

# Final Project

## Introduction

### The Data

We will be using data that has extensive information on secondary school students in their math class.

```
data <- read.csv("data/student-mat.csv")
```

### Creation of New Variables

In order to provide more insight, we saw room to create informative variables based upon the given data.

The given variables Medu and Fedu give information about the student's parents education history. Using this, we created a new variable "first\_gen\_college" that indicates if the student would be a first generation college student if they decided to pursue higher education. This will give more tangible and clear insight to how parental education impacts student's performance.

```
data <- data %>%  
  mutate(first_gen_college = case_when(  
    Medu < 4 & Fedu < 4 ~ "yes",  
    TRUE ~ "no"  
  ))
```

Additionally, many variables are self reported ratings from the students on a scale of 1-5. We decided that instead factoring these variables so that scores of 1-3 would be "low" and scores of 4-5 would be "high" would be beneficial to our analysis as it would be more interpretable in context.

```
data <- data %>%  
  mutate(famrel = case_when(  
    famrel == 1 ~ "low",  
    famrel == 2 ~ "low",  
    famrel == 3 ~ "low",  
    famrel == 4 ~ "high",  
    famrel == 5 ~ "high"  
  ))  
  
data <- data %>%  
  mutate(freetime = case_when(  
    freetime == 1 ~ "low",  
    freetime == 2 ~ "low",  
    freetime == 3 ~ "low",  
    freetime == 4 ~ "high",  
    freetime == 5 ~ "high"  
  ))  
  
data <- data %>%  
  mutate(goout = case_when(  
    goout == 1 ~ "low",  
    goout == 2 ~ "low",  
    goout == 3 ~ "low",
```

```

    goout == 4 ~ "high",
    goout == 5 ~ "high"
  ))

data <- data %>%
  mutate(Dalc = case_when(
    Dalc == 1 ~ "low",
    Dalc == 2 ~ "low",
    Dalc == 3 ~ "low",
    Dalc == 4 ~ "high",
    Dalc == 5 ~ "high"
  ))

data <- data %>%
  mutate(Walc = case_when(
    Walc == 1 ~ "low",
    Walc == 2 ~ "low",
    Walc == 3 ~ "low",
    Walc == 4 ~ "high",
    Walc == 5 ~ "high"
  ))

data <- data %>%
  mutate(health = case_when(
    health == 1 ~ "low",
    health == 2 ~ "low",
    health == 3 ~ "low",
    health == 4 ~ "high",
    health == 5 ~ "high"
  ))

```

Additionally, using information from the famsup and internet variables, we created a variable called “stable\_learning\_env”. If famsup is “yes” and internet is “yes”, then stable\_learning\_env is “yes”, otherwise “no”.

```

data <- data %>%
  mutate(stable_learning_env = case_when(
    internet == "yes" & famsup == "yes" ~ "yes",
    TRUE ~ "no"
  ))

```

Also, we created a new variable “high\_freq\_absent”, which if absences  $\geq 10$  for a student, we considered them a highly frequent student.

```

data <- data %>%
  mutate(high_freq_absent = case_when(
    absences >= 10 ~ "yes",
    TRUE ~ "no"
  ))

```

We also created a “failed” variable, which was “yes” if failures  $> 0$ , and “no” otherwise.

```

data <- data %>%
  mutate(failed = case_when(
    failures > 0 ~ "yes",
    TRUE ~ "no"
  ))

```

))

## Exploratory Data Analysis

```
summary(data)
```

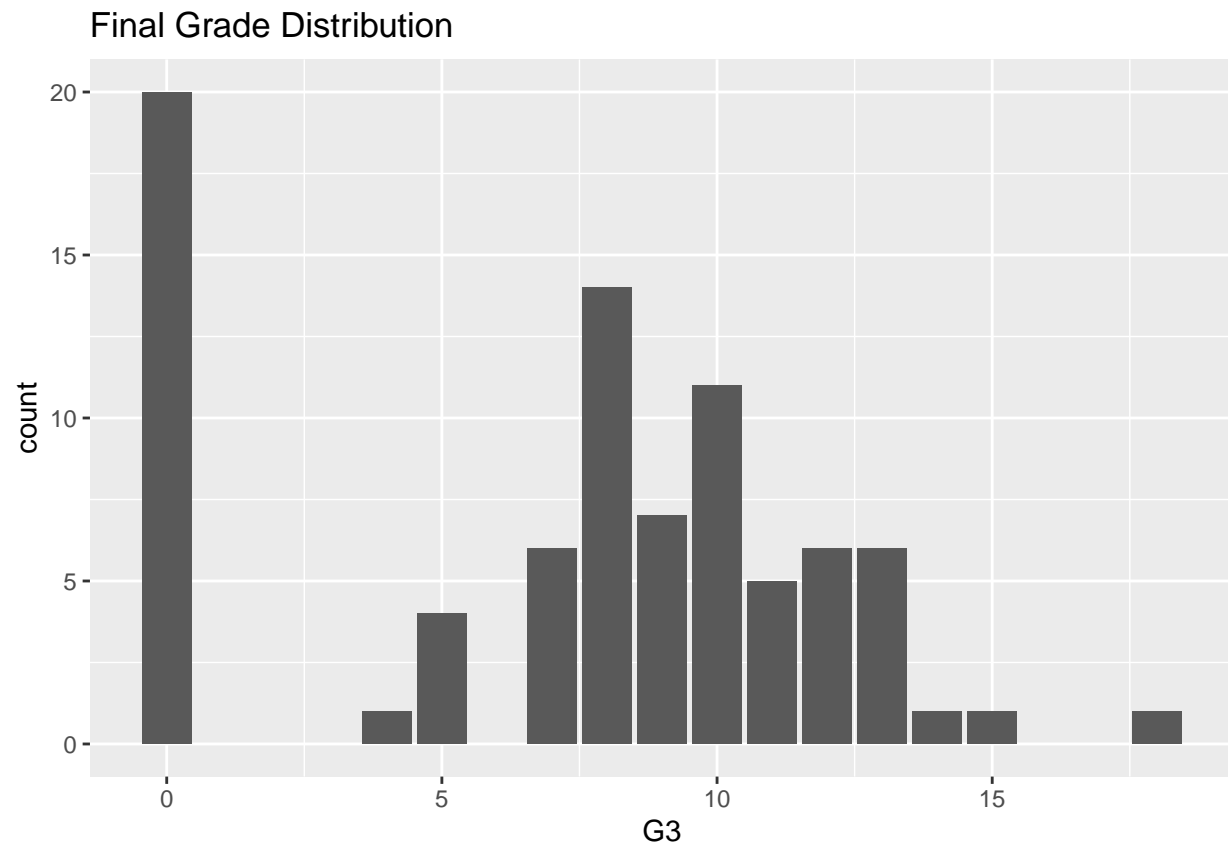
```
##      school      sex      age      address
## Length:395      Length:395      Min.   :15.0      Length:395
## Class :character Class :character 1st Qu.:16.0      Class :character
## Mode  :character Mode  :character Median :17.0      Mode  :character
##                                     Mean  :16.7
##                                     3rd Qu.:18.0
##                                     Max.   :22.0
##      famsize      Pstatus      Medu      Fedu
## Length:395      Length:395      Min.   :0.000      Min.   :0.000
## Class :character Class :character 1st Qu.:2.000      1st Qu.:2.000
## Mode  :character Mode  :character Median :3.000      Median :2.000
##                                     Mean  :2.749      Mean  :2.522
##                                     3rd Qu.:4.000      3rd Qu.:3.000
##                                     Max.   :4.000      Max.   :4.000
##      Mjob      Fjob      reason      guardian
## Length:395      Length:395      Length:395      Length:395
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##
##      traveltime      studytime      failures      schoolsup
## Min.   :1.000      Min.   :1.000      Min.   :0.0000      Length:395
## 1st Qu.:1.000      1st Qu.:1.000      1st Qu.:0.0000      Class :character
## Median :1.000      Median :2.000      Median :0.0000      Mode  :character
## Mean   :1.448      Mean   :2.035      Mean   :0.3342
## 3rd Qu.:2.000      3rd Qu.:2.000      3rd Qu.:0.0000
## Max.   :4.000      Max.   :4.000      Max.   :3.0000
##      famsup      paid      activities      nursery
## Length:395      Length:395      Length:395      Length:395
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##
##      higher      internet      romantic      famrel
## Length:395      Length:395      Length:395      Length:395
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##
##      freetime      goout      Dalc      Walc
## Length:395      Length:395      Length:395      Length:395
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
```

```
##
##
##   health      absences      G1      G2
## Length:395      Min.    : 0.000  Min.    : 3.00  Min.    : 0.00
## Class :character 1st Qu.: 0.000  1st Qu.: 8.00  1st Qu.: 9.00
## Mode  :character Median : 4.000  Median :11.00  Median :11.00
##                Mean   : 5.709  Mean   :10.91  Mean   :10.71
##                3rd Qu.: 8.000  3rd Qu.:13.00  3rd Qu.:13.00
##                Max.    :75.000  Max.    :19.00  Max.    :19.00
##      G3      first_gen_college stable_learning_env high_freq_absent
## Min.    : 0.00  Length:395      Length:395      Length:395
## 1st Qu.: 8.00  Class :character Class :character Class :character
## Median :11.00  Mode  :character Mode  :character Mode  :character
## Mean   :10.42
## 3rd Qu.:14.00
## Max.    :20.00
##   failed
## Length:395
## Class :character
## Mode  :character
##
##
##
```

First, I will start off with univariate and bivariate plots of the response variable and key predictors I see being important.

