STAT 355 Project

Shri Atluri

November 25, 2024

1 Introduction

Managing personal finances is a critical skill for students, yet their financial behaviors often remain underexplored. This project analyzes a dataset on student spending habits to uncover patterns, identify influential factors, and draw meaningful insights. The dataset contains a variety of variables, ranging from demographic attributes to detailed spending categories, allowing for a comprehensive analysis of student financial behavior.
Dataset: Student Spending Dataset

The structure of the project follows as below:

1. **Variables Exploration:** The first section focuses on exploring the dataset to gain a deeper understanding of its variables. Each variable is described in context, supplemented with summary statistics and visualizations. This preliminary analysis helps identify key variables for further investigation.

2. **Research Questions and Statistical Analysis** In the second section, three research questions are developed based on the findings from the exploration phase. Each question addresses a specific aspect of student spending and is analyzed using appropriate statistical methods. The methods applied include regression, ANOVA, and correlation analysis, among others. The results are interpreted in the context of student financial behavior.

3. **Discussion and Implications**: This section synthesizes the results of the analyses, discussing their implications for understanding student spending habits. Practical recommendations and insights are provided based on the findings.

4. **Conclusion** The final section summarizes the key takeaways from the project and highlights potential areas for future research.

By combining exploratory data analysis and statistical techniques, this project aims to uncover meaningful insights into the spending habits of students, providing a foundation for better understanding their financial behaviors.

2 Variables Exploration

The variables below were chosen based off their correlation with the questions in the next part

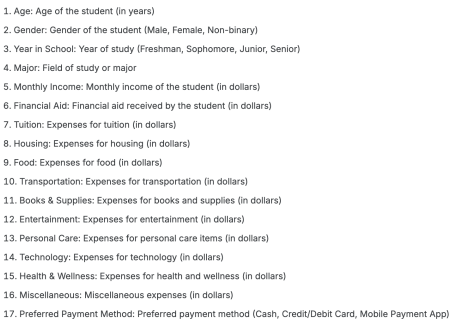

Figure 1: 17 total variables in the dataset

1. Age

Description: The age of the student, measured in years.
Relevance: Provides demographic context for analysis. While not directly tied to a specific research question, age may influence other variables like income or expenses.

2. Gender

Description: The gender of the student, categorized as Male, Female, or Non-binary.
Relevance: Used to explore potential differences in monthly income for Research Question 1.

3. Year in School

Description: The academic year of the student, categorized as Freshman, Sophomore, Junior, or Senior.
Relevance: Used to examine differences in housing expenses across academic years for Research Question 3.

4. Monthly Income

Description: The total income earned by the student each month, measured in dollars.
Relevance: A key variable for Research Questions 1 and 2, providing insights into the financial resources available to students.

5. Housing

Description: The amount of money spent on housing by the student each month, measured in dollars.
Relevance: Used to investigate spending differences across academic years for Research Question 3.

6. Entertainment

Description: The amount of money spent on entertainment by the student each month, measured in dollars.
Relevance: Used to analyze the relationship between income and discretionary spending for Research Question 2.

7. Food

Description: The amount of money spent on food by the student each month, measured in dollars.
Relevance: Provides context for overall spending habits and is relevant to understanding students' financial priorities.

8. Transportation

Description: The amount of money spent on transportation by the student each month, measured in dollars.
Relevance: Contextualizes students' spending on necessary expenses and may relate to housing or other financial behaviors.

9. Financial Aid

Description: The total amount of financial aid received by the student, measured in dollars.
Relevance: Highlights external support that might influence spending behaviors or savings patterns.

10. Savings (Derived) - user made

Description: Calculated as the difference between monthly income and total expenses (Housing, Food, Entertainment, etc.), measured in dollars.
Relevance: Provides insight into students' ability to save money, tying into broader financial behaviors.

3 Research Questions and Statistical Analysis

The next section will be the research questions we chose, with each question having the analysis and outcome provided, along with the test that was made.

1. Does the average monthly income differ significantly between male and female students?

Method: Two-sample t-test for means.
Hypotheses:

- H_0: There is no significant difference in average monthly income between male and female students.
- H_a: There is a significant difference in average monthly income between male and female students.

Analysis: A two-sample t-test was conducted to compare the average monthly income between male and female students. The assumptions of normality and equal variances were tested and validated prior to conducting the test.
Outcome: The t-test yielded a t-statistic of **-0.713** with 677 degrees of freedom and a p-value of **0.476**. The 95% confidence interval for the difference in means was [**-60.32, 28.18**]. The mean monthly income for males was $1008.26, while for females it was $1024.33.

Figure 2: Male vs Female monthly income

Conclusion: There is no significant difference in average monthly income between male and female students. The observed difference in means may be attributed to random variation.

2. Is there a relationship between monthly income and entertainment expenses?

Method: Simple linear regression.
Model: $y = \beta_0 + \beta_1 x + \epsilon$, where y is Entertainment Expenses and x is Monthly Income.
Hypotheses:

- H_0: There is no relationship between monthly income and entertainment expenses.
- H_a: There is a relationship between monthly income and entertainment expenses.

