

# STUDENT PERFORMANCE ANALYSIS



Mitchell Brogan

# INTRODUCTION

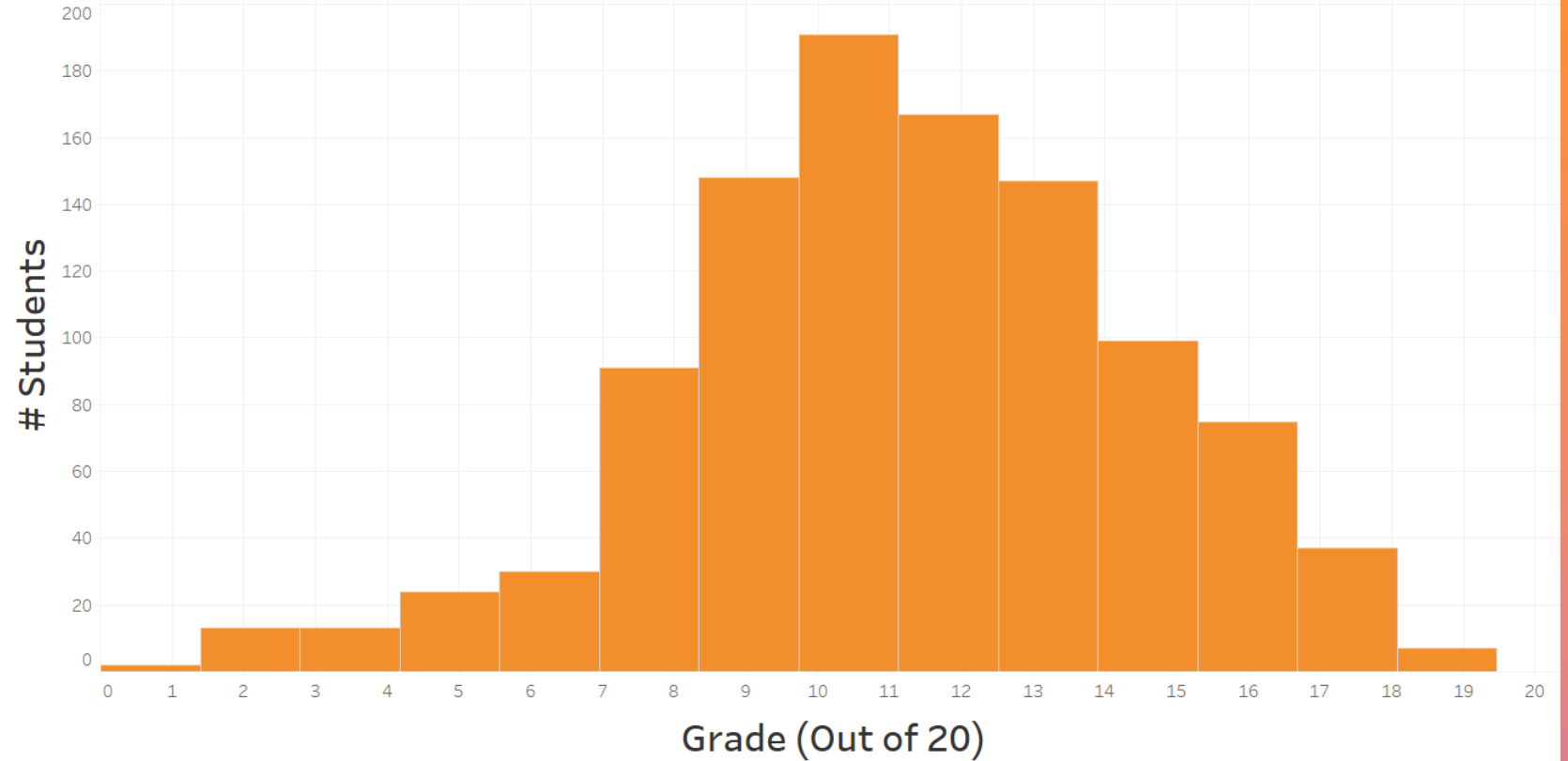
- **Motivation:** Identify risk factors correlating to low grades for high school students
- **Objective:** Perform an EDA to determine high risk factors, rank them, and calculate a weighted total risk metric
- **Result:** Determined factors that have a high impact on student success

