Survey Analysis: Math 3210 Project 1

Noah G. Gallego

November 12, 2024

Abstract

This paper presents an in-depth analysis of financial habits among classmates at California State University, Bakersfield (CSUB), focusing on budgeting, spending, saving, and debt management. Using descriptive statistics, inferential analysis, and visualizations, we explore relationships between various financial behaviors and the factors that influence financial decision-making. Key findings suggest a positive relationship between employment status and saving habits, with employed students saving for future expenses at a mean rate of 0.636 compared to non-employed students. Additionally, notable correlations include the relationship between student loans and debt concern (correlation coefficient: 0.57). Inferential analyses further validate these findings, identifying significant relationships such as the impact of saving for retirement on financial confidence.

1 Introduction

In this project, we analyze survey data collected from classmates at CSUB to gain insights into their financial habits, such as budgeting methods, spending patterns, saving behaviors, and debt concerns. We conducted data cleaning and exploratory data analysis (EDA) to identify trends and patterns in financial practices, aiming to provide a better understanding of the financial challenges among our peers. Additionally, we plotted the locations of surveyed CSUB students to provide context for the demographic distribution.

2 Methodology

The survey was designed to gather quantitative data on financial behaviors among classmates. The sample included responses from 11 students. Data was collected on various factors such as budgeting methods, financial confidence, student loans, and saving habits. Each response was processed to transform qualitative answers (e.g., "yes" or "no") into binary values for easier statistical analysis.

2.1 Data Cleaning

Several data cleaning steps were implemented:

- Columns were renamed for clarity (e.g., `Question.One.` to `Budget_Management_Method`).
- Binary responses were converted to 1/0 format.
- Missing values were replaced with the mode of each respective column.

3 Data Analysis and Results

This section presents descriptive statistics, inferential analysis, and visualizations generated in R.

3.1 Descriptive Statistics

Descriptive statistics, including mean, median, mode, and standard deviation, were calculated for numerical columns. The summary statistics provided insights into general trends and variability in financial behaviors.

3.2 Summary Statistics

Table 1: Summary Statistics for Key Variables

Variable	Min	1st Quartile	Median	Mean	Max
Date Recorded	19999	20014	20016	20013	20018
Latitude	33.70	35.30	35.39	36.20	45.85
Longitude	-119.7	-119.1	-119.1	-119.0	-117.9
Financial Confidence	2.000	3.000	4.000	3.727	5.000
Student Loans	0.000	0.000	0.000	0.182	1.000
Student Debt Concern	0.000	2.000	4.000	3.091	5.000
Employment Status	0	0	0	0.000	0
Saving for Future Expenses	0.000	0.000	1.000	0.636	1.000
Saving for Retirement	0.000	0.000	0.000	0.455	1.000
Financial Independence Confidence	0.000	0.500	1.000	0.727	1.000
Job Confidence Post Graduation	0.000	0.000	1.000	0.636	1.000

3.3 Financial Confidence Distribution

The distribution of financial confidence among respondents is shown in Figure 1. Most students rated their confidence between moderate and high levels.

