DataFest 2022

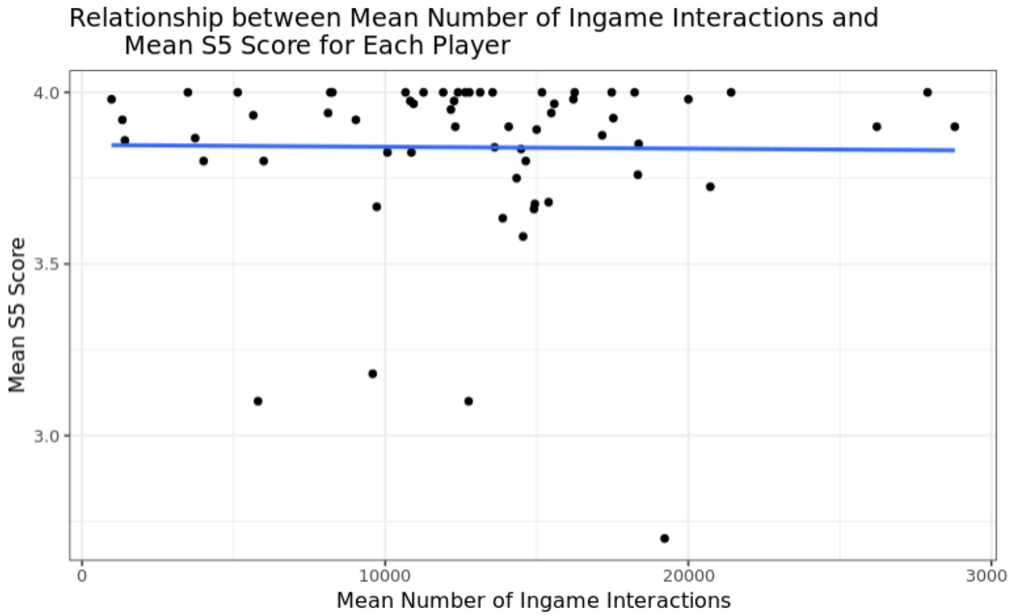
Team FishSwish

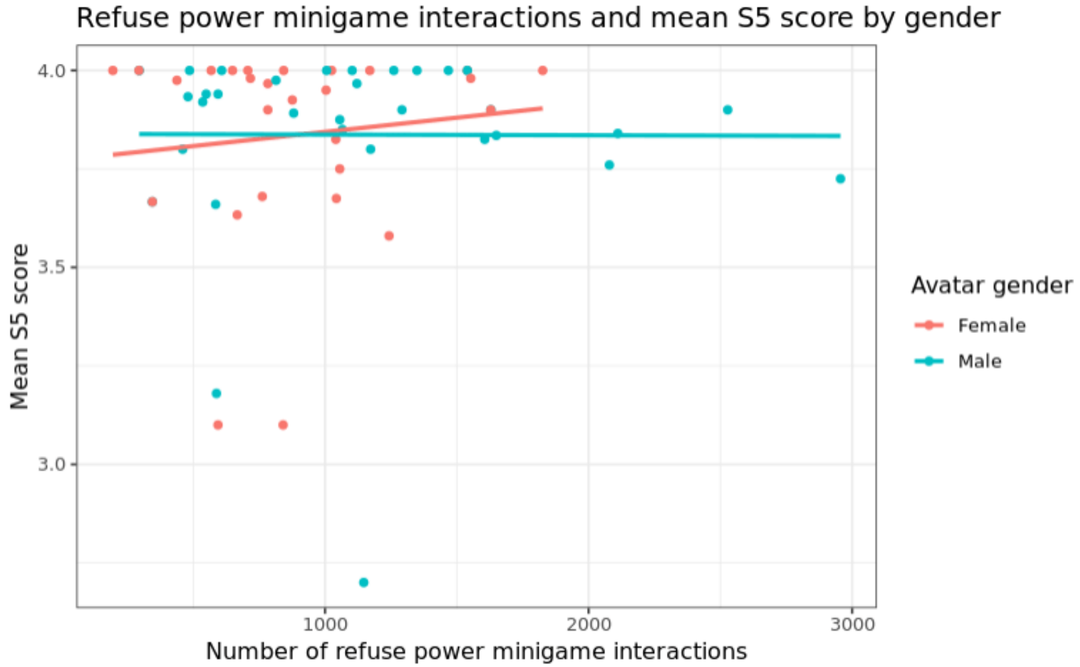
Athena Yao, Emily Du, Jenna Boguslavsky, Helene Gu, and Rebecca Combs

Introduction

Our question:
Can we use *PlayForward: Elm City Stories* game data as a predictive tool to identify at-risk adolescents?

Methodology:
We conducted our statistical analysis using R, which allowed us to perform logistic regression analysis, create plots of the data, and conduct hypothesis testing to identify variables in the logs dataset that could correlate with significant differences in drug use efficacy, as measured by S5 scores.