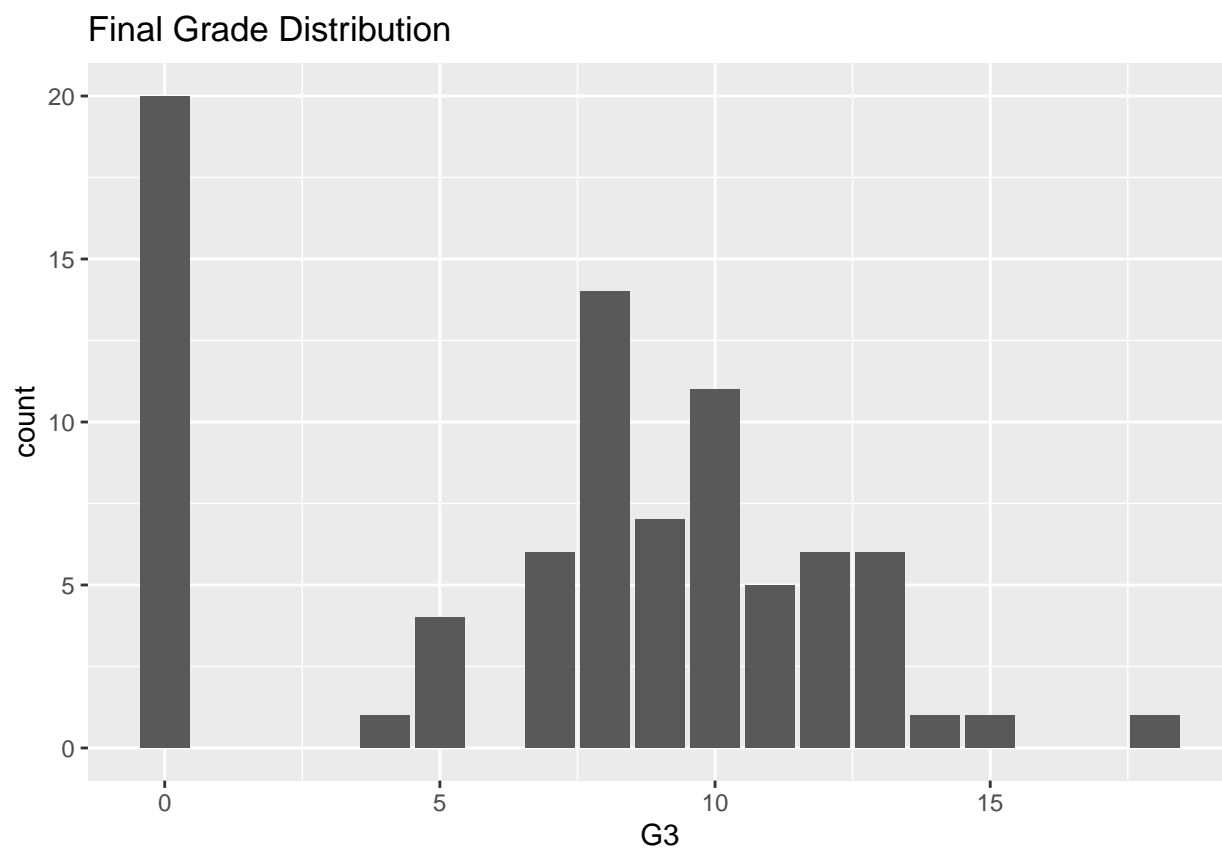
```
data %>%
  filter(failed == "yes") %>%
  ggplot(aes(G3)) +
  geom_histogram(stat = "count") +
  labs(title="Final Grade Distribution")
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



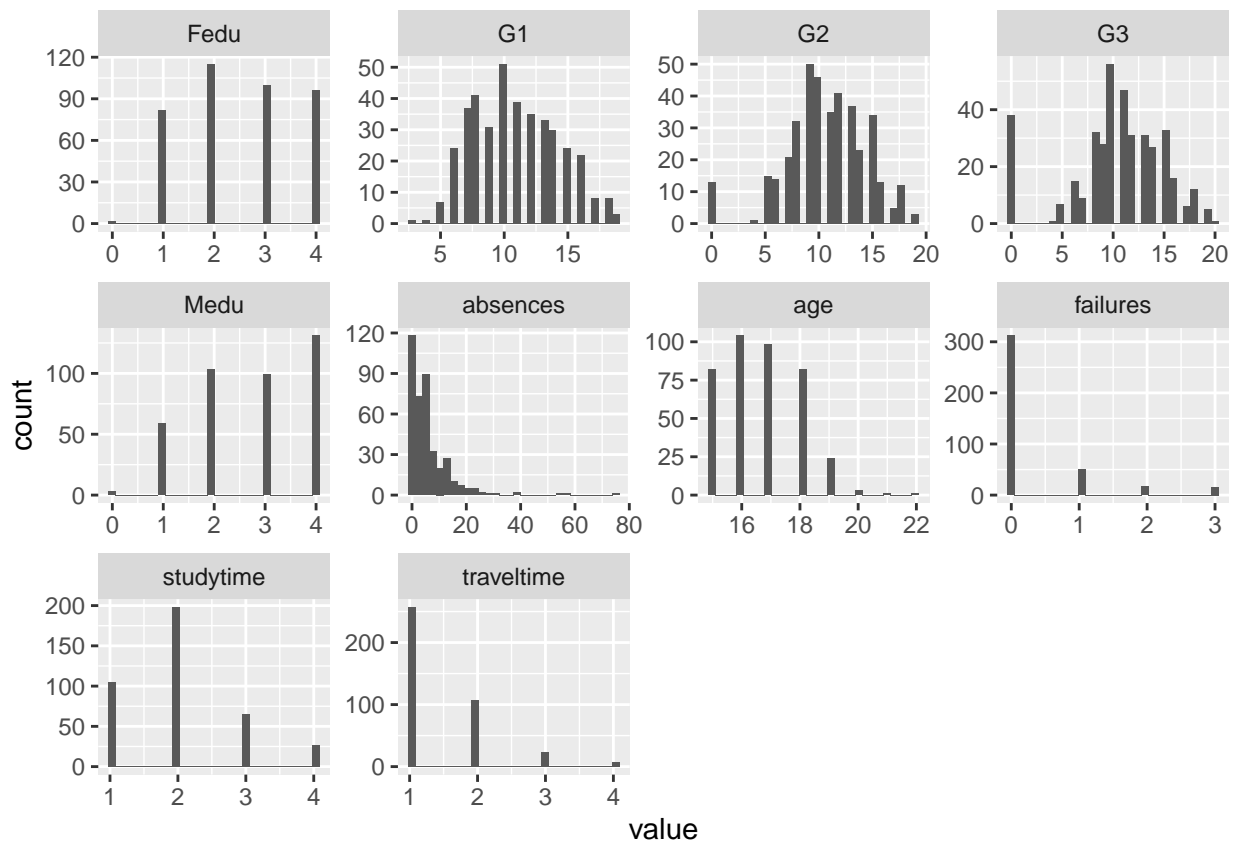
```
data %>%  
  filter(failed == "yes") %>%  
  ggplot(aes(G3)) +  
  geom_histogram(stat = "count") +  
  labs(title = "Final Grade Distribution")
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



