

ABSTRACT

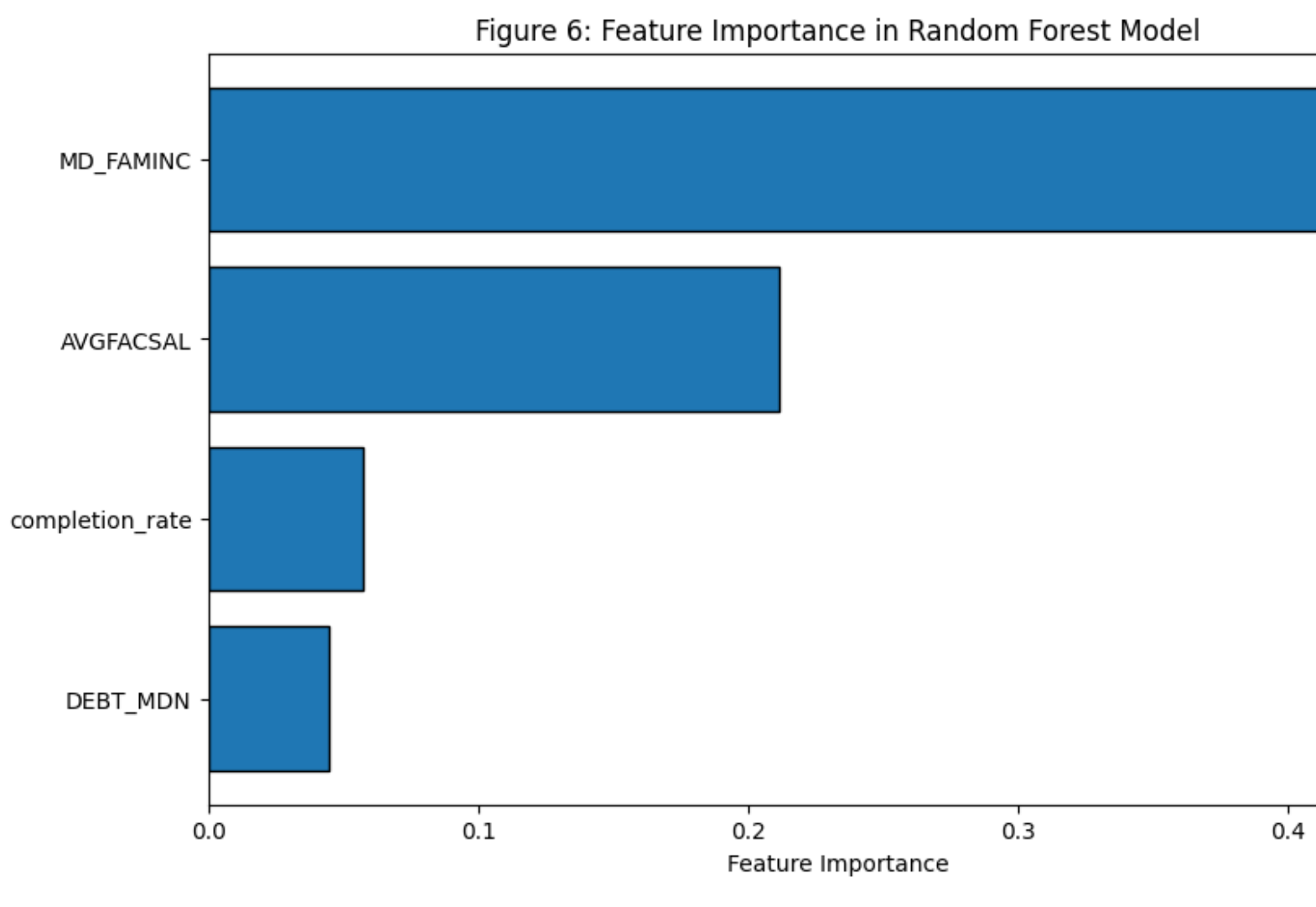
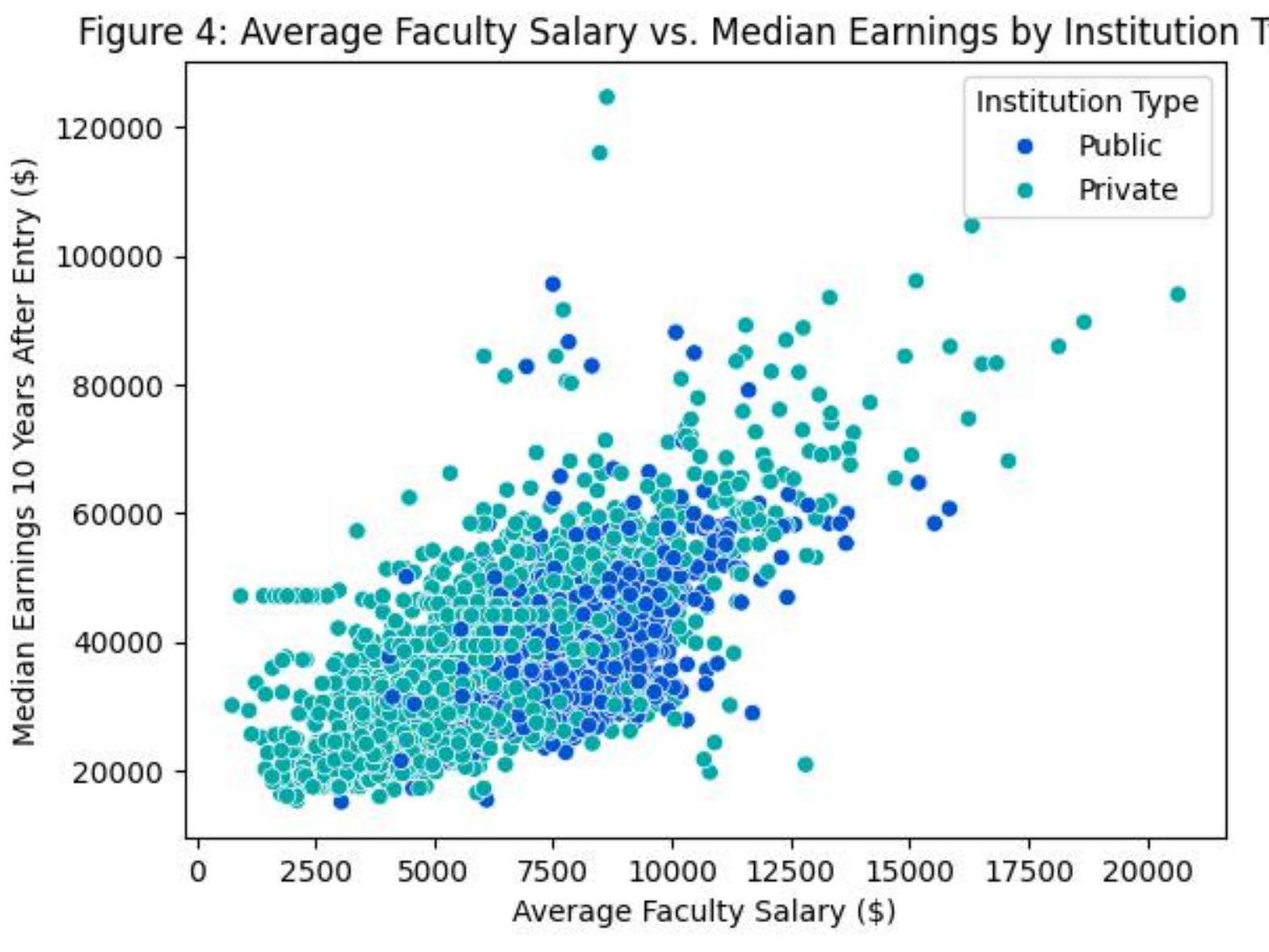
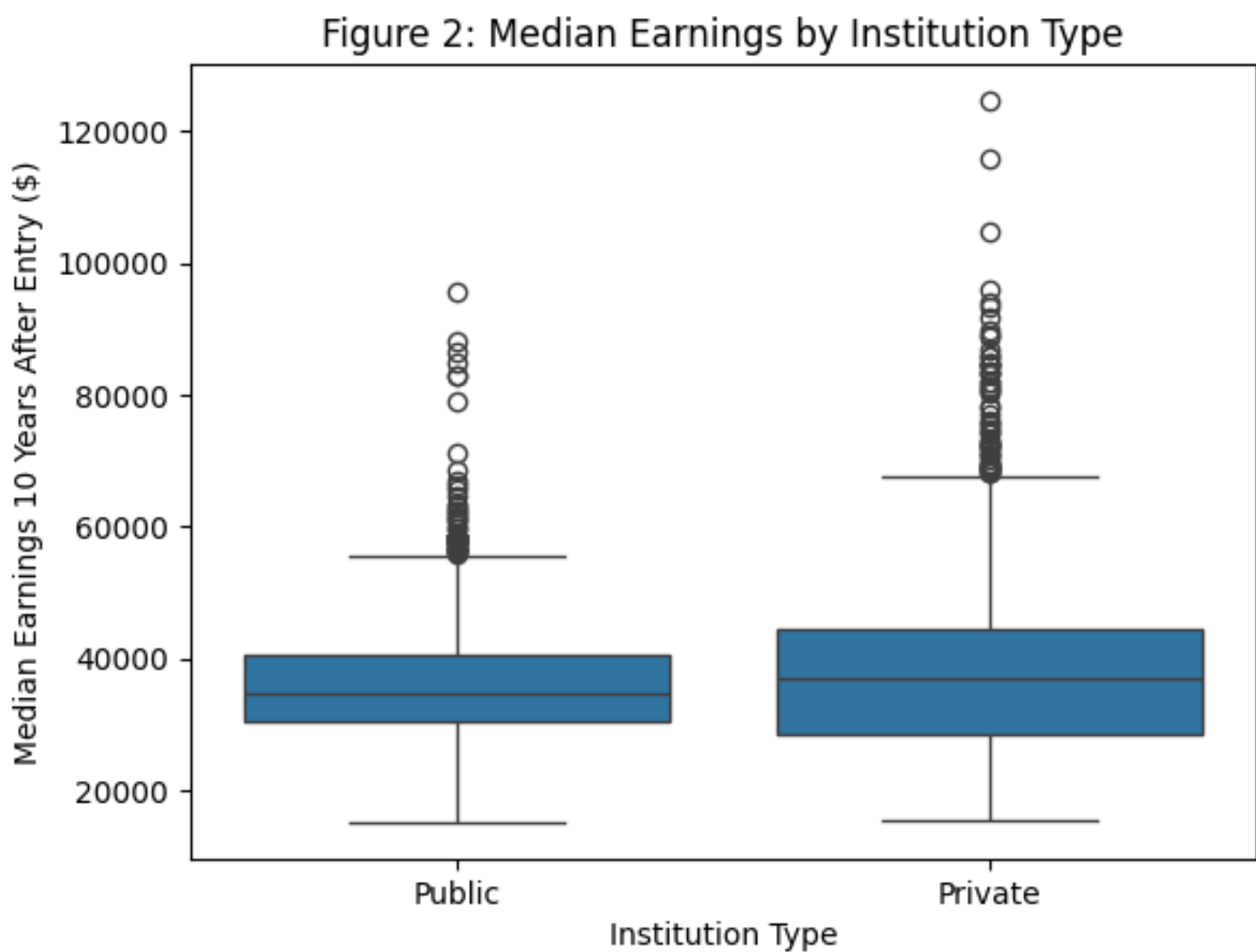
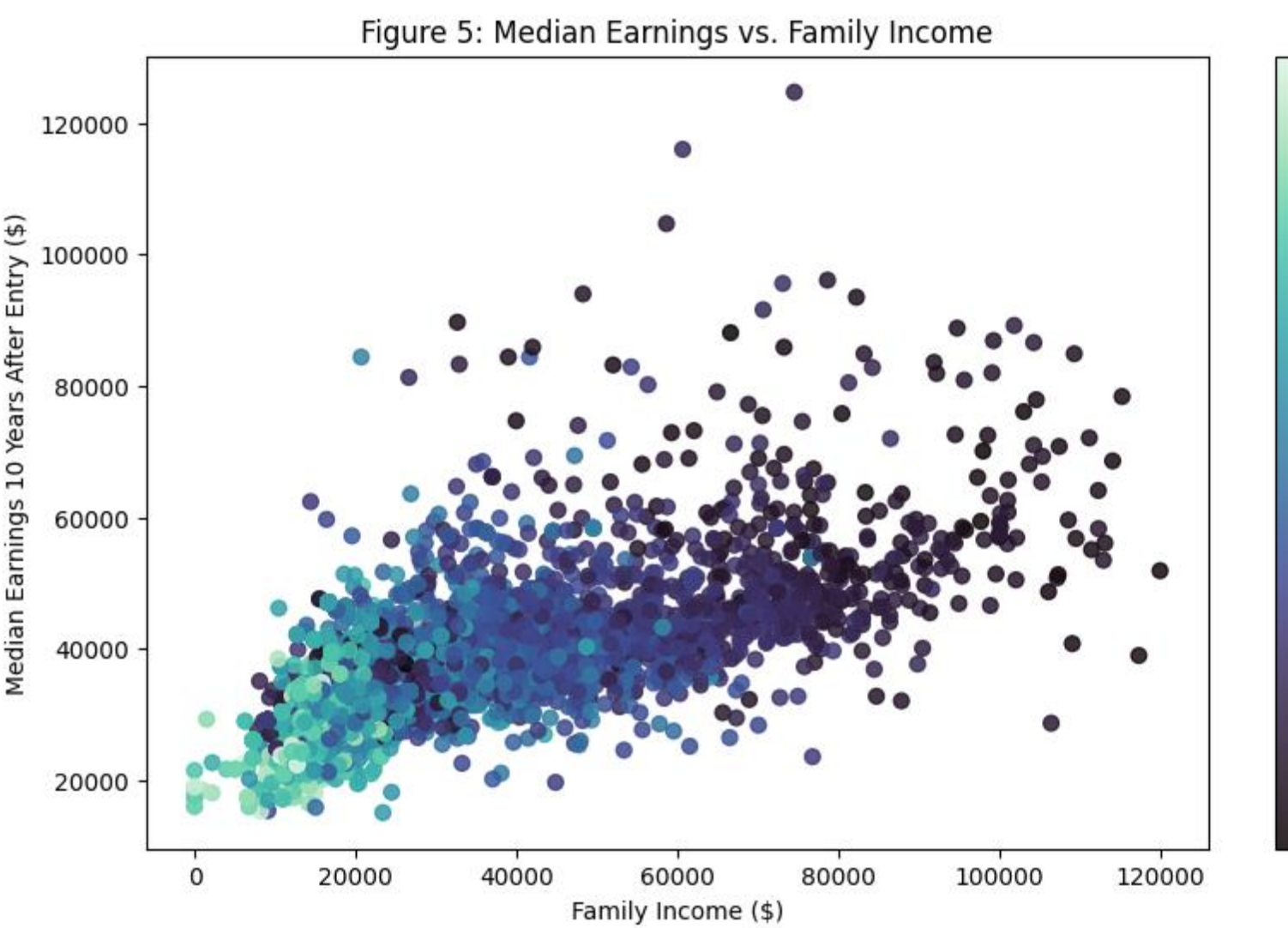
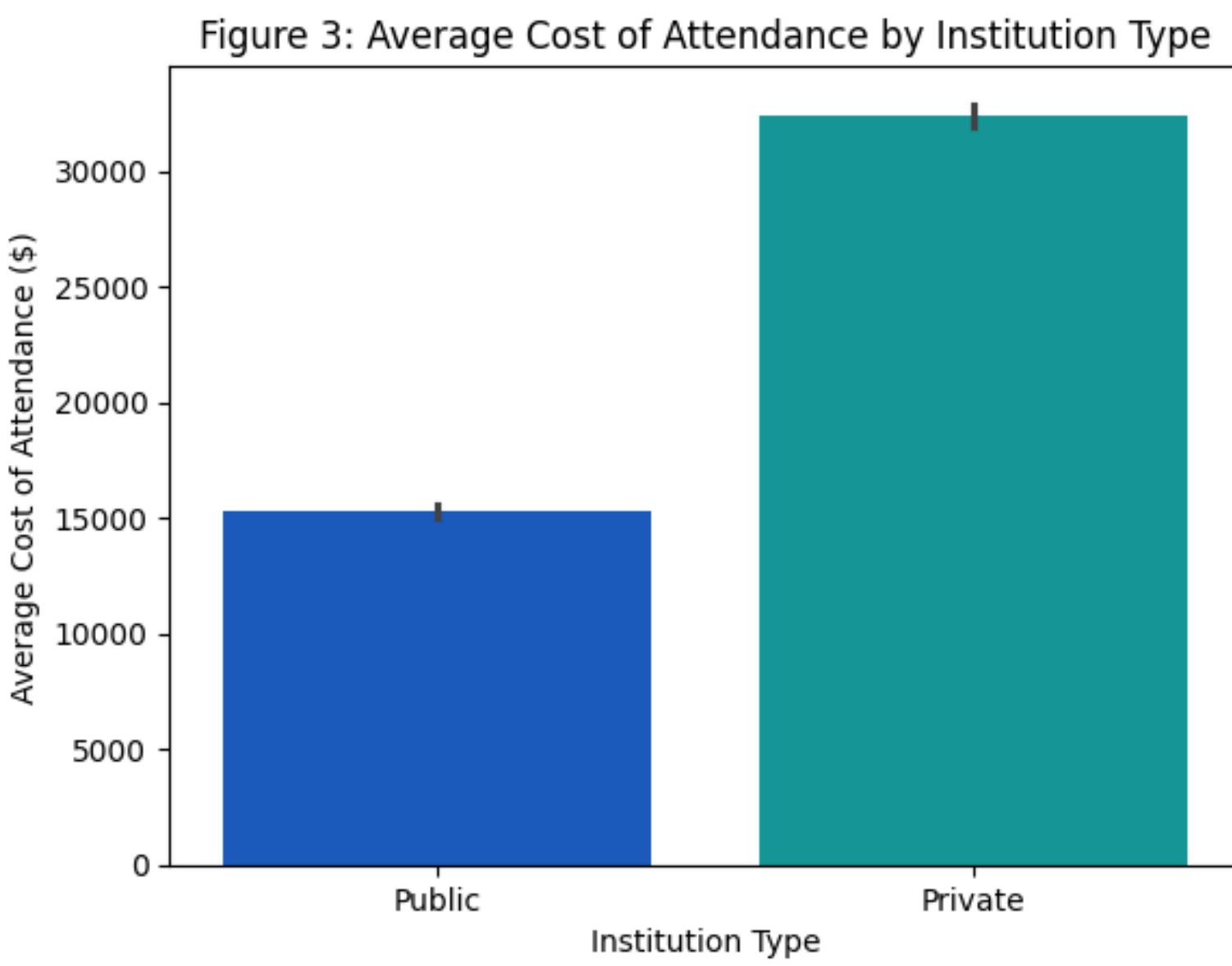
