Influences on Adolescent Well-being: A Multi-Dimensional Analysis

M. Ahmadi[1], D. Barnes[1], S. Khan[1], Z. Xu[1]
[1] Brown University, Data Science Institute, USA

GitHub Link:
https://github.com/GunnerForever/DATA-2020-Final-Project

BROWN Data Science Institute

Introduction

- The Future of Families and Child Wellbeing Study (FFCWS)[1]
  - 7 waves over 22 years
  - Insights into long-term impacts on adolescent wellbeing
- Focus of this project: 6th wave, Year 15
- Idea: Examine adolescent outcomes in 4 sections
  - a) Mental Health   b) Lifestyle
  - c) Education   d) Finances

Methods

Mental Health:
- Logistic regression model with c.v. of k = 5 folds
- Goal: Classify diagnoses of depression and anxiety
- Variables: Empathy, friendliness, communication

Lifestyle:
- Logistic regression model with c.v. of k = 5 folds
- Goal: Identify environmental predictors that influence whether or not focal child has been stopped by police
- Variables: Lifestyle, home, crime, family life

Education:
- Random Forest[3] multiclass classification model
- Goal: Analyze focal child's academic grades for each subject
- Variables: Academic vars, family income, health conditions
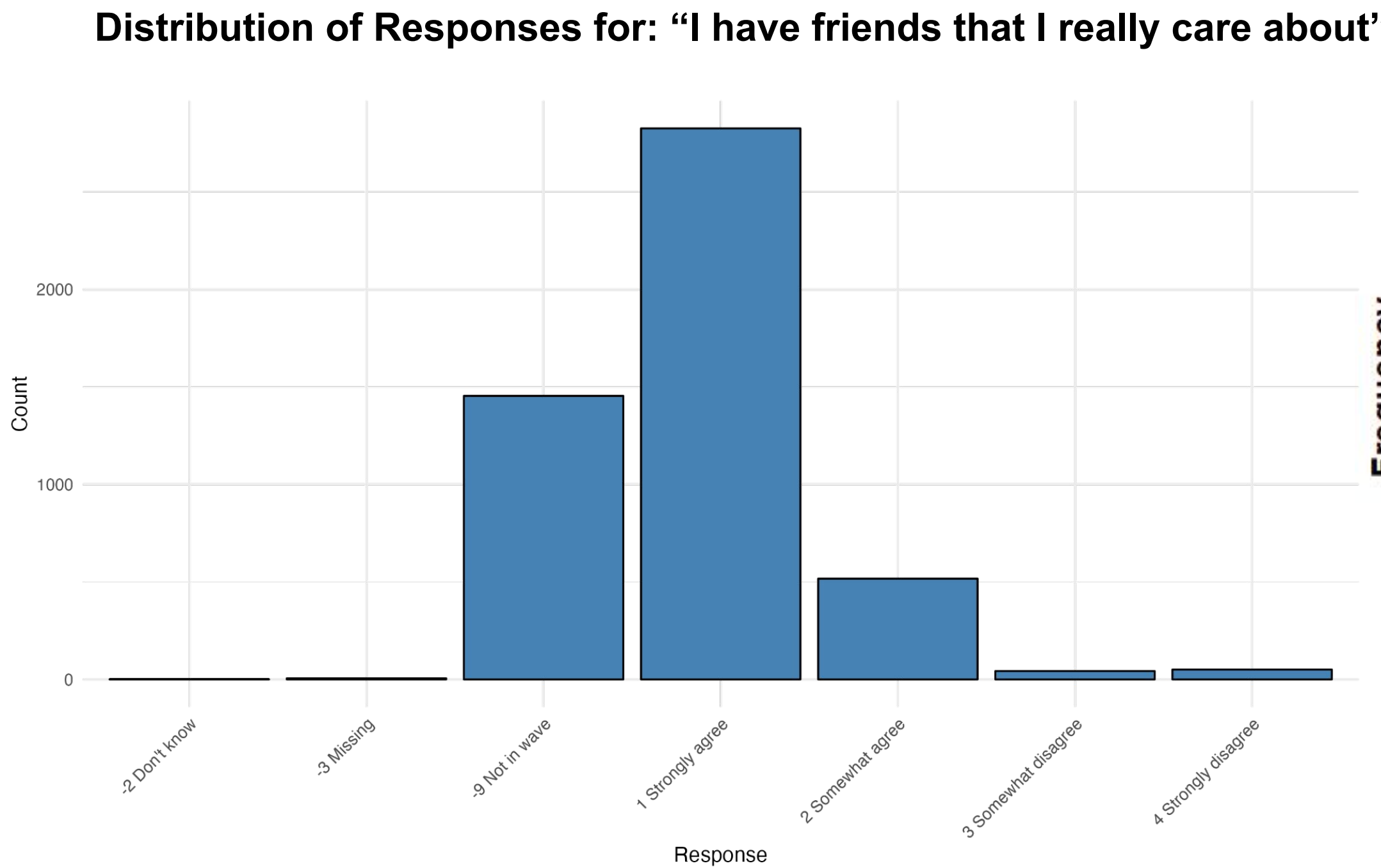
Finances:
- Logistic regression model with c.v. of k = 5 folds
- Goal: Examine the impact of government and social assistance programs on poverty alleviation
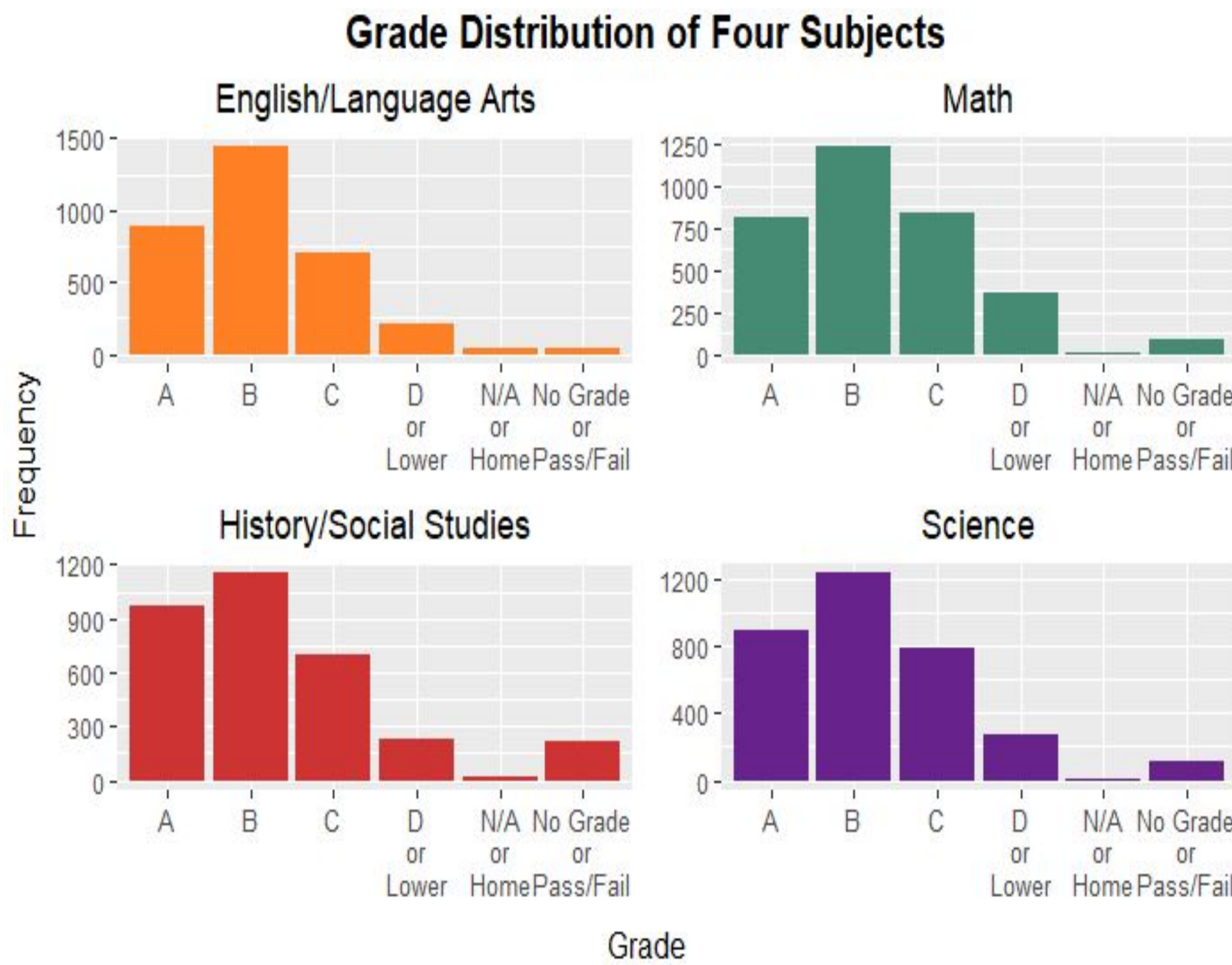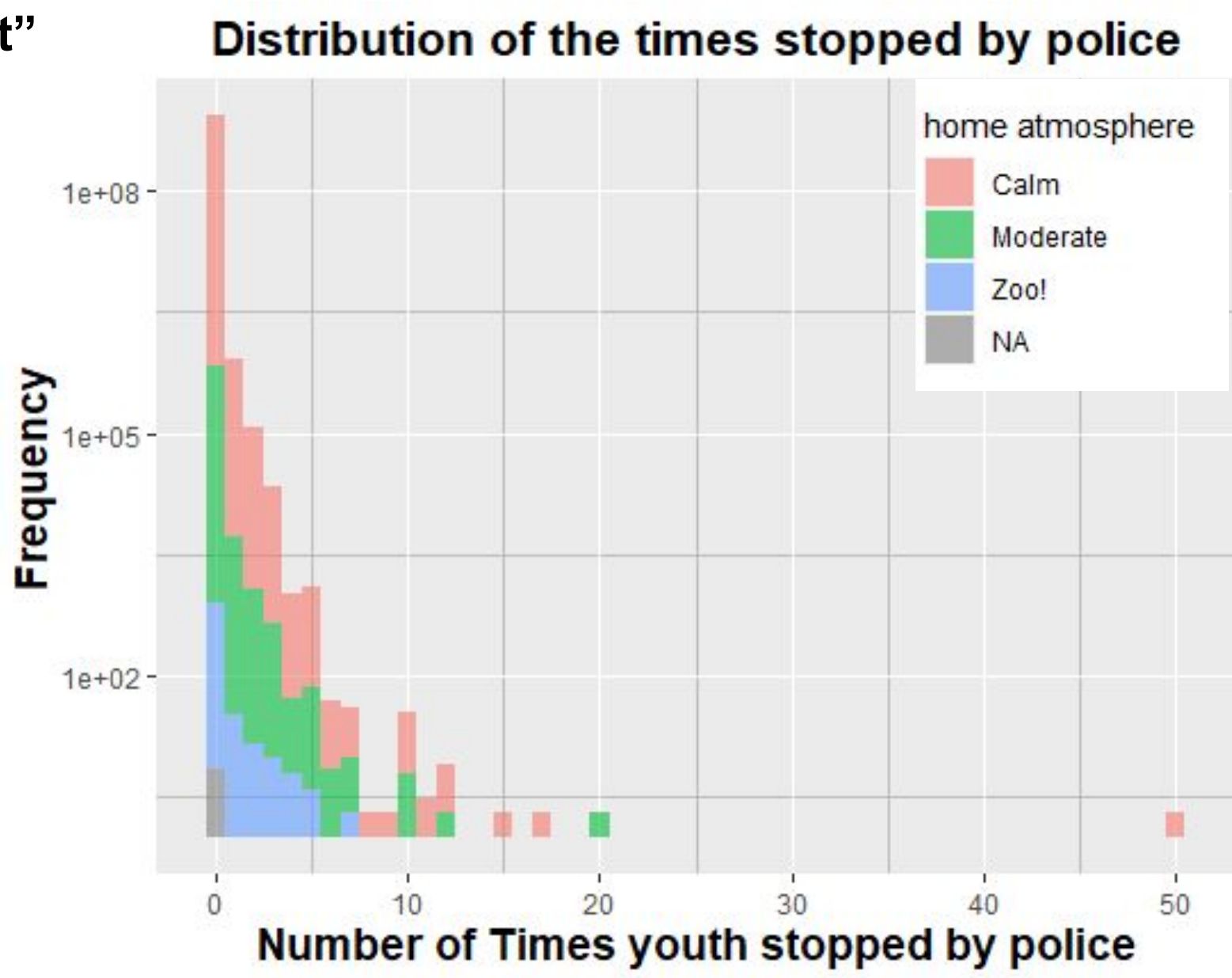- Variables: Family income, poverty status, aid assistance