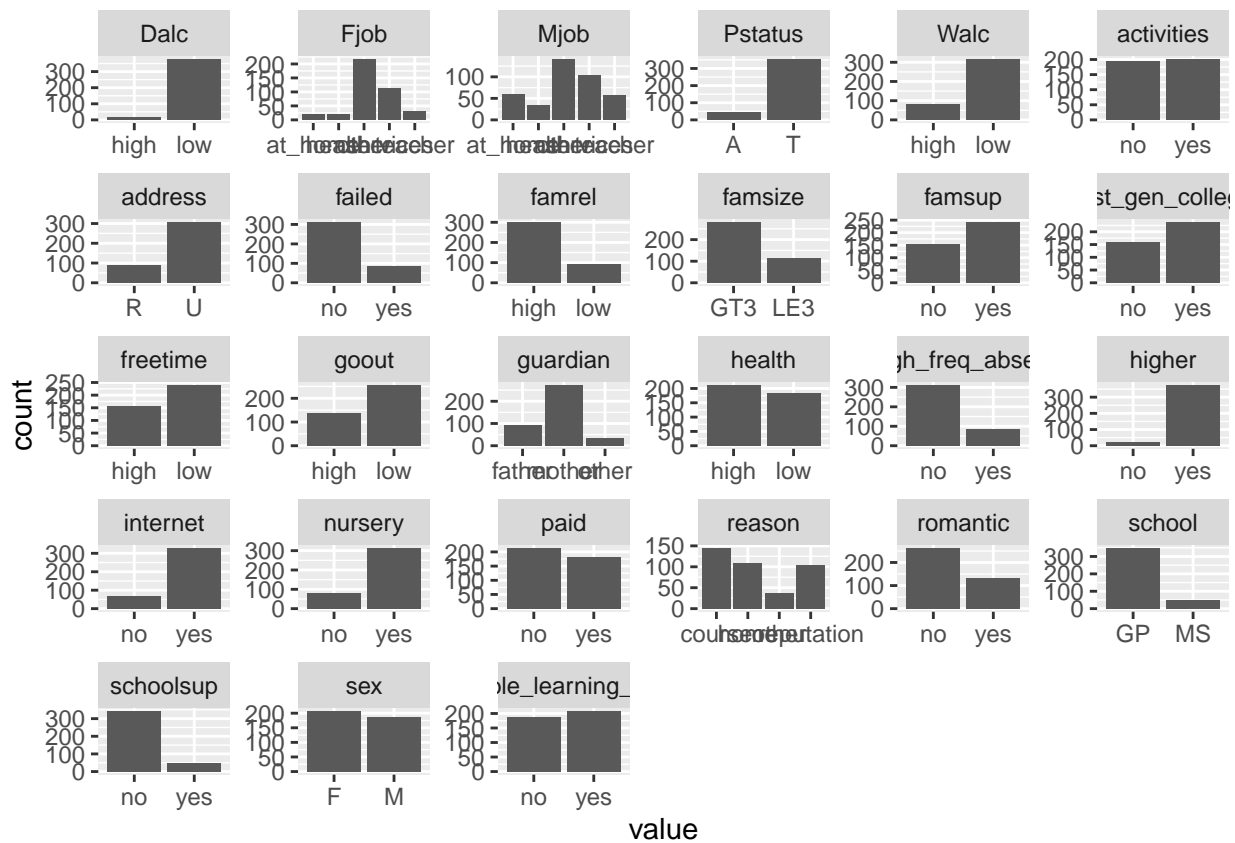
```
data %>%  
  keep(is.numeric) %>%  
  gather() %>%  
  ggplot(aes(value)) +  
    facet_wrap(~ key, scales = "free") +  
    geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
data %>%
  keep(is.character) %>%
  gather() %>%
  ggplot(aes(value)) +
    facet_wrap(~ key, scales = "free") +
    geom_histogram(stat="count")
```

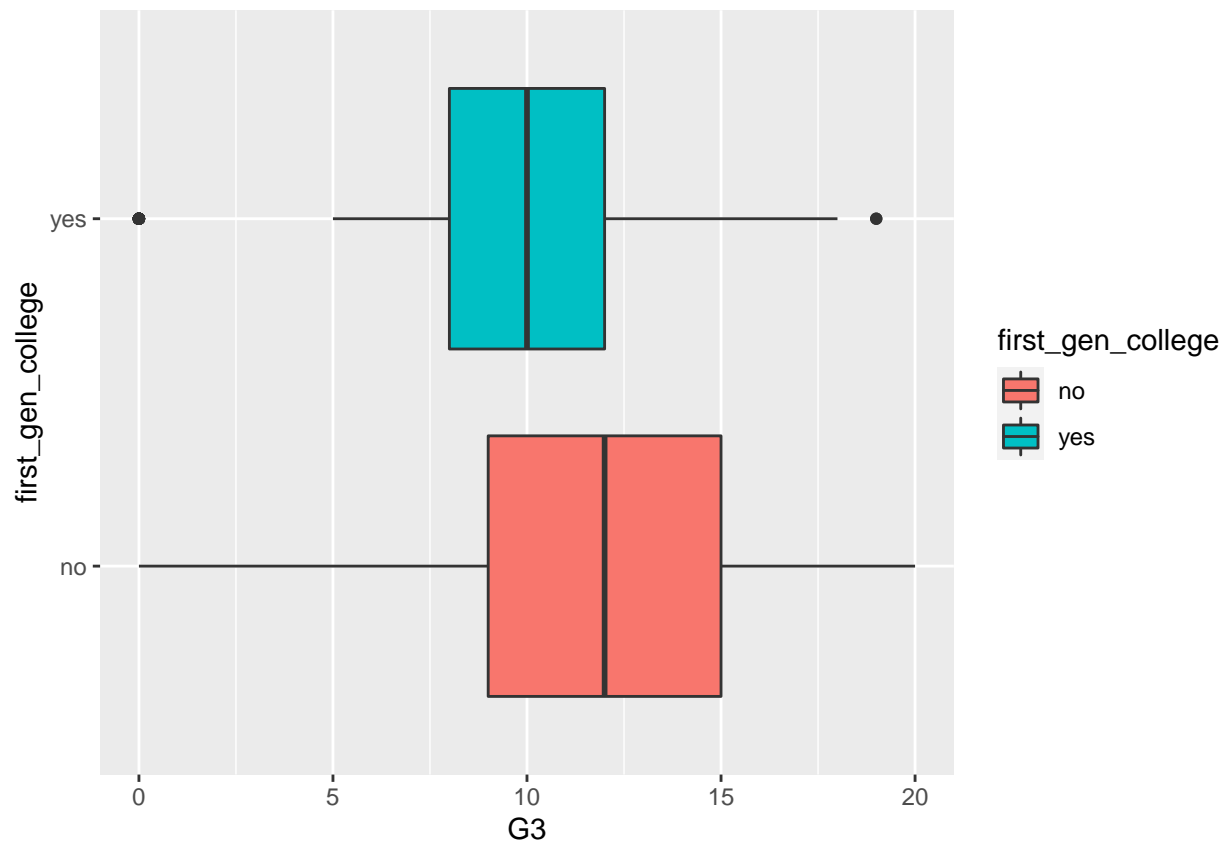
```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



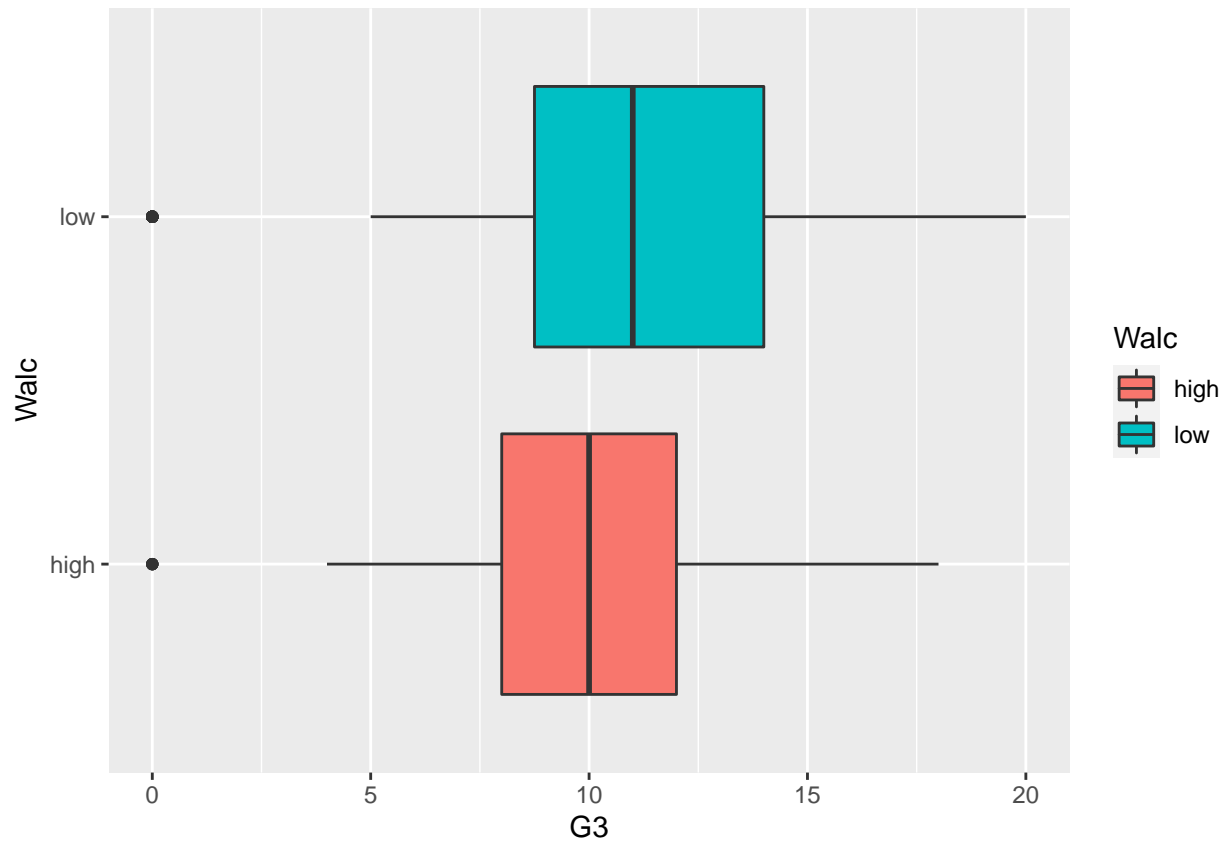
Above we see that the response variable, G3, is pretty normally distributed, thus no transformation is necessary,