# METHODOLOGY

- Data:
  - 1044 students
  - 33 features
  - <https://archive.ics.uci.edu/ml/datasets/student+performance>
- Metrics:
  - Personal habits
  - family environment
- Tools:
  - Pandas, numpy, matplotlib, Seaborn, statsmodels, Excel, Tableau

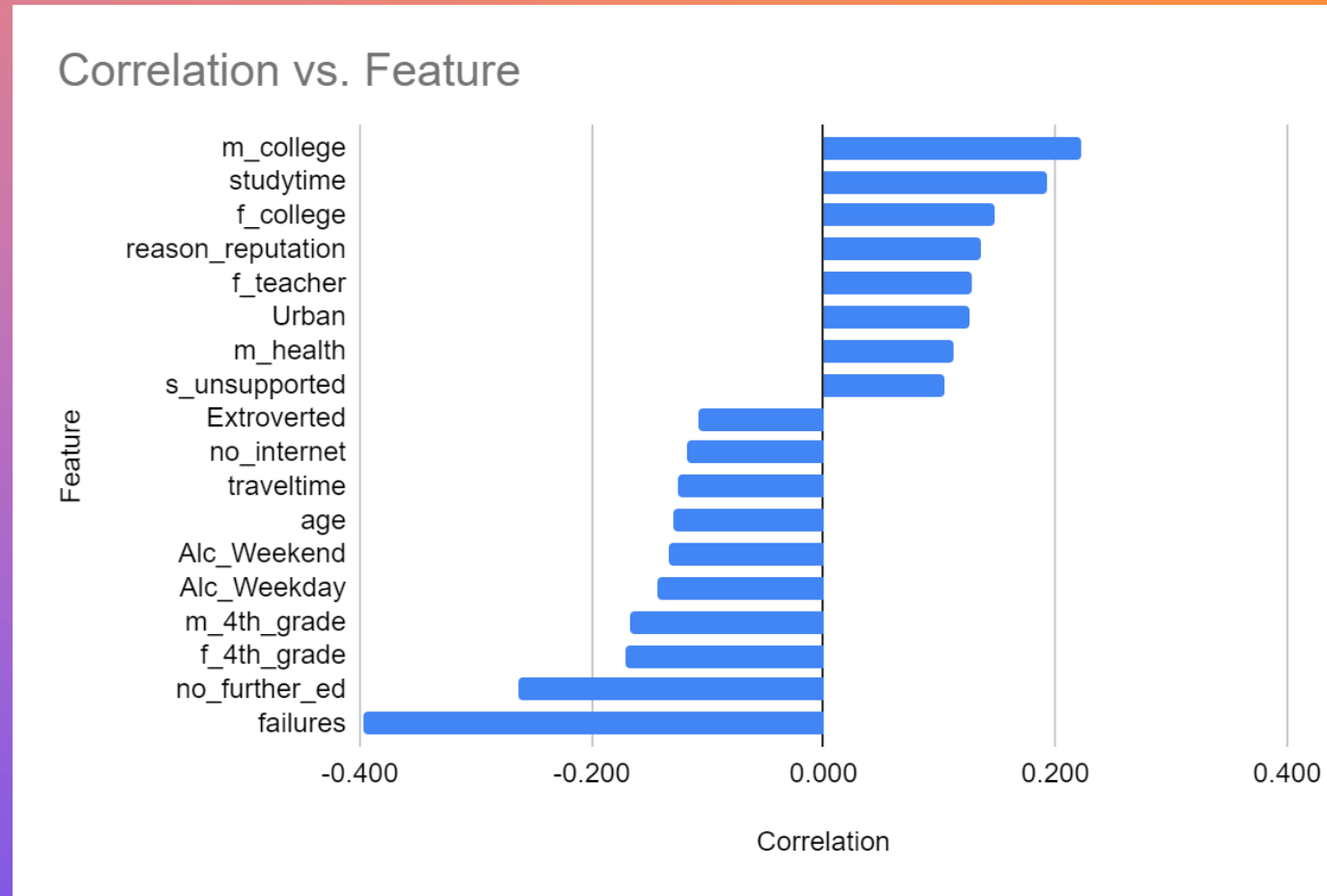
# RESULTS

Grade Distribution



- Insight:
  - About 1/3 of students did not pass

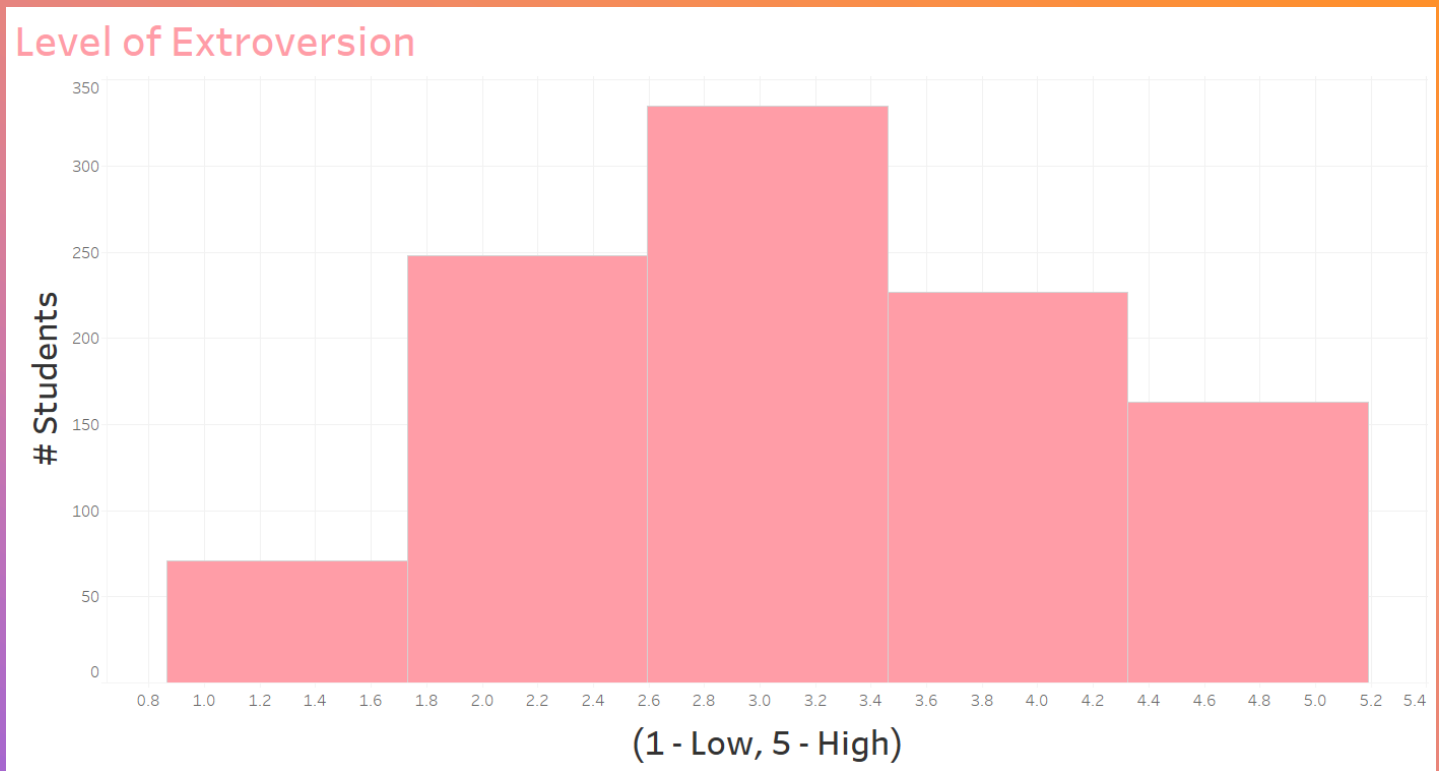
# RESULTS CONT.



