

# Statistical Analysis Report - Student Dropout Prediction

Name: Andy Ndubuisi Joseph

Company: 3SignetLtd

Position: Data Science Intern

## 1. Descriptive Statistics

In examining our dataset of 4,424 student entries, here's what I found:

- Marital Status: The majority of students are single, with an average value of 1.
- Application Mode: Students used a variety of application modes, with an average of 18 and a maximum of 57.
- Previous Qualifications: Students come from diverse educational backgrounds, with an average qualification level of 4 and a maximum of 43.
- Admission Grade: The average admission grade is 126, with a standard deviation of 14.
- Age at Enrollment: Students generally enroll around the age of 23, though ages range from 17 to 70.
- Curricular Units: Most students completed a few units, with the highest number being 26 in the first semester.
- Socioeconomic Indicators: The unemployment rate averages 11, and the inflation rate is quite low at 1.

This analysis highlights the diversity among students and points to areas that might benefit from further exploration, such as how socioeconomic factors impact educational outcomes.

## 2. Correlation Matrix Heatmap

While I encountered some clustering issues with the full correlation matrix, here are a few insights:

- Age at Enrollment vs. Admission Grade: There's minimal correlation here, suggesting that a student's age doesn't significantly affect their admission grades.
- Scholarship Holder Status vs. Admission Grade: There seems to be a positive correlation, indicating that students with higher admission grades might be more likely to receive scholarships.

### 3. Results and Interpretation of Hypothesis Tests

#### 3.1. Socioeconomic Status and Dropout Rates

- Pearson Correlation: 0.018

- P-value: 0.224

- Interpretation: The correlation is very weak, and the result isn't statistically significant. This means there's no strong evidence that socioeconomic status (based on "Father's qualification") influences dropout rates in our dataset.

#### 3.2. Admission Grades and Dropout Rates

- Pearson Correlation: -0.096

- P-value:  $< 0.05$

- Interpretation: There's a slight negative correlation here, which is statistically significant. This suggests that students with higher admission grades are somewhat less likely to drop out.

#### 3.3. Financial Aid/Scholarships and Dropout Rates

- T-statistic: -16.83

- P-value:  $1.18e-61$

- Interpretation: There's a significant difference in dropout rates between students who receive scholarships and those who don't. Scholarship holders tend to have much lower dropout rates, and this result is highly statistically significant.