

# Alcohol Effects On Academic Performance

Group 13

Amber Szulc, Shril Mody, Yuzhao Chen, Liyang Ru, Linfeng Wen

# Contents

- Introduction
- Methodology
- Factors & Academic Performance Correlation
- Feature Importance Analysis
- Conclusion

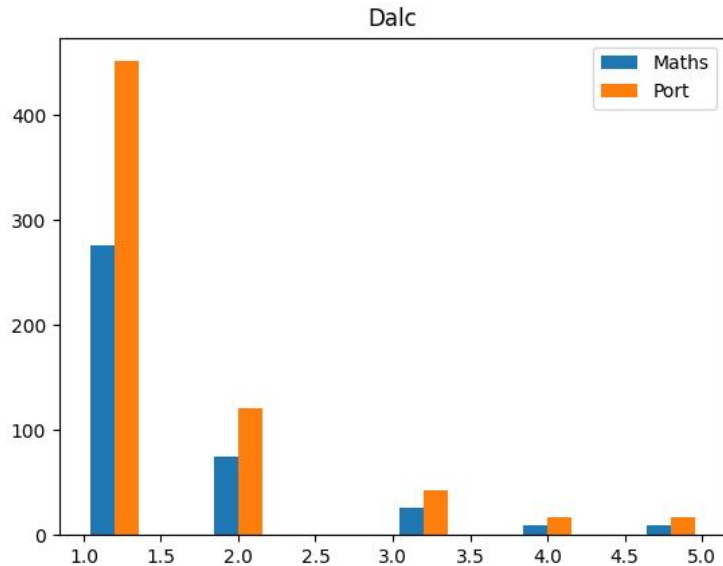


# Introduction

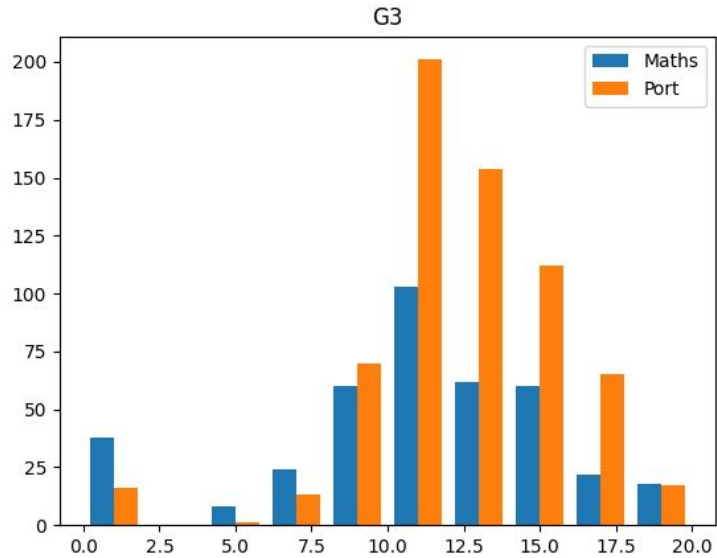
# Dataset

- Two datasets from two Portuguese schools (900+ students):
  - Math
  - Portuguese
- **30** Features
  - Alcohol Consumption
  - Other factors, including education, habit, etc.
- **3** Targets
  - G1, G2 -> **Grades** of two separate period
  - G3 -> **Final grade**

# Alcohol Usage Distribution - Weekday Alcohol Consumption



# G3 - Final Grades Distribution



## We want to answer...

- How does **alcohol consumption** affect students' grades?
  - How does **non-alcohol factors** affect students' final grades?
  - What is the **importance** of every feature?
- 
- Observe the difference between Math class and Portuguese class.



