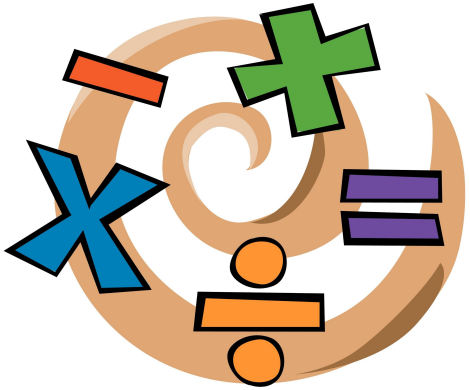
Predicting Student Grades in Math & Portuguese

Luke Barthelmess
Ernst van Bruggen
Chris Byrne
Teresa Ciaccio
Josh Landay
Yurika Tokuzumi

Math is a language with numbers instead of letters

The relationship between learning math and foreign languages has long been studied to see if there is a correlation between performance in these subjects. Some studies have concluded that yes, learning math and a foreign language require the same mechanics and that excelling in one is most likely correlated with high performance in the other. However, some say the evidence is not yet conclusive.



Our dataset includes a number of student performance measures

We pulled a comprehensive data set from:
https://archive.ics.uci.edu/ml/datasets/student+performance

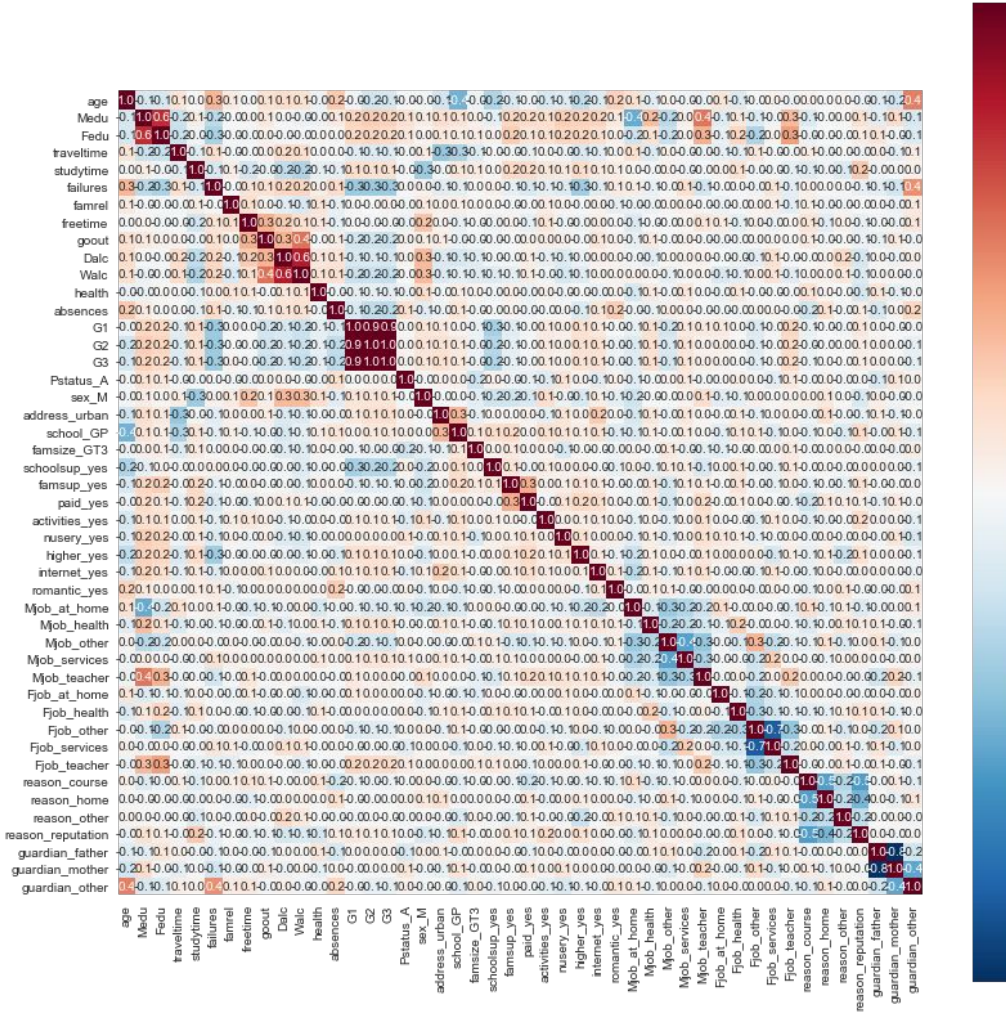
- 343 students (aged 15-22)
- 42 independent variables
- Dependent variable will be final grades for Portuguese and Math class

Using these data, we are hopeful to develop a regression model that can explain final grades with a high degree of confidence.

