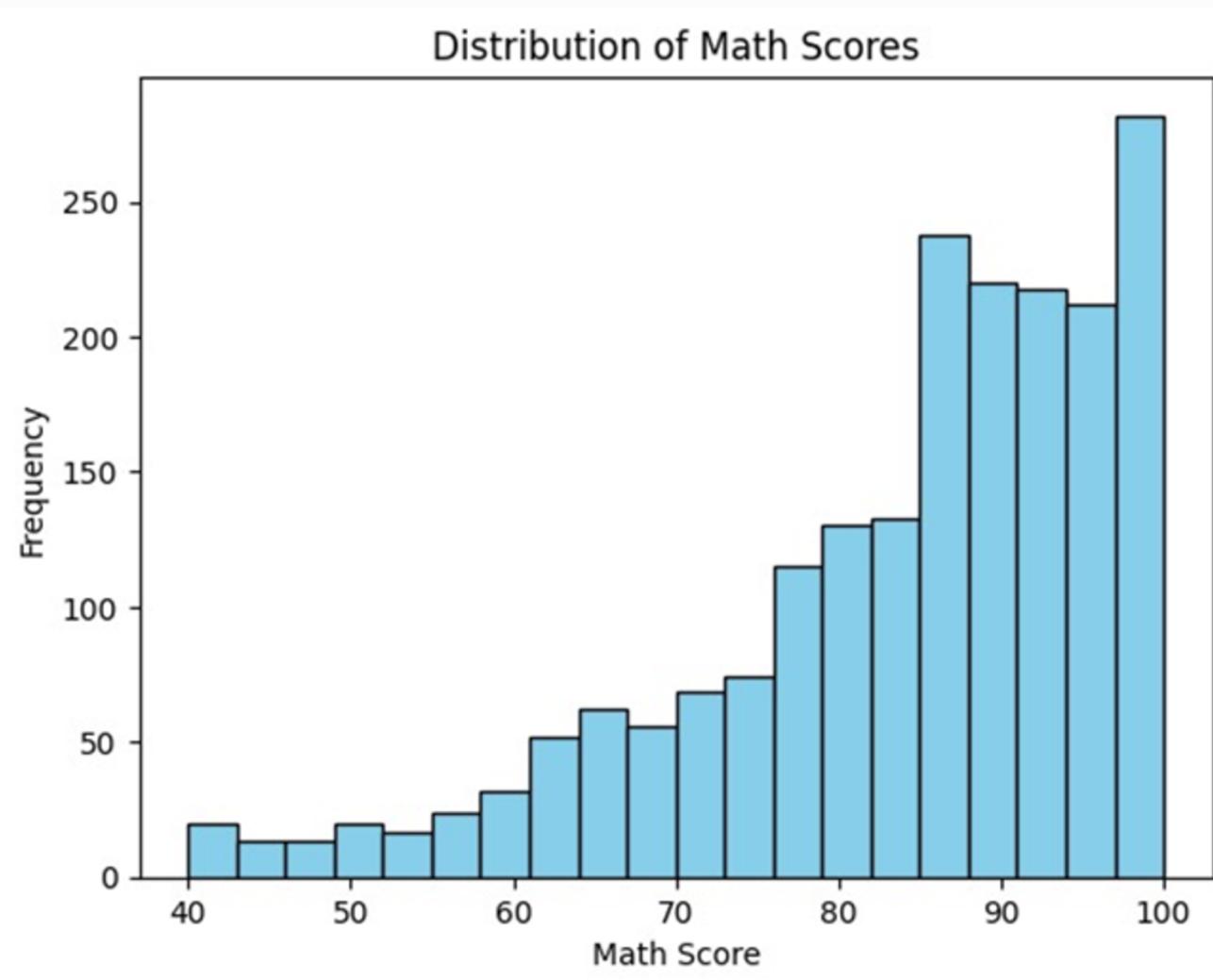
There is a significant difference in mean scores between students with and without part-time jobs.