```
ggplot(data = data, aes(x = G3, y = first_gen_college, fill=first_gen_college)) +  
  geom_boxplot()
```

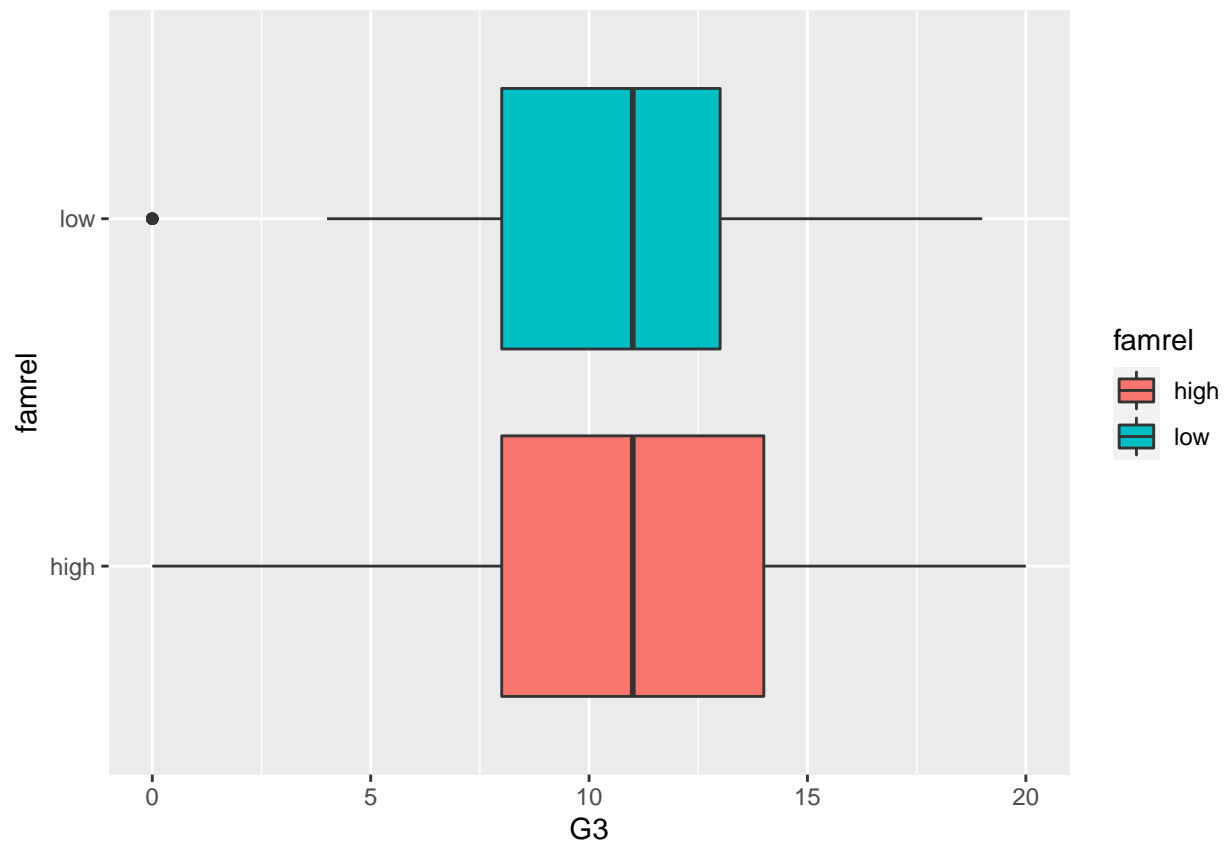




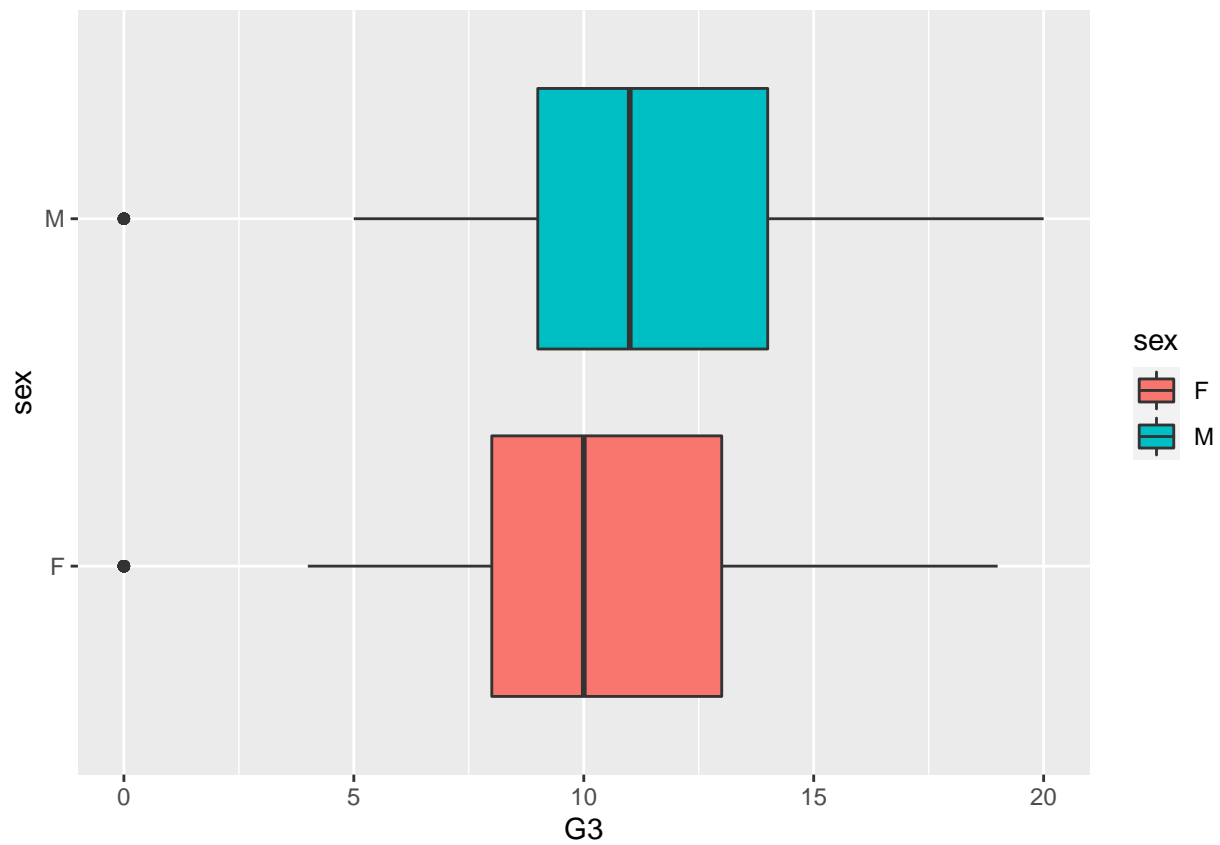
```
ggplot(data = data, aes(x = G3, y = Walc, fill = Walc)) +  
  geom_boxplot()
```



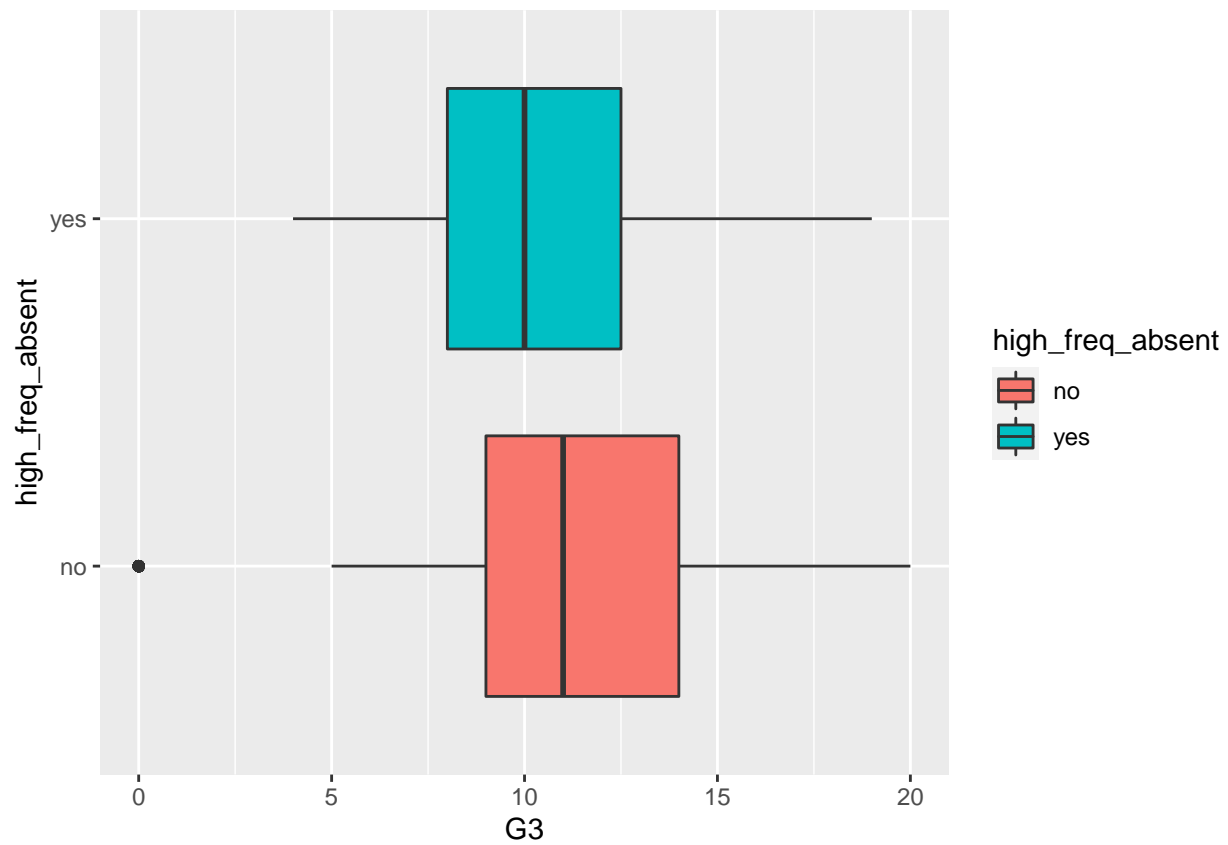
```
ggplot(data = data, aes(x = G3, y = famrel, fill = famrel)) +  
  geom_boxplot()
