

Predicting Failure

DATA 151 Project Part 2

Josh Nibbe
Ethan Hawk
Ashley Darnell

THE DATA

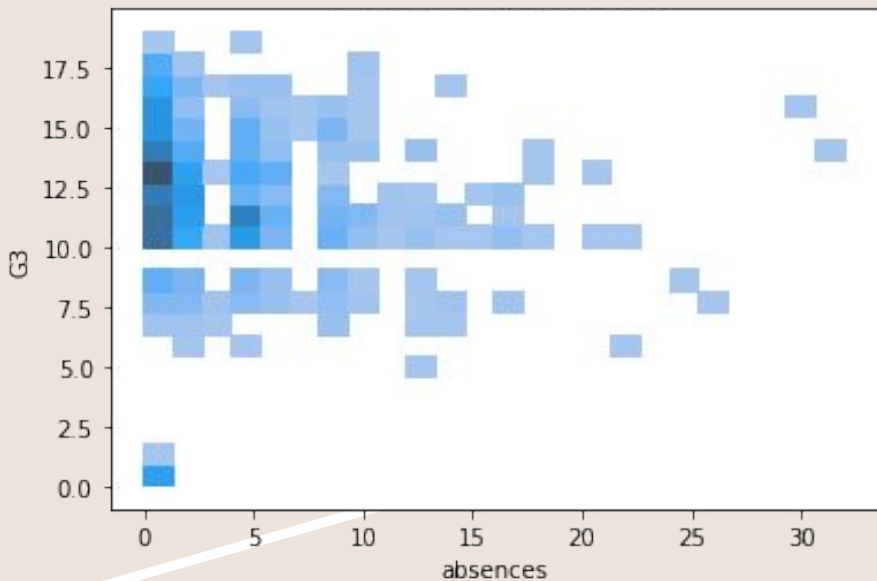
- The data comes from The UCI Machine Learning Repository
- 33 variables
 - 16 numerical & 17 categorical
- The target variable is **failures**.
- Problems reading in the data in Jupyter Notebook
 - Columns separated by semicolons
 - Make new table and save it

CORRELATIONS

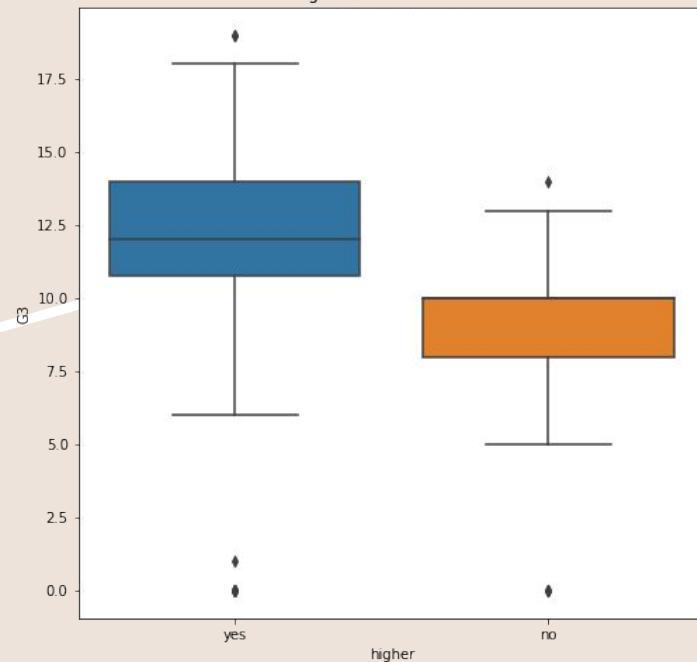
	Absences	Age	Failures	Studytime	G3
Absences	1	.1499	.1227	-.1183	-.0914
Age	.1499	1	.3199	-.0084	-.1065
Failures	.1227	.3199	1	-.1474	-.3933
Studytime	-.1183	-.0084	-.1474	1	.2498
G3	-.0914	-.1065	-.3933	.2498	1

INTERESTING RELATIONSHIPS

Absences & Final Grade

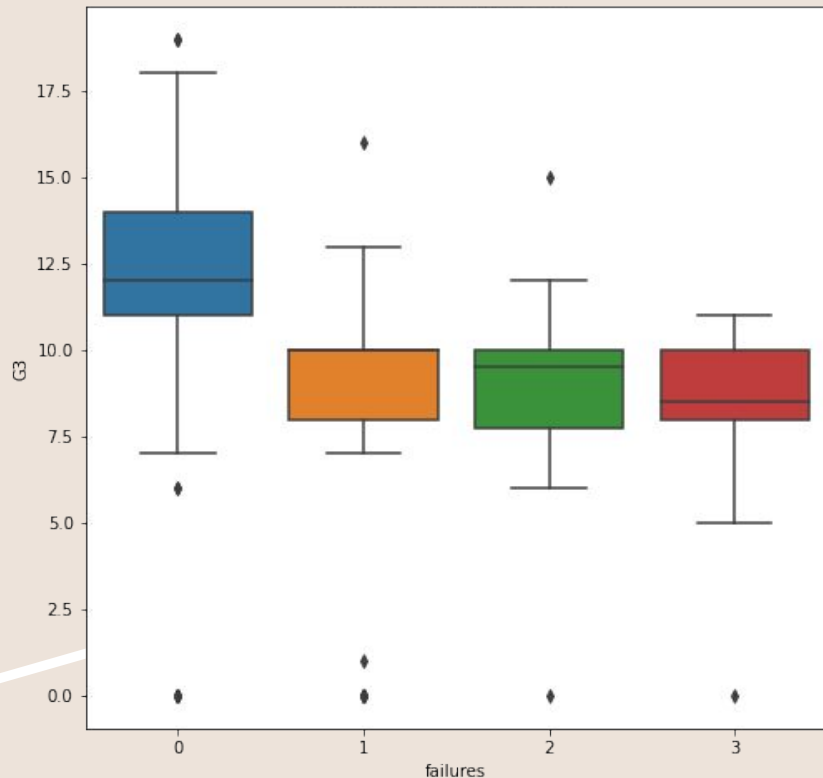


Higher Ed & Final Grade



MORE RELATIONSHIPS

Failures vs. Final Grade



Weekly Alcohol vs Failures

