

Student Outcomes Study

N = 40 · Dataset: survey.csv

Statistical Tests Used

Objective	Test	Variables	Why this test
To describe the age, study_hours, gpa, and satisfaction of the respondents.	Descriptive Statistics	age, study_hours, gpa	Summary statistics for the continuous variables.
To describe the age, study_hours, gpa, and satisfaction of the respondents.	Frequency & Percentage Distribution	satisfaction	Frequency and percentage distribution for the categorical variables.
To determine the relationship between study_hours and gpa.	Spearman Rank Correlation	study_hours, gpa	
To determine the relationship between age and gpa.	Pearson Correlation	age, gpa	Both 'age' and 'gpa' are continuous; Pearson correlation measures their linear association.
To determine whether gpa differs across the program groups.	One-Way ANOVA	gpa, program	Comparing mean 'gpa' across the 3 independent groups of 'program'.
To determine whether study_hours differs across the program groups.	One-Way ANOVA	study_hours, program	Comparing mean 'study_hours' across the 3 independent groups of 'program'.
To determine the association between gender and civil_status.	Chi-Square Test of Independence	gender, civil_status	Both 'gender' and 'civil_status' are categorical, so correlation is not appropriate. The association between two categorical variables is tested with the Chi-Square Test of Independence.
To determine the association between gender and program.	Chi-Square Test of Independence	gender, program	Both 'gender' and 'program' are categorical, so correlation is not appropriate. The association between two categorical variables is tested with the Chi-Square Test of Independence.

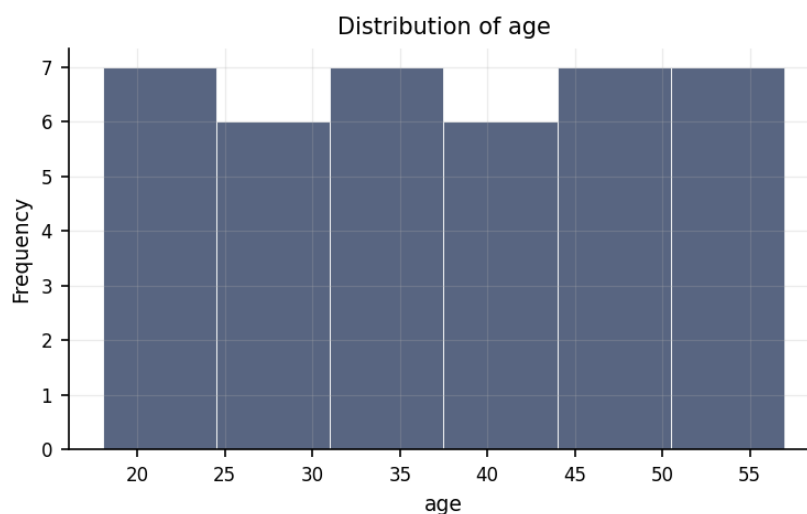
Methods

All statistical analyses were performed in Python using the SciPy and statsmodels libraries via thericerca's deterministic analysis engine (v0.1.0). Statistical significance was evaluated at $\alpha = .05$ (a 95% confidence level). Descriptive statistics were computed to summarise the study's variables. To assess monotonic associations between variables, a Spearman rank-order correlation was used (Spearman, 1904). To assess associations between continuous variables, a Pearson product-moment correlation was used (Pearson, 1895). To assess differences in a continuous outcome across three or more groups, a one-way analysis of variance (ANOVA) was used (Fisher, 1925). To assess associations between categorical variables, a Pearson chi-square test of independence was used (Pearson, 1900). For each inferential test, the appropriate effect size is reported with its 95% confidence interval, alongside the achieved statistical power to detect the observed effect (Cohen, 1988).

Results

To describe the age, study_hours, gpa, and satisfaction of the respondents.

This simply summarises the data — the average, spread, and range of each variable are in the table above.