03

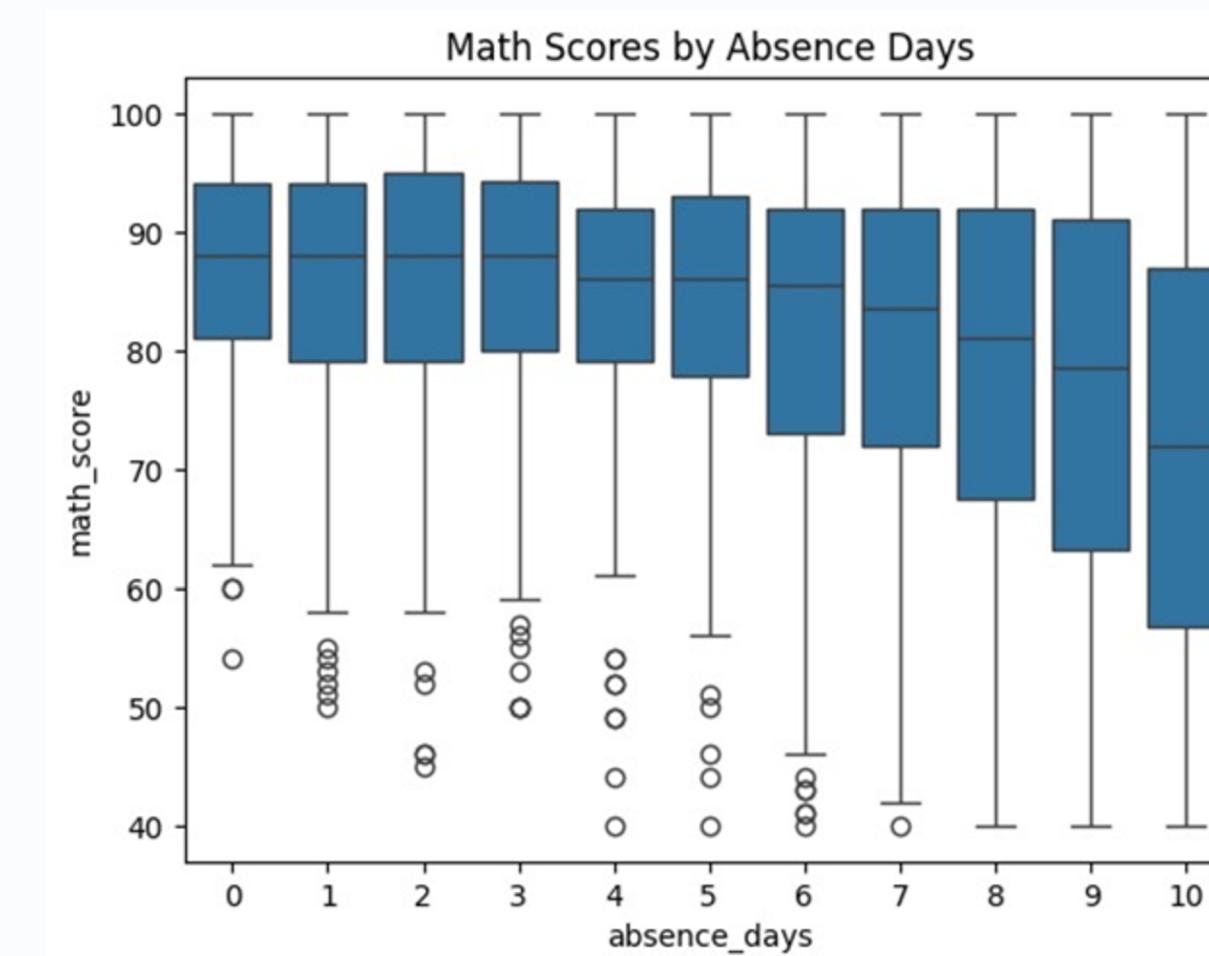
H3

Logistic regression can accurately predict whether a student will pass or fail based on the selected features.