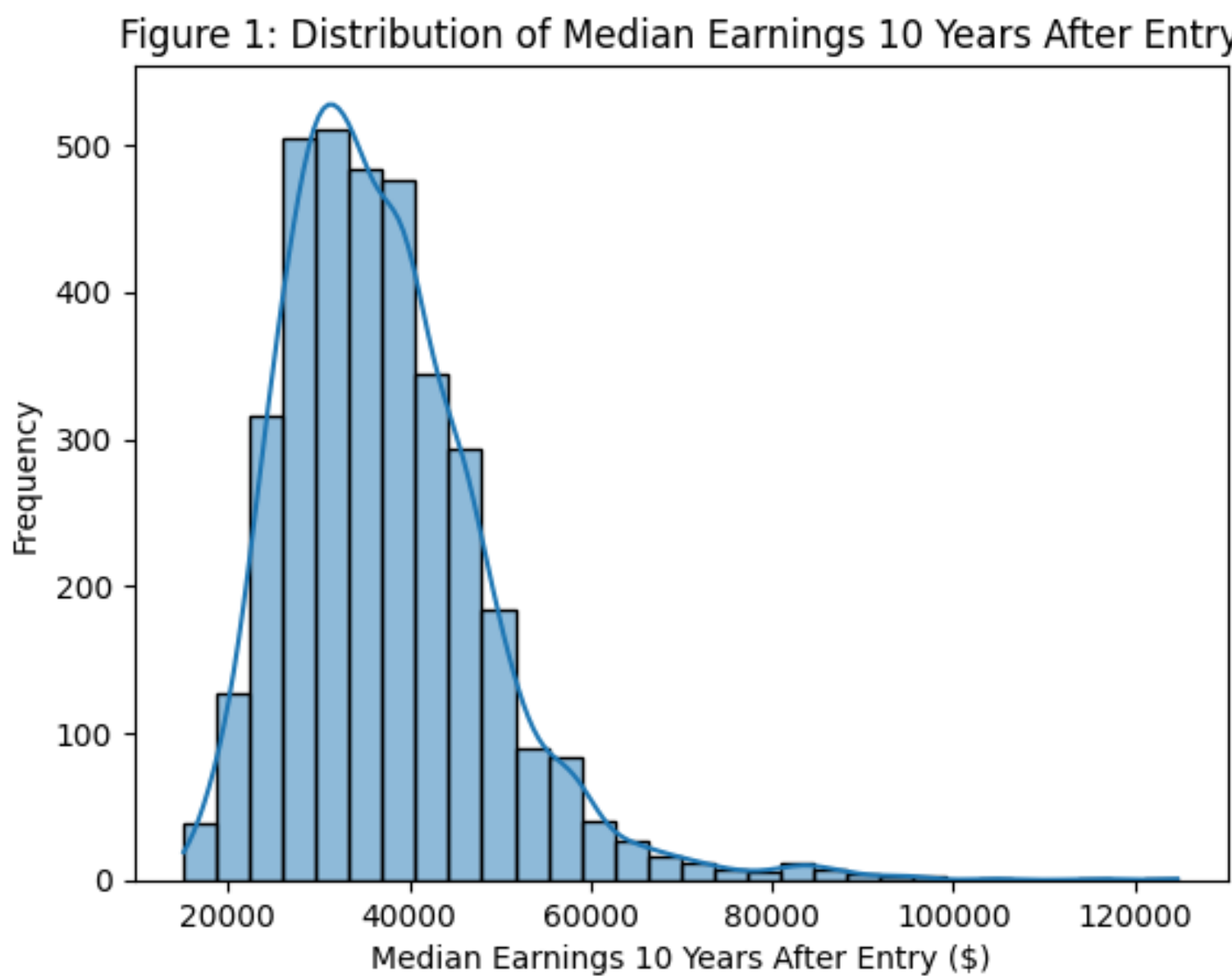
- This study examines how specific institutional characteristics influence the median earnings of college graduates ten years after enrollment.
- Using data from the U.S. Department of Education's College Scorecard, each entry represents an individual institution's attributes and outcomes.
- A Multiple Linear Regression model was used for interpretive insights, while a Random Forest model provided enhanced predictive accuracy.
- Key findings indicate that institutions with higher faculty salaries, higher median family incomes, greater completion rates, and fewer students receiving Pell Grants are associated with graduates earning higher wages. This suggests that socioeconomic background and resources play a significant role in post-graduation financial success.
- These insights aim to guide institutions, policymakers, and prospective students in making data-driven decisions that enhance educational return on investment and address disparities in economic outcomes.

