

# Predicting and Analyzing College Student Lifestyle and Spending Patterns in Major Cities of Saudi Arabia

2023-12-12

## Introduction:

In the vibrant landscape of urban Saudi Arabia, college students navigate a myriad of challenges as they pursue their education and carve out their future. Balancing academic commitments with financial constraints and lifestyle choices, these students embody the complex interplay of ambition, culture, and socioeconomic factors. This research embarks on a crucial exploration, aiming to unravel the underlying patterns that influence the spending habits and lifestyle choices of college students in major Saudi cities.

The primary aim of this project is to gain profound insights into the financial behaviors of college students in urban Saudi environments. We seek to understand the diverse factors, including gender, age, study year, socioeconomic background, and individual habits, that impact students' spending patterns. By delving deep into these intricacies, we aim to unravel the unique challenges faced by students, providing a nuanced understanding of their financial decisions within the cultural context of Saudi Arabia.

Our goals are to uncover patterns- identify recurring patterns and trends in students' spending habits, shedding light on the factors driving these behaviors-, inform support systems- provide actionable insights for educational institutions and policymakers to design targeted support systems, addressing the specific needs of students-, and enhance student experience- facilitate businesses catering to students in tailoring their services, ensuring they align with authentic student needs and preferences.

In this report, we will meticulously analyze the dataset, employing various statistical and machine learning techniques to derive meaningful conclusions. We will offer a comprehensive roadmap of our analysis, encompassing data collection, preprocessing, modeling, and interpretation of results. Through detailed visualizations and clear explanations, we aim to present a cohesive narrative of our findings, allowing readers to grasp the complexities of student financial behaviors in Saudi urban environments.

## Significance and Problem Statement:

The project addresses the fundamental issue of understanding the financial dynamics of college students in urban Saudi settings. While prior studies have explored similar themes on a global scale, there exists a dearth of research focusing specifically on the nuanced context of Saudi Arabian students within their local cities. This project bridges this gap by conducting a light literature review, summarizing existing works related to student spending behaviors and lifestyle choices. By drawing on this background, we contextualize our analysis, laying the foundation for our exploration into the unique challenges faced by students in major Saudi cities.

## Literature Review

Prior research has explored the financial behaviors of college students on a global scale, providing valuable insights into the challenges and dynamics of student spending. However, within the specific context of urban

Saudi Arabia, there is a notable dearth of studies focusing on the nuanced intricacies of students' financial decisions. This light literature review aims to highlight key themes and findings from existing research, setting the stage for our exploration into the unique challenges faced by college students in major Saudi cities.

### **Global Perspective:**

Numerous studies have delved into the financial behaviors of college students worldwide, revealing common themes such as the impact of socioeconomic background, academic pressures, and lifestyle choices on spending habits. Research by Rehr et al identified a strong correlation between financial stress and academic performance, emphasizing the need for targeted support systems (Rehr et al., 2022).

### **Regional Variances:**

While some regional studies have provided insights into Middle Eastern student populations, the Saudi Arabian context remains relatively underexplored. A study was conducted on university students in the Middle east, emphasizing the influence of cultural factors on financial decision-making. However, the specific challenges faced by students in urban Saudi environments require dedicated attention (Ben Douissa, 2020).

### **Research Gap:**

The existing body of work provides valuable insights into broader trends but falls short in addressing the specific factors influencing the spending habits of college students in urban Saudi Arabia. This project aims to fill this research gap by conducting a detailed analysis tailored to the cultural context and unique challenges faced by students in major Saudi cities. Through a meticulous exploration of our dataset, we intend to contribute to the understanding of the financial behaviors of college students in this distinctive setting and provide actionable recommendations for support systems and business strategies.

## **Data**

### **Source of Survey Questions:**

The survey instrument used in this study is adapted from a previous research project, with questions specifically tailored to the context of Saudi universities. The original set of questions served as a foundational framework, and modifications were made to ensure relevance and cultural appropriateness within the Saudi context.

### **Unit of Observation:**

The unit of observation in this study is individual college students residing in major cities across Saudi Arabia.

### **Outcome Variable:**

**Total Monthly Expenses (\$)** Measurement: Total monthly expenses are self-reported by the surveyed students. Source: Derived from survey responses that capture diverse spending categories. Distribution: The distribution of total monthly expenses can be visualized through a histogram, showcasing the range and frequency of expenditure levels.