- Insight:
  - Individually, each category had a small influence on performance

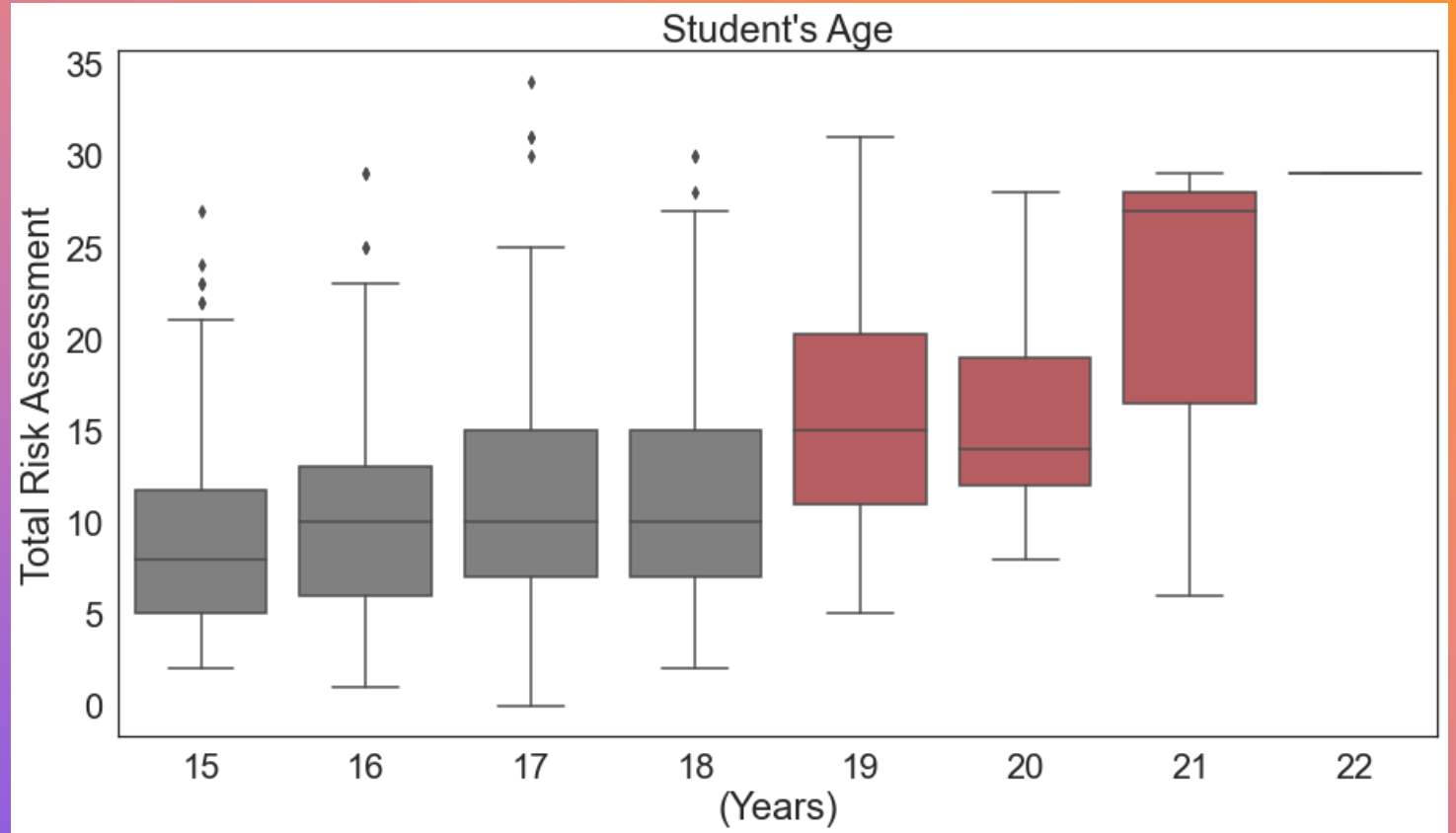
# RESULTS CONT.