Visual Examples

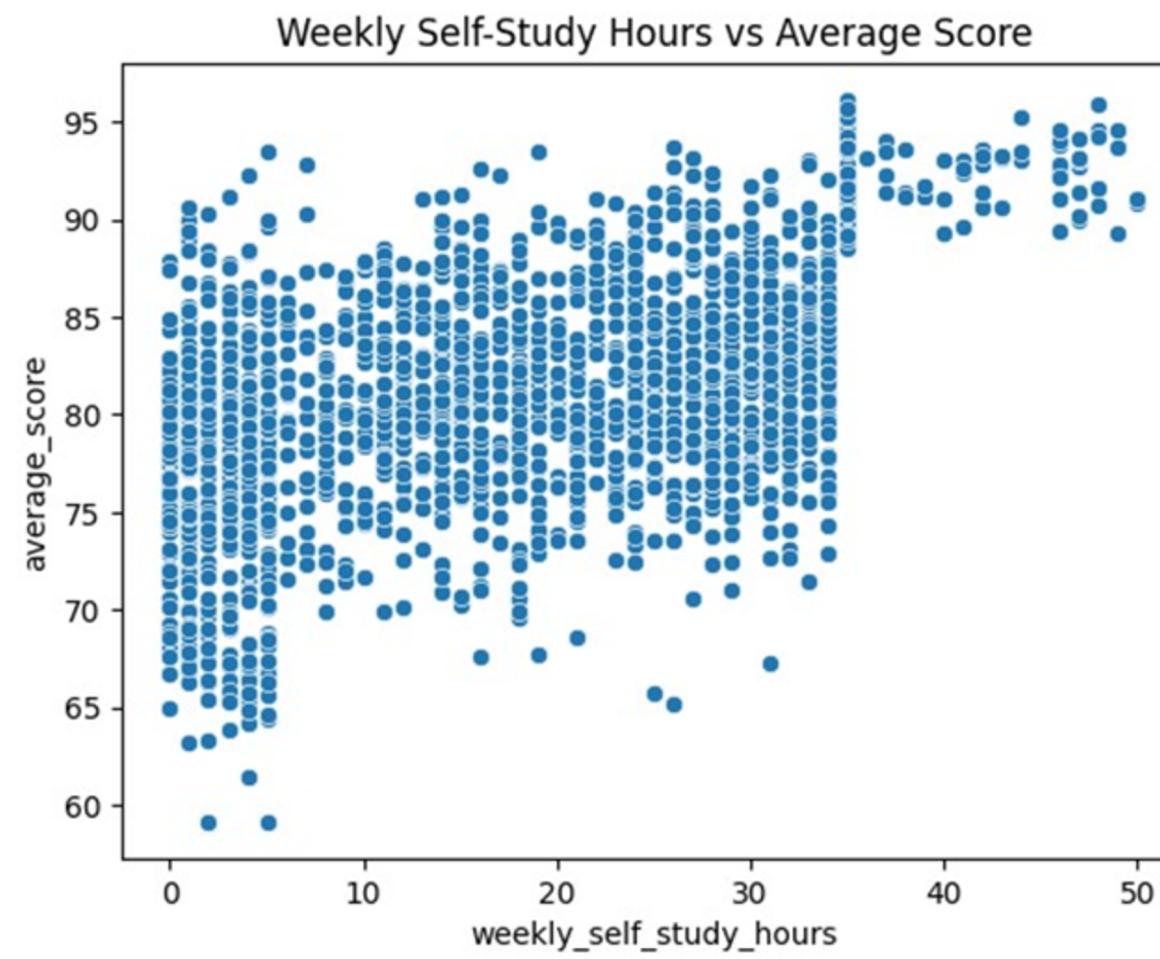


Histogram

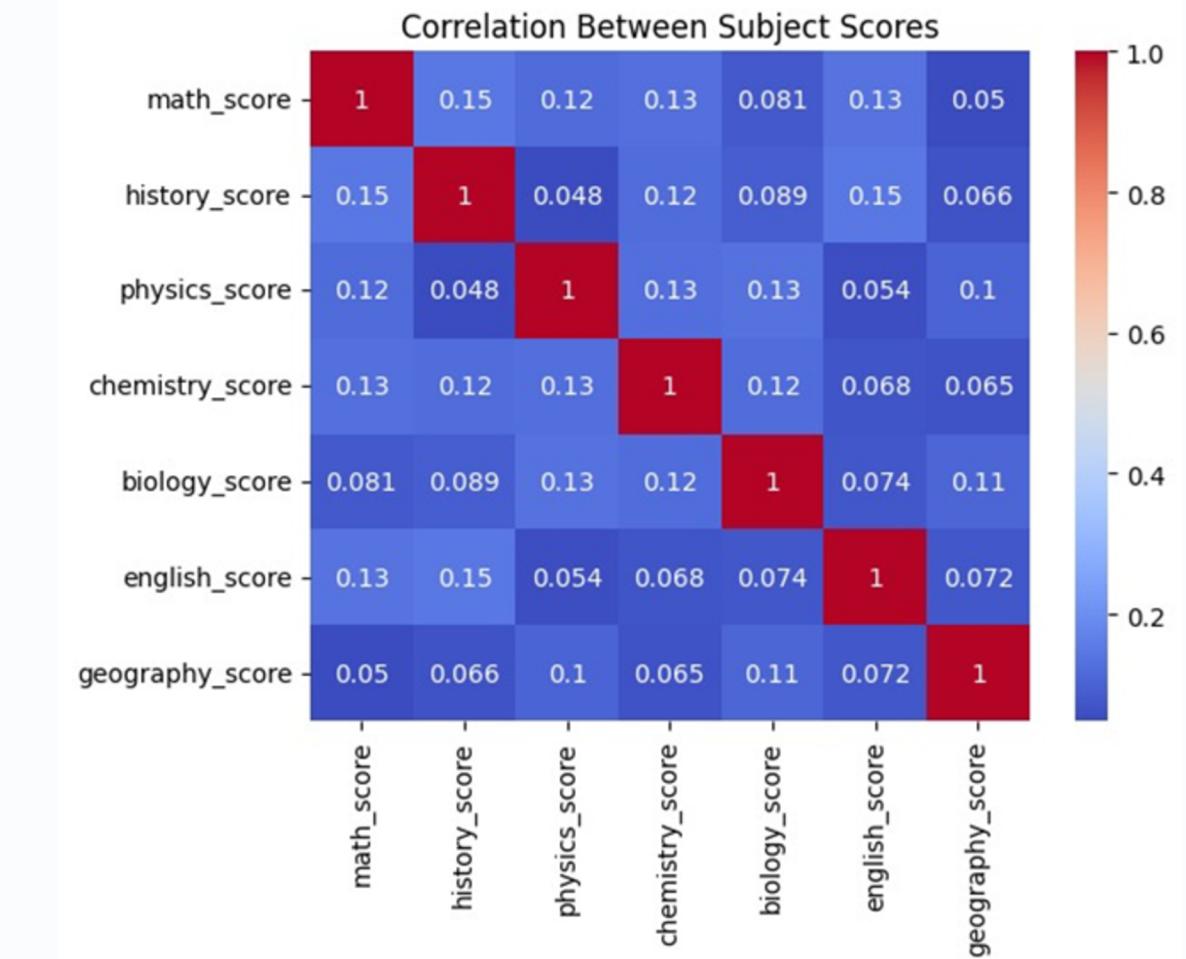