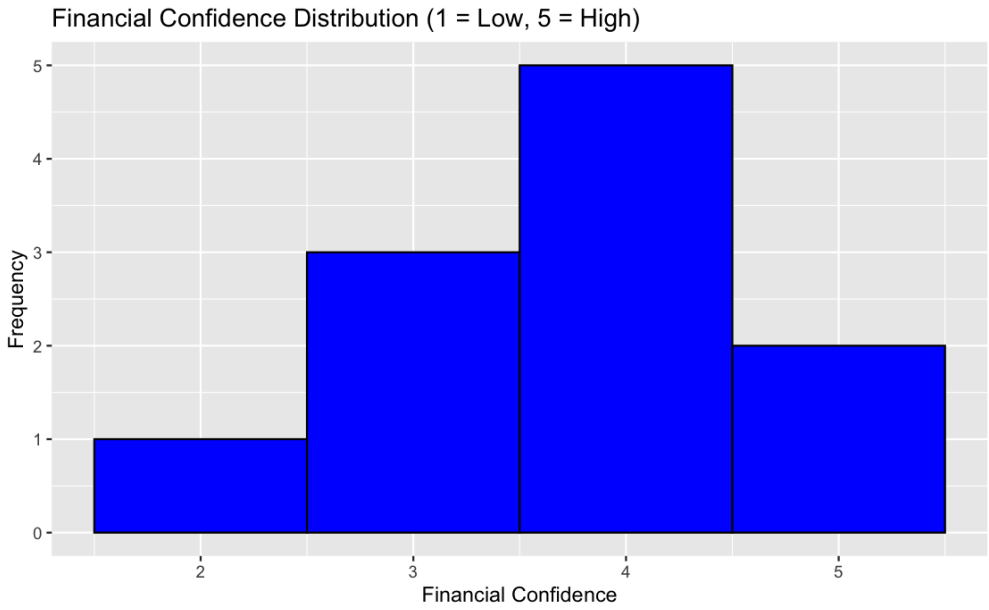

Figure 1: Financial Confidence Distribution (1 = Low, 5 = High)

3.4 Employment Status Distribution

The distribution of employment status is depicted in Figure 2. The majority of students are employed, which may influence financial confidence and saving behaviors.

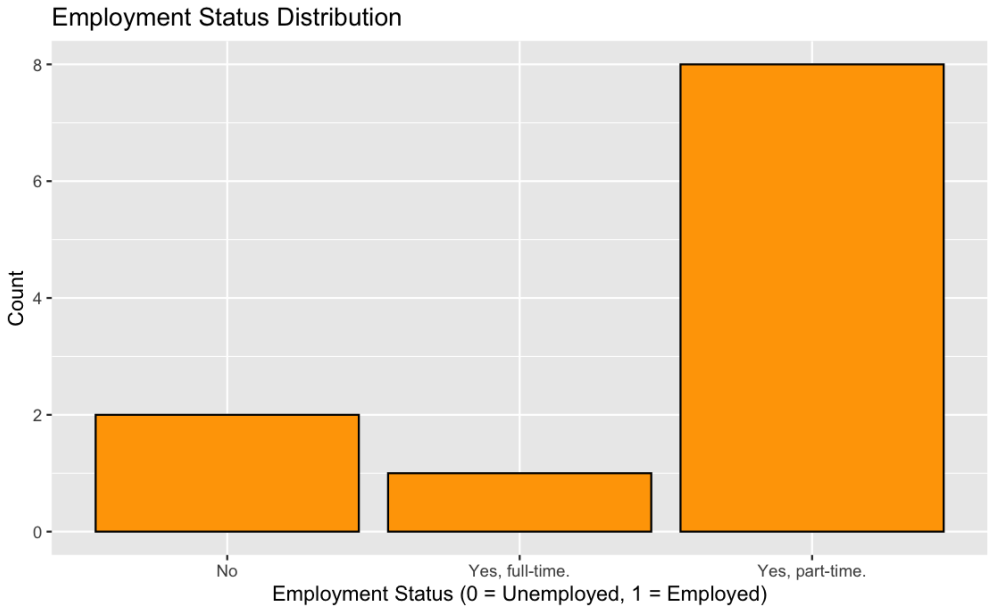

Figure 2: Employment Status Distribution

3.5 Saving for Future Expenses by Employment Status

As shown in Figure 3, employed students are more likely to save for future expenses, indicating a link between employment and financial planning.

