

Intelligence

Jeffrey Cheng

5/20/2019

INTRODUCTION: This project analyzes a data set of 582 high school students who speak Portuguese. Students took both a Portuguese and a Mathematics course over the year and their quarterly and final grades are displayed, alongside various lifestyle statistics that the students were polled on. The question I seek to answer is whether intelligence is general (i.e. a higher math score would correlate with a higher Portuguese score) or whether intelligence is more sectionalized (i.e. a higher math score would not necessarily be correlated with a higher Portuguese score).

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.2.1 --

## v ggplot2 3.1.0      v purrr  0.3.0
## v tibble  2.0.1      v dplyr  0.8.0.1
## v tidyr   0.8.2      v stringr 1.3.1
## v readr   1.3.1      v forcats 0.3.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(dplyr)
library(ggplot2)

#load data
data_math <- read.table("~/Desktop/student/student-mat.csv",sep=";",header=TRUE)
data_por <- read.table("~/Desktop/student/student-por.csv",sep=";",header=TRUE)
data <- merge(data_math,data_por,by=c("school","sex","age","address","famsize","Pstatus","Medu","Fedu"),

#examine data
head(data)

##   school sex age address famsize Pstatus Medu Fedu   Mjob   Fjob
## 1    GP   F  15      R    GT3      T    1    1 at_home  other
## 2    GP   F  15      R    GT3      T    1    1  other  other
## 3    GP   F  15      R    GT3      T    2    2 at_home  other
## 4    GP   F  15      R    GT3      T    2    4 services health
## 5    GP   F  15      R    GT3      T    3    3 services services
## 6    GP   F  15      R    GT3      T    3    4 services health
##   reason nursery internet guardian.x traveltime.x studytime.x
## 1      home     yes      yes      mother          2            4
## 2 reputation    no      yes      mother          1            2
## 3 reputation    yes     no      mother          1            1
## 4   course     yes     yes      mother          1            3
## 5 reputation    yes     yes     other          2            3
## 6   course     yes     yes      mother          1            3
##   failures.x schoolsup.x famsup.x paid.x activities.x higher.x romantic.x
## 1          1          yes      yes    yes          yes      yes      no
## 2          2          yes      yes    no          no      yes      yes
## 3          0          yes      yes    yes          yes      yes      no
```

```

## 4      0      yes      yes      yes      yes      yes      no
## 5      2      no       yes      yes      yes      yes      yes
## 6      0      yes      yes      yes      yes      yes      no
##   famrel.x freetime.x goout.x Dalc.x Walc.x health.x absences.x G1.x G2.x
## 1      3      1       2       1       1       1       2       7      10
## 2      3      3       4       2       4       5       2       8       6
## 3      4      3       1       1       1       2       8      14      13
## 4      4      3       2       1       1       5       2      10       9
## 5      4      2       1       2       3       3       8      10      10
## 6      4      3       2       1       1       5       2      12      12
##   G3.x guardian.y traveltime.y studytime.y failures.y schoolsup.y famsup.y
## 1   10     mother           2           4           0       yes      yes
## 2    5     mother           1           2           0       yes      yes
## 3   13     mother           1           1           0       yes      yes
## 4    8     mother           1           3           0       yes      yes
## 5   10      other           2           3           0       no       yes
## 6   11     mother           1           3           0       yes      yes
##   paid.y activities.y higher.y romantic.y famrel.y freetime.y goout.y
## 1   yes           yes      yes           no       3           1       2
## 2   no            no       yes           yes       3           3       4
## 3   no            yes      yes           no       4           3       1
## 4   no            yes      yes           no       4           3       2
## 5   yes           yes      yes           yes       4           2       1
## 6   no            yes      yes           no       4           3       2
##   Dalc.y Walc.y health.y absences.y G1.y G2.y G3.y
## 1     1     1       1           4    13    13    13
## 2     2     4       5           2    13    11    11
## 3     1     1       2           8    14    13    12
## 4     1     1       5           2    10    11    10
## 5     2     3       3           2    13    13    13
## 6     1     1       5           2    11    12    12