```



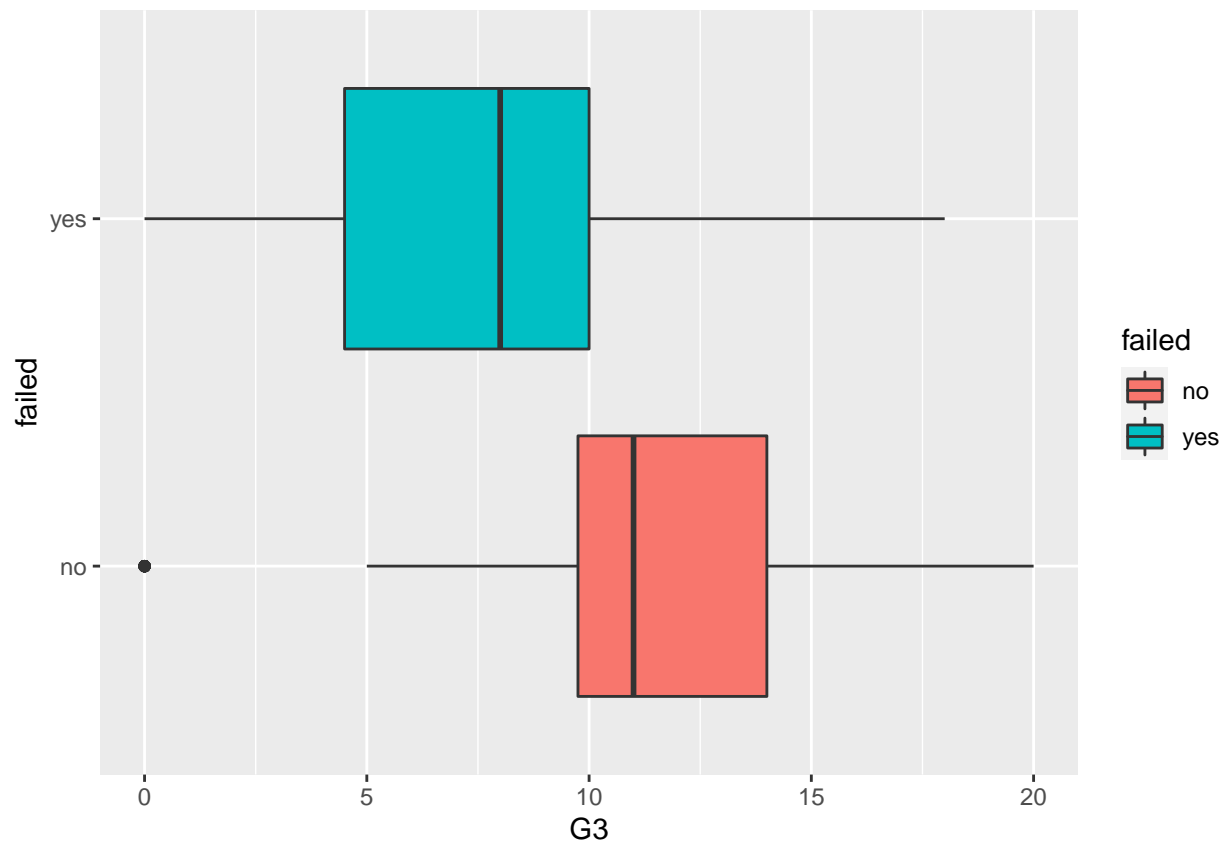
```
ggplot(data = data, aes(x = G3, y= sex, fill = sex)) +  
  geom_boxplot()
```



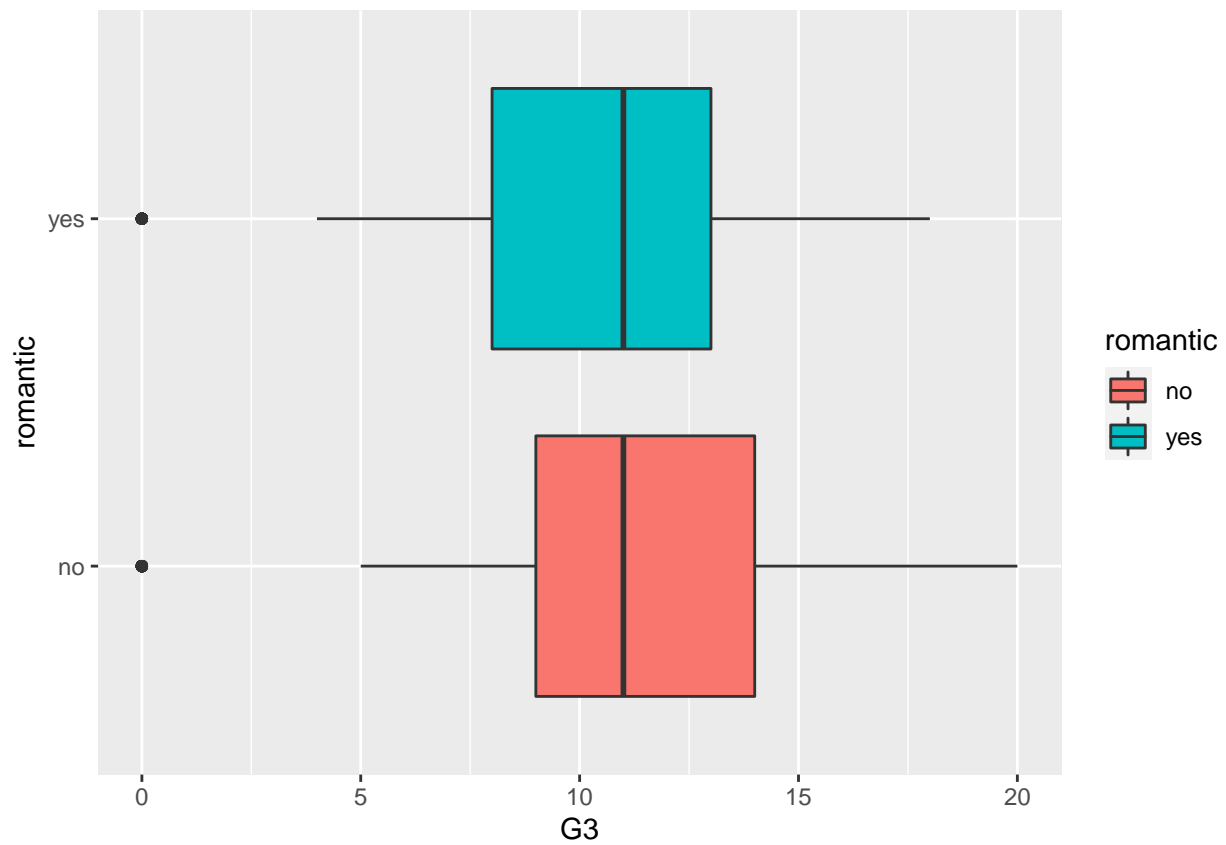
```
ggplot(data = data, aes(x = G3, y = high_freq_absent, fill = high_freq_absent)) +  
  geom_boxplot()
```



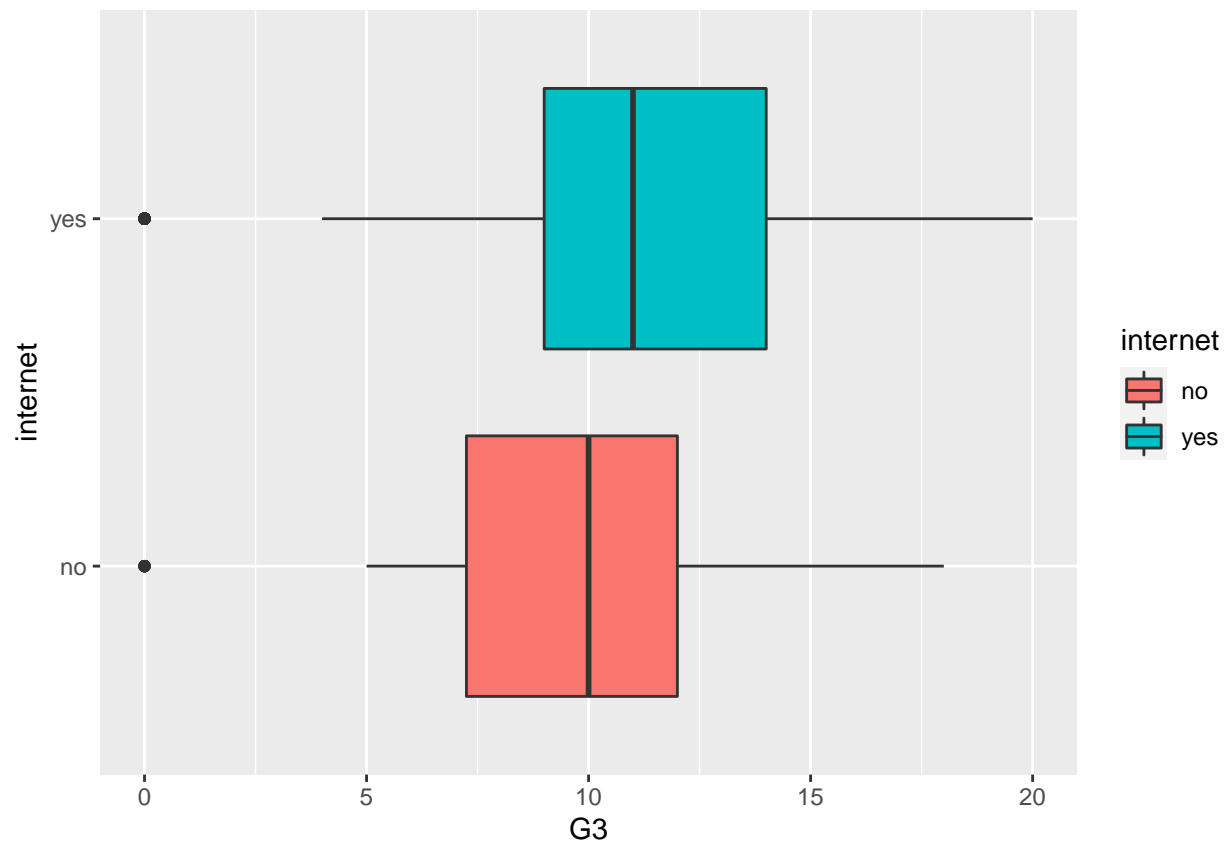
```
ggplot(data = data, aes(x = G3, y=failed, fill = failed)) +  
  geom_boxplot()
```