Extroverted Count	% of students
4	21.74%
3	32.09%
2	23.75%
1	6.80%
5	15.61%



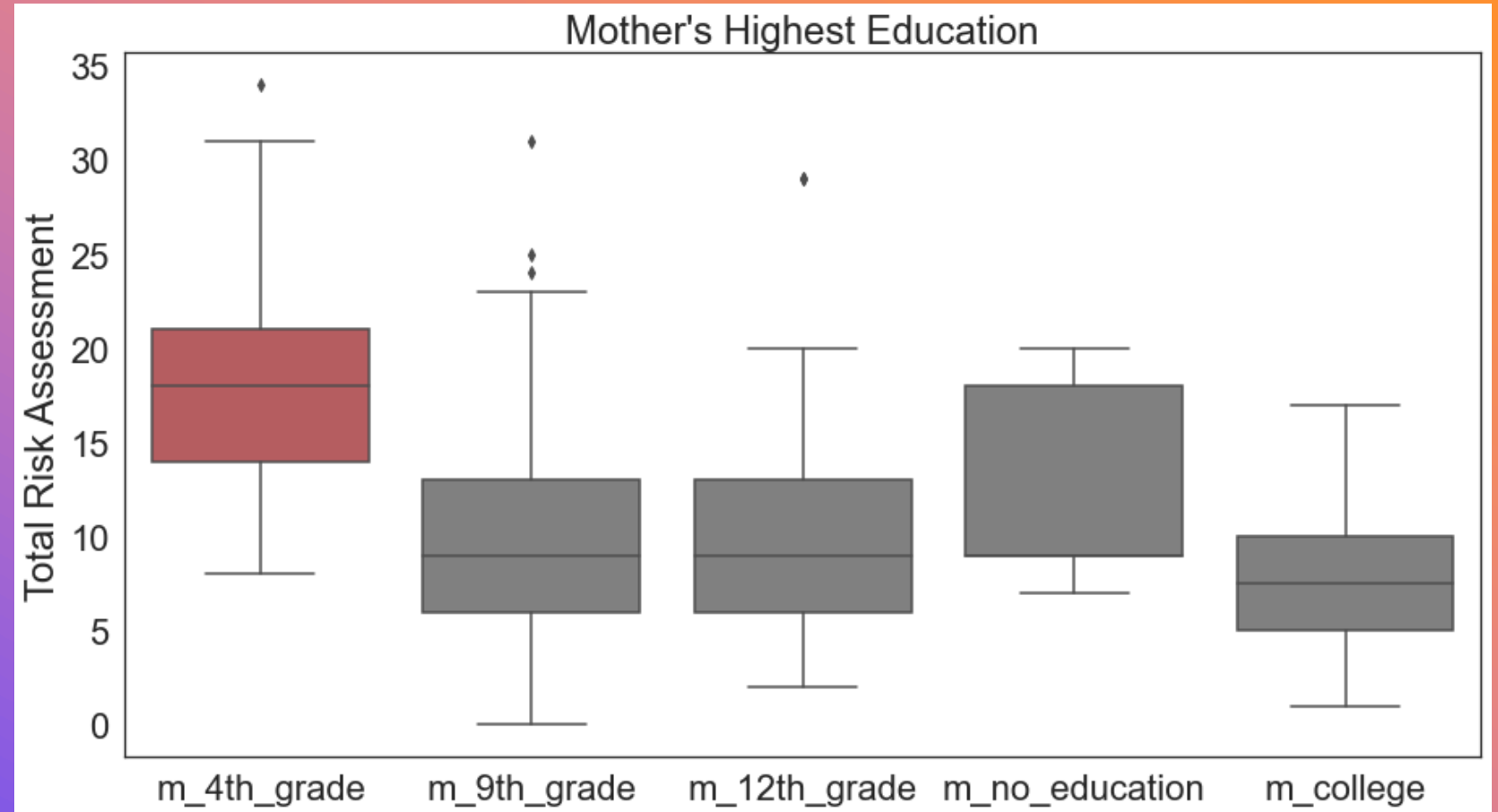
- Insight:
  - Students that are more extroverted performed poorly