Correlation Matrix

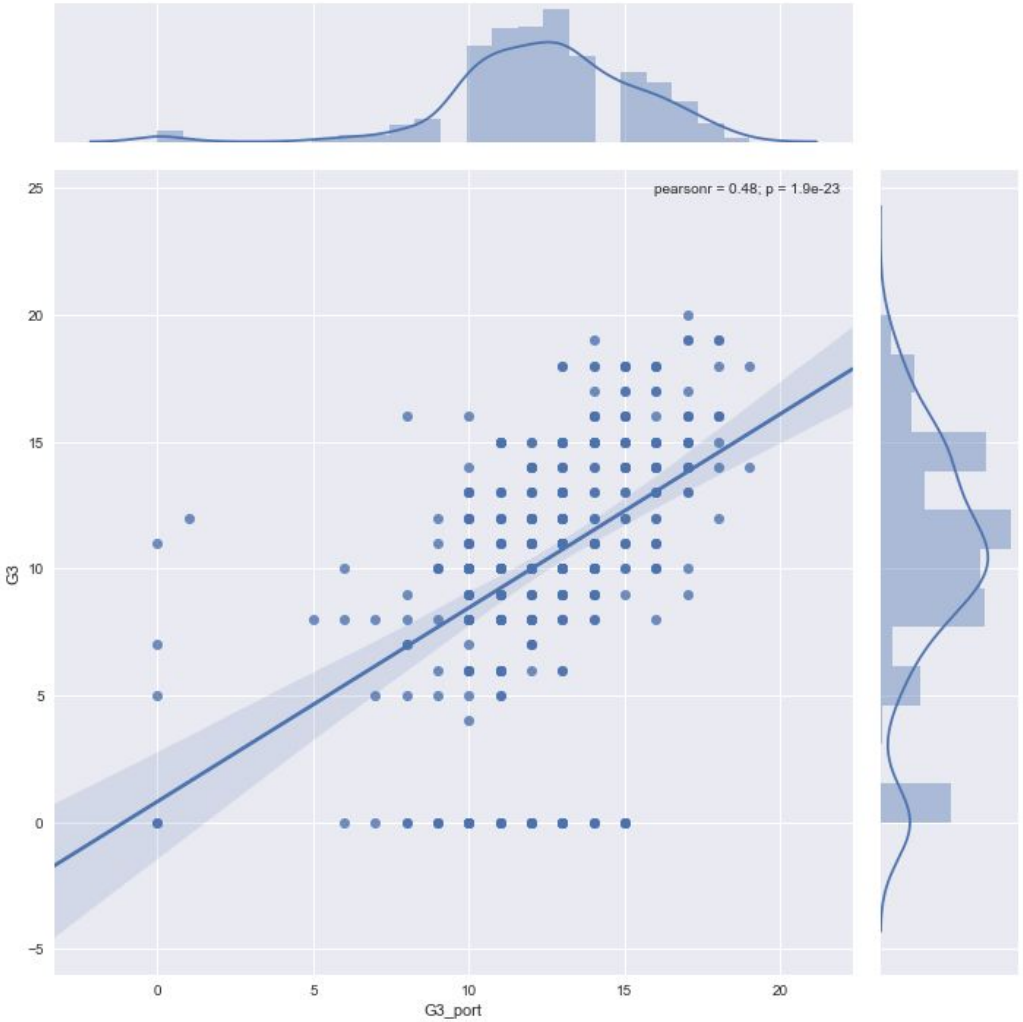
- G1, G2, and G3 are highly correlated (no surprise there)
- Dalc & Walc (weekday and weekend alcohol consumption) also quite correlated
- Medu and Fedu (parents' education levels)

Net/net: proceed with some caution



Inclusion / exclusion of outliers from dataset

To increase the model's R^2, we elected to eliminate some outlier scores from our dataset; specifically, those Portuguese scores less than 2.



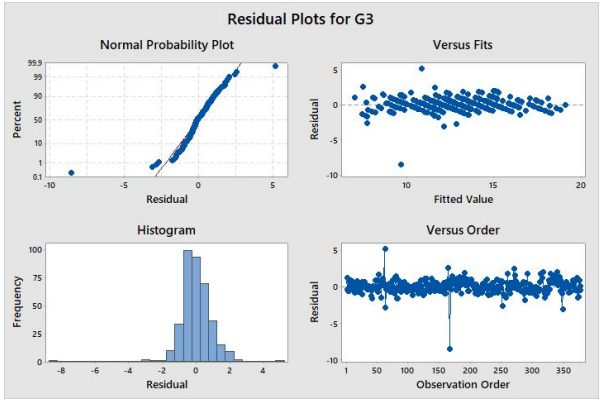
Full Model - Portuguese Final Grade

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.999964	86.70%	85.03%	82.49%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-1.57	1.24	-1.26	0.208	
age	0.1101	0.0569	1.94	0.054	1.65
sex_M	-0.135	0.128	-1.05	0.294	1.55
higher_yes	0.410	0.289	1.42	0.156	1.35
G3_math	0.0082	0.0149	0.55	0.583	1.80
failures	-0.231	0.121	-1.91	0.057	1.40
Dalc	-0.0503	0.0858	-0.59	0.558	2.10
Medu	-0.0287	0.0831	-0.35	0.730	3.04
Fedu	-0.0463	0.0707	-0.65	0.513	2.24
traveltime	0.1535	0.0859	1.79	0.075	1.35
studytime	-0.0017	0.0719	-0.02	0.982	1.39
famrel	0.0720	0.0612	1.18	0.241	1.16
freetime	-0.0442	0.0596	-0.74	0.458	1.30
goout	-0.0866	0.0556	-1.56	0.120	1.50
Walc	0.0264	0.0613	0.43	0.667	2.32
health	-0.0333	0.0406	-0.82	0.412	1.23
absences	-0.0085	0.0117	-0.73	0.464	1.24
G1	0.1513	0.0474	3.19	0.002	5.31
G2	0.7885	0.0533	14.80	0.000	6.20
Pstatus_A	0.281	0.183	1.53	0.126	1.15
address_urban	-0.020	0.150	-0.14	0.892	1.38
school_GP	0.276	0.208	1.33	0.186	1.41
famsize_GT3	0.183	0.124	1.47	0.142	1.15
schoolsup_yes	-0.204	0.173	-1.18	0.238	1.29
famsup_yes	-0.008	0.117	-0.07	0.947	1.21
paid_yes	-0.027	0.225	-0.12	0.905	1.19
activities_yes	0.094	0.113	0.83	0.406	1.20
nusery_yes	-0.047	0.141	-0.33	0.741	1.14
internet_yes	0.210	0.164	1.28	0.201	1.28
romantic_yes	0.158	0.118	1.34	0.183	1.14
Mjob_at_home	0.057	0.185	0.31	0.758	1.55
Mjob_health	0.331	0.233	1.42	0.156	1.63
Mjob_services	0.258	0.149	1.73	0.085	1.59
Mjob_teacher	0.431	0.205	2.11	0.036	2.17
Fjob_at_home	0.612	0.279	2.20	0.029	1.19
Fjob_health	-0.139	0.282	-0.49	0.622	1.29
Fjob_services	-0.163	0.134	-1.21	0.226	1.38
Fjob_teacher	-0.019	0.230	-0.08	0.934	1.46
reason_course	0.199	0.212	0.94	0.348	3.94
reason_home	0.269	0.218	1.23	0.218	3.68
reason_reputation	0.174	0.220	0.79	0.431	3.52
guardian_father	-0.196	0.292	-0.67	0.503	5.80
guardian_mother	-0.168	0.277	-0.61	0.543	5.79