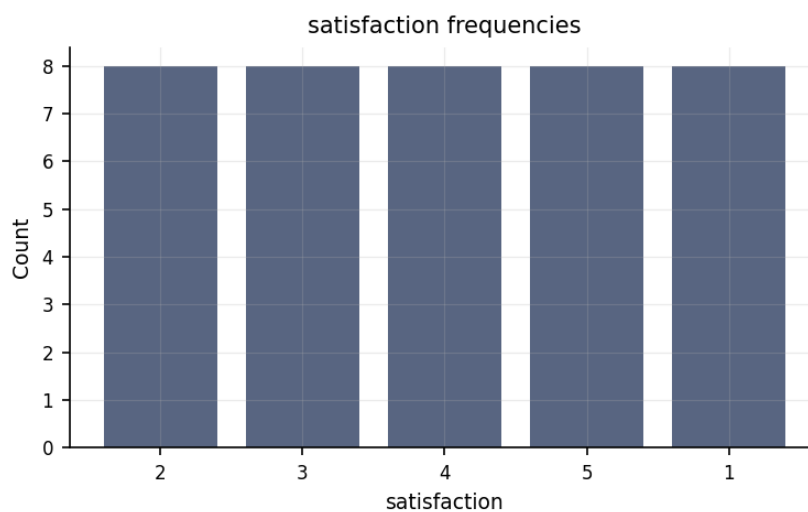


Distribution of age.

max	mean	median	min	mode	n	range	std	variable	variance
57	37.5	37.5	18	18	40	39	11.69	age	136.667
32	18.5	18.5	5	5	40	27	8.727	study_hours	76.154
3.97	2.728	2.72	1.57	1.57	40	2.4	0.748	gpa	0.559

To describe the age, study_hours, gpa, and satisfaction of the respondents.

This shows how the responses are distributed: how many people fall into each category, and what share of the total that is.



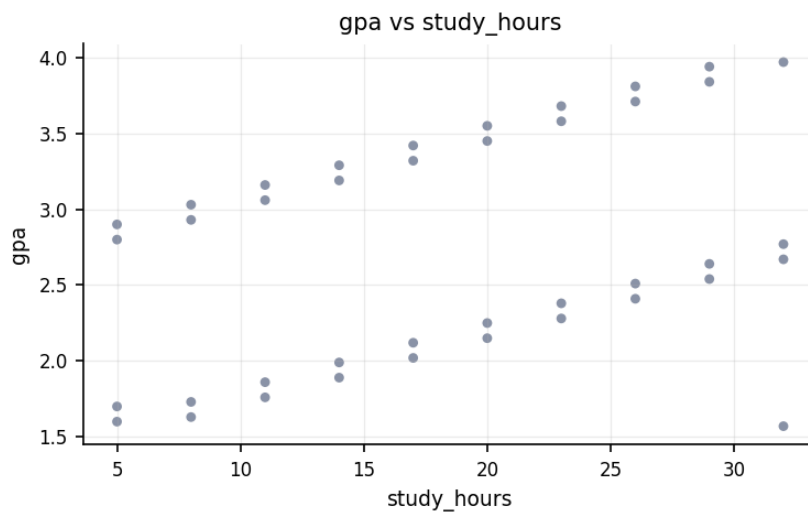
Frequency of each satisfaction category.

count	percent	value	variable
8	20	2	satisfaction
8	20	3	satisfaction
8	20	4	satisfaction
8	20	5	satisfaction
8	20	1	satisfaction

To determine the relationship between study_hours and gpa.

We looked at whether study hours and GPA move together. They do — as one goes up, the other tends to go up too. The strength of this pattern is medium.

Statistically: $\rho = .36$, $p = .022$, achieved power = 0.63



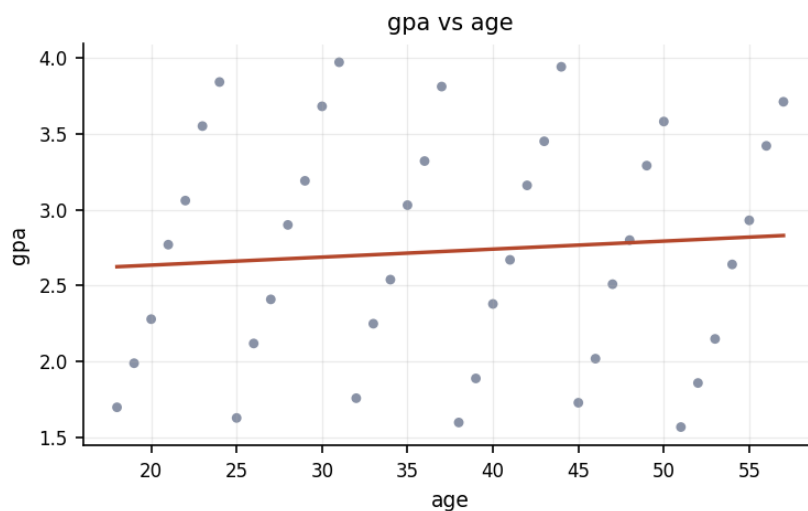
Relationship between study_hours and gpa.