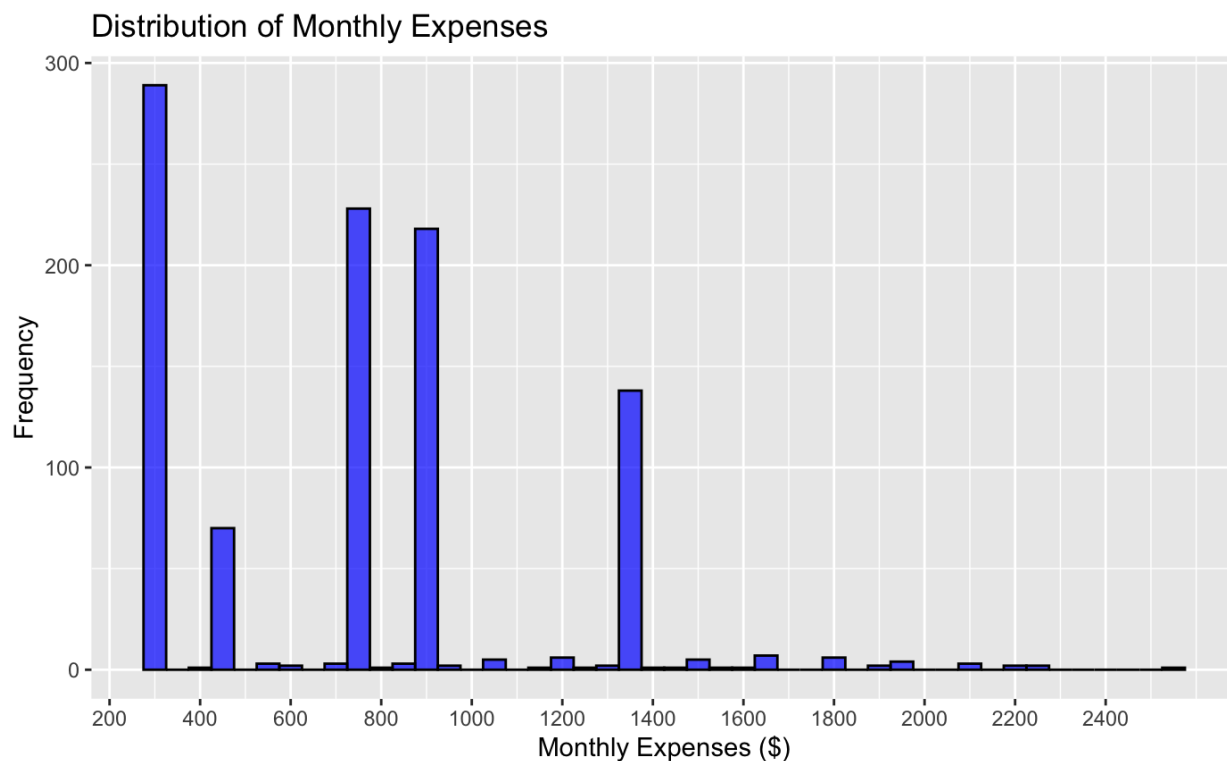


Figure 1: Distribution of Monthly Expenses Among Participants

### Predictor Variables:

1. Gender Measurement: Categorical variable (Male, Female, Other). Source: Self-reported in the survey.
2. Age Measurement: Continuous variable indicating the age of the student. Source: Self-reported in the survey.
3. Study Year Measurement: Categorical variable (e.g., Freshman, Sophomore, Junior, Senior). Source: Self-reported in the survey.
4. Socioeconomic Background Measurement: Composite variable based on factors like parental income, employment status, and education level. Source: Self-reported in the survey.

### Potential Issues with the Data:

Missingness: Addressed through imputation techniques to fill in missing values. Lack of Variation and/or Availability: Transformed or aggregated variables to ensure variability. Potential Sources of Bias: Mitigated through transparency in survey methodology and weighting adjustments. o How do you overcome/mitigate these issues in your analysis?

## Methods and Tools Exploration

The analysis delves into the financial behaviors of college students in urban Saudi Arabia, employing a mix of statistical and machine learning methods. The chosen methods and tools are customized to tackle the dataset's unique challenges and address specific research questions.