# Methodology



# Methodology

- Heatmap
- Linear Regression & Bar Charts & Box Charts
- Modeling (Tree-based Classifier)
- Feature Importance Analysis

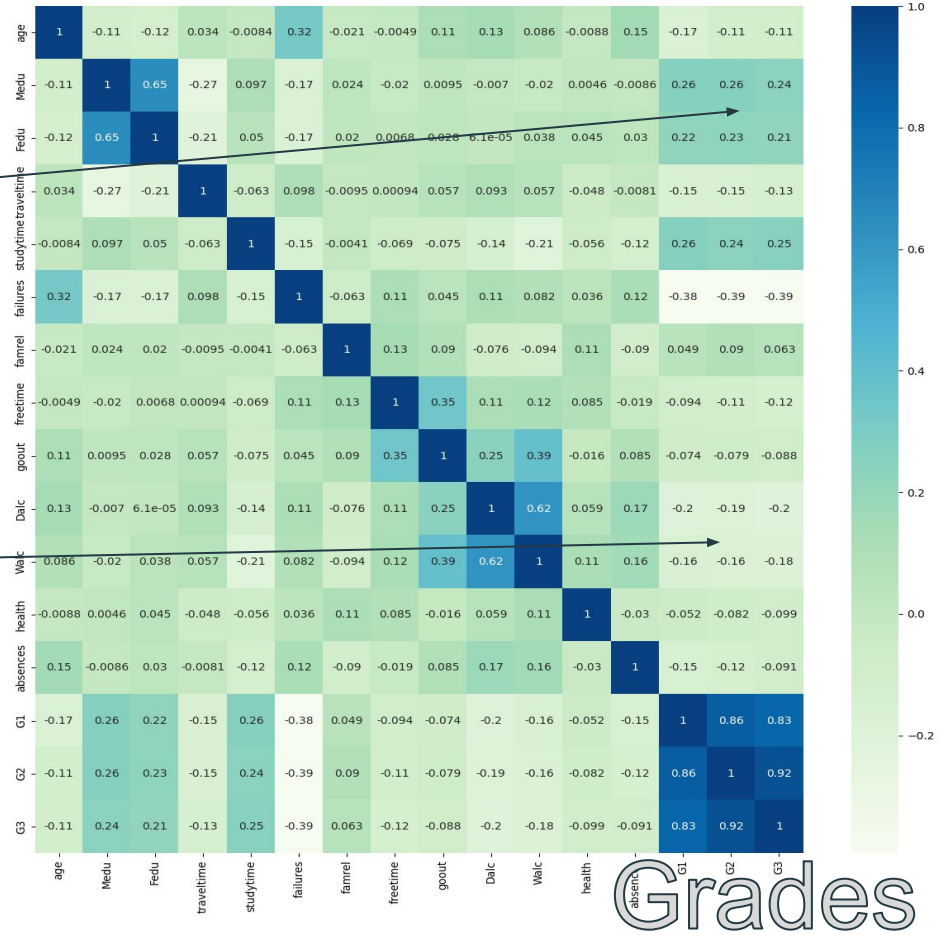


# Factors & Academic Performance Correlation

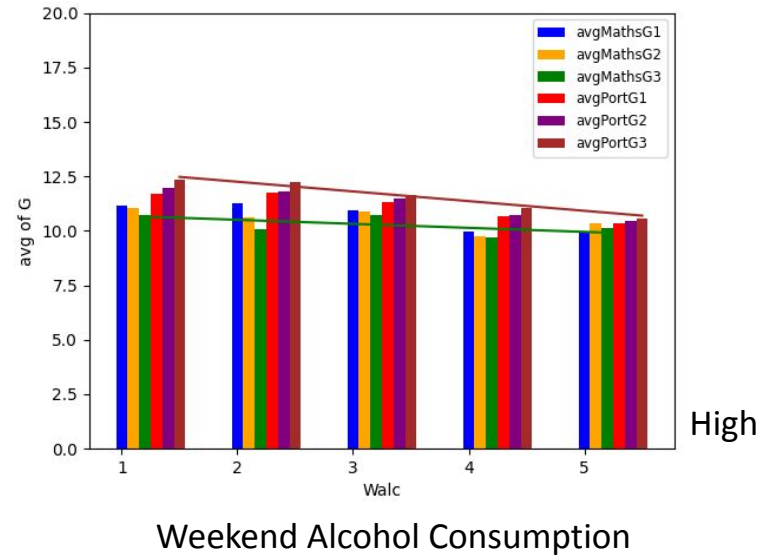
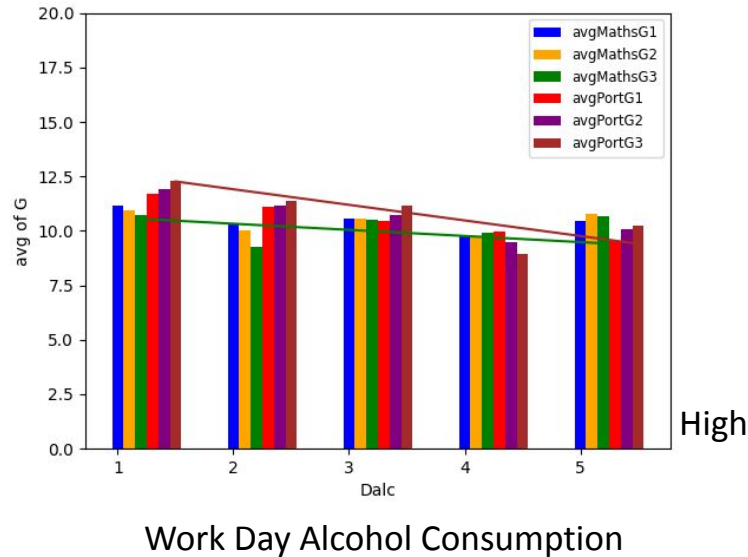
# Heatmap