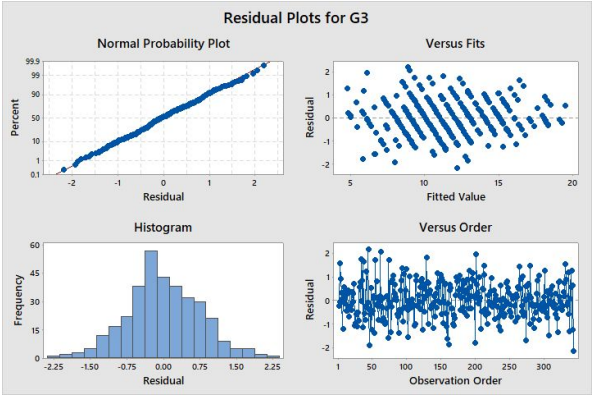
Full Model - Math Final Grade

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.807505	94.69%	93.95%	93.03%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-1.27	1.10	-1.15	0.251	
age	0.0406	0.0483	0.84	0.402	1.63
Medu	-0.0619	0.0683	-0.91	0.366	2.88
Fedu	-0.0310	0.0601	-0.52	0.606	2.22
traveltime	0.0206	0.0736	0.28	0.780	1.33
studytime	0.0191	0.0618	0.31	0.757	1.41
failures	0.1194	0.0858	1.39	0.165	1.57
famrel	0.1883	0.0524	3.59	0.000	1.17
freetime	0.0092	0.0509	0.18	0.857	1.35
goout	-0.0808	0.0495	-1.63	0.104	1.56
Dalc	0.0257	0.0699	0.37	0.713	2.16
Walc	0.0056	0.0543	0.10	0.919	2.58
health	-0.0787	0.0344	-2.29	0.023	1.24
absences	-0.01053	0.00658	-1.60	0.111	1.39
G1	0.0957	0.0358	2.68	0.008	7.14
G2	0.8795	0.0367	23.94	0.000	7.31
G3_port	0.0429	0.0227	1.89	0.060	2.17
Pstatus_A	0.189	0.155	1.22	0.224	1.19
sex_M	0.055	0.110	0.50	0.617	1.60
address_urban	0.060	0.127	0.48	0.635	1.40
school_GP	0.114	0.179	0.64	0.525	1.54
famsize_GT3	0.131	0.103	1.27	0.206	1.15
schoolsup_yes	-0.014	0.144	-0.10	0.920	1.35
famsup_yes	0.115	0.103	1.12	0.263	1.30
paid_yes	-0.2122	0.0994	-2.14	0.034	1.30
activities_yes	-0.0572	0.0973	-0.59	0.557	1.24
nusery_yes	-0.168	0.121	-1.38	0.169	1.16
higher_yes	-0.100	0.294	-0.34	0.733	1.41
internet_yes	0.057	0.138	0.42	0.678	1.26
romantic_yes	0.000	0.102	0.00	0.998	1.15
Mjob_at_home	0.188	0.162	1.16	0.246	1.53
Mjob_health	0.354	0.195	1.81	0.071	1.65
Mjob_services	0.118	0.127	0.93	0.352	1.63
Mjob_teacher	0.380	0.175	2.17	0.031	2.21
Fjob_at_home	-0.432	0.249	-1.74	0.083	1.19
Fjob_health	0.039	0.231	0.17	0.865	1.32
Fjob_services	-0.117	0.114	-1.03	0.304	1.37
Fjob_teacher	0.002	0.197	0.01	0.991	1.49
reason_course	-0.021	0.172	-0.12	0.902	3.56
reason_home	0.133	0.177	0.75	0.453	3.36
reason_reputation	0.020	0.181	0.11	0.913	3.39
guardian_father	0.316	0.269	1.18	0.240	6.91
guardian_mother	0.285	0.258	1.10	0.271	6.98



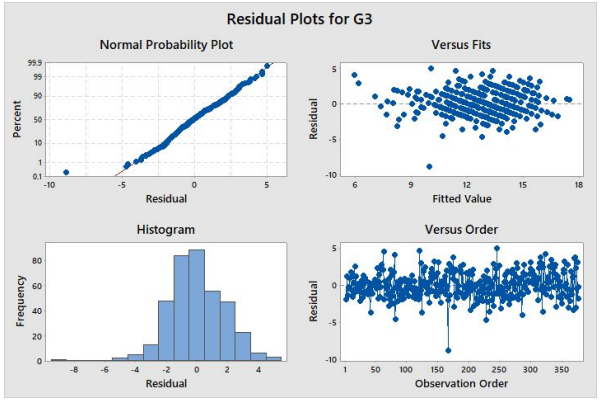
Full Model - Portuguese Final Grade - No G1/G2

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.87051	53.20%	47.63%	40.74%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.70	2.29	0.30	0.761	
age	0.449	0.103	4.37	0.000	1.54
sex_M	-0.963	0.234	-4.12	0.000	1.47
higher_yes	1.710	0.533	3.21	0.001	1.32
G3_math	0.2330	0.0238	9.80	0.000	1.31
failures	-0.816	0.223	-3.65	0.000	1.36
Dalc	-0.256	0.160	-1.60	0.111	2.09
Medu	0.109	0.155	0.70	0.484	3.03
Fedu	-0.009	0.132	-0.07	0.945	2.24
traveltime	0.212	0.160	1.32	0.186	1.34
studytime	0.256	0.133	1.92	0.056	1.36
famrel	0.064	0.114	0.56	0.578	1.15
freetime	-0.078	0.111	-0.70	0.483	1.30
goout	-0.201	0.103	-1.94	0.053	1.48
Walc	-0.031	0.115	-0.27	0.785	2.32
health	-0.1678	0.0754	-2.22	0.027	1.21
absences	-0.0486	0.0216	-2.25	0.025	1.22
Pstatus_A	0.420	0.341	1.23	0.219	1.14
address_urban	0.522	0.277	1.88	0.060	1.34
school_GP	0.810	0.386	2.10	0.036	1.38
famsize_GT3	0.028	0.231	0.12	0.904	1.14
schoolsup_yes	-1.108	0.314	-3.53	0.000	1.22
famsup_yes	0.195	0.219	0.89	0.375	1.20
paid_yes	-0.542	0.420	-1.29	0.198	1.18
activities_yes	0.546	0.209	2.61	0.009	1.18
nusery_yes	-0.348	0.263	-1.32	0.188	1.13
internet_yes	0.062	0.306	0.20	0.838	1.28
romantic_yes	0.022	0.221	0.10	0.919	1.14
Mjob_at_home	0.197	0.343	0.57	0.567	1.53
Mjob_health	0.697	0.435	1.60	0.110	1.63
Mjob_services	0.147	0.279	0.53	0.597	1.59
Mjob_teacher	0.525	0.383	1.37	0.171	2.17
Fjob_at_home	0.953	0.520	1.83	0.068	1.18
Fjob_health	-0.176	0.528	-0.33	0.739	1.29
Fjob_services	-0.166	0.251	-0.66	0.509	1.37
Fjob_teacher	0.802	0.427	1.88	0.062	1.44
reason_course	0.002	0.396	0.00	0.997	3.93
reason_home	0.340	0.408	0.83	0.405	3.68
reason_reputation	-0.016	0.412	-0.04	0.968	3.52
guardian_father	0.323	0.545	0.59	0.554	5.78
guardian_mother	0.195	0.517	0.38	0.706	5.78