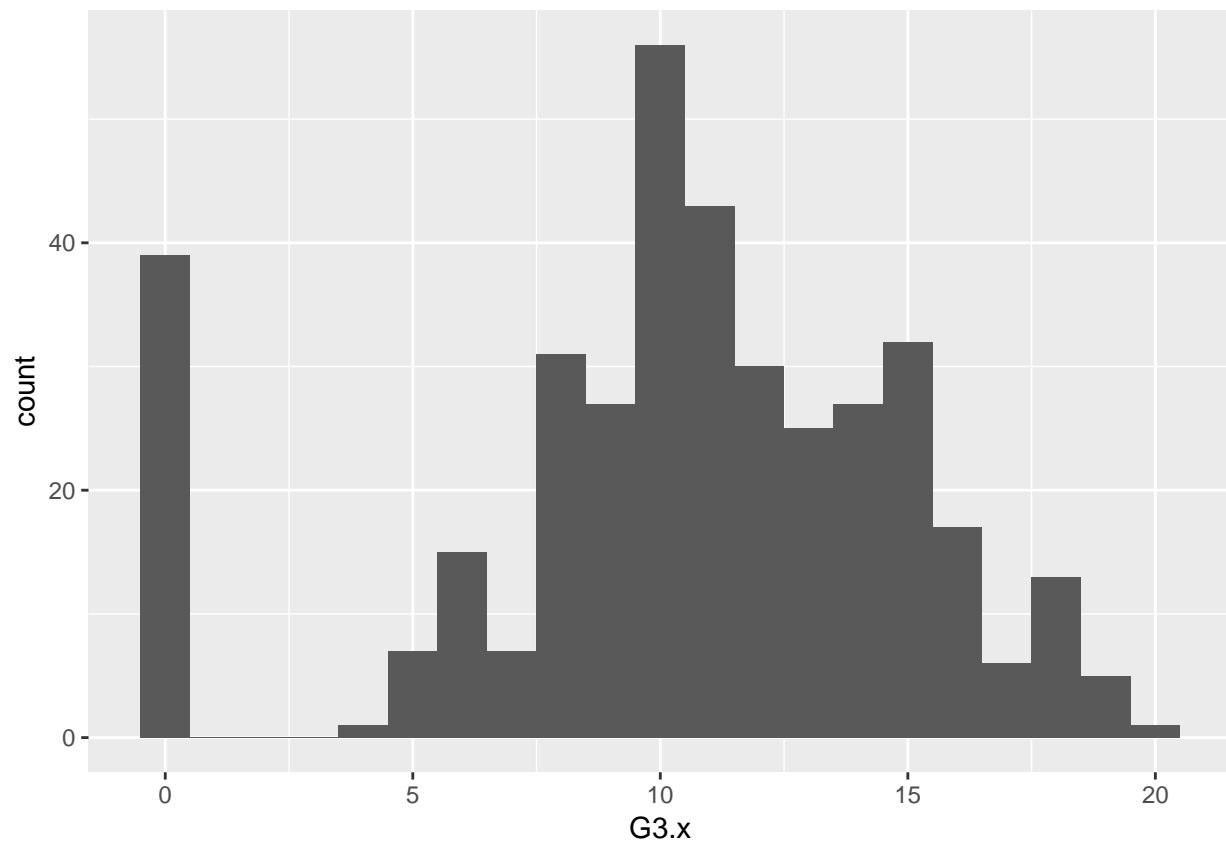
```

```
dim(data)
```

```
## [1] 382 53
```

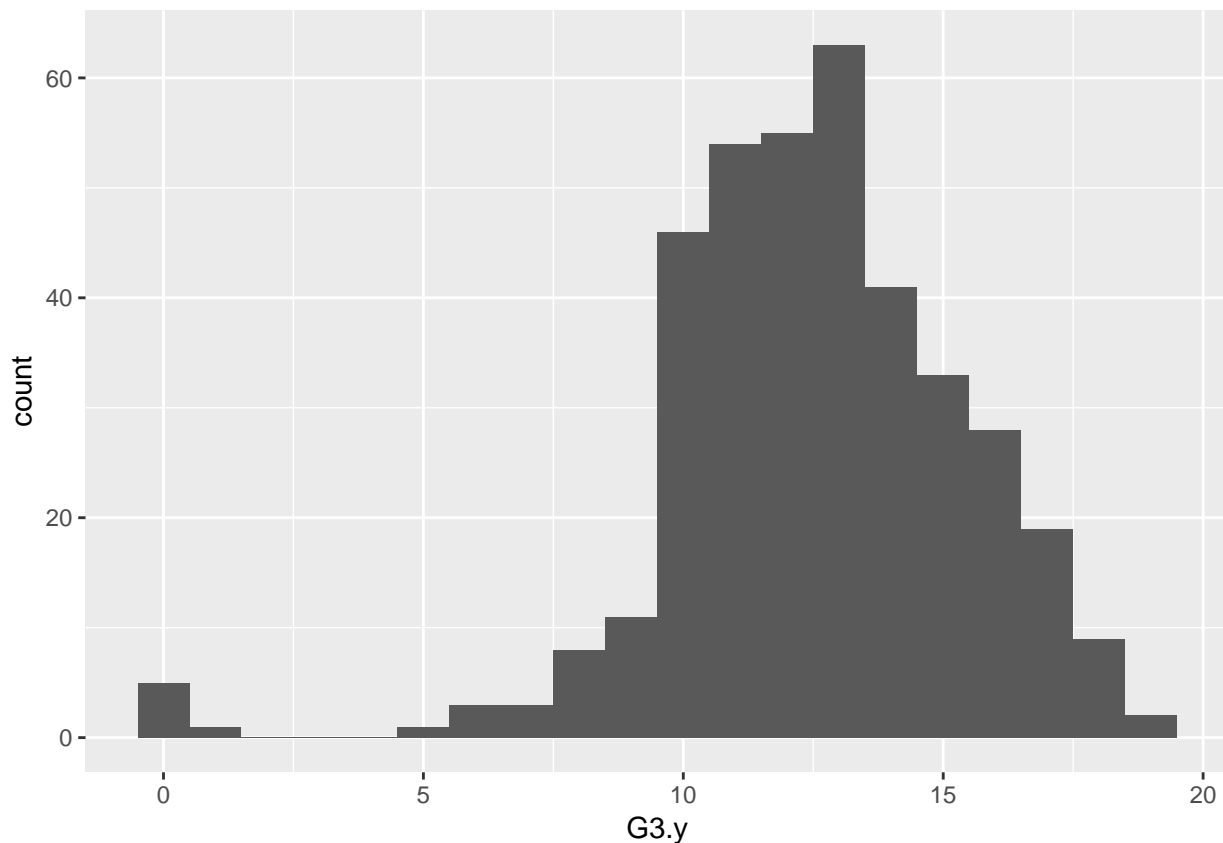
```
ggplot(data, aes(x=G3.x)) + geom_bar(binwidth = 1)
```

```
## Warning: `geom_bar()` no longer has a `binwidth` parameter. Please use
## `geom_histogram()` instead.
```



```
ggplot(data, aes(x=G3.y)) + geom_bar(binwidth = 1)
```

```
## Warning: `geom_bar()` no longer has a `binwidth` parameter. Please use  
## `geom_histogram()` instead.
```



Math scores are generally lower than Portugese scores, with more failures as well.