# RESULTS CONT.



- Insight:
  - Students above normal high school age tend to perform poorly

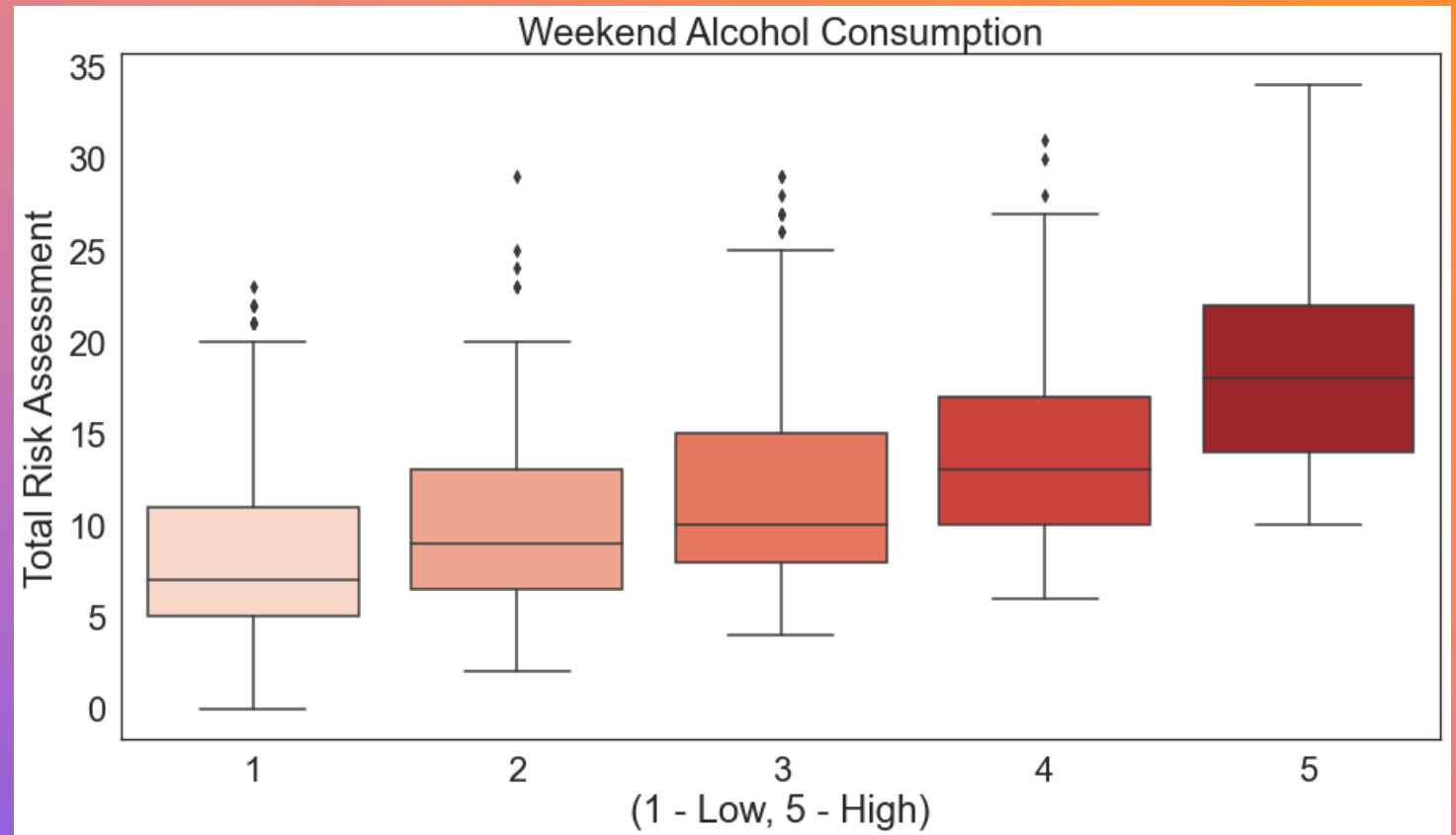
# RESULTS CONT.



- Insight:
  - Students who had parents only graduating 4<sup>th</sup> grade is a risk factor