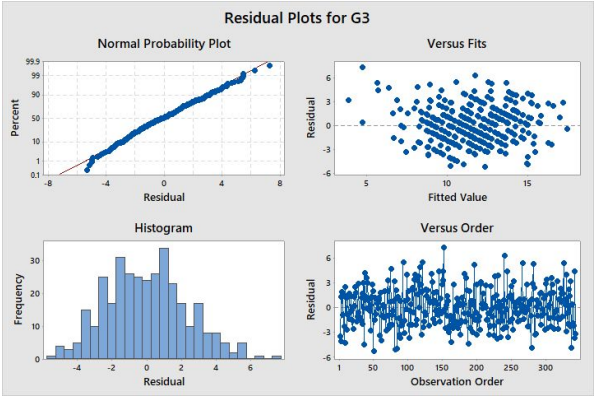
Full Model - Math Final Grade - No G1/G2

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
2.46550	50.17%	43.57%	34.96%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	12.17	3.23	3.77	0.000	
age	-0.441	0.143	-3.08	0.002	1.54
Medu	0.190	0.207	0.92	0.359	2.85
Fedu	0.102	0.183	0.56	0.579	2.21
traveltime	-0.044	0.225	-0.20	0.845	1.33
studytime	0.221	0.188	1.18	0.241	1.40
failures	-0.477	0.258	-1.85	0.066	1.52
famrel	0.121	0.160	0.75	0.452	1.16
freetime	0.104	0.155	0.67	0.503	1.34
goout	-0.261	0.151	-1.73	0.085	1.55
Dalc	0.423	0.212	1.99	0.047	2.13
Walc	-0.247	0.165	-1.50	0.136	2.56
health	-0.087	0.105	-0.83	0.409	1.24
absences	-0.0339	0.0198	-1.71	0.088	1.36
G3_port	0.5938	0.0605	9.82	0.000	1.65
Pstatus_A	-0.285	0.472	-0.60	0.547	1.18
sex_M	1.136	0.329	3.45	0.001	1.53
address_urban	-0.157	0.387	-0.41	0.685	1.40
school_GP	-0.262	0.543	-0.48	0.630	1.53
famsize_GT3	-0.168	0.314	-0.54	0.592	1.14
schoolsup_yes	-1.503	0.428	-3.51	0.001	1.29
famsup_yes	-0.603	0.309	-1.95	0.052	1.27
paid_yes	-0.289	0.303	-0.95	0.341	1.30
activities_yes	-0.093	0.297	-0.31	0.754	1.24
nusery_yes	-0.185	0.370	-0.50	0.618	1.16
higher_yes	-1.468	0.891	-1.65	0.100	1.39
internet_yes	0.460	0.418	1.10	0.272	1.25
romantic_yes	-0.218	0.311	-0.70	0.484	1.14
Mjob_at_home	0.282	0.491	0.57	0.566	1.52
Mjob_health	0.843	0.593	1.42	0.156	1.63
Mjob_services	0.941	0.384	2.45	0.015	1.60
Mjob_teacher	-0.656	0.527	-1.25	0.214	2.14
Fjob_at_home	0.193	0.757	0.25	0.799	1.18
Fjob_health	-0.046	0.704	-0.07	0.948	1.32
Fjob_services	-0.116	0.347	-0.33	0.739	1.37
Fjob_teacher	1.464	0.592	2.47	0.014	1.44
reason_course	-0.025	0.525	-0.05	0.962	3.53
reason_home	0.070	0.539	0.13	0.896	3.33
reason_reputation	-0.062	0.546	-0.11	0.910	3.32
guardian_father	0.267	0.819	0.33	0.744	6.88
guardian_mother	0.127	0.787	0.16	0.872	6.95



Best Subsets - Portuguese vs. Math

Portuguese

Response is G3

Vars	R-Sq	R-Sq (adj)	R-Sq (pred)	Mallows Cp	S	G3_math	failures	sex_M	schoolsup_yes	Dalc	age	higher_yes	Medu	absences	address_R	Fjob_at_home	Fjob_other	guardian_mother	studytime	health	Mjob_services	Pstatus_A	famsup_yes	paid_yes	traveltime
1	22.9	22.7	22.0	178.5	2.2726	X																			
2	29.9	29.5	28.6	130.5	2.1699	X	X																		
3	35.1	34.5	33.5	95.4	2.0910	X	X	X																	
4	37.6	36.9	35.7	79.3	2.0526	X	X	X	X																
5	40.1	39.3	37.7	63.5	2.0140	X	X			X	X	X													
6	42.3	41.4	39.6	49.7	1.9791	X	X	X		X	X	X													
7	43.5	42.4	40.4	43.0	1.9608	X	X	X	X	X	X	X													
8	44.6	43.4	41.2	37.1	1.9441	X	X	X	X	X	X	X	X												
9	45.6	44.3	41.8	31.8	1.9287	X	X	X	X	X	X	X	X	X											
10	46.3	44.9	42.2	28.8	1.9189	X	X	X	X	X	X	X	X	X			X								
11	47.1	45.5	42.9	25.6	1.9087	X	X	X	X	X	X	X			X	X			X	X					
12	47.9	46.1	43.5	22.0	1.8970	X	X	X	X	X	X	X			X	X		X	X	X					
13	48.4	46.6	43.7	19.8	1.8890	X	X	X	X	X	X	X		X	X	X		X	X	X					
14	49.0	47.0	44.0	17.9	1.8814	X	X	X	X	X	X	X	X	X	X	X		X	X	X					
15	49.6	47.5	44.2	15.7	1.8731	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X					
16	49.8	47.5	44.2	16.3	1.8721	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X				X	
17	49.9	47.6	44.1	17.0	1.8714	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	X	
18	50.1	47.6	43.9	18.0	1.8715	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	X	X
19	50.2	47.5	43.8	19.3	1.8721	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X
20	50.2	47.4	43.5	21.0	1.8740	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