```
correlation <- lm(G3.x ~ G3.y, data)
summary(correlation)
```

```
##
## Call:
## lm(formula = G3.x ~ G3.y, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.2864  -1.5220   0.4491   2.5356  10.4153
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.8203     0.9205   0.891   0.373
## G3.y           0.7644     0.0716  10.676 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.116 on 380 degrees of freedom
## Multiple R-squared:  0.2307, Adjusted R-squared:  0.2287
## F-statistic: 114 on 1 and 380 DF, p-value: < 2.2e-16
confint(correlation)

##              2.5 %   97.5 %
## (Intercept) -0.9896571 2.630338
## G3.y         0.6236251 0.905189
```

We can see here that math scores and Portuguese scores are clearly positively correlated. The confidence interval shows us that we are extremely confident that the correlation between math and Portuguese scores is actually positive. However, the R^2 value is fairly low, showing that this may be due to factors other than a “general intelligence” (i.e. this correlation could be due to other factors). We are going to investigate this further by looking at “splitters” (students with a good score for one course and a low score for another).

```
lms <- data %>%
  filter(G3.x < 6, G3.y > 10)
dim(lms)
```

```
## [1] 26 53
```

```
head(lms)
```

```
##   school sex age address famsize Pstatus Medu Fedu   Mjob   Fjob
## 1    GP   F  15      R    GT3      T    1    1   other   other
## 2    GP   F  15      R    GT3      T    3    4 services teacher
## 3    GP   F  15      U    GT3      T    1    1 at_home   other
## 4    GP   F  15      U    GT3      T    3    2 health services
## 5    GP   F  15      U    GT3      T    4    4 services at_home
## 6    GP   F  16      U    GT3      T    1    3 at_home services
##   reason nursery internet guardian.x traveltime.x studytime.x
## 1 reputation      no      yes      mother          1          2
## 2   course      yes      yes      father          2          3
## 3   course      no      yes      mother          3          1
## 4    home      yes      yes      father          1          2
## 5   course      yes      yes      mother          1          3
## 6    home      no      yes      mother          1          2
##   failures.x schoolsup.x famsup.x paid.x activities.x higher.x romantic.x
## 1          2          yes      yes      no          no      yes      yes
## 2          2          no      yes      no          no      yes      yes
## 3          0          no      yes      no          yes      yes      yes
## 4          3          no      yes      no          no      yes      no
## 5          0          no      yes      no          yes      yes      yes
## 6          3          no      no      no          yes      yes      yes
##   famrel.x freetime.x goout.x Dalc.x Walc.x health.x absences.x G1.x G2.x
## 1          3          3          4          2          4          5          2      8      6
## 2          4          2          2          2          2          5          0     12      0
## 3          4          3          3          1          2          4          0      8      0
## 4          3          3          2          1          1          3          0      6      7
## 5          4          3          3          1          1          5          0     11      0
## 6          4          3          5          1          1          3          0      8      7
##   G3.x guardian.y traveltime.y studytime.y failures.y schoolsup.y famsup.y
## 1    5      mother          1          2          0          yes      yes
## 2    0      father          2          3          0          no      yes
## 3    0      mother          3          1          0          no      yes
## 4    0      father          1          2          1          no      yes
## 5    0      mother          1          3          0          no      yes
## 6    0      mother          1          2          0          no      no
##   paid.y activities.y higher.y romantic.y famrel.y freetime.y goout.y
## 1    no          no      yes      yes          3          3          4
## 2    no          no      yes      yes          4          2          2
## 3    no          yes      yes      yes          4          3          3
## 4    no          no      yes      no          3          3          2
## 5    no          yes      yes      yes          4          3          3
## 6    no          yes      yes      yes          4          3          5
```

```
## Dalc.y Walc.y health.y absences.y G1.y G2.y G3.y
## 1      2      4      5      2 13 11 11
## 2      2      2      5      0 10 11 12
## 3      1      2      4      6 11 12 13
## 4      1      1      3      2 11 11 11
## 5      1      1      5      4 13 14 15
## 6      1      1      3      0 14 13 13
```