Parent's  
Education Level

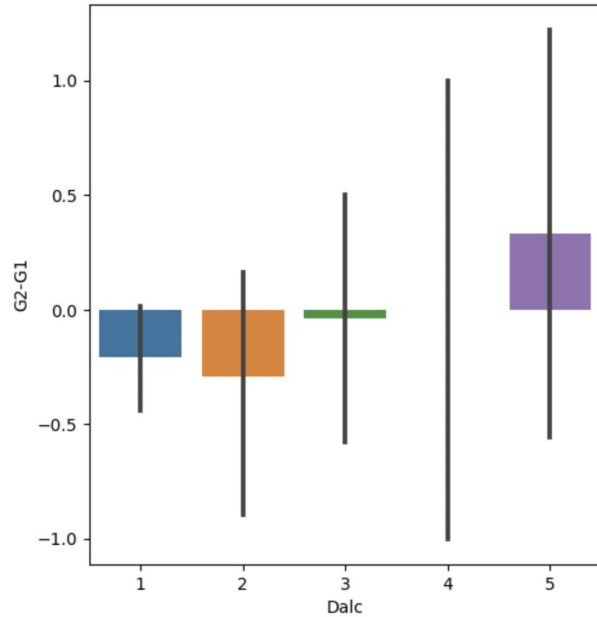
Alcohol  
Consumption



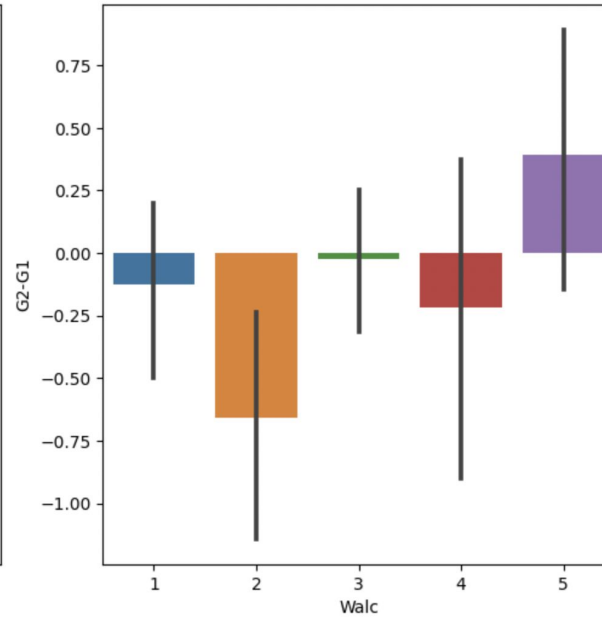
# Alcohol's Effect on Academic Performance