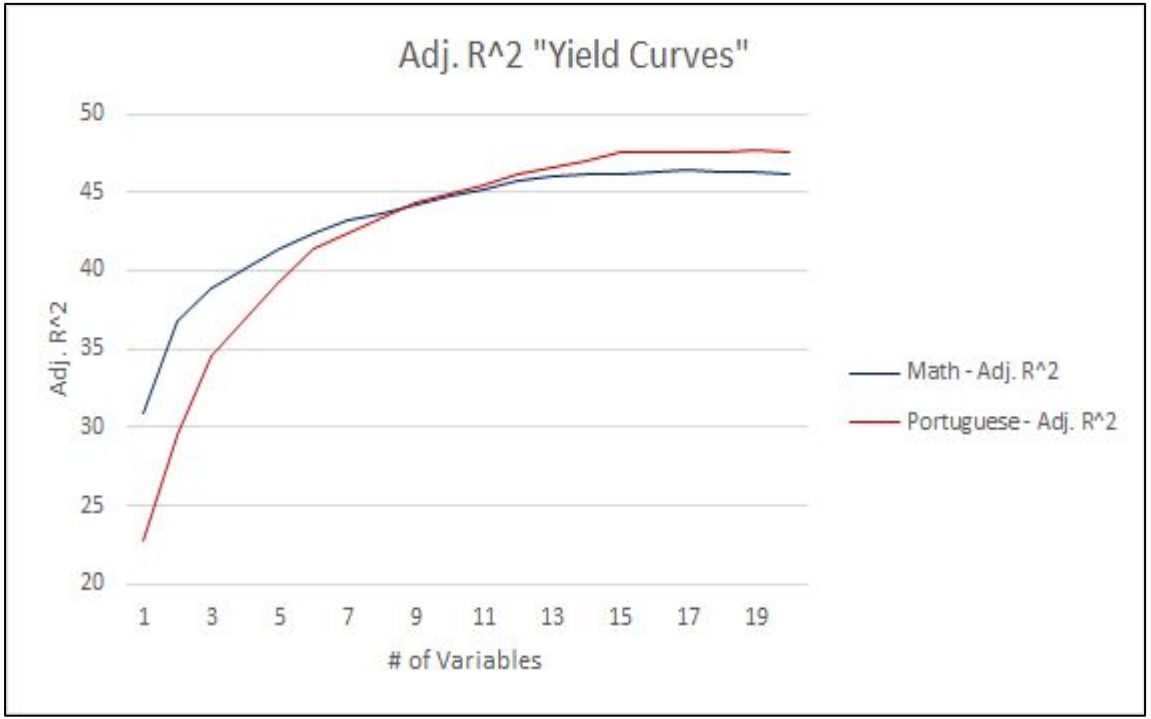
Vars	traveltime
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	X
19	X
20	X

Math

Response is G3

Vars	R-Sq	R-Sq (adj)	R-Sq (pred)	Mallows Cp	S	G3_portuguese	sex_M	age	schoolsup_yes	Fjob_teacher	Mjob_services	absences	Dalc	Walc	famsup_yes	Mjob_health	higher_yes	goout	internet_yes	Mjob_teacher	studytime	Medu	Fedu	health
1	31.1	30.9	29.9	98.8	2.7284	X																		
2	37.2	36.8	35.6	62.1	2.6086	X	X																	
3	39.3	38.8	37.3	50.5	2.5678	X	X	X																
4	40.9	40.2	38.6	42.7	2.5385	X	X	X	X															
5	42.3	41.4	39.5	35.8	2.5121	X	X	X	X	X														
6	43.4	42.4	40.3	30.5	2.4906	X	X	X	X	X	X													
7	44.3	43.2	40.7	26.7	2.4740	X	X	X	X	X	X	X												
8	45.1	43.7	41.1	24.1	2.4618	X	X	X	X	X	X	X	X											
9	45.7	44.2	41.7	22.0	2.4508	X	X	X	X	X	X			X	X	X								
10	46.4	44.8	42.1	19.7	2.4392	X	X	X	X	X	X			X	X	X	X							
11	47.0	45.2	42.3	17.8	2.4289	X	X	X	X	X	X	X		X	X	X	X							
12	47.6	45.7	42.5	16.0	2.4189	X	X	X	X	X	X	X	X	X	X	X	X							
13	48.1	46.0	42.8	15.0	2.4115	X	X	X	X	X	X	X	X	X	X	X	X	X						
14	48.4	46.2	42.9	15.0	2.4079	X	X	X	X	X	X	X	X	X	X	X	X	X	X					
15	48.6	46.2	42.8	15.8	2.4071	X	X	X	X	X	X	X	X	X	X	X	X	X	X					
16	48.8	46.3	42.8	16.2	2.4050	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X		
17	49.0	46.4	42.7	16.9	2.4038	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
18	49.2	46.3	42.5	18.1	2.4044	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
19	49.3	46.3	42.3	19.3	2.4053	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X
20	49.3	46.2	42.0	21.0	2.4079	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Best Subsets - Graphical Adjusted R^2



- **Math**: Regression with Math final grade as dependent variable, excludes G1/G2 as explanatory variables.
- **Portuguese**: Regression with Portuguese final grade as dependent variable, excludes G1/G2 as explanatory variables.

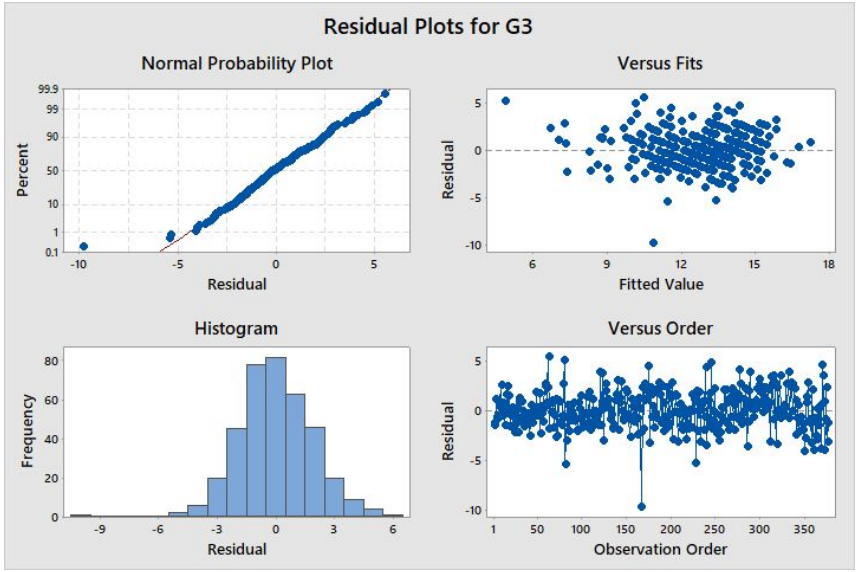
Best Model - Portuguese Final Score

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.91887	46.35%	44.88%	42.15%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	3.10	1.78	1.74	0.083	
G3_math	0.2381	0.0229	10.38	0.000	1.16
failures	-0.877	0.217	-4.04	0.000	1.22
sex_M	-0.990	0.215	-4.59	0.000	1.19
schoolsup_yes	-0.952	0.311	-3.06	0.002	1.14
Dalc	-0.454	0.121	-3.76	0.000	1.14
age	0.3648	0.0944	3.86	0.000	1.24
higher_yes	1.859	0.522	3.56	0.000	1.20
Medu	0.2918	0.0961	3.03	0.003	1.10
absences	-0.0566	0.0209	-2.71	0.007	1.08
Fjob_at_home	1.079	0.494	2.18	0.030	1.01