Exploratory Data Analysis

Mental Health

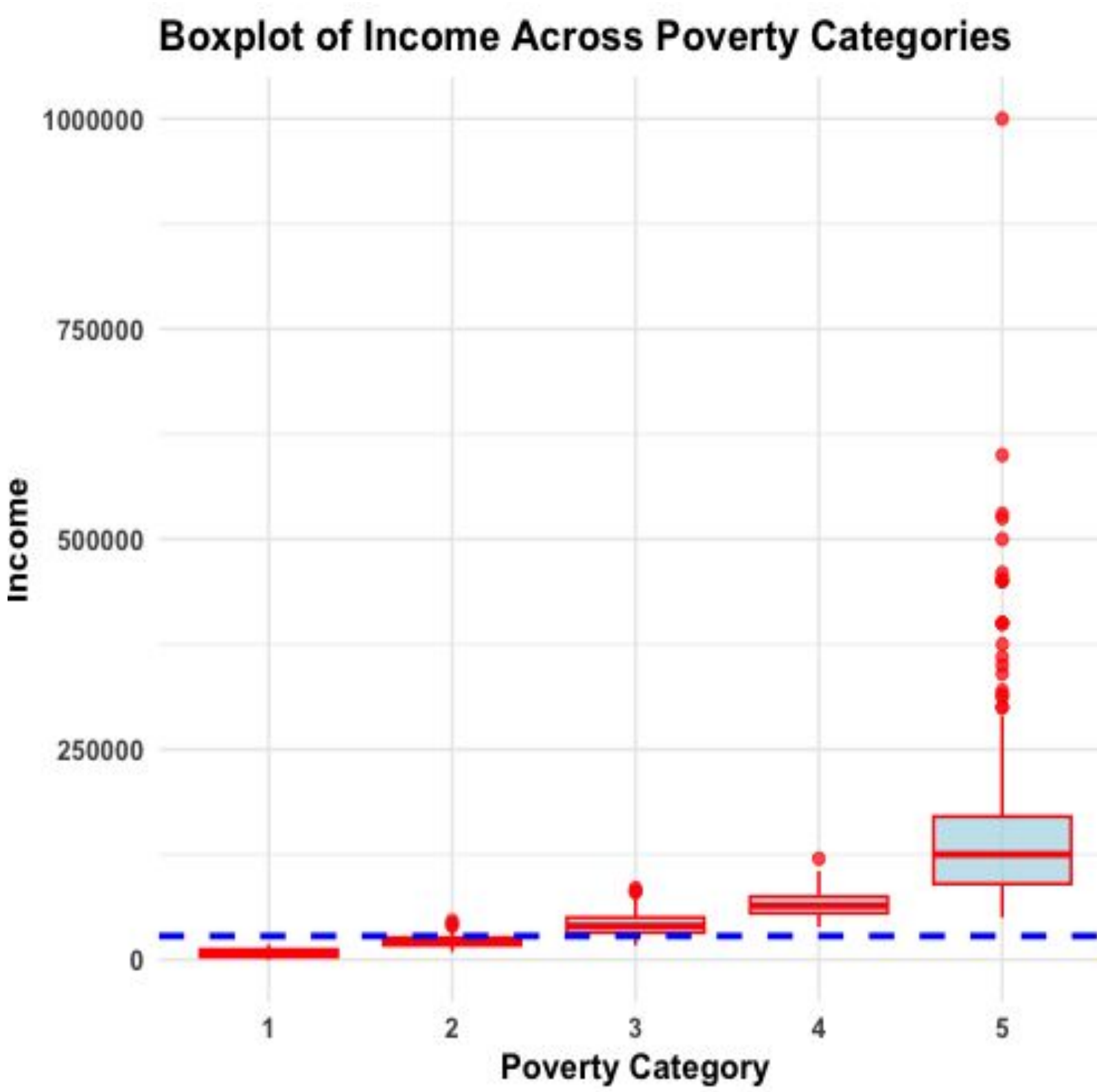Distribution of Responses for: "I have friends that I really care about"