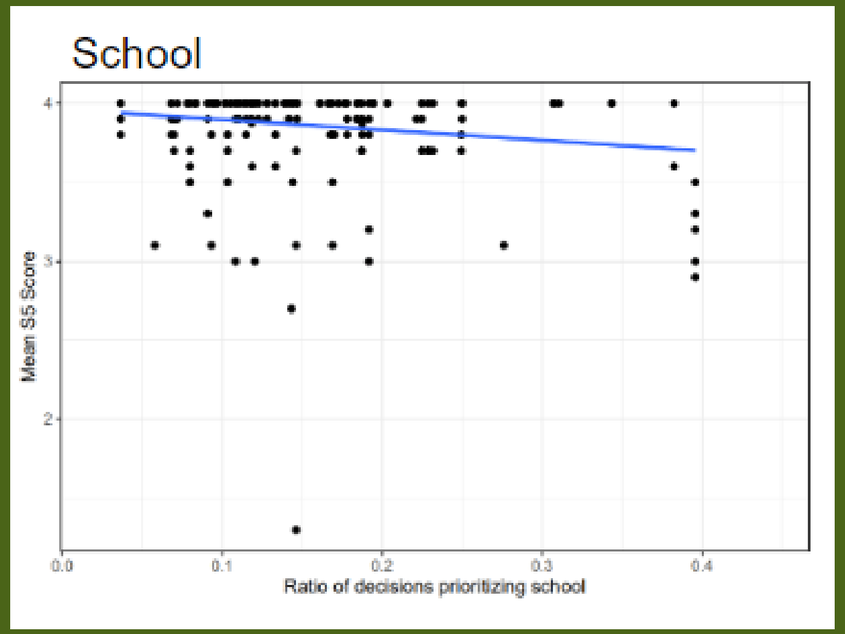


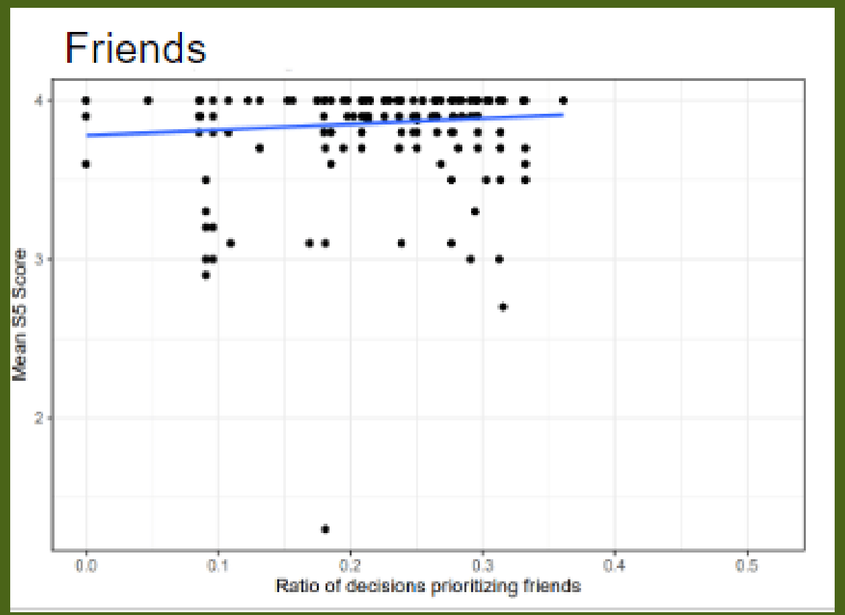


Priority Minigame Trends

We graphed S5 scores vs percentage of times a player prioritized each value and saw significant p-values for data on **school** and **friends**.

Priority	Correlation with S5 score	Slope	P value
School	***Negative***	***- 0.656***	***0.0004***
Health	Negative	- 0.395	0.141
Family	Positive	0.0668	0.817
Money	Positive	0.167	0.498
Happiness	Positive	0.188	0.271
Friends	***Positive***	***0.361***	***0.0238***





Conclusions and Recommendations

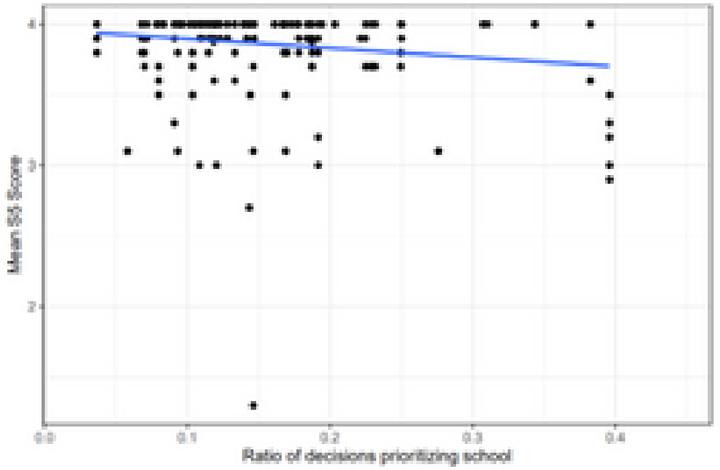
- Interacting with the game more, meaning with more interactions and more choices, correlates slightly with a higher score, indicating a lower likelihood of taking drugs. However, these effects are too small to be statistically significant.
- Analysis of minigame data can yield statistically significant correlations to S5 scores to identify at-risk adolescents.