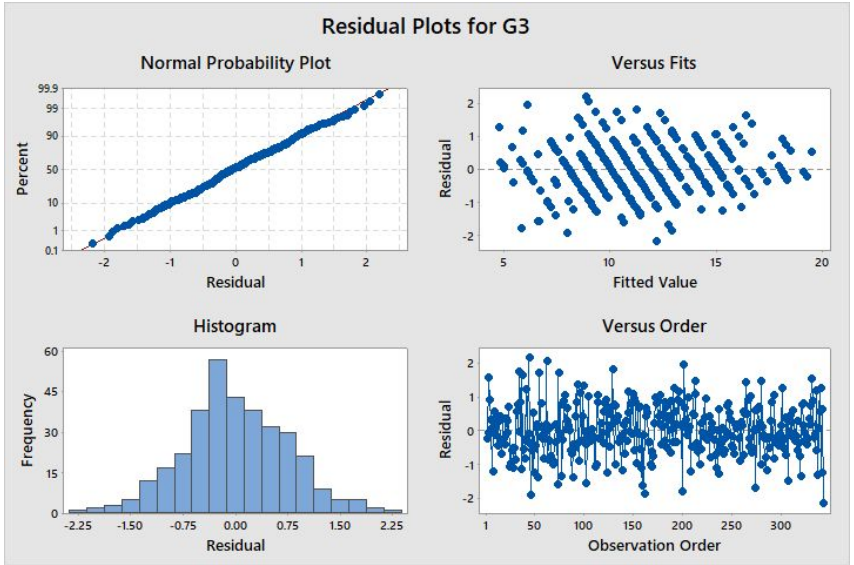
Best Model - Math Final Score

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
2.45397	45.73%	44.10%	41.05%

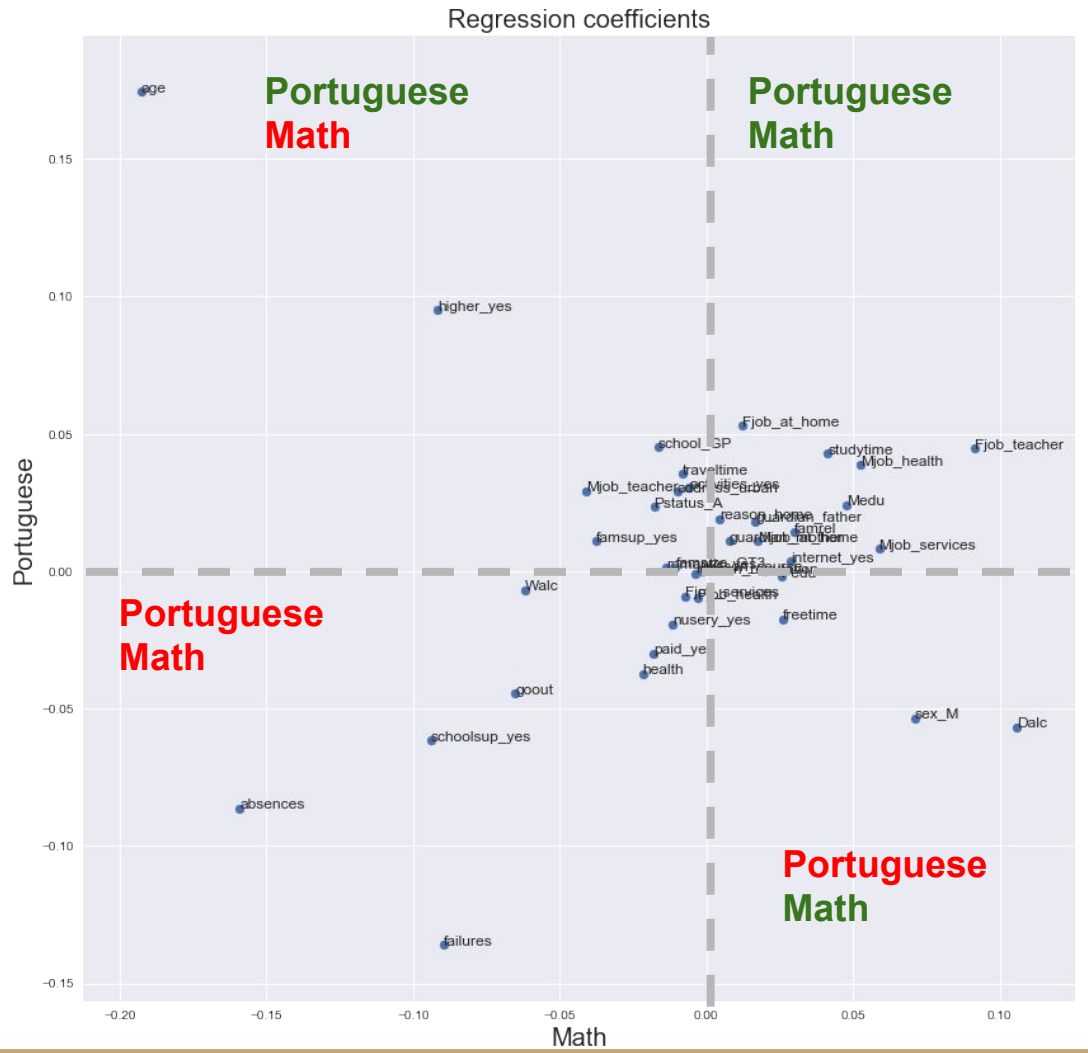
Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	10.87	2.18	4.99	0.000	
G3_port	0.6016	0.0540	11.14	0.000	1.33
sex_M	1.126	0.290	3.88	0.000	1.20
age	-0.430	0.123	-3.49	0.001	1.15
schoolsup_yes	-1.360	0.406	-3.35	0.001	1.17
Fjob_teacher	1.470	0.498	2.95	0.003	1.02
Mjob_services	0.925	0.307	3.02	0.003	1.03
absences	-0.0413	0.0177	-2.34	0.020	1.08
failures	-0.486	0.227	-2.14	0.033	1.18
Dalc	0.101	0.158	0.64	0.522	1.19
famsup_yes	-0.548	0.280	-1.96	0.051	1.05