# Academic Improvement - Math

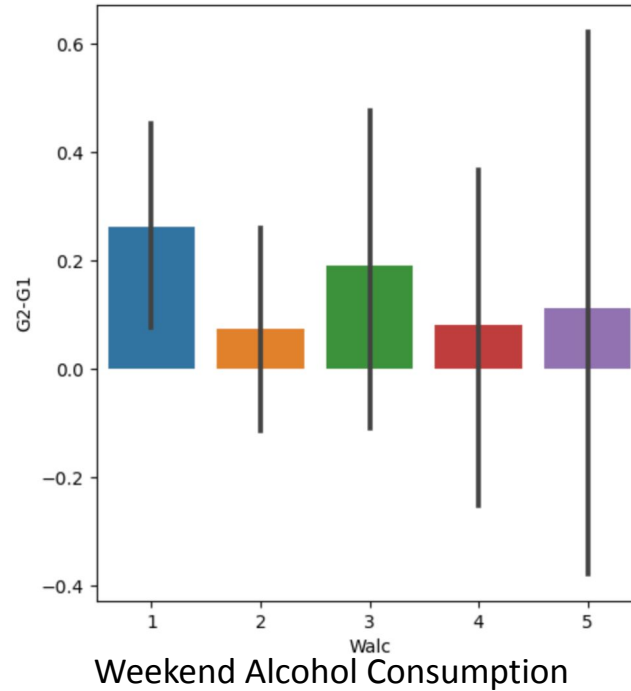
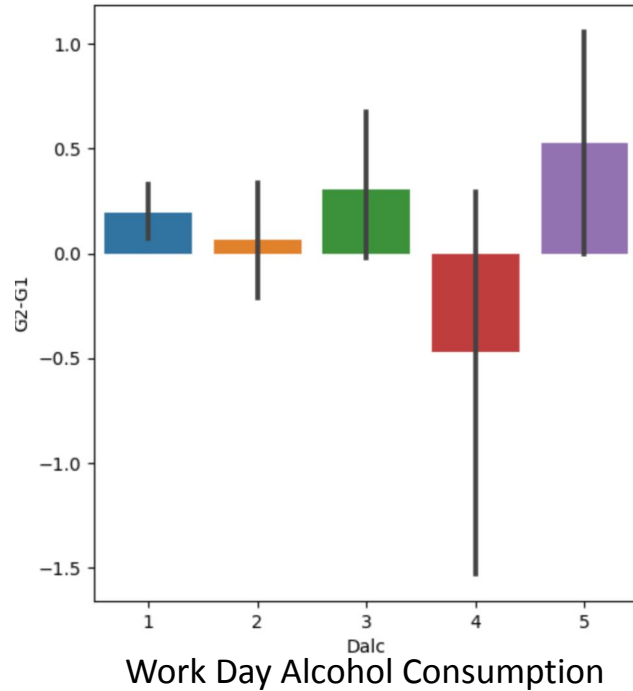


Work Day Alcohol Consumption

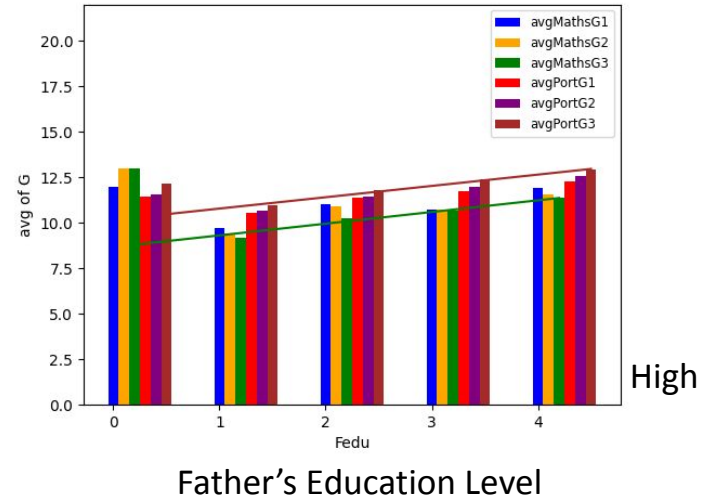
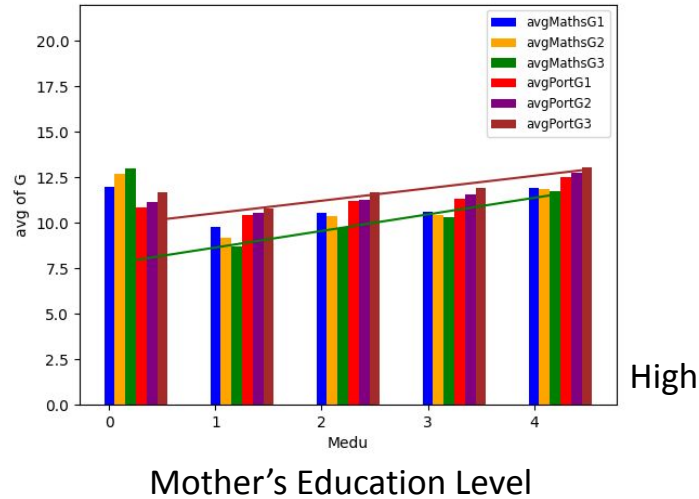