Box Plot

Visual Examples

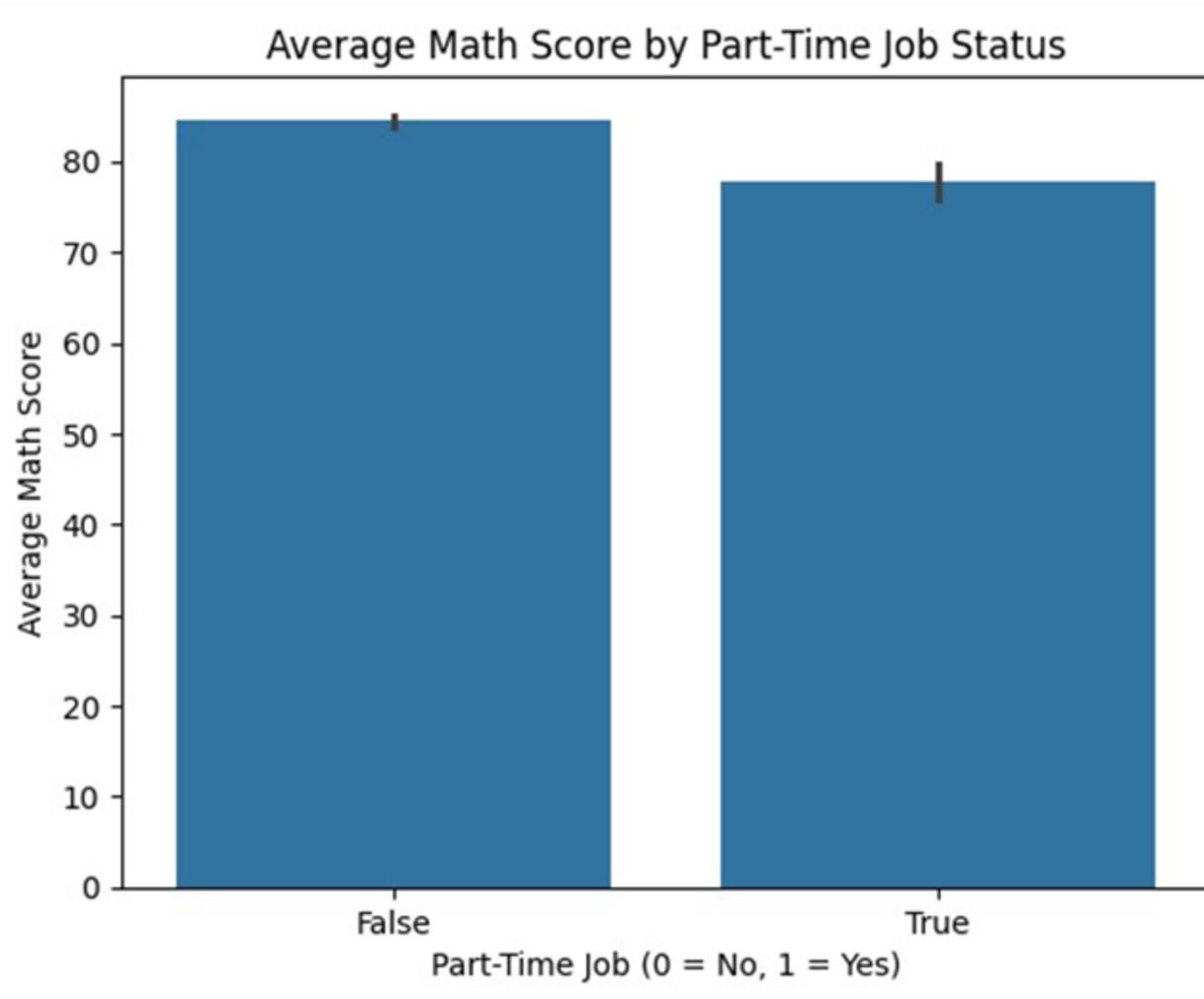


Scatter Plot