Lifestyle

Distribution of the times stopped by police

Education

Grade Distribution of Four Subjects

Finances

Boxplot of Income Across Poverty Categories

Results

Mental Health

Confusion Matrix Heatmap for k6d1c - Fold 5
Communication Accuracy: 0.5256

Confusion Matrix Heatmap for k6d1a - Fold 5
Empathy Accuracy: 0.6471

Confusion Matrix Heatmap for k6d1e - Fold 5
Friendliness Accuracy: 0.6047

Lifestyle

Expected vs Observed Proportion
Accuracy: 0.915

Influence of F&F Background on Substance Usage by Ethnicity
Effect of Friends / Effect of Family
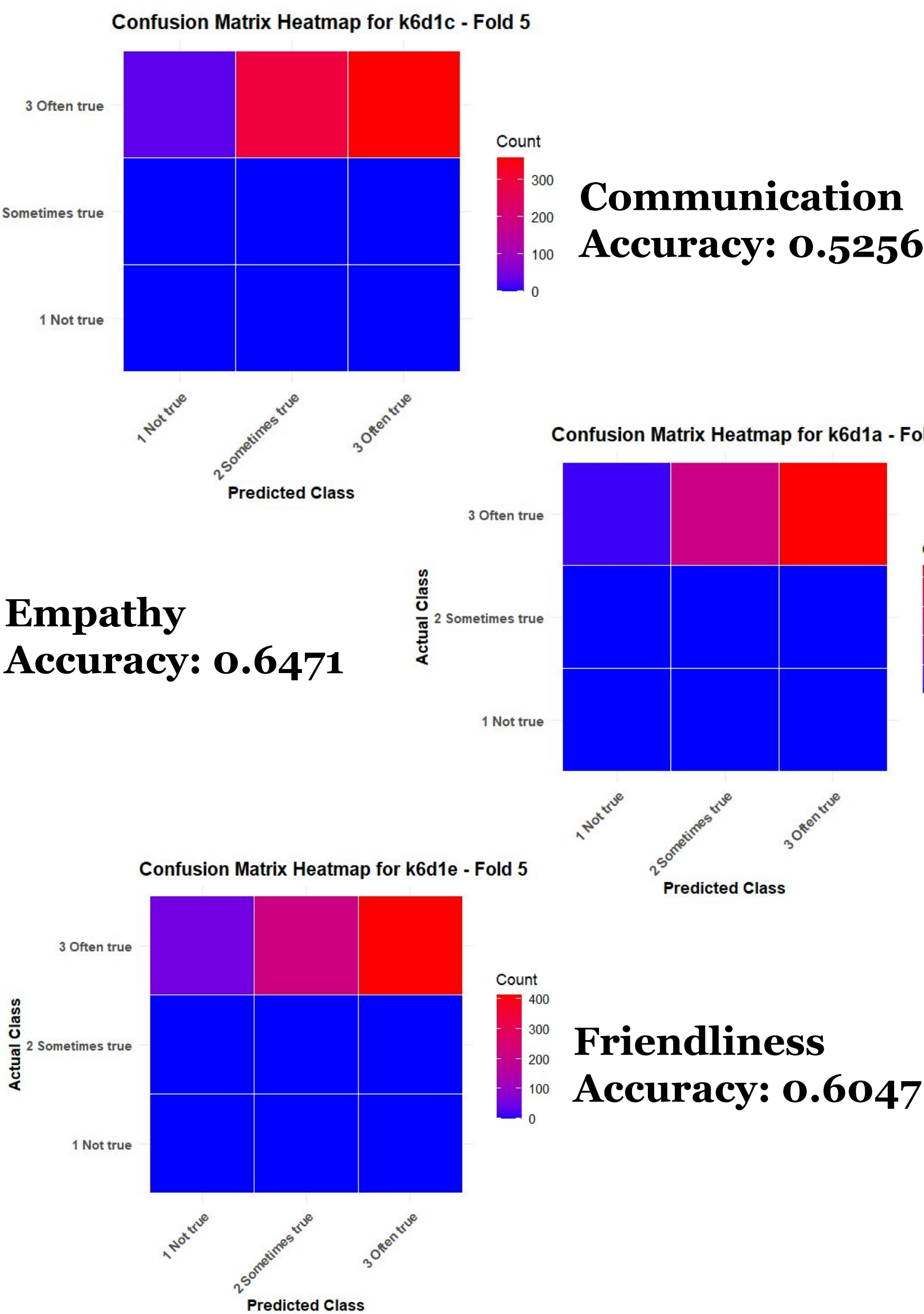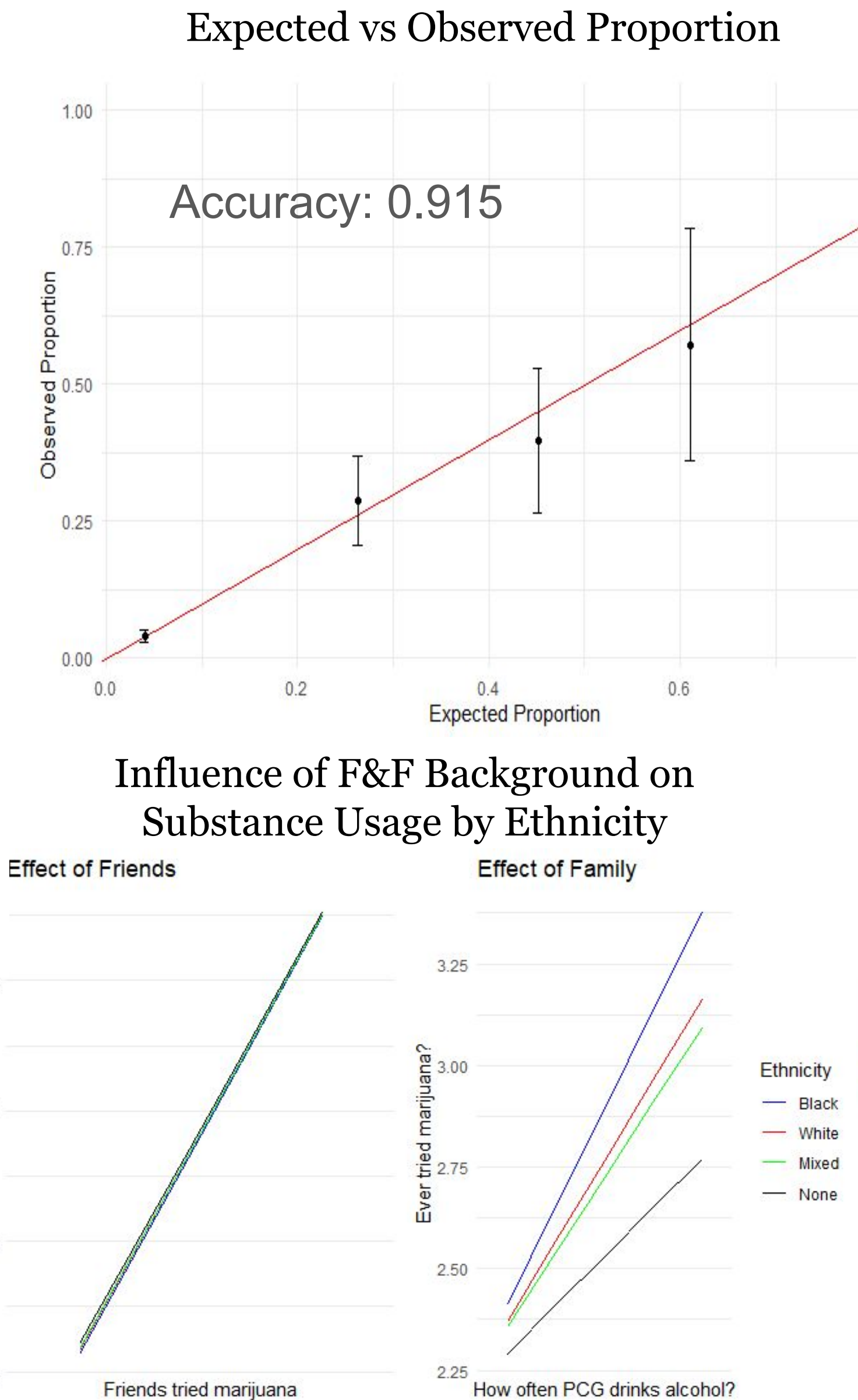
Education