Weekend Alcohol Consumption

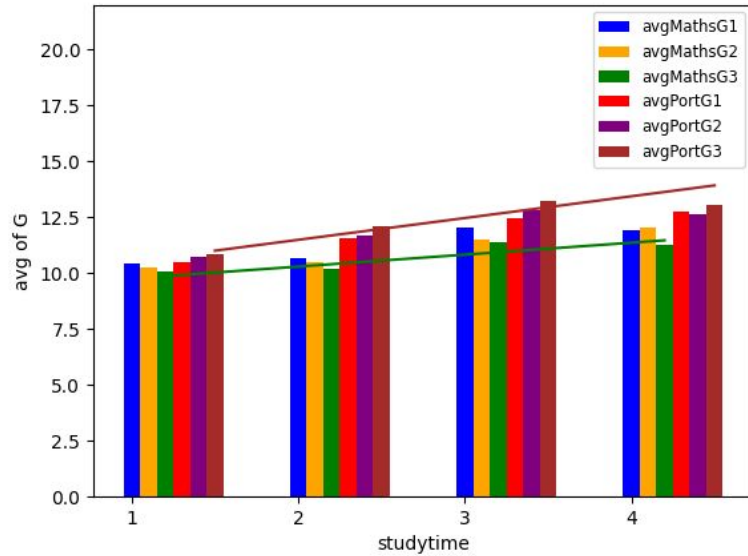
# Academic Improvement - Portuguese



# Non-Alcohol Effect on Academic Performance



# Non-Alcohol Effect on Academic Performance



Study Time per Week

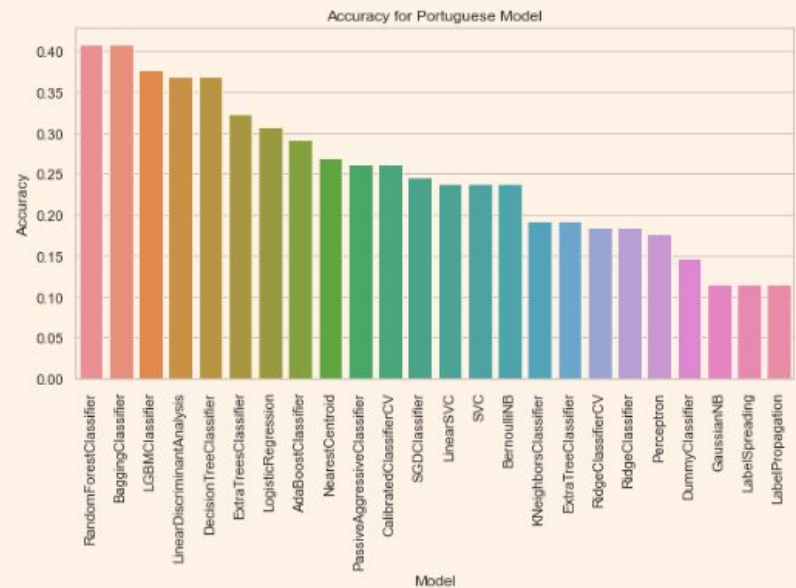
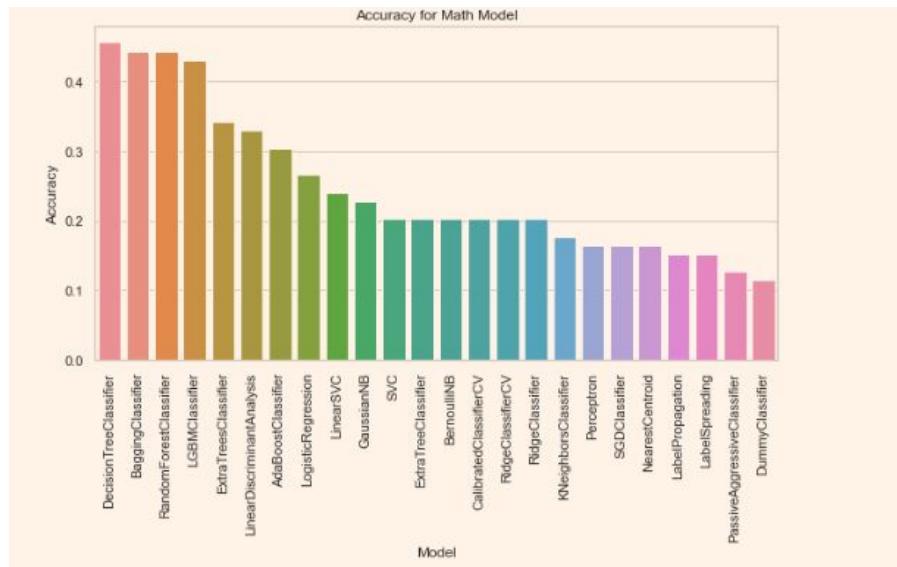