```
ggplot(data = data, aes(x = G3, y=romantic, fill = romantic)) +  
  geom_boxplot()
```

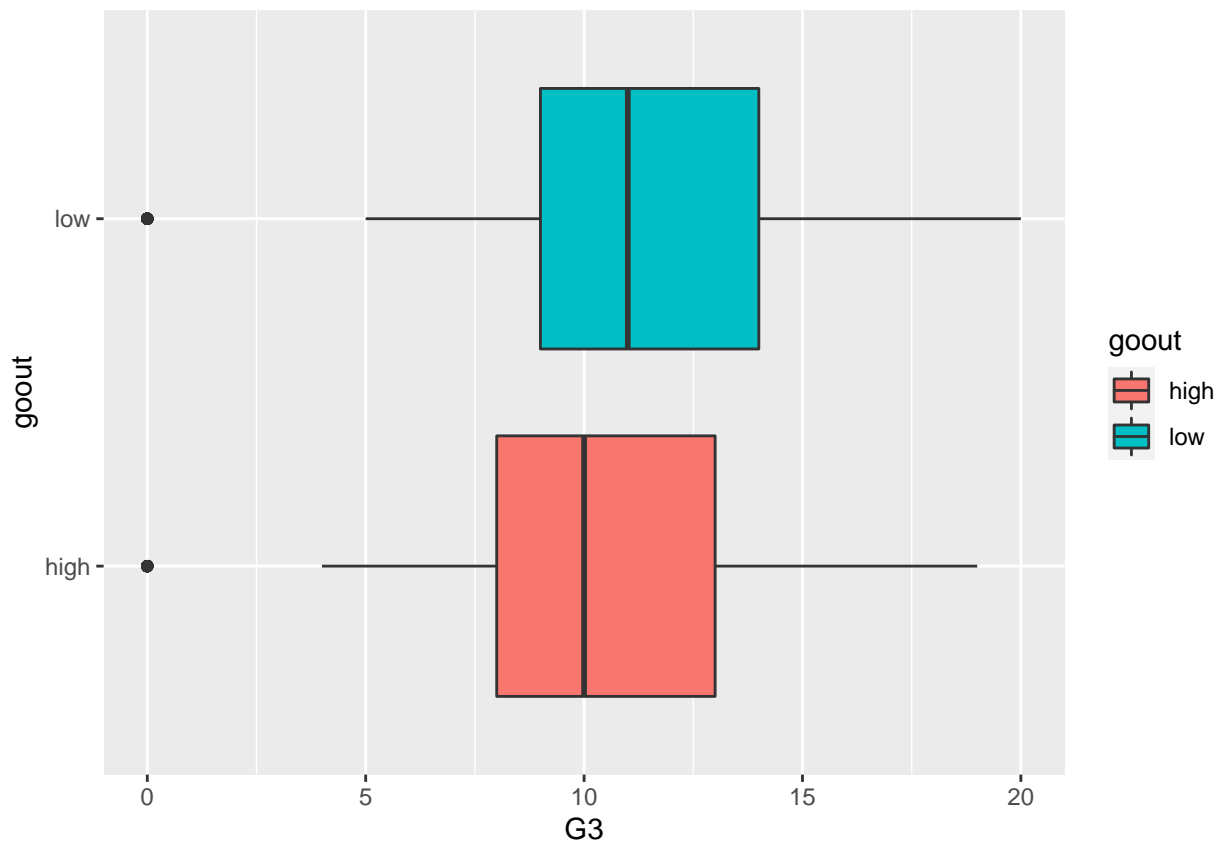