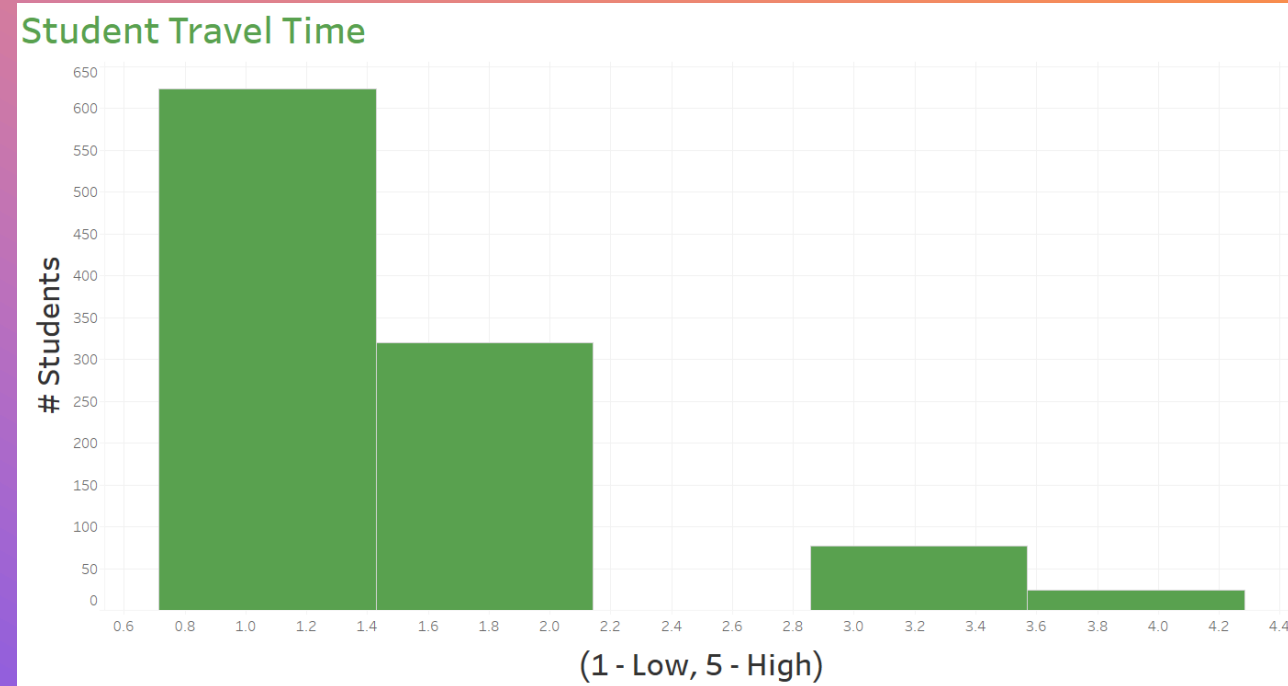
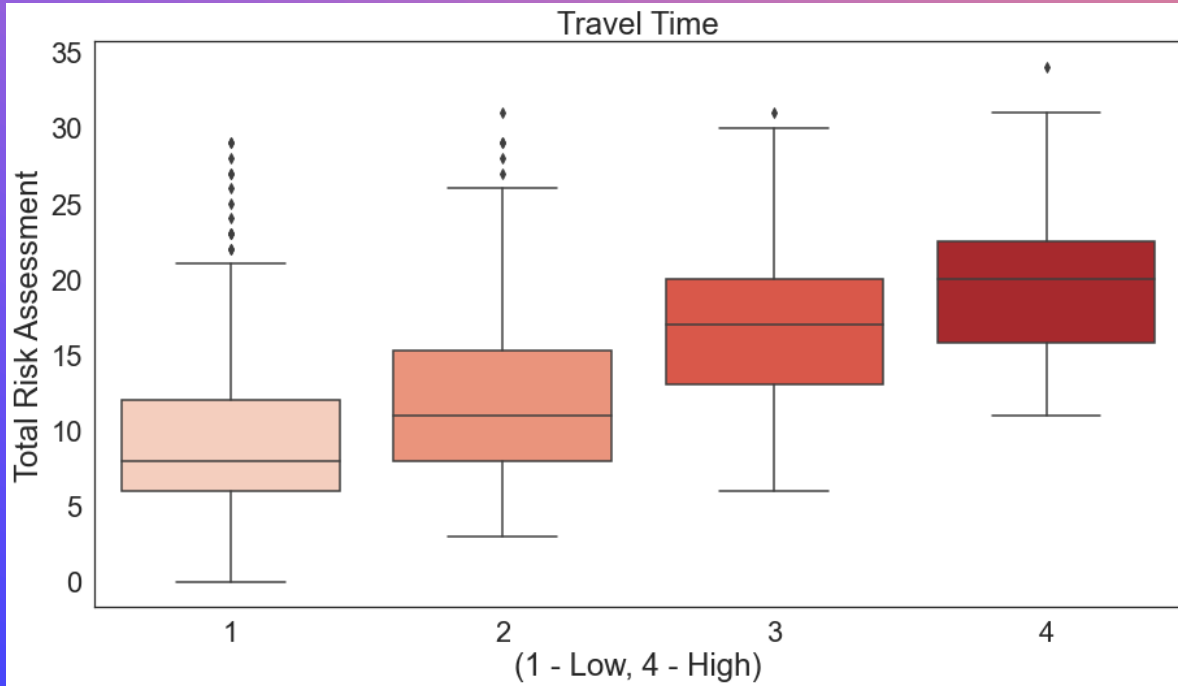