Limitations

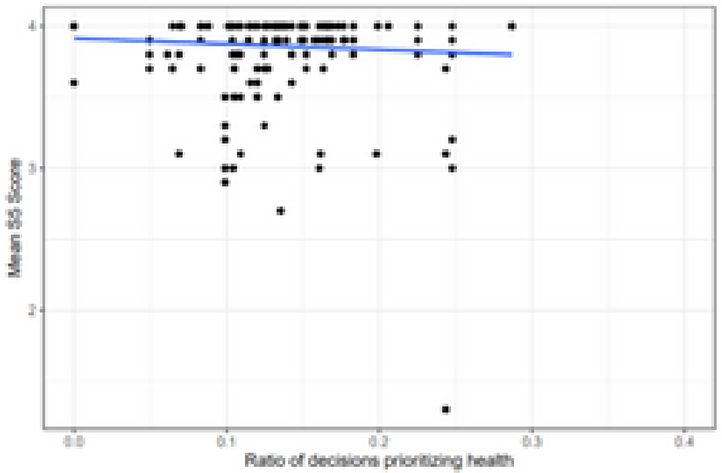
- Our analysis relied on S5 means, which were derived from self-reported data and showed little variability
- Gameplay decisions may vary from real-life decisions
- S5 scores of each player might not be true to their real-life actions
- Optimism bias, curiosity, peer pressure, etc.

Appendix

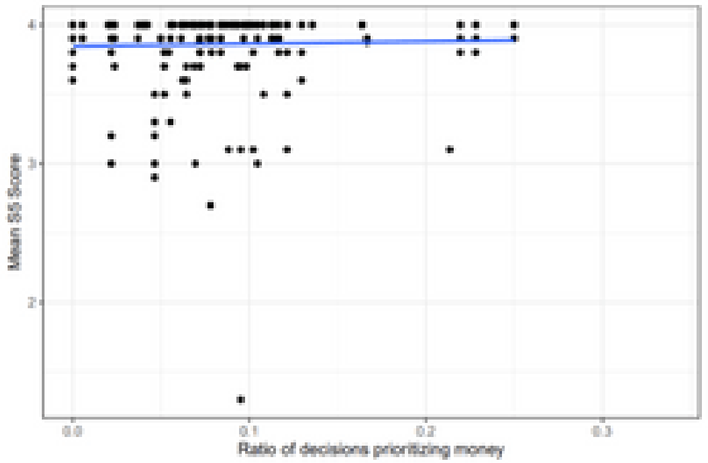
School



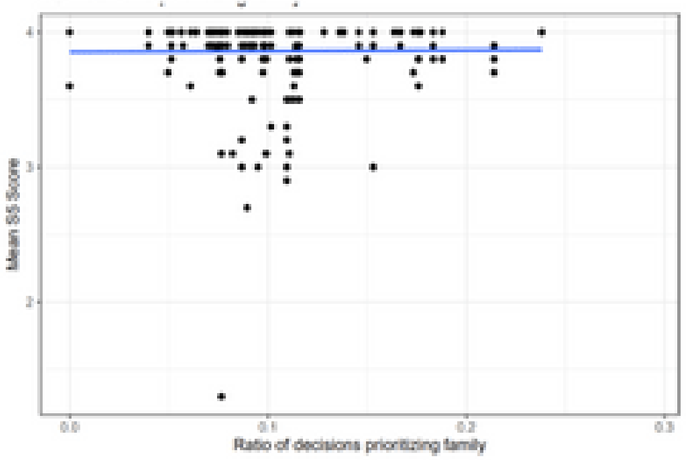
Health



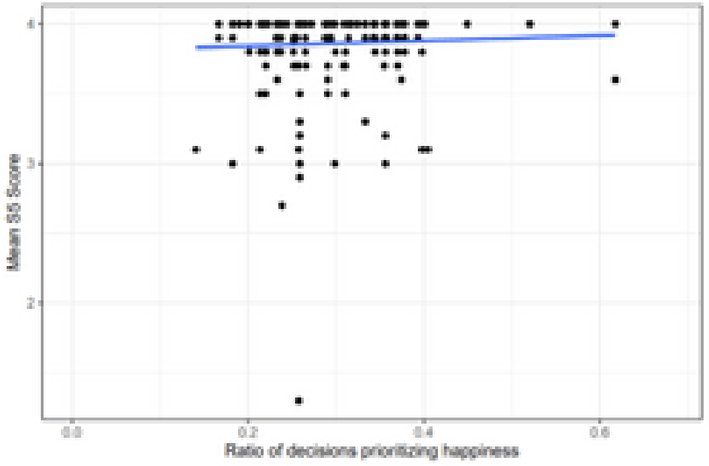
Money



Family



Happiness



Friends