```
ggplot(data = data, aes(x = G3, y=romantic, fill = romantic)) +  
  geom_boxplot()
```



```
ggplot(data = data, aes(x = G3, y=goout, fill = goout)) +  
  geom_boxplot()
```





From the initial explorations above, we can see a few possible trends. Students who had at least one of the following traits: failed a class previously, were a highly frequent absent student, frequently went out, without internet, were frequent drinkers on the weekend, were in romantic relationships, and were first generation students, on average had lower final grades than their counterparts.

```
names(data)
```

```
## [1] "school"      "sex"         "age"
## [4] "address"    "famsize"     "Pstatus"
## [7] "Medu"       "Fedu"       "Mjob"
## [10] "Fjob"       "reason"     "guardian"
## [13] "traveltime" "studytime"  "failures"
## [16] "schoolsup"  "famsup"     "paid"
## [19] "activities" "nursery"    "higher"
## [22] "internet"   "romantic"   "famrel"
## [25] "freetime"   "goout"      "Dalc"
## [28] "Walc"       "health"     "absences"
## [31] "G1"         "G2"         "G3"
## [34] "first_gen_college" "stable_learning_env" "high_freq_absent"
## [37] "failed"
```

```
num_cols <- unlist(lapply(data, is.numeric))
quant_vars <- data[,num_cols]
cor(quant_vars)
```

```
##           age           Medu           Fedu  traveltime  studytime
## age      1.000000000 -0.16365842 -0.163438069  0.07064072 -0.004140037
## Medu     -0.163658419  1.00000000  0.623455112 -0.17163930  0.064944137
## Fedu     -0.163438069  0.62345511  1.000000000 -0.15819405 -0.009174639
```

```
## traveltime 0.070640721 -0.17163930 -0.158194054 1.00000000 -0.100909119
## studytime -0.004140037 0.06494414 -0.009174639 -0.10090912 1.000000000
## failures 0.243665377 -0.23667996 -0.250408444 0.09223875 -0.173563031
## absences 0.175230079 0.10028482 0.024472887 -0.01294378 -0.062700175
## G1 -0.064081497 0.20534100 0.190269936 -0.09303999 0.160611915
## G2 -0.143474049 0.21552717 0.164893393 -0.15319796 0.135879999
## G3 -0.161579438 0.21714750 0.152456939 -0.11714205 0.097819690
## failures absences G1 G2 G3
## age 0.24366538 0.17523008 -0.06408150 -0.1434740 -0.16157944
## Medu -0.23667996 0.10028482 0.20534100 0.2155272 0.21714750
## Fedu -0.25040844 0.02447289 0.19026994 0.1648934 0.15245694
## traveltime 0.09223875 -0.01294378 -0.09303999 -0.1531980 -0.11714205
## studytime -0.17356303 -0.06270018 0.16061192 0.1358800 0.09781969
## failures 1.00000000 0.06372583 -0.35471761 -0.3558956 -0.36041494
## absences 0.06372583 1.00000000 -0.03100290 -0.0317767 0.03424732
## G1 -0.35471761 -0.03100290 1.00000000 0.8521181 0.80146793
## G2 -0.35589563 -0.03177670 0.85211807 1.0000000 0.90486799
## G3 -0.36041494 0.03424732 0.80146793 0.9048680 1.00000000
```

```
#library(corr)
#quant_vars %>% correlate() %>% network_plot(min_cor=0.2)
```

## Splitting data into training and testing sets

```
data <- data %>%
  mutate(ord_g3 = factor(G3, ordered=T))
data$ord_g3
```

```
## [1] 6 6 10 15 10 15 11 6 19 15 9 12 14 11 16 14 14 10 5 10 15 15 16 12 8
## [26] 8 11 15 11 11 12 17 16 12 15 6 18 15 11 13 11 12 18 11 9 6 11 20 14 7
## [51] 13 13 10 11 13 10 15 15 9 16 11 11 9 9 10 15 12 6 8 16 15 10 5 14 11
## [76] 10 10 11 10 5 12 11 6 15 10 8 6 14 10 7 8 18 6 10 14 10 15 10 14 8
## [101] 5 17 14 6 18 11 8 18 13 16 19 10 13 19 9 16 14 13 8 13 15 15 13 13 8
## [126] 12 11 9 0 18 0 0 12 11 0 0 0 0 12 15 0 9 11 13 0 11 0 11 0 10
## [151] 0 14 10 0 12 8 13 10 15 12 0 7 0 10 7 12 10 16 0 14 0 16 10 0 9
## [176] 9 11 6 9 11 8 12 17 8 12 11 11 15 9 10 13 9 8 10 14 15 16 10 18 10
## [201] 16 10 10 6 11 9 7 13 10 7 8 13 14 8 10 15 4 8 8 10 6 0 17 13 14
## [226] 7 15 12 9 12 14 11 9 13 6 10 13 12 11 0 12 12 0 12 0 18 13 8 5 15
## [251] 8 10 8 8 12 8 13 11 14 0 18 8 12 9 0 17 10 11 10 0 9 14 11 14 10
## [276] 12 9 9 8 10 8 10 12 10 11 11 19 12 14 15 11 15 13 18 14 11 0 8 14 16
## [301] 11 10 14 18 13 12 18 8 12 10 0 13 11 11 13 11 0 9 10 11 13 9 11 15 15
## [326] 11 16 10 9 14 8 14 0 0 0 15 13 0 17 10 11 0 15 0 10 14 16 9 15 13
## [351] 8 13 8 8 11 9 13 11 10 16 13 12 10 15 12 10 13 0 10 11 9 12 11 5 19
## [376] 10 15 10 15 10 14 7 10 0 5 10 6 0 8 0 9 16 7 10 9
## 18 Levels: 0 < 4 < 5 < 6 < 7 < 8 < 9 < 10 < 11 < 12 < 13 < 14 < 15 < ... < 20
```