△ Spearman Rank Correlation was used instead of the Pearson Correlation because the latter's assumption(s) were not met (Normality of study_hours (Shapiro-Wilk)).

To determine the relationship between age and gpa.

We looked at whether age and GPA are related. We found no clear connection: in this sample, knowing one tells us little about the other.

Statistically: $r(38) = .08$, $p = .614$, 95% CI $[-.24, .38]$, achieved power = 0.08

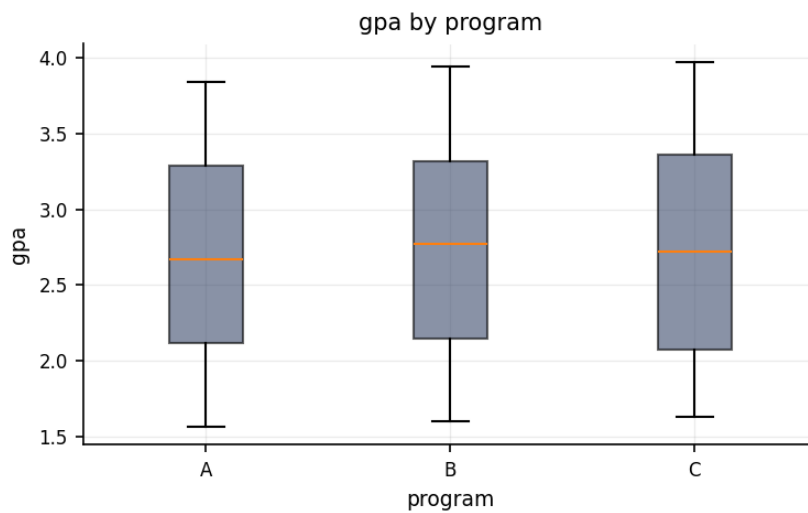


Relationship between age and gpa.

To determine whether gpa differs across the program groups.

We compared GPA across the program groups. They were broadly similar — no group clearly stood out.

Statistically: $F(2, 37) = 0.02$, $p = .977$, $\eta^2 = .00$, achieved power = 0.05



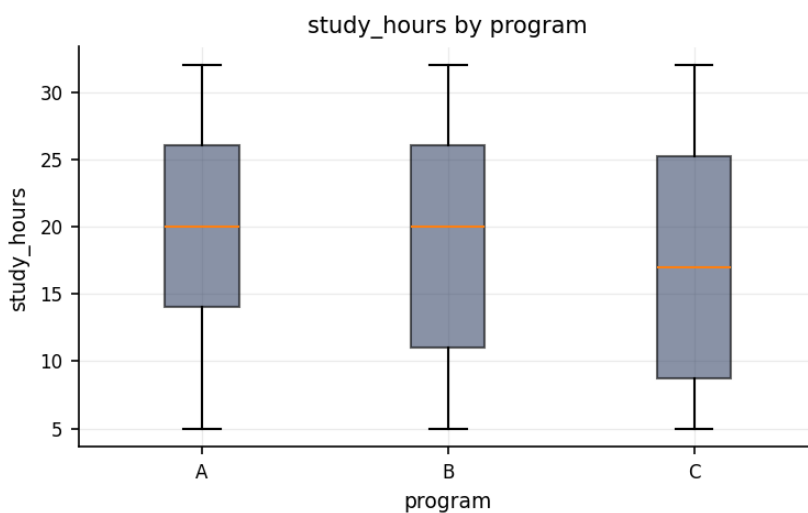
Distribution of gpa across program groups.

group	mean	n	std
A	2.692	13	0.751
B	2.754	13	0.76
C	2.736	14	0.788

To determine whether study_hours differs across the program groups.

We compared study hours across the program groups. They were broadly similar — no group clearly stood out.

Statistically: $F(2, 37) = 0.24$, $p = .784$, $\eta^2 = .01$, achieved power = 0.09



Distribution of study_hours across program groups.