INTRODUCTION

- Rising education costs have increased scrutiny on the return on investment (ROI) of college degrees for students and policymakers alike.
- Median earnings of graduates serve as a critical measure of college value, reflecting the economic mobility higher education can offer.
- This study investigates how select institutional characteristics—such as faculty salary, median family income, and the percentage of Pell Grant recipients—influence graduate earnings ten years after enrollment.
- By analyzing these factors, we aim to identify the institutional attributes most closely linked to economic success, recognizing that students from lower-income backgrounds may face additional financial challenges post-graduation.
- Using a combination of Multiple Linear Regression (MLR) and Random Forest models, this research balances interpretability and predictive accuracy to provide actionable insights for institutions, policymakers, and prospective students.

METHODS

- **Transformations:** Logarithmic transformations were applied to select features in the Multiple Linear Regression model to address skewness; no transformations were applied to the Random Forest model.
- **Plots:** Scatter plots explored relationships between median earnings and various factors. Box plots highlighted differences in median earnings between public and private institutions.
- **Correlation:** A correlation heatmap examined relationships among key variables.
- **Multiple Linear Regression:** Used to assess the significance of individual predictors with log-transformed variables.
- **Random Forest:** Selected for predictive accuracy without transformations, using feature importance to identify key factors.
- **Evaluation:** Models were compared using R-squared and Mean Squared Error to assess accuracy.
- **Software:** Analysis conducted with Python using pandas, matplotlib, seaborn, and scikit-learn.



Random Forest Model Results	
Test Set Root Mean Squared Error	5443.8017
Test Set R-squared	0.7631
Cross-Validation R <sup>2</sup> Scores	[0.7502491 0.75319308 0.70259163 0.79613627 0.52901419]
Mean R <sup>2</sup> Score	0.7062

Multiple Linear Regression Results	
Median Family Income (Logged)	0.11
Percentage of Pell Grant Recipients	-0.29
Average Attendance (Logged)	0.09
Median Debt (Logged)	0.07

RESULTS

Random Forest Model:

- Performance: Achieved an R-squared of 0.7631 on the test set, indicating strong predictive accuracy for estimating graduate median earnings.
- Important Predictors:
- **Median Family Income:** The top predictor, showing that higher family income is strongly associated with higher post-graduation earnings.
  - **Average Faculty Salary:** Second most important feature, suggesting that institutions investing more in faculty salaries tend to have graduates with better earning outcomes.
  - **Completion Rate:** Ranked as an important predictor. This metric reflects support systems that aid student persistence, suggesting that institutions with higher completion rates may better support students through to graduation.
  - **Median Debt:** Slightly negatively associated with earnings, which may indicate that students graduating with higher debt face financial constraints affecting post-graduation success.

Multiple Linear Regression Model:

- Performance: Achieved an R-squared of 0.703, allowing interpretability of individual predictors.
- Key Coefficients:
- **Median Family Income (Log-Transformed):** Significant positive association with earnings, reinforcing that students from higher-income backgrounds have stronger financial outcomes.
  - **Percentage of Pell Grant Recipients:** Negative coefficient, suggesting that institutions with higher proportions of low-income students may face challenges in raising their graduates' earning potential.
  - **Average Attendance (Log-Transformed):** Positive association with earnings, suggesting that larger institutions may provide more resources or opportunities that could contribute to better post-graduation financial outcomes for students.
  - **Median Debt (Log-Transformed):** Positive association in MLR but weaker than other factors, indicating that, while higher debt may correlate with earnings, it's a less reliable predictor.

DISCUSSION

- **Support Low-Income Students:** Colleges with more Pell Grant recipients have lower post-graduation earnings. Institutions can help close this gap by offering career support, financial literacy, and mentorship programs. Low-income students should look for schools with robust support systems.
- **Invest in Faculty:** Higher faculty salaries correlate with better graduate earnings. Institutions should prioritize faculty resources, and students might consider faculty quality when selecting a college.
- **Enhance Retention Programs:** Schools with higher completion rates tend to produce graduates with stronger financial outcomes. Institutions can boost retention through advising and mental health resources, while students should look for colleges with strong support for graduation.
- **Reduce Financial Barriers:** Although debt has a modest impact on earnings, minimizing student debt relieves financial stress. Colleges should expand need-based aid, and students should explore grants and work-study to limit debt.