

Egypt University of Informatics

Computer and Information Systems

Data Analysis Course

The Analysis of the Performance of Data Analysis Students

Abdullah moussa 22-101114

29/3/2023

# Introduction

# This survey examines how high school curriculum influences college success. Your participation provides valuable insights into the relationship between high school education and college outcomes.

# Research Question

Is there any relation between the high school curriculum and the college outcomes?

# Hypothesis

# Population of Interest:

EUI students

# Sampling Method:

Simple random sampling method

# Bias Identification:

to ensure fairness, we'll consider potential biases related to age, major, and gender:

1. Age Bias: Different age groups may perceive their high school experiences and college outcomes differently.
2. Major Bias: The choice of major can influence perceptions of the relevance and effectiveness of high school curriculum.
3. Gender Bias: Gender differences may impact how individuals perceive and respond to questions about their educational experiences. By acknowledging and addressing these biases, we aim to gather diverse perspectives on the impact of high school curriculum on college success.

# Survey Questions:

Which high school program did you graduate from?

Gender?

How would you rate the difficulty of adapting to the college Curriculum?

What is your major?

What is your age?

What is your CGPA?

Online survey link: https://docs.google.com/forms/d/e/1FAIpQLSetqE8r-8ZYvOtBPImCpOi2DKOGLdyoPlpvxqmlh-KYZfXhRQ/viewform?usp=sf\_link

Number of samples collected: 36.

# Analysis:

The Gender distribution graph illustrates a survey response distribution of 21 males and 15 females out of a total of 36 respondents.

This imbalance suggests a higher representation of males in the survey sample, which could impact the generalizability of findings and underscores the importance of considering gender diversity in data interpretation.

A blue and pink pie chart

Description automatically generated

The analysis shows that among the respondent, 20 graduated from Thanawya Amma, 14 from the IG program, and 2 from the American program, indicating a higher prevalence of Thanawya Amma graduated.

A graph with blue squares and white text

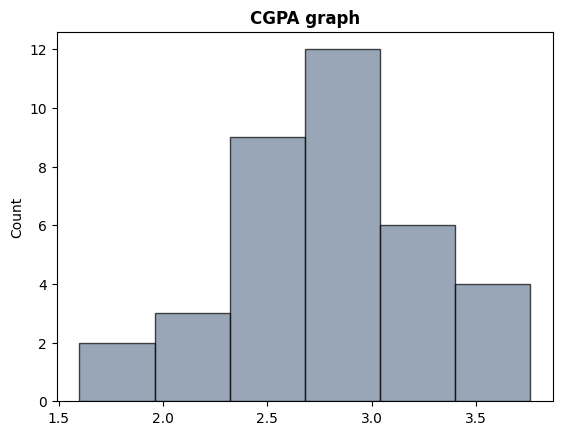
Description automatically generated

the difficulty in adapting to college graph demonstrates that a significant portion of respondents (16) reported a neutral experience, while fewer reported difficulties (10) or simplicity (8). Very few respondents found the adaptation either very difficult (1) or very simple (1).

A graph of different colored bars

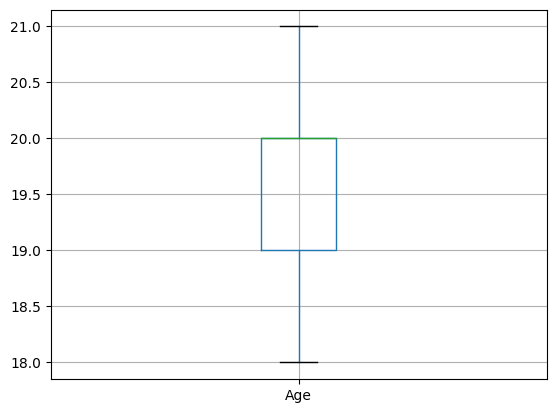
Description automatically generated

The analysis of GPAs indicates a central distribution, with a mean of around 2.77 and a median of 2.76. Notably, the data exhibits a bimodal pattern, with prominent peaks at both 2.4 and 2.79 GPA, indicating two distinct clusters in the respondents' academic performance.



The analysis of ages reveals a mean of approximately 19.66 years, with a predominant mode at 20 years based on 16 responses.

While no median is provided due to the even distribution, the concentration of responses around 20 suggests a central tendency in the age distribution of the survey respondents.



The correlation between age and GPA is found to be 0.193629, indicating a very weak positive relationship between these variables.

# Conclusion

The survey reveals a strong negative correlation (-0.636) between adapting to college and CGPA, implying that students struggling with college adaptation tend to have lower GPAs.

However, the correlation between high school attended and CGPA is negligible (0.023), indicating little impact of school on GPA.

In summary, while college adaptation strongly influences GPA, the high school attended has minimal effect.

# Any potential issues

1. Sampling Bias: If the survey respondents are not representative of the entire population, the results may not accurately reflect the broader context. ( 69.4% CIS students , however there was only 27.8 engineering students)
2. Response Bias: Respondents may provide inaccurate or biased information due to factors such as social desirability bias or misunderstanding of survey questions.
3. Limited Scope: The survey may not cover all relevant factors influencing the relationships being studied, leading to incomplete or biased conclusions. ( the link was sent to the WhatsApp group and not even half of them filled out the form )
4. Making some correlations by one-hot encoding