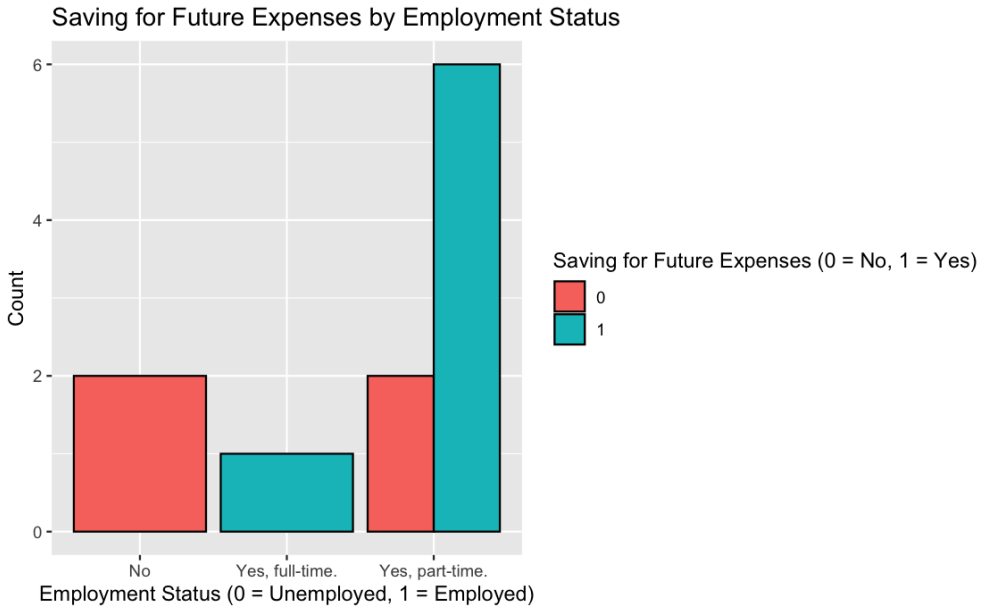

Figure 3: Saving for Future Expenses by Employment Status

3.6 Budget Management Method Distribution

Figure 4 shows the budget management methods used by students, revealing the variety in approaches to budgeting among the respondents.

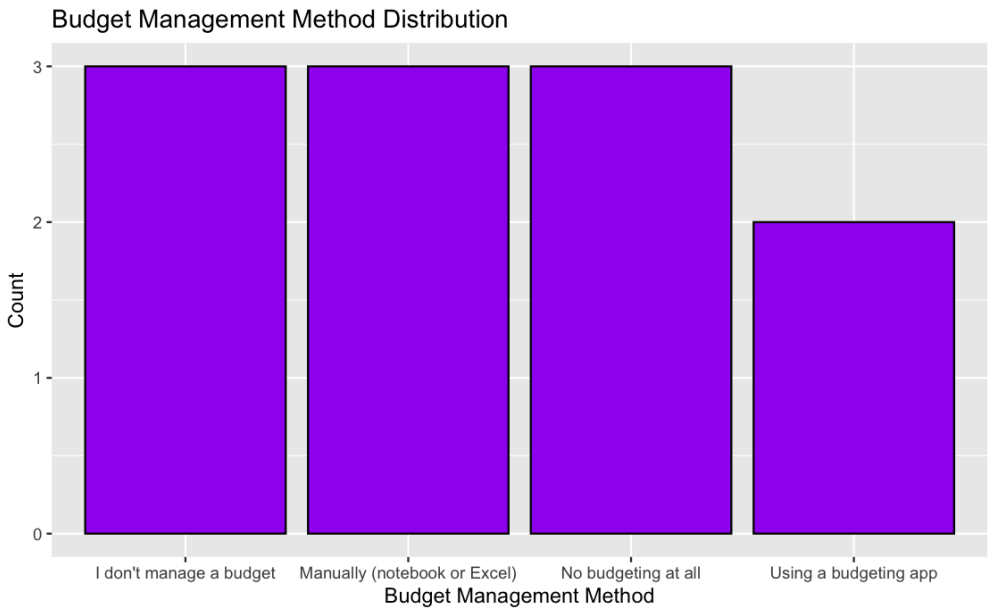

Figure 4: Budget Management Method Distribution

3.7 Correlation Analysis

The correlation heatmap in Figure 5 highlights relationships between financial factors. For instance, student loans show a notable correlation with debt concern.