```
lps <- data %>%
  filter(G3.x > 10, G3.y < 6)
dim(lps)
```

```
## [1] 2 53
```

```
head(lps)
```

```
## school sex age address famsize Pstatus Medu Fedu Mjob Fjob reason
## 1 GP M 16 U GT3 T 3 3 other services course
## 2 MS F 18 R GT3 T 4 4 other teacher other
## nursery internet guardian.x traveltime.x studytime.x failures.x
## 1 yes yes father 1 2 1
## 2 no yes father 3 2 0
## schoolsup.x famsup.x paid.x activities.x higher.x romantic.x famrel.x
## 1 no yes yes no yes yes 4
## 2 no yes yes no yes yes 3
## freetime.x goout.x Dalc.x Walc.x health.x absences.x G1.x G2.x G3.x
## 1 5 5 4 4 5 4 10 12 12
## 2 2 2 4 2 5 10 14 12 11
## guardian.y traveltime.y studytime.y failures.y schoolsup.y famsup.y
## 1 father 1 2 1 no yes
## 2 father 3 2 0 no yes
## paid.y activities.y higher.y romantic.y famrel.y freetime.y goout.y
## 1 no no yes yes 4 5 5
## 2 no no yes yes 3 2 2
## Dalc.y Walc.y health.y absences.y G1.y G2.y G3.y
## 1 4 4 5 0 10 10 1
## 2 4 2 5 0 7 5 0
```

There are only 2 students with a high math score and a low Portuguese score. Both of these students failed to take the final, leaving them with failing scores for the class. One student was doing strongly before the final, but other other was actually what we would define as a splitter based off of his quarterly grades. However, this is only the case for 1/382 students and his/her result cannot be extrapolated. What is more interesting to look at is the low math splitters, of which there are 26.

```
failcorr_por <- lm(G3.y ~ failures.y, data)
summary(failcorr_por)
```

```
##
## Call:
## lm(formula = G3.y ~ failures.y, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.7944  -1.7944   0.2056   2.1984   6.2056
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) 12.7944      0.1470   87.03 < 2e-16 ***
## failures.y  -1.9713      0.2765   -7.13 5.1e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.77 on 380 degrees of freedom
## Multiple R-squared:  0.118, Adjusted R-squared:  0.1157
## F-statistic: 50.83 on 1 and 380 DF, p-value: 5.1e-12

failcorr_lms_por <- lm(G3.y ~ failures.y, lms)
summary(failcorr_lms_por)
```

```
##
## Call:
## lm(formula = G3.y ~ failures.y, data = lms)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.72  -0.72   0.14   1.03   2.28
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   12.720      0.280  45.429 <2e-16 ***
## failures.y    -1.720      1.428  -1.205    0.24
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.4 on 24 degrees of freedom
## Multiple R-squared:  0.05702, Adjusted R-squared:  0.01773
## F-statistic: 1.451 on 1 and 24 DF, p-value: 0.2401

failcorr_math <- lm(G3.x ~ failures.x, data)
summary(failcorr_math)
```

```
##
## Call:
## lm(formula = G3.x ~ failures.x, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.0983  -2.0983  -0.0983   2.9017   9.3481
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   11.0983      0.2391  46.425 < 2e-16 ***
## failures.x    -2.4464      0.3048  -8.027 1.26e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.34 on 380 degrees of freedom
## Multiple R-squared:  0.145, Adjusted R-squared:  0.1427
## F-statistic: 64.43 on 1 and 380 DF, p-value: 1.264e-14

failcorr_lms_math <- lm(G3.x ~ failures.x, lms)
summary(failcorr_lms_math)
```

```
##
## Call:
## lm(formula = G3.x ~ failures.x, data = lms)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.6759 -0.4276 -0.3034 -0.3034  4.6966
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.3034     0.3201   0.948   0.353
## failures.x    0.1241     0.2614   0.475   0.639
##
## Residual standard error: 1.38 on 24 degrees of freedom
## Multiple R-squared:  0.00931,    Adjusted R-squared:  -0.03197
## F-statistic: 0.2255 on 1 and 24 DF,  p-value: 0.6391
```