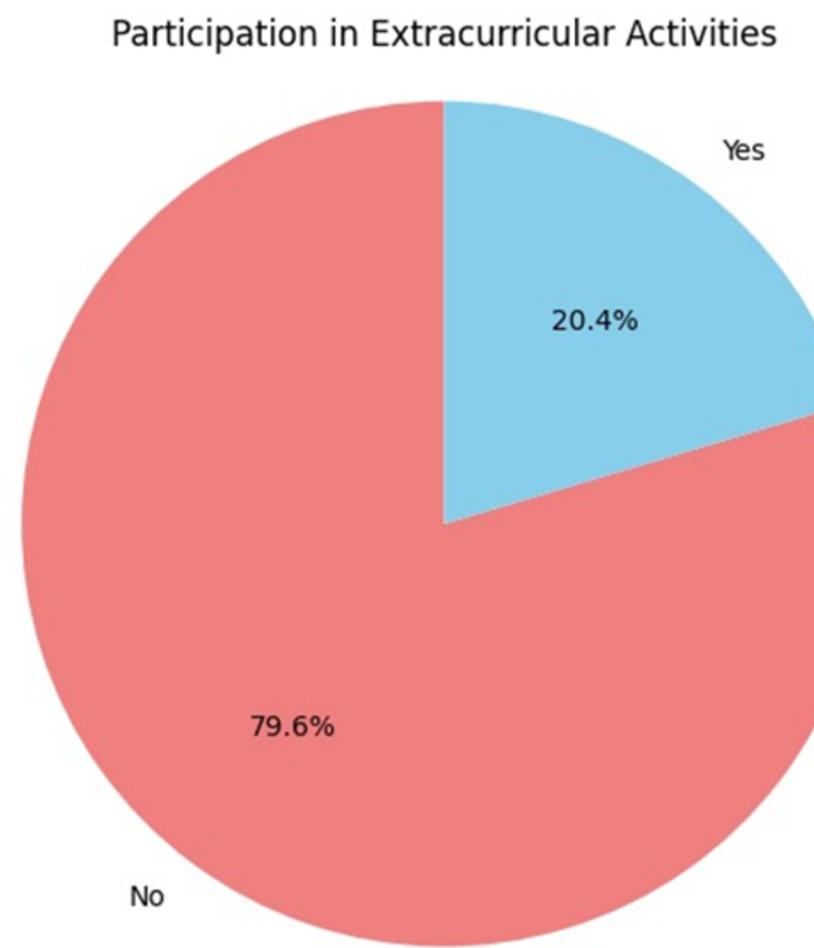


Heatmap

Visual Examples

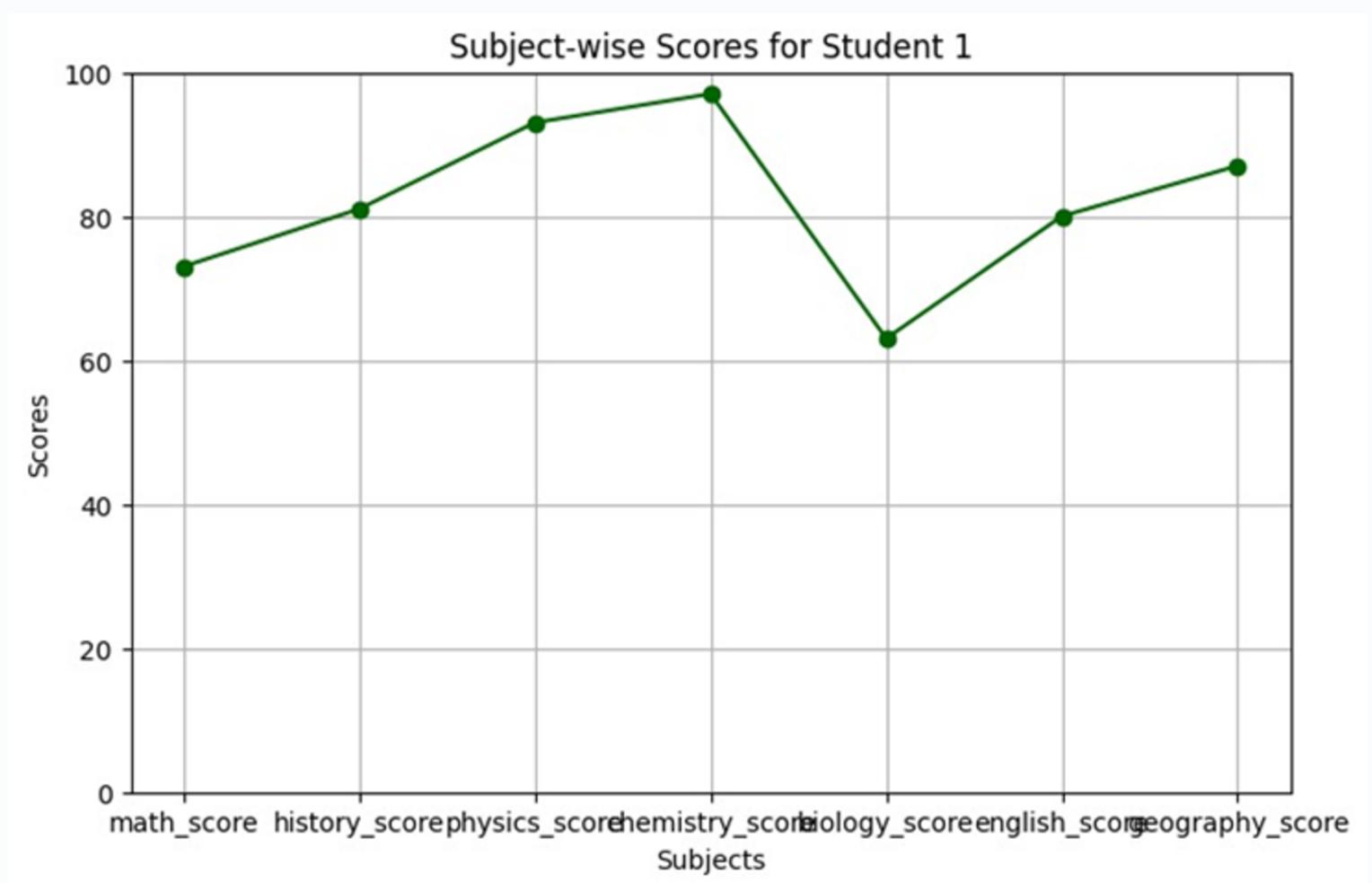