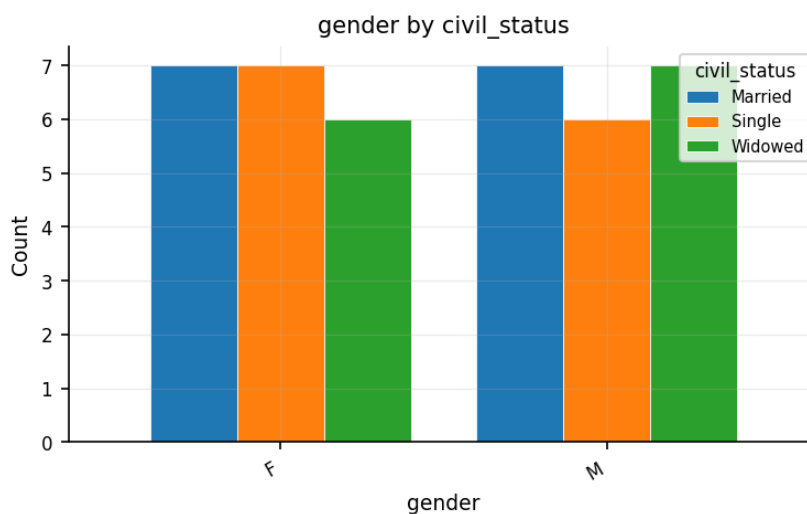
group	mean	n	std
A	19.538	13	8.903

group	mean	n	std
B	18.846	13	8.707
C	17.214	14	9.074

To determine the association between gender and civil_status.

We checked whether gender and civil status are linked — for example, whether knowing someone's gender tells you anything about their civil status. In this sample they appear unrelated.

Statistically: $\chi^2(2, N = 40) = 0.15$, $p = .926$, $V = .06$, achieved power = 0.06



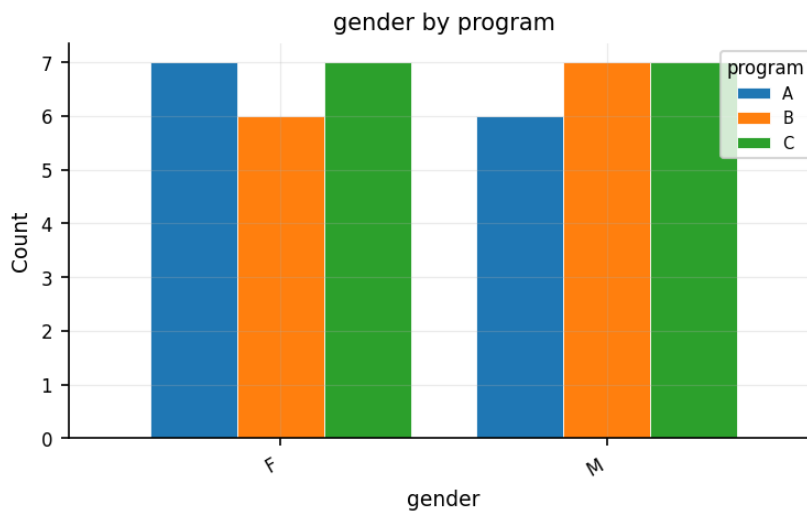
Counts of gender broken down by civil_status.

Married	Single	Widowed	row
7	7	6	F
7	6	7	M

To determine the association between gender and program.

We checked whether gender and program are linked — for example, whether knowing someone's gender tells you anything about their program. In this sample they appear unrelated.

Statistically: $\chi^2(2, N = 40) = 0.15$, $p = .926$, $V = .06$, achieved power = 0.06



Counts of gender broken down by program.

A	B	C	row
7	6	7	F
6	7	7	M

Discussion

The relationship between study hours and GPA. The analysis of study hours and GPA revealed a statistically reliable medium positive relationship. In practical terms, higher values of study hours tend to accompany higher values of GPA. An association of this strength indicates the two measures share a substantial amount of common variation, though a correlation on its own cannot establish that one variable causes the other.

The relationship between age and GPA. The analysis of age and GPA did not reveal a statistically reliable relationship. The two measures appear to vary largely independently in this sample, and the small association observed sits well within the range expected by chance. This is not evidence that no relationship exists — only that the present data give no clear sign of one.

Differences in GPA across program. Comparing GPA across the program groups did not reveal a statistically reliable overall difference. The groups were broadly similar on GPA, so no single group stands out and no pairwise follow-up interpretation is warranted.

Differences in study hours across program. Comparing study hours across the program groups did not reveal a statistically reliable overall difference. The groups were broadly similar on study hours, so no single group stands out and no pairwise follow-up interpretation is warranted.

The association between gender and civil status. Gender and civil status were not significantly associated. The two categorical variables appear to be independent in this sample, so a case's gender tells us little about its civil status.

The association between gender and program. Gender and program were not significantly associated. The two categorical variables appear to be independent in this sample, so a case's gender tells us little about its program.