```
print(lms)
```

```
##      school sex age address famsize Pstatus Medu Fedu      Mjob      Fjob
## 1      GP   F  15      R      GT3      T    1    1    other    other
## 2      GP   F  15      R      GT3      T    3    4  services  teacher
## 3      GP   F  15      U      GT3      T    1    1  at_home    other
## 4      GP   F  15      U      GT3      T    3    2  health  services
## 5      GP   F  15      U      GT3      T    4    4  services  at_home
## 6      GP   F  16      U      GT3      T    1    3  at_home  services
## 7      GP   F  16      U      GT3      T    2    2    other    other
## 8      GP   F  16      U      GT3      T    3    4  at_home    other
## 9      GP   F  17      U      GT3      A    2    2  at_home  at_home
## 10     GP   F  17      U      GT3      T    1    1  at_home    other
## 11     GP   F  17      U      GT3      T    3    2    other    other
## 12     GP   F  17      U      LE3      T    2    2  services  services
## 13     GP   F  18      R      GT3      T    2    2  at_home    other
## 14     GP   F  18      U      GT3      T    2    1    other    other
## 15     GP   F  18      U      GT3      T    2    1  services    other
## 16     GP   F  18      U      GT3      T    2    2  at_home  services
## 17     GP   F  18      U      GT3      T    3    3  services  services
## 18     GP   F  18      U      LE3      T    2    2    other    other
## 19     GP   F  19      U      GT3      T    4    4  health    other
## 20     GP   M  15      R      GT3      T    3    4  at_home  teacher
## 21     GP   M  15      U      GT3      T    4    3  teacher  services
## 22     GP   M  16      U      GT3      T    4    4  teacher  teacher
## 23     GP   M  16      U      LE3      T    4    3  teacher    other
## 24     GP   M  18      U      GT3      T    4    4  teacher  services
## 25     GP   M  18      U      GT3      T    4    4  teacher  services
## 26     GP   M  19      U      GT3      T    3    2  services  at_home
##
##      reason nursery internet guardian.x traveltime.x studytime.x
## 1  reputation      no      yes    mother              1              2
## 2    course      yes      yes    father              2              3
## 3    course      no      yes    mother              3              1
## 4      home      yes      yes    father              1              2
## 5    course      yes      yes    mother              1              3
## 6      home      no      yes    mother              1              2
## 7      home      no      yes    mother              1              2
## 8    course      yes      yes    mother              1              2
```