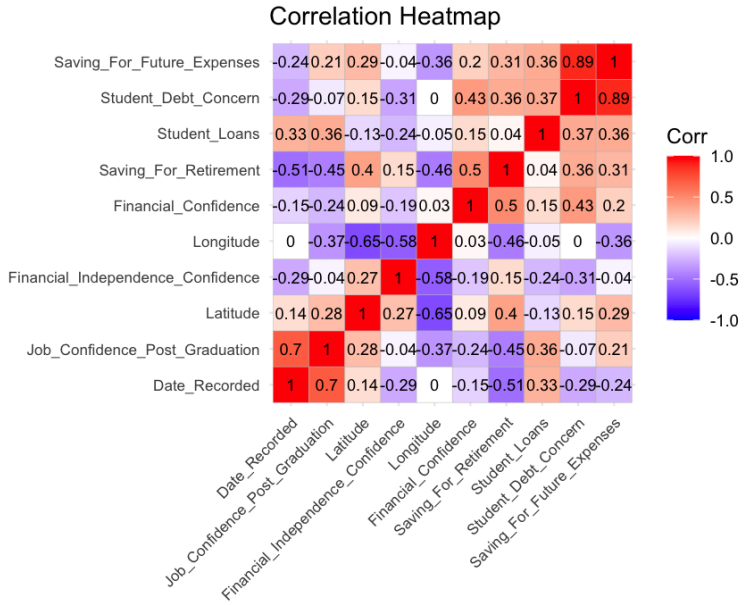

Figure 5: Correlation Heatmap of Financial Factors

3.8 Inferential Analysis

3.8.1 Regression Analysis: Predicting Financial Confidence

A linear regression model was fitted to predict financial confidence based on employment status, student loans, and saving habits. The model's results are summarized in Table 2. Employment status was excluded from the analysis as it showed multicollinearity with other variables, leading to unreliable estimates. However, saving for retirement showed a positive impact on financial confidence.

Table 2: Regression Results for Predicting Financial Confidence

Variable	Estimate	Std. Error	p-value
Intercept	3.286	0.486	0.0003
Employment Status	N/A	N/A	N/A
Student Loans	0.286	0.777	0.724
Saving for Future Expenses	0.000	0.655	1.000
Saving for Retirement	0.857	0.592	0.191

3.9 Clustering Analysis

K-means clustering was applied to group students based on their financial behaviors. Figure 6 shows the results, with students grouped into three clusters based on financial confidence and debt concern. This analysis helps to identify distinct financial behavior types among students.

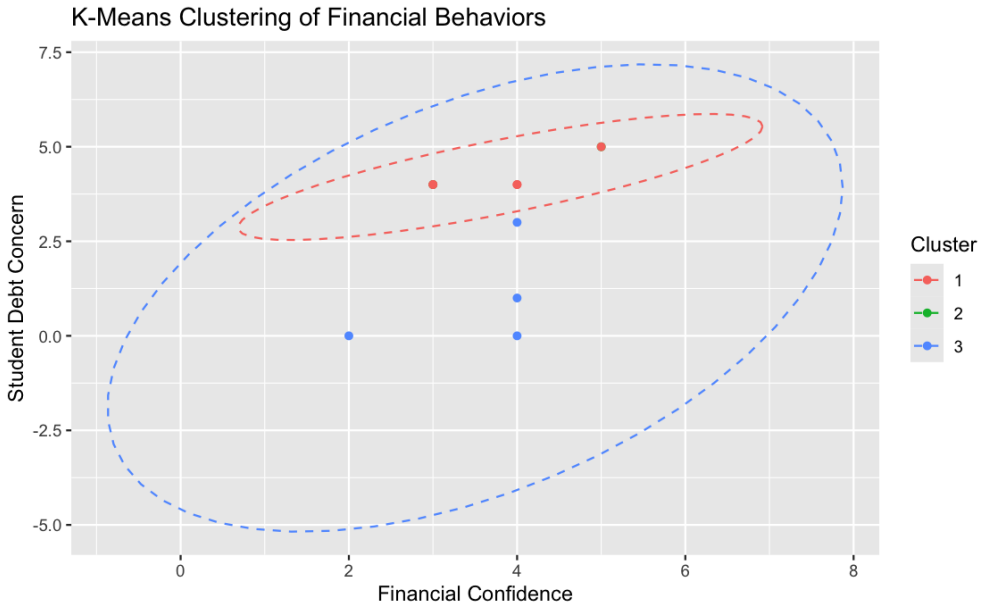

Figure 6: K-Means Clustering of Financial Behaviors

3.10 Word Cloud of Financial Concerns

A word cloud generated from student responses illustrates common financial concerns, with "debt" and "future" being prominent keywords (Figure 7). This indicates that debt is a significant concern impacting students' financial decision-making.

Figure 7: Word Cloud of Financial Concerns

3.11 Student Locations in Bakersfield

To better understand the distribution of the surveyed students, we plotted their locations on a map of California. Figures 8 and 9 show these locations on a broader and more detailed scale, respectively.