The investigation revealed meaningful insights into what influences academic performance, showing a clear distinction between factors that relate to student grades and those that do not. Most notably, the inquiry established a reliable link between the number of hours students spend studying and their overall grade point average. In practical terms, this demonstrates that dedicating more time to independent academic preparation is associated with better academic outcomes. In statistical terms, this finding is considered statistically significant, meaning the observed connection is strong enough that it is very unlikely to be the result of random chance. In contrast, several other variables examined in the research showed no clear link or detectable pattern. Specifically, there was no noticeable relationship between a student's age and their grade point average, suggesting that maturity or age alone does not dictate academic success in this context. Furthermore, when comparing students across different academic programs, the analysis revealed no meaningful differences in grade point averages or in the average number of hours students spent studying; students across distinct fields of study appear to perform and allocate study time at similar levels. Finally, the analysis evaluated demographic background characteristics and found no clear association between gender and civil status, nor between gender and program choice. These non-significant findings mean that any subtle differences observed in the sample were too small to be distinguished from normal, random variation. Importantly, a lack of statistical evidence for these relationships should be interpreted as an absence of detectable

pattern in this specific group rather than absolute proof that no relationship exists in broader settings. Consequently, the firm link between study hours and performance stands as the primary substantiated finding from this dataset.

Conclusions

In summary, the evidence collected yields a well-defined picture regarding academic achievement, supporting some initial expectations while finding no support for others. The firmly established finding—that study hours are positively tied to grade point average—provides a trustworthy, evidence-based insight that can be confidently integrated into educational discussions and future academic planning. This key finding highlights personal study time as a key area for supporting student achievement. Conversely, the absence of detectable effects regarding age, academic program, gender, and civil status indicates that these specific demographic and structural factors did not meaningfully differentiate student performance or study habits within this group. Maintaining a clear distinction between confirmed relationships and unproven ones is essential for drawing measured, defensible conclusions that neither overstate the influence of demographic background nor ignore the demonstrated value of dedicated study effort.

Recommendations

Based on these outcomes, several practical and research-focused steps are recommended. First, because the reliable connection between study time and academic performance offers the strongest basis for action, future research should aim to confirm this pattern in independent and larger student groups to ensure it applies broadly. To better understand how study habits influence grades, researchers should consider longitudinal designs that track students over extended periods or experimental interventions that test specific study strategies, allowing for a clearer understanding of cause and effect rather than simple association. Second, the areas that showed no clear link—such as program-level differences, age, and demographic distributions—should not be discarded prematurely; they warrant re-examination in larger, more diverse samples to determine whether very small, subtle patterns

exist that this evaluation could not detect. Throughout future work, emphasis should be placed on evaluating the real-world significance and practical magnitude of every finding, ensuring that educational decisions are guided by meaningful impacts rather than mathematical probability alone. Finally, maintaining consistent data collection methods and screening for inconsistencies will preserve data integrity in subsequent studies.

Limitations

Several important boundaries should be kept in mind when interpreting these conclusions. First, no major data-quality issues were identified during processing, which supports the internal consistency of the findings. However, the study design is purely observational, meaning it captures patterns as they occur naturally without direct intervention. As a result, even the statistically reliable relationship between study hours and grade point average cannot prove that increased studying directly causes higher grades; unmeasured background factors, such as individual motivation, prior academic preparation, or effective learning techniques, could also account for the observed success. Second, these insights describe the specific group of individuals evaluated in this study. Extending these conclusions to broader student populations or different institutional contexts depends entirely on how closely this group reflects those wider populations.

References

1. Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd.
2. Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58, 240-242.
3. Pearson, K. (1900). On the criterion that a given system of deviations... *Philosophical Magazine*, 50(302), 157-175.
4. Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15(1), 72-101.
5. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.
6. Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with Python. *Proceedings of the 9th Python in Science Conference*, 92-96.

7. Virtanen, P., Gommers, R., Oliphant, T. E., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17, 261–272.

Python Code Appendix

Descriptive Statistics

```
df[['study_hours', 'gpa']].describe()
```

Frequency & Percentage Distribution

```
df['satisfaction'].value_counts(normalize=True)
```

Spearman Rank Correlation

```
r, p = stats.spearmanr(df['study_hours'], df['gpa'])
```

Pearson Correlation

```
r, p = stats.pearsonr(df['age'], df['gpa'])
```

One-Way ANOVA

```
groups = by_level(df, 'gpa', 'program')  
F, p = stats.f_oneway(*groups)
```

Chi-Square Test of Independence

```
table = pd.crosstab(df['gender'], df['civil_status'])  
chi2, p, dof, expected = stats.chi2_contingency(table)
```

Numeric-claim validation: PASSED (44 numbers checked). Engine v0.1.0 · Registry v0.2.0. Prose source: llm:gemini.