Subject	Best Train Accuracy	Best Test Accuracy
English & Language Arts	0.84356	0.47910
Math	0.85618	0.42176
History & Social Studies	0.67113	0.40881
Science	0.91733	0.43910

Top 10 Most Important Variables
[Target: Science grade]
[Metric: MeanDecreaseAccuracy]

Finances

Coefficients of Government and Social Aid
* Significant P-Value
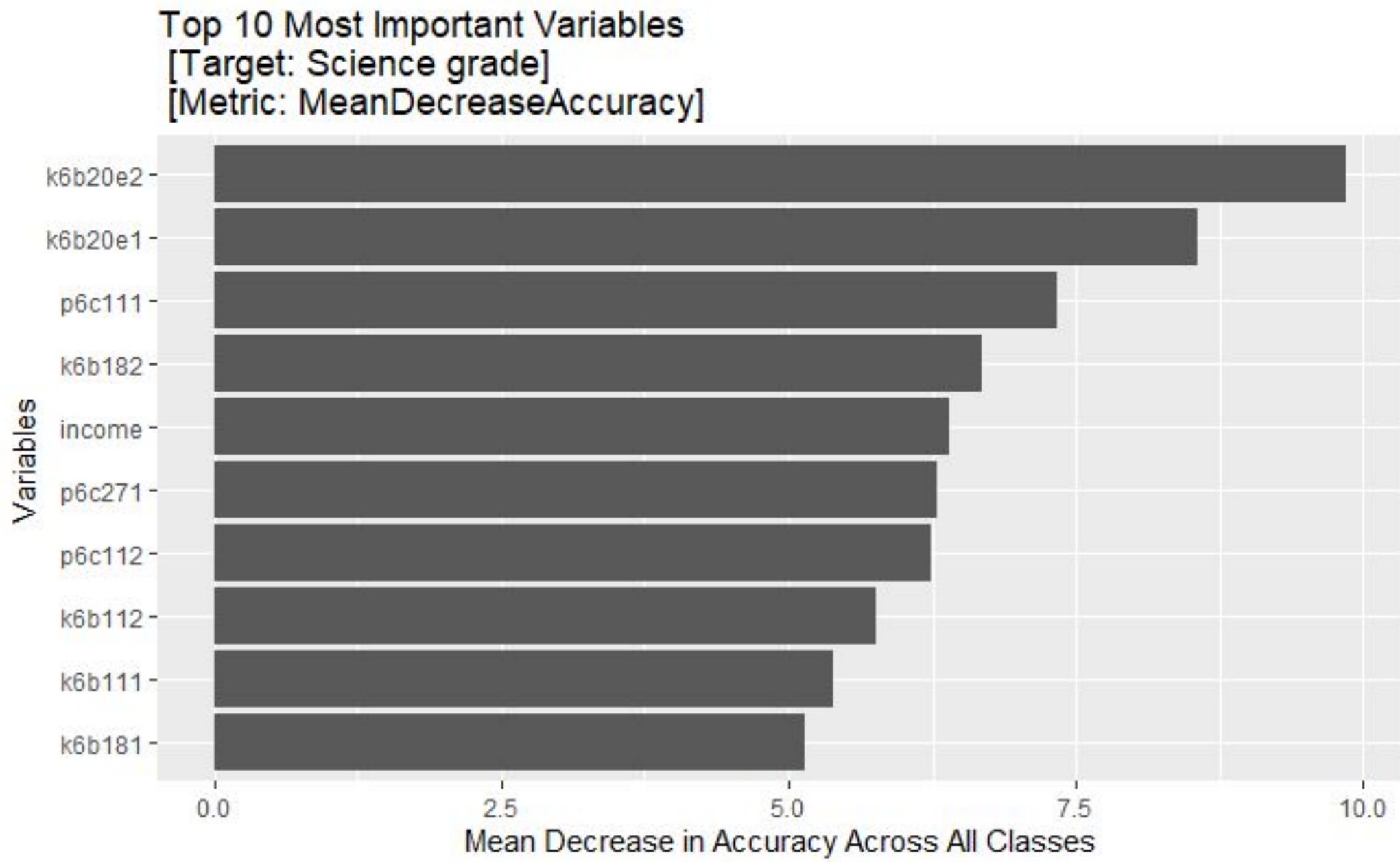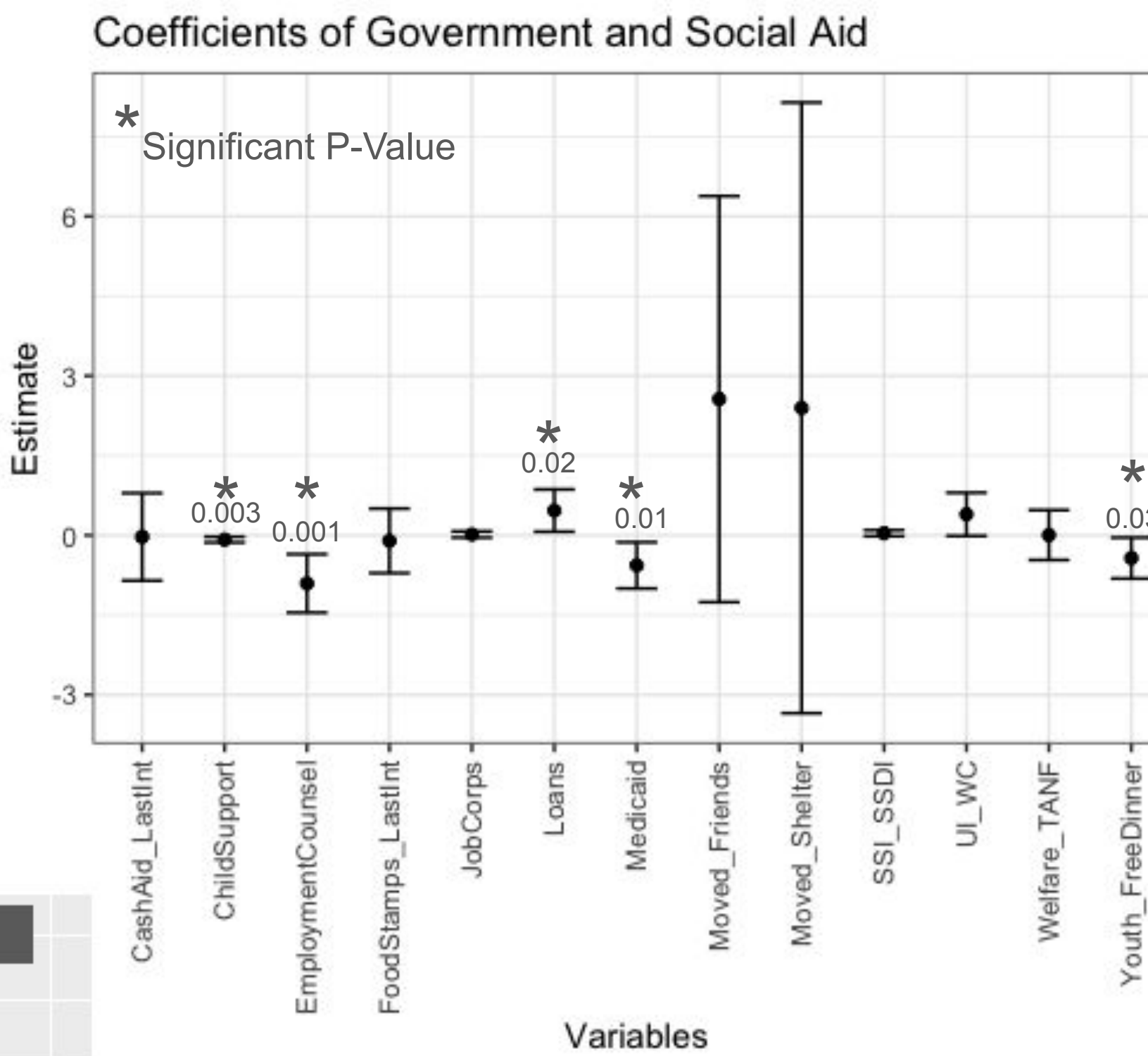
Accuracy	0.820
Precision	0.887
Recall	0.863
Specificity	0.750
F1 Score	0.875

References

1. McLanahan, Sara, Garfinkel, Irwin, Waldfogel, Jane, and Edin, Kathryn. The Future of Families and Child Wellbeing Study (FFCWS), Public Use, United States, 1998-2017. Inter-university Consortium for Political and Social Research [distributor], 2019-09-26. https://doi.org/10.3886/ICPSR31622.v2
2. R Core Team (2024). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
3. A. Liaw and M. Wiener (2002). Classification and Regression by randomForest. R News 2(3), 18--22.
4. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Contributions

M. Ahmadi - Intro, EDA, Methods, Results of lifestyle variables; D. Barnes - Intro, EDA, Methods, Results of financial variables ; S. Khan - Intro, EDA, Methods, Results of mental health variables; Z. Xu - Intro, EDA, Methods, Results of educational variables