## 9	home	yes	yes	father	1	2		
## 10	reputation	yes	no	mother	1	3		
## 11	home	yes	yes	mother	1	2		
## 12	course	yes	yes	father	1	4		
## 13	course	yes	no	mother	2	4		
## 14	course	no	yes	other	2	3		
## 15	course	yes	yes	mother	2	2		
## 16	home	yes	yes	mother	1	3		
## 17	home	yes	yes	mother	1	2		
## 18	home	no	yes	other	1	2		
## 19	reputation	yes	yes	other	2	2		
## 20	course	yes	no	mother	4	2		
## 21	course	yes	yes	father	2	4		
## 22	course	yes	yes	mother	1	1		
## 23	course	no	yes	mother	1	1		
## 24	home	yes	yes	father	1	2		
## 25	home	yes	yes	father	1	2		
## 26	home	yes	yes	mother	1	1		
##	failures.x	schoolsup.x	famsup.x	paid.x	activities.x	higher.x	romantic.x	
## 1	2	yes	yes	no	no	yes	yes	
## 2	2	no	yes	no	no	yes	yes	
## 3	0	no	yes	no	yes	yes	yes	
## 4	3	no	yes	no	no	yes	no	
## 5	0	no	yes	no	yes	yes	yes	
## 6	3	no	no	no	yes	yes	yes	
## 7	0	no	yes	yes	no	yes	no	
## 8	0	no	yes	no	no	yes	no	
## 9	1	no	yes	no	no	yes	yes	
## 10	1	no	yes	no	yes	yes	yes	
## 11	0	no	yes	yes	no	yes	yes	
## 12	0	no	no	yes	yes	yes	yes	
## 13	0	no	no	no	yes	yes	no	
## 14	0	no	yes	yes	no	yes	yes	
## 15	0	no	yes	yes	yes	yes	no	
## 16	0	no	yes	yes	yes	yes	yes	
## 17	0	no	no	no	yes	yes	no	
## 18	0	no	no	no	yes	yes	yes	
## 19	0	no	yes	yes	yes	yes	no	
## 20	0	no	yes	no	no	yes	yes	
## 21	0	yes	yes	no	no	yes	no	
## 22	0	no	yes	no	no	no	yes	
## 23	0	no	no	no	yes	yes	no	
## 24	1	no	yes	no	yes	yes	no	
## 25	1	no	yes	no	yes	yes	no	
## 26	3	no	yes	no	no	no	yes	
##	famrel.x	freetime.x	goout.x	Dalc.x	Walc.x	health.x	absences.x	G1.x G2.x
## 1	3	3	4	2	4	5	2 8 6	
## 2	4	2	2	2	2	5	0 12 0	
## 3	4	3	3	1	2	4	0 8 0	
## 4	3	3	2	1	1	3	0 6 7	
## 5	4	3	3	1	1	5	0 11 0	
## 6	4	3	5	1	1	3	0 8 7	
## 7	5	1	5	1	1	4	0 6 7	
## 8	2	4	3	1	2	3	12 5 5	

## 9	3	3	1	1	2	4	0	9	8
## 10	4	3	4	1	1	5	0	6	5
## 11	4	3	2	2	3	2	0	7	8
## 12	3	4	1	1	1	2	0	10	9
## 13	4	4	4	1	1	4	0	10	9
## 14	4	4	4	1	1	3	0	7	0
## 15	5	3	3	1	2	1	0	8	8
## 16	4	3	3	1	1	3	0	9	10
## 17	5	3	4	1	1	4	0	7	0
## 18	4	3	3	1	1	2	0	8	8
## 19	2	3	4	2	3	2	0	10	9
## 20	5	3	3	1	1	5	0	9	0
## 21	2	2	2	1	1	3	0	7	9
## 22	3	3	2	2	1	5	0	7	6
## 23	5	4	5	1	1	3	0	6	0
## 24	4	3	3	2	2	2	0	10	10
## 25	4	3	3	2	2	2	0	10	10
## 26	4	5	4	1	1	4	0	5	0

##	G3.x	guardian.y	traveltime.y	studytime.y	failures.y	schoolsup.y
----	------	------------	--------------	-------------	------------	-------------

## 1	5	mother	1	2	0	yes
## 2	0	father	2	3	0	no
## 3	0	mother	3	1	0	no
## 4	0	father	1	2	1	no
## 5	0	mother	1	3	0	no
## 6	0	mother	1	2	0	no
## 7	0	mother	1	2	0	no
## 8	5	mother	1	2	0	no
## 9	0	father	1	2	0	no
## 10	0	mother	1	3	0	no
## 11	0	mother	1	2	0	no
## 12	0	father	1	4	0	no
## 13	0	mother	2	4	0	no
## 14	0	other	2	3	0	no
## 15	0	mother	2	2	0	no
## 16	0	mother	1	3	0	no
## 17	0	mother	1	2	0	no
## 18	0	other	1	2	0	no
## 19	0	other	2	2	0	no
## 20	0	mother	4	2	0	no
## 21	0	father	2	4	0	yes
## 22	0	father	1	2	0	no
## 23	0	mother	1	1	0	no
## 24	0	father	1	2	0	no
## 25	0	mother	2	1	0	no
## 26	0	mother	1	1	0	