# RESULTS CONT.



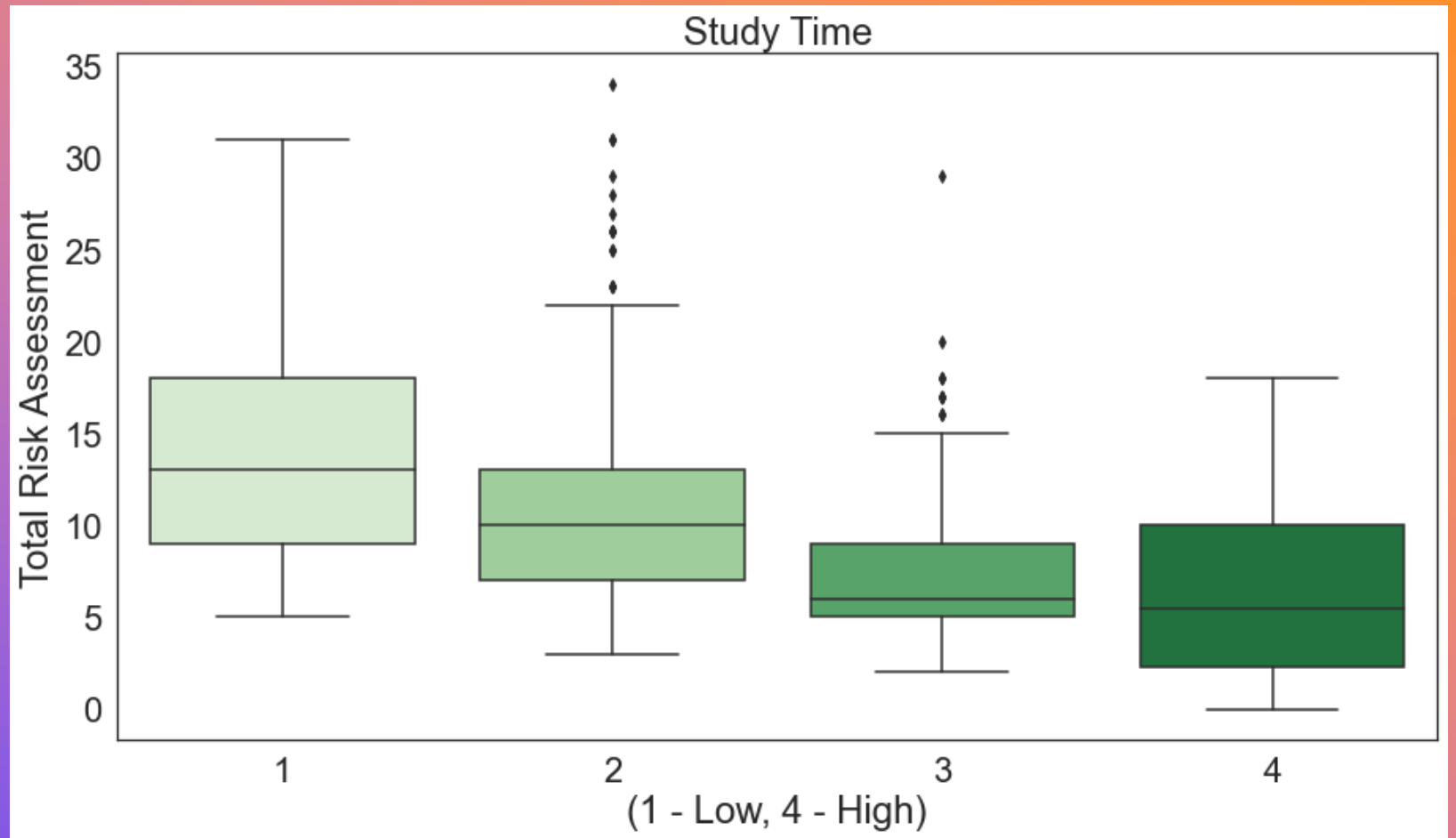
- Insight:
  - Students that drink more alcohol is a risk factor

# RESULTS CONT.



- Insight:
  - Students with higher travel times is a risk factor

# RESULTS CONT.



- Insight:
  - Students that study longer score better

# CONCLUSIONS

- This project performed an EDA which revealed the family and personal factors that have the highest impact on student grades
- Schools can reason that students that drink more, have less educated parents, live farther away, etc. are likely to require more academic support



# FUTURE WORK

- Create and refine a predictive model
- Explore further correlations between other features
- Find out why 4<sup>th</sup> grade education is worse than none
- Find additional features



# APPENDIX

age	Environment	M_Education	F_Education	traveltime	studytime	failures	School_Support	Higher_Ed	internet	Extroverted	Alc_Weekday	Alc_Weekend	Grades_AVG
17	Rural	m_4th_grade	f_4th_grade	4	2	3	s_unsupported	no_further_ed	no_internet	5	1	5	6.7
17	Urban	m_9th_grade	f_9th_grade	1	4	0	s_unsupported	continued_ed	internet	1	1	1	6.3