Bar Plot

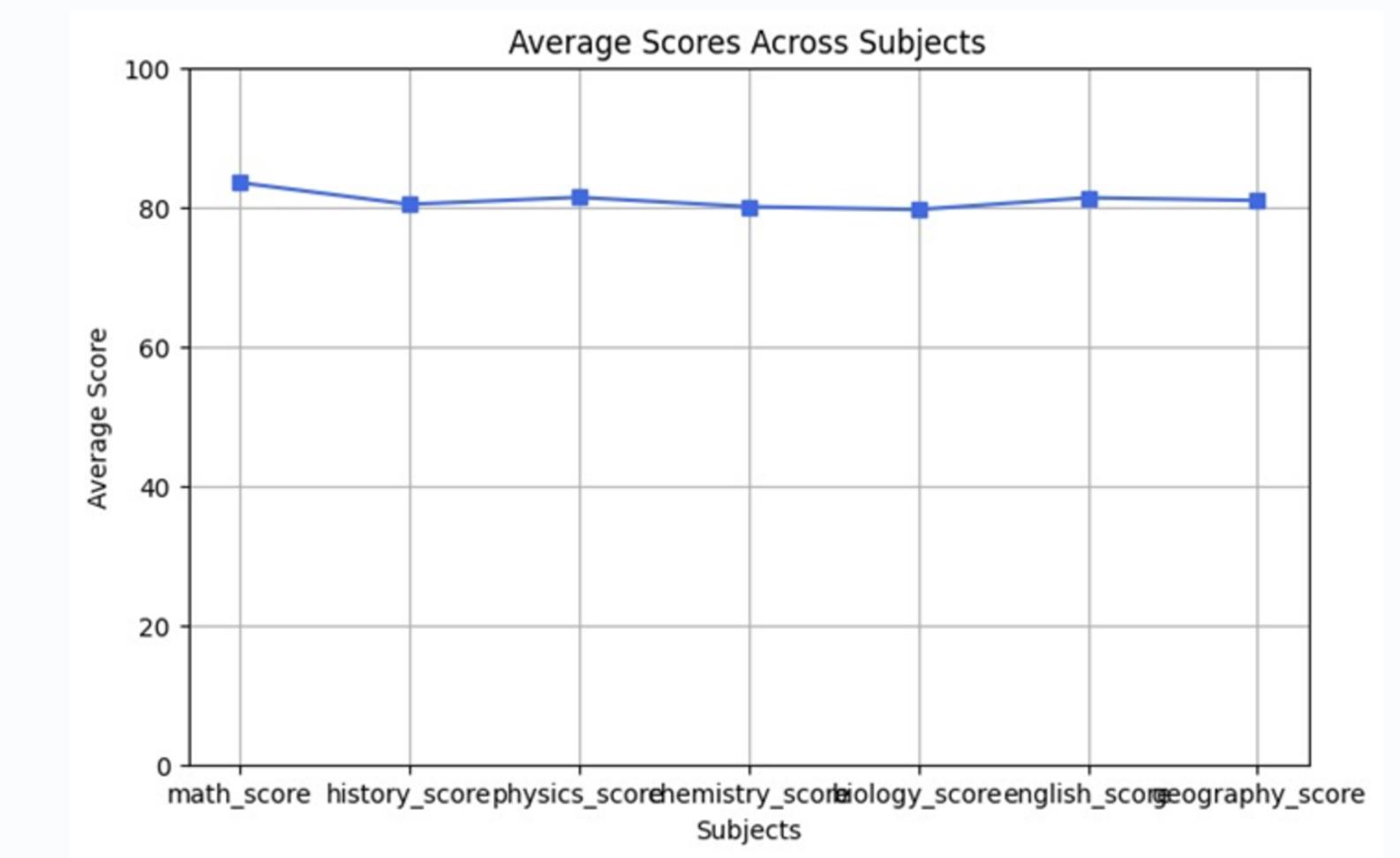


Pie Chart

Visual Examples



Line Plot:
Subject wise score for one student



Line Plot:
Average scores across subjects

Thank You