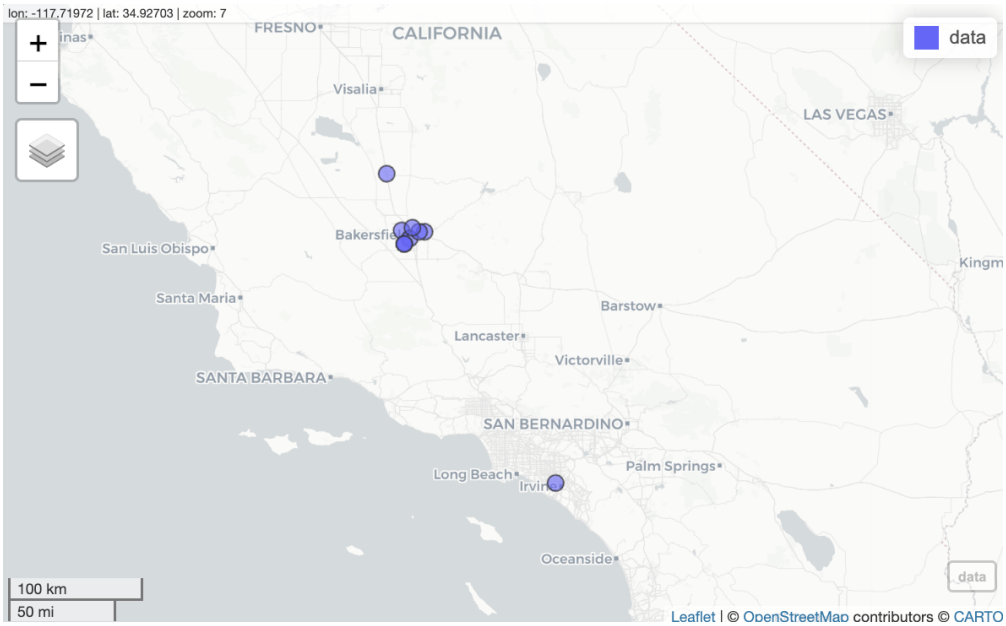

Figure 8: Locations of Surveyed CSUB Students (Zoomed Out View)

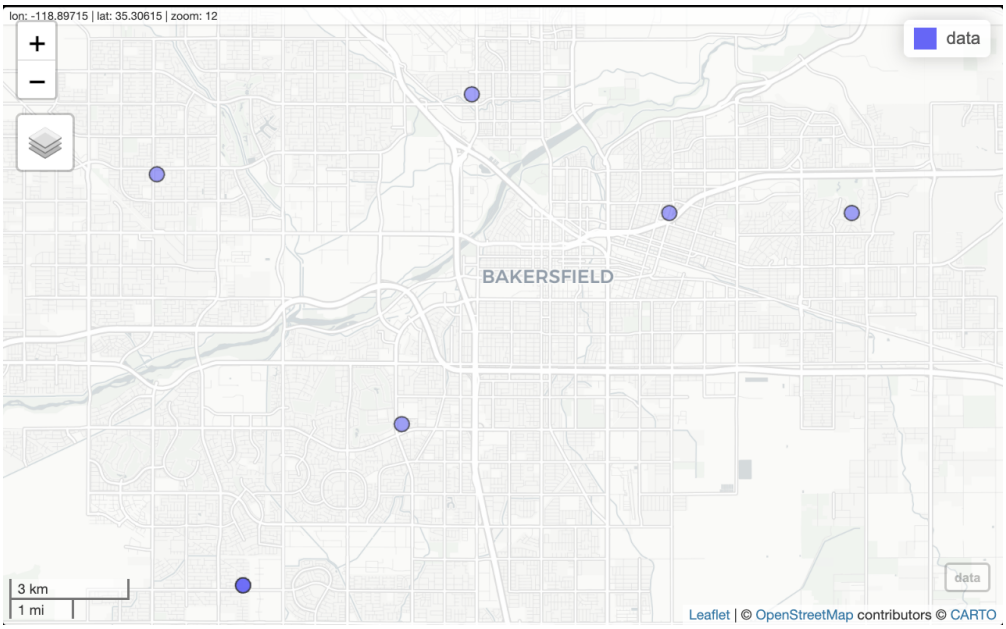

Figure 9: Locations of Surveyed Students in Bakersfield, CA (Zoomed In View)

4 Conclusion

The survey analysis reveals insights into students' financial behaviors, particularly highlighting employment's influence on saving habits and financial confidence. Students who are employed show better financial planning, as evidenced by a higher mean saving rate for future expenses (0.636). Inferential analysis suggests a significant relationship between employment and saving for retirement, although employment status was excluded from the regression model due to multicollinearity concerns. Limitations of this study include the small sample size, potential biases in survey design, and the reliance on self-reported data, which may lead to inaccuracies. Further research could expand on these findings with a larger, more diverse sample and by addressing potential biases through more robust survey design.