# Feature Importance Analysis

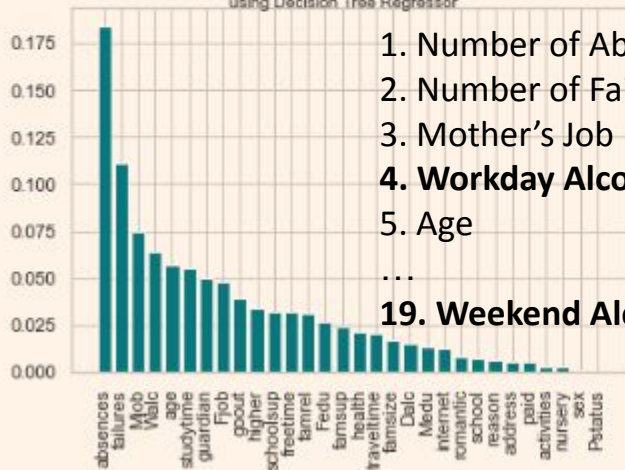
# Modeling: Determining the Most Accurate Model

Use lazy classifier to determine most accurate model for predicting data out of set of 29 models



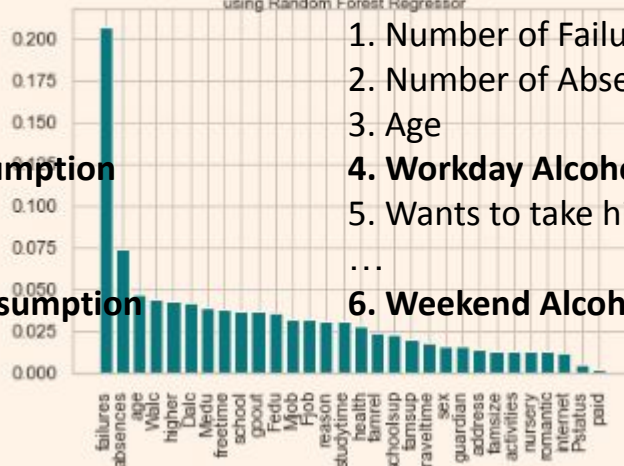
# Feature Importance Analysis

Math Grades: Feature Importance  
How important each feature is in determining MathGrade 3  
using Decision Trees Regressor



1. Number of Absences
2. Number of Failures
3. Mother's Job
4. Workday Alcohol Consumption
5. Age
- ...
19. Weekend Alcohol Consumption

Portuguese Grades: Feature Importance  
How important each feature is in determining PortugueseGrade 3  
using Random Forest Regressor



1. Number of Failures
2. Number of Absences
3. Age
4. Workday Alcohol Consumption
5. Wants to take higher Education
- ...
6. Weekend Alcohol Consumption



# Conclusion

# Conclusion

- More Alcohol Consumption -> Lower Grades
- Study Hard, Better Parents -> Higher Grades
- Compare to non-alcoholic features
- For prediction task, failures & absences are more important.