no

##	famsup.y	paid.y	activities.y	higher.y	romantic.y	famrel.y	freetime.y
## 1	yes	no	no	yes	yes	3	3
## 2	yes	no	no	yes	yes	4	2
## 3	yes	no	yes	yes	yes	4	3
## 4	yes	no	no	yes	no	3	3
## 5	yes	no	yes	yes	yes	4	3
## 6	no	no	yes	yes	yes	4	3
## 7	yes	no	no	yes	no	5	1
## 8	yes	no	no	yes	no	2	4

## 9	yes	no	no	yes	yes	3	3	
## 10	yes	no	yes	yes	yes	4	3	
## 11	yes	no	no	yes	yes	4	3	
## 12	no	no	yes	yes	yes	3	4	
## 13	no	no	yes	yes	no	4	4	
## 14	yes	no	no	yes	yes	4	4	
## 15	yes	no	yes	yes	no	5	3	
## 16	yes	no	yes	yes	yes	4	3	
## 17	no	no	yes	yes	no	5	3	
## 18	no	no	yes	yes	yes	4	3	
## 19	yes	no	yes	yes	no	2	3	
## 20	yes	no	no	yes	yes	5	3	
## 21	yes	no	no	yes	no	2	2	
## 22	yes	no	yes	yes	no	5	4	
## 23	no	no	yes	yes	no	5	4	
## 24	yes	no	yes	yes	no	4	3	
## 25	no	no	yes	yes	no	3	2	
## 26	yes	no	no	no	yes	4	5	
##	goout.y	Dalc.y	Walc.y	health.y	absences.y	G1.y	G2.y	G3.y
## 1	4	2	4	5	2	13	11	11
## 2	2	2	2	5	0	10	11	12
## 3	3	1	2	4	6	11	12	13
## 4	2	1	1	3	2	11	11	11
## 5	3	1	1	5	4	13	14	15
## 6	5	1	1	3	0	14	13	13
## 7	5	1	1	4	0	12	12	13
## 8	3	1	2	3	14	12	11	11
## 9	1	1	2	4	18	10	12	14
## 10	4	1	1	5	12	12	12	12
## 11	2	2	3	2	0	12	13	15
## 12	1	1	1	2	2	10	11	12
## 13	4	1	1	4	6	14	13	14
## 14	4	1	1	3	10	12	10	11
## 15	3	1	2	1	2	12	12	15
## 16	3	1	1	3	0	11	12	13
## 17	4	1	1	4	8	10	11	12
## 18	3	1	1	2	0	10	9	12
## 19	4	2	3	2	2	14	13	13
## 20	3	1	1	5	2	12	11	11
## 21	2	1	1	3	6	9	11	11
## 22	4	1	2	5	6	16	14	14
## 23	5	1	1	3	7	14	14	15
## 24	3	2	2	2	0	12	12	13
## 25	4	1	4	3	6	11	12	12
## 26	4	1	1	4	6	11	9	11

What's interesting to note here is that for the entire population, the number of previous classes failed is negatively correlated to both math and Portuguese scores. However, low math splitters show no correlation between number of past courses failed and math/Portuguese scores. This implies that these students' struggle is not with the Portuguese content or study habits in general, but with this mathematics content in particular. So we have a good ~10% of this population who have generally good study habits and high scores in Portuguese, but are struggling with this mathematics content.

I'd suspect that this is due to students implicitly practicing Portuguese with friends/family: the low math splitters mostly showed larger family sizes and family support for their studies. The limitation of the study

time metric in this survey is that it does not differentiate between time spent on math and time spent on Portuguese. Perhaps the amount of time studying and interacting with mathematics is simply less because students either a) prefer to study Portuguese or b) practice Portuguese implicitly. So these differences in scores potentially speak to not how “sectionally intelligent” a student is, but to how much work they put into each subject respectively.