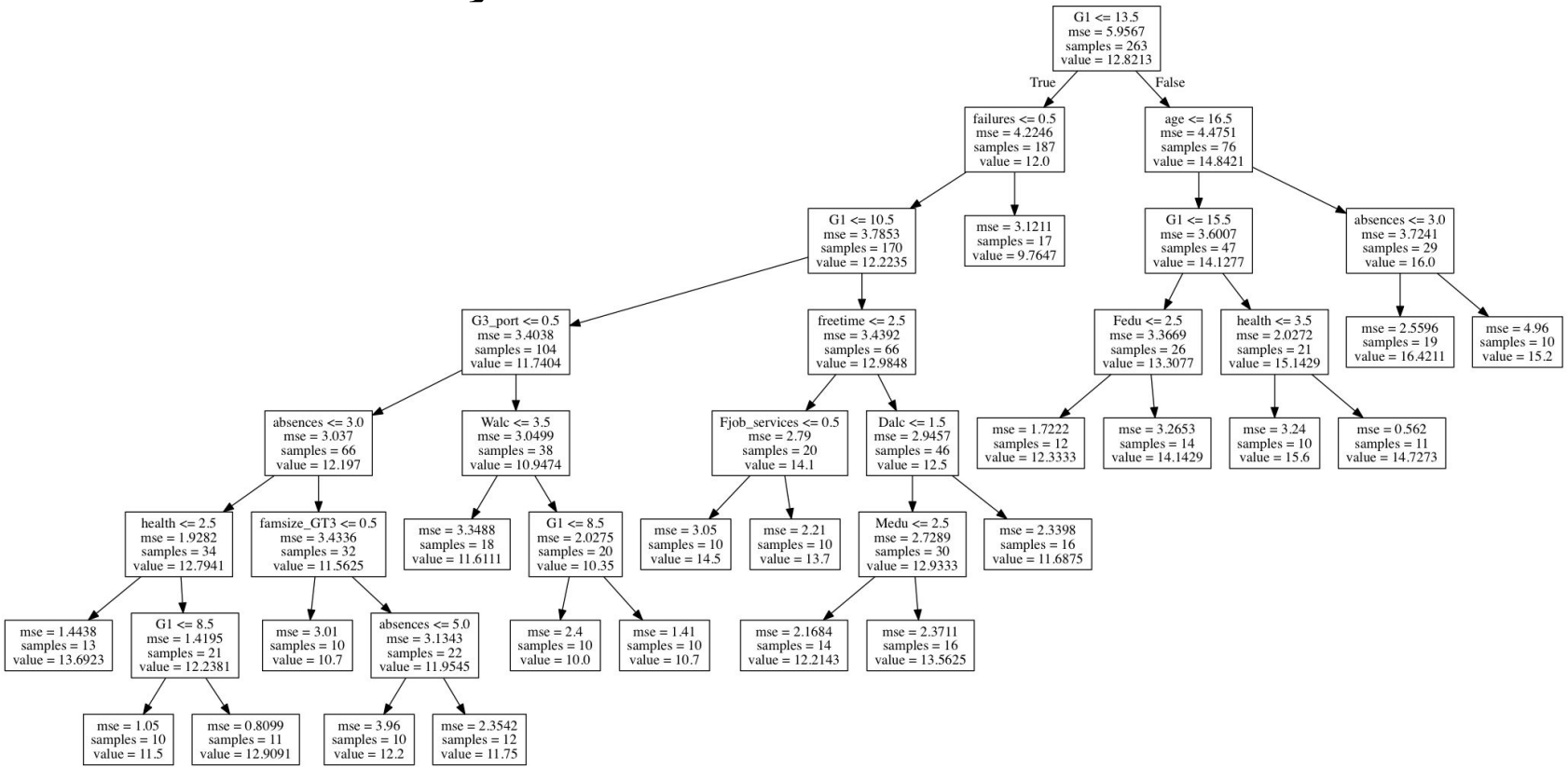


Normalizing Inde-pendent Variables

- Wanna rock at Math?
  - Drink during the week
  - Have more free time
  - Be a man
- Wanna rock at Portuguese?
  - Be old
  - Have academic aspirations
  - Be a woman
- In any case:
  - Don’t miss class!
  - Do well in previous classes
  - Make sure you don’t need additional school support



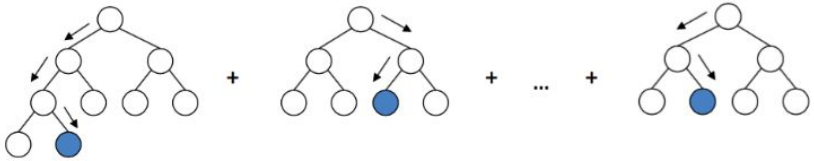
Predictive analysis - Decision Trees



Why use Gradient Boosting Regression?

https://gormanalysis.com/gradient-boosting-explained/

Gradient-Boosting Regression

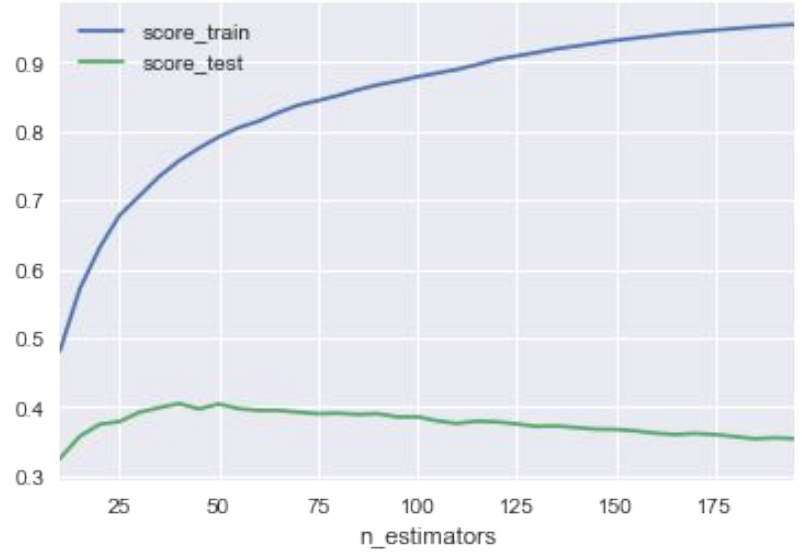


$$D(\mathbf{x}) = d_{\text{tree }1}(\mathbf{x}) + d_{\text{tree }2}(\mathbf{x}) + \ldots$$