Analysis: A simple linear regression model was fitted to examine the relationship between monthly income and entertainment expenses. The model yielded an intercept of **81.59** and a slope of **0.0032**. The p-value for the slope was **0.44**, and the R-squared value was **0.0006**.
Outcome: The regression analysis indicates that there is no significant relationship between monthly income and entertainment expenses. The weak R-squared value further suggests that monthly income has little to no explanatory power for entertainment expenses.

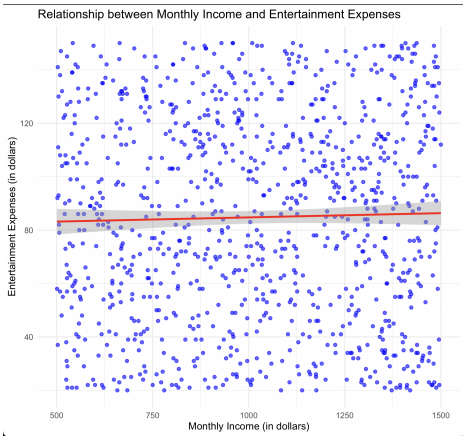

Figure 3: Relationship between monthly income and entertainment expenses

Conclusion: There is no statistically significant relationship between monthly income and entertainment expenses. Entertainment spending appears to be largely independent of income level. This is what out null hypothesis states as well.

3. Do students in different years of study (Freshman, Sophomore, Junior, Senior) spend significantly different amounts on housing?

Method: One-way ANOVA.
Hypotheses:

- H_0: There is no significant difference in mean housing expenses among students in different years of study.
- H_a: At least one group has a significantly different mean housing expense.

Analysis: A one-way ANOVA was conducted to compare housing expenses across the four academic years. The results yielded an F-statistic of **3.436** and a p-value of **0.0165**. The assumptions of equal variances and normality were tested and validated.
Outcome: The analysis indicates that **there is** a significant difference in housing expenses among students in different years of study. Due to this analysis with the p value being less than 0.05, we reject the null hypothesis.

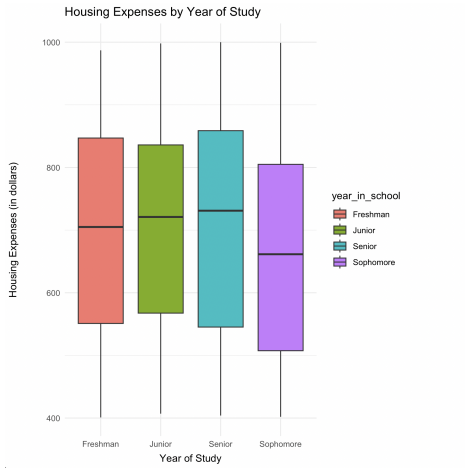

Figure 4: Housing prices for all 4 grade levels

Conclusion: Students' housing expenses differ significantly based on their year of study. Freshmen and Sophomores may spend more on housing due to dormitory costs, while Juniors and Seniors likely transition to less expensive off-campus housing.

4 Discussion and Implications

Recap of Key Findings

This study explored three key questions about student spending habits:

- **Monthly Income by Gender:** No significant difference in average monthly income was found between male and female students, suggesting financial opportunities and support are similarly distributed across genders.
- **Income and Entertainment Expenses:** No statistically significant relationship was observed between monthly income and entertainment expenses. Students appeared to maintain consistent discretionary spending habits regardless of income.
- **Housing Expenses by Year of Study:** The ANOVA test revealed a statistically significant difference in mean housing expenses across years of study ($p = 0.0165$). Further analysis was taken by viewing the box plot, which determined that housing expenses do not defer as much as the ANOVA indicated.

Broader Implications

The findings highlight consistent spending patterns among students, particularly in housing and entertainment. Universities could use these insights to better understand financial challenges faced by students and develop targeted financial education programs. The lack of a strong relationship between income and entertainment spending suggests students prioritize certain categories regardless of financial constraints.

Limitations

This analysis has several limitations:

- The dataset may not represent all students, as it could reflect a specific population or region.
- The data is self-reported, introducing potential biases in responses, along with the fact that it is fictional due to privacy regulations.
- The ANOVA analysis assumes equal variances and normality, which may not have been fully met in this dataset.

Future Research Directions

- Investigating other factors influencing student spending, such as parental support or employment status.
- Analyzing datasets from diverse student populations to have no skewed data sets.

5 Conclusion

This study analyzed student spending habits using statistical methods, including t-tests, regression analysis, and ANOVA. The results revealed no significant difference in monthly income by gender and no relationship between income and entertainment expenses, suggesting consistent financial patterns. An important difference in housing expenses across academic years was identified, though the practical differences were minimal.

These findings provide insights into student financial behaviors and can inform targeted financial literacy programs. Future studies should explore additional factors, such as parental support and regional cost-of-living differences, for a more comprehensive understanding.