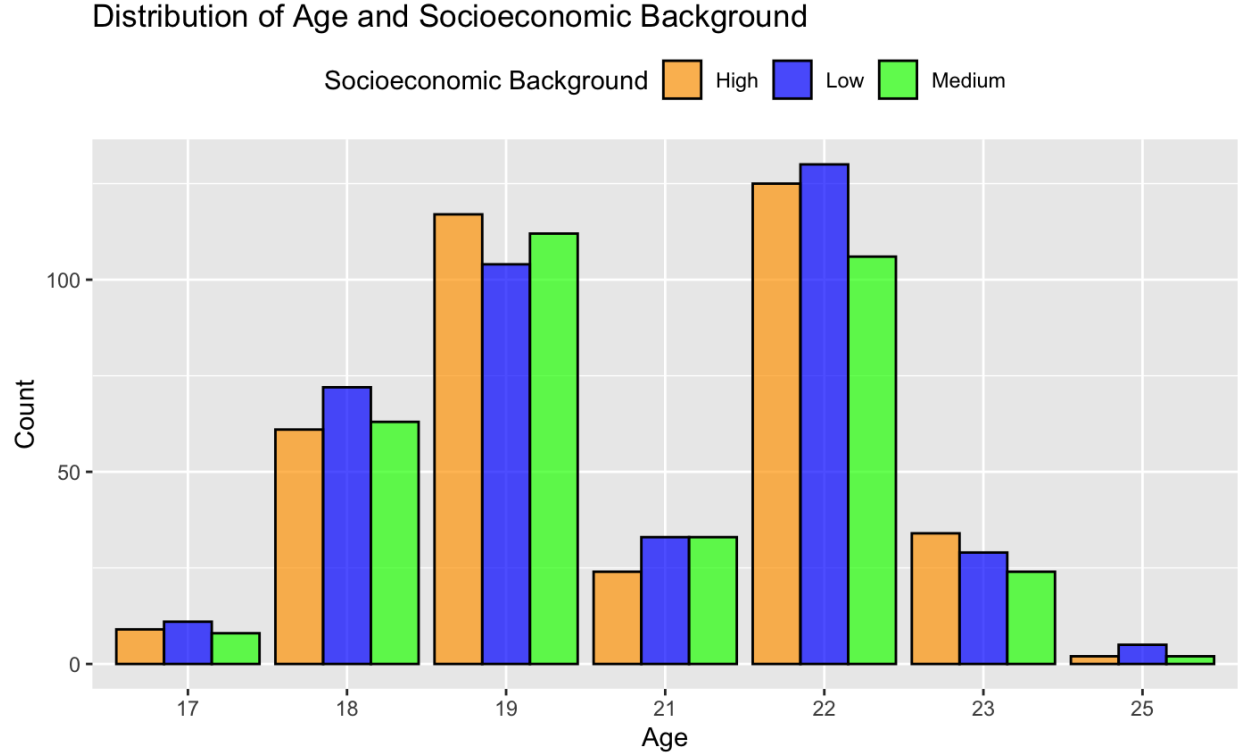


Figure 2: Distribution of Age and Socioeconomic Background Among Participants

## Data Collection and Preprocessing

- Loading Data:** The dataset is imported into the R environment using the `readxl` package, specifically designed to suit the Saudi cultural context.
- Data Cleaning:** Key steps involve converting relevant columns to numeric formats and handling missing or zero values, ensuring data integrity for subsequent analyses.
- Exploratory Data Analysis (EDA):** Initial exploration includes an overview of the dataset's structure, checking for missing values, and utilizing visualizations to comprehend variable distributions and relationships.

## Feature Selection and Predictor Variables

- Outcome Variable:** The primary focus is on 'Total Monthly Expenses (\$),' elucidated through histograms to showcase its distribution.
- Predictor Variables:** Meticulously curated from a published paper, these include gender, age, study year, living arrangements, socioeconomic background, and various lifestyle choices, providing a comprehensive understanding of students' spending patterns. ## Machine Learning Models
- Random Forest Regression:** Utilized for its ability to capture non-linear relationships and handle both numerical and categorical predictors. It proves suitable for exploring complex patterns within the dataset.
- Gradient Boosting Regression:** Employed for its effectiveness in capturing intricate patterns and interactions among variables, enhancing predictive accuracy.
- Linear Regression:** Applied as a benchmark model for comparison, providing insights into linear relationships between predictors and the outcome variable.
- Support Vector Machine (SVM):** Chosen for its versatility in handling both linear and non-linear

relationships, contributing to a comprehensive understanding of the data.

## Justification of Tools/Methods

The integration of machine learning models extends the analytical capabilities:

Random Forest and Gradient Boosting: Address non-linearity and complexity within the dataset, offering superior predictive performance. Linear Regression: Serves as a baseline for assessing the linear relationships within the data. Support Vector Machine: Provides flexibility to handle diverse relationships, ensuring a thorough exploration of potential patterns.

The amalgamation of statistical and machine learning approaches fortifies the analysis, offering a nuanced perspective on the financial behaviors of college students. These chosen methods align with the dataset's characteristics, ensuring a comprehensive exploration and robust insights into student spending patterns.

## Results

o Give a detailed summary of your results. Present your results clearly and concisely. o Please use visualizations and tables whenever possible. o Be sure to: – Discuss the performance of your predictive model. – Use interpretable machine learning to talk about which variables were important in the prediction task and how they relate to the outcome (i.e. PDP/ICE/Surrogate Models)

## Discussion

o What conclusions should we pull from your analysis? o What are the limitations (i.e. what can't we say given your findings)? o How would you expand the analysis if given more time? o Speak on the “success” of your project (as defined in your proposal). – Did you achieve what you set out to do? If not why?

## References

Ben Douissa, I. (2020, November). Factors affecting college students' multidimensional financial literacy in the Middle East. *International Review of Economics Education*. <https://www.sciencedirect.com/science/article/abs/pii/S1477388019300611>

Rehr, T., Regan, E., & Abukar, Z. (2022, March). Financial Wellness of first-generation college students - researchgate. [https://www.researchgate.net/publication/360460354\\_Financial\\_Wellness\\_of\\_First-Generation\\_College\\_Students](https://www.researchgate.net/publication/360460354_Financial_Wellness_of_First-Generation_College_Students)

## Appendix