$$D(\mathbf{x}) = d_{\text{tree }1}(\mathbf{x}) + d_{\text{tree }2}(\mathbf{x}) + d_{\text{tree }3}(\mathbf{x})$$

$$D(\mathbf{x}) + d_{\text{tree }4}(\mathbf{x}) = f(\mathbf{x}).$$

$$R(\mathbf{x}) = f(\mathbf{x}) - D(\mathbf{x}).$$



	n_estimators	score_test	score_train
6	40	41%	76%
8	50	40%	79%
5	35	40%	73%

Feature importance

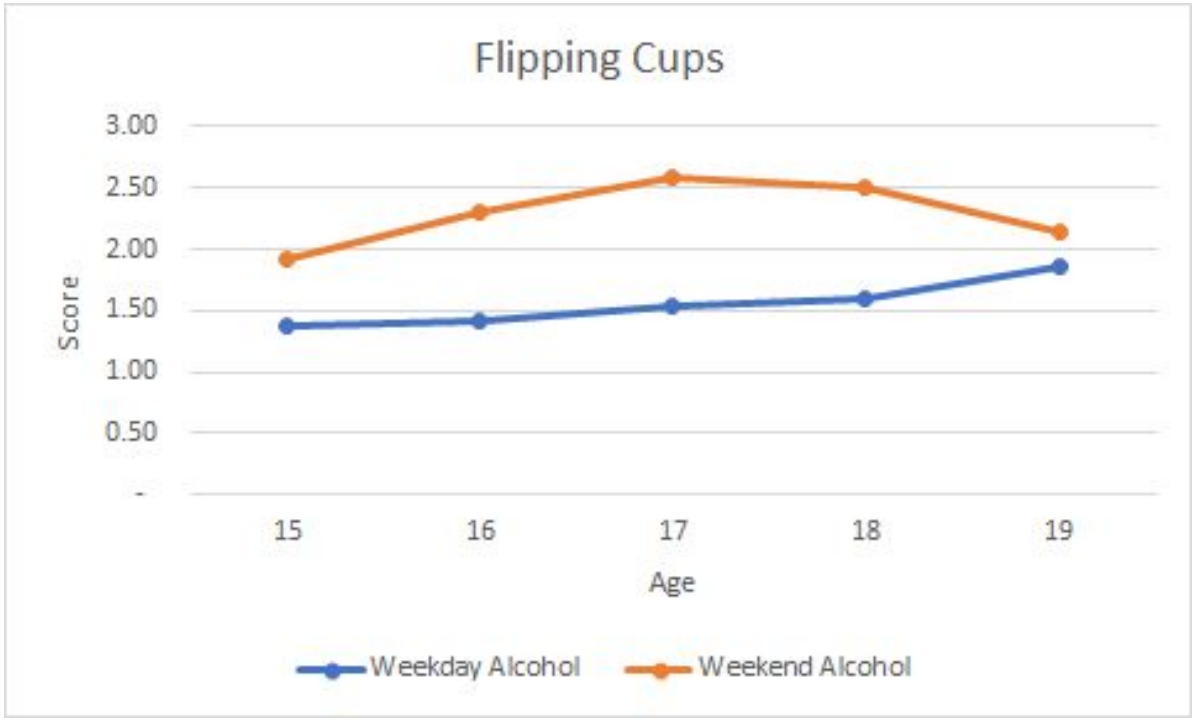
- Portuguese / Math grade excluded
- Important for both:
  - Absence
  - Age
  - Failures
  - School support
  - Sex
- Portuguese:
  - Academic ambition
  - Weekday drinking
  - Health
- Math:
  - Family relations

Summary

Team 7's Valuable Lessons:

1. Performance in one class might be the best indicator for performance in another class.
2. If you want to do well, go to class.
3. Past failures are a good predictor of future performance.
4. Math & Portuguese students aren't all that different.
5. I still can't speak math.

Appendix A: 15-year-olds Drink?



Appendix B: Definitions

Attributes for both student-mat.csv (Math course) and student-por.csv (Portuguese language course) datasets:
1 school - student's school (binary: "GP" - Gabriel Pereira or "MS" - Mousinho da Silveira)
2 sex - student's sex (binary: "F" - female or "M" - male)
3 age - student's age (numeric: from 15 to 22)
4 address - student's home address type (binary: "U" - urban or "R" - rural)
5 famsize - family size (binary: "LE3" - less or equal to 3 or "GT3" - greater than 3)
6 Pstatus - parent's cohabitation status (binary: "T" - living together or "A" - apart)
7 Medu - mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 – 5th to 9th grade, 3 – secondary education or 4 – higher education)
8 Fedu - father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 – 5th to 9th grade, 3 – secondary education or 4 – higher education)
9 Mjob - mother's job (nominal: "teacher", "health" care related, civil "services" (e.g. administrative or police), "at_home" or "other")
10 Fjob - father's job (nominal: "teacher", "health" care related, civil "services" (e.g. administrative or police), "at_home" or "other")
11 reason - reason to choose this school (nominal: close to "home", school "reputation", "course" preference or "other")
12 guardian - student's guardian (nominal: "mother", "father" or "other")
13 traveltime - home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)
14 studytime - weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)
15 failures - number of past class failures (numeric: n if 1<=n<3, else 4)
16 schoolsup - extra educational support (binary: yes or no)
17 famsup - family educational support (binary: yes or no)

Appendix B: Ctd. Definitions

18 paid - extra paid classes within the course subject (Math or Portuguese) (binary: yes or no)
19 activities - extra-curricular activities (binary: yes or no)
20 nursery - attended nursery school (binary: yes or no)
21 higher - wants to take higher education (binary: yes or no)
22 internet - Internet access at home (binary: yes or no)
23 romantic - with a romantic relationship (binary: yes or no)
24 famrel - quality of family relationships (numeric: from 1 - very bad to 5 - excellent)
25 freetime - free time after school (numeric: from 1 - very low to 5 - very high)
26 goout - going out with friends (numeric: from 1 - very low to 5 - very high)
27 Dalc - workday alcohol consumption (numeric: from 1 - very low to 5 - very high)
28 Walc - weekend alcohol consumption (numeric: from 1 - very low to 5 - very high)
29 health - current health status (numeric: from 1 - very bad to 5 - very good)
30 absences - number of school absences (numeric: from 0 to 93)