```
attach(data)
set.seed(3)
train_ind <- sample(x = nrow(data), size = 0.8 * nrow(data))
test_ind_neg <- -train_ind
training <- data[train_ind, ]
testing <- data[test_ind_neg, ]
training
```

##	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob
## 261	GP	F	18	U	GT3	T	4	3	services	other
## 186	GP	M	17	U	GT3	T	3	3	services	services
## 140	GP	F	15	U	GT3	T	4	4	teacher	teacher
## 36	GP	F	15	U	GT3	T	2	3	other	other
## 394	MS	M	18	R	LE3	T	3	2	services	other
## 363	MS	F	18	U	GT3	T	3	3	services	services
## 276	GP	F	17	U	LE3	T	2	2	services	other
## 330	GP	F	17	U	GT3	T	4	4	teacher	teacher
## 183	GP	F	17	U	GT3	T	2	4	services	services
## 168	GP	F	16	U	GT3	T	4	2	health	services
## 48	GP	M	16	U	GT3	T	4	3	health	services
## 104	GP	F	15	U	GT3	T	3	2	services	other
## 136	GP	F	15	U	GT3	T	4	4	services	at_home
## 37	GP	M	15	U	LE3	T	4	3	teacher	services
## 108	GP	M	16	U	GT3	T	3	3	services	other
## 395	MS	M	19	U	LE3	T	1	1	other	at_home
## 165	GP	M	17	R	LE3	T	1	1	other	services
## 137	GP	M	17	R	GT3	T	3	4	at_home	other
## 256	GP	M	17	U	LE3	T	1	1	health	other
## 376	MS	F	18	R	GT3	T	1	1	other	other
## 374	MS	F	17	R	GT3	T	1	2	other	other
## 274	GP	M	17	R	GT3	T	1	2	at_home	at_home
## 12	GP	F	15	U	GT3	T	2	1	services	other
## 40	GP	F	15	R	GT3	T	2	2	at_home	other
## 22	GP	M	15	U	GT3	T	4	4	health	health
## 70	GP	F	15	R	LE3	T	3	1	other	other
## 296	GP	M	17	U	GT3	T	3	3	health	other
## 15	GP	M	15	U	GT3	A	2	2	other	other
## 388	MS	F	19	R	GT3	T	2	3	services	other
## 75	GP	F	16	U	GT3	T	3	3	other	services
## 29	GP	M	16	U	LE3	A	3	4	services	other
## 343	GP	M	18	U	LE3	T	3	4	services	other
## 131	GP	F	15	R	GT3	T	3	4	services	teacher
## 62	GP	F	16	U	GT3	T	1	1	services	services
## 127	GP	F	15	U	LE3	A	3	4	other	other
## 101	GP	M	16	U	GT3	T	4	4	services	services
## 247	GP	M	17	U	GT3	T	2	3	other	other
## 182	GP	M	16	U	GT3	T	3	3	services	other
## 383	MS	M	17	U	GT3	T	2	3	other	services
## 138	GP	F	16	U	GT3	A	3	3	other	other
## 68	GP	F	16	U	GT3	T	3	1	services	other
## 33	GP	M	15	R	GT3	T	4	3	teacher	at_home
## 275	GP	F	17	U	GT3	T	2	4	at_home	health
## 166	GP	M	16	U	GT3	T	3	2	services	services
## 185	GP	F	16	U	GT3	T	3	2	other	other
## 237	GP	M	17	U	LE3	T	2	2	other	other
## 192	GP	F	17	U	GT3	T	1	1	at_home	services
## 391	MS	M	20	U	LE3	A	2	2	services	services
## 65	GP	F	15	U	LE3	T	4	3	services	services
## 9	GP	M	15	U	LE3	A	3	2	services	other
## 241	GP	M	17	U	LE3	T	4	3	health	other
## 105	GP	M	15	U	GT3	A	3	4	services	other
## 197	GP	M	17	U	GT3	T	4	4	services	teacher

## 128	GP	F	19	U	GT3	T	0	1	at_home	other
## 198	GP	M	16	R	LE3	T	3	3	teacher	other
## 317	GP	F	18	U	GT3	T	2	1	services	other
## 171	GP	M	16	U	GT3	T	3	4	other	other
## 392	MS	M	17	U	LE3	T	3	1	services	services
## 334	GP	F	18	U	LE3	T	2	2	other	other
## 73	GP	F	15	R	GT3	T	1	1	other	other
## 195	GP	M	16	U	GT3	T	2	3	other	other
## 204	GP	F	17	R	GT3	T	2	2	other	other
## 50	GP	F	15	U	GT3	T	4	4	services	teacher
## 145	GP	M	17	U	GT3	T	2	1	other	other
## 176	GP	M	17	U	LE3	T	4	3	teacher	other
## 164	GP	M	17	U	GT3	T	1	3	at_home	services
## 352	MS	M	17	U	GT3	T	3	3	health	other
## 225	GP	F	16	U	GT3	T	4	4	teacher	services
## 258	GP	M	19	U	LE3	A	4	3	services	at_home
## 117	GP	M	15	U	GT3	T	4	4	other	teacher
## 88	GP	F	15	U	GT3	T	4	2	other	other
## 359	MS	M	18	U	LE3	T	1	1	other	services
## 71	GP	M	16	U	GT3	T	3	1	other	other
## 266	GP	M	18	R	LE3	A	3	4	other	other
## 102	GP	M	16	U	GT3	T	4	4	services	teacher
## 302	GP	M	17	U	LE3	T	4	4	other	teacher
## 19	GP	M	17	U	GT3	T	3	2	services	services
## 47	GP	F	16	U	LE3	A	3	3	other	services
## 161	GP	M	17	R	LE3	T	2	1	at_home	other
## 66	GP	F	16	U	LE3	T	4	3	teacher	services
## 245	GP	F	18	U	GT3	T	2	1	other	other
## 303	GP	F	17	U	GT3	T	4	2	other	other
## 139	GP	M	16	U	LE3	T	1	1	services	other
## 218	GP	M	18	U	LE3	T	3	3	services	health
## 362	MS	M	18	R	LE3	T	1	1	at_home	other
## 4	GP	F	15	U	GT3	T	4	2	health	services
## 16	GP	F	16	U	GT3	T	4	4	health	other
## 326	GP	M	18	U	GT3	T	4	4	other	other
## 262	GP	M	18	U	GT3	T	4	3	teacher	other
## 309	GP	M	19	R	GT3	T	3	3	other	services
## 99	GP	F	16	U	GT3	T	4	4	other	other
## 393	MS	M	21	R	GT3	T	1	1	other	other
## 123	GP	F	16	U	LE3	T	2	4	other	health
## 365	MS	F	17	R	GT3	T	1	2	other	services
## 10	GP	M	15	U	GT3	T	3	4	other	other
## 28	GP	M	15	U	GT3	T	4	2	health	services
## 351	MS	M	19	R	GT3	T	1	1	other	services
## 91	GP	F	16	U	GT3	T	3	3	other	other
## 114	GP	M	15	U	LE3	T	4	2	teacher	other
## 265	GP	F	18	U	GT3	T	2	2	at_home	services
## 313	GP	M	19	U	GT3	T	1	2	other	services
## 368	MS	F	17	R	GT3	T	1	1	other	services
## 233	GP	M	17	U	GT3	T	4	4	teacher	teacher
## 206	GP	F	17	U	GT3	T	3	4	at_home	services
## 7	GP	M	16	U	LE3	T	2	2	other	other
## 332	GP	F	17	R	GT3	T	2	4	at_home	other
## 222	GP	F	17	U	GT3	T	1	1	at_home	other

## 109	GP	M	15	R	GT3	T	4	4	other	other
## 252	GP	M	16	U	GT3	T	3	3	at_home	other
## 348	GP	M	18	U	GT3	T	4	3	teacher	other
## 232	GP	M	17	R	GT3	T	2	2	other	other
## 346	GP	F	18	U	GT3	T	3	2	other	services
## 55	GP	F	15	U	LE3	A	3	3	other	other
## 259	GP	M	18	U	GT3	T	2	1	other	other
## 373	MS	F	17	U	GT3	T	2	2	other	at_home
## 193	GP	M	17	U	GT3	T	1	2	at_home	services
## 112	GP	F	16	R	GT3	T	3	3	services	other
## 44	GP	M	15	U	GT3	T	2	2	services	services
## 354	MS	M	19	R	GT3	T	1	1	other	other
## 53	GP	M	15	U	LE3	A	4	2	health	health
## 13	GP	M	15	U	LE3	T	4	4	health	services
## 110	GP	F	16	U	LE3	T	4	4	health	health
## 14	GP	M	15	U	GT3	T	4	3	teacher	other
## 377	MS	F	20	U	GT3	T	4	2	health	other
## 85	GP	F	15	U	GT3	T	1	1	other	other
## 179	GP	M	16	R	GT3	T	4	2	teacher	services
## 163	GP	M	16	U	LE3	T	1	2	other	other
## 272	GP	F	18	U	GT3	T	2	3	other	services
## 205	GP	F	16	R	GT3	T	2	2	services	services
## 113	GP	F	16	U	GT3	T	2	2	at_home	other
## 228	GP	M	17	U	LE3	T	2	3	services	services
## 170	GP	F	16	U	GT3	T	4	4	health	health
## 8	GP	F	17	U	GT3	A	4	4	other	teacher
## 327	GP	M	17	U	GT3	T	3	3	other	services
## 319	GP	F	17	R	GT3	T	3	4	at_home	services
## 201	GP	F	16	U	GT3	T	4	3	health	other
## 74	GP	M	16	U	GT3	T	3	1	other	other
## 49	GP	M	15	U	GT3	T	4	2	teacher	other
## 151	GP	M	18	U	LE3	T	1	1	other	other
## 338	GP	F	17	U	GT3	T	3	2	other	other
## 253	GP	M	18	U	GT3	T	2	1	services	services
## 35	GP	M	16	U	GT3	T	3	2	other	other
## 308	GP	M	19	U	GT3	T	4	4	teacher	services
## 321	GP	F	17	U	GT3	A	4	3	services	services
## 125	GP	F	16	U	GT3	T	2	2	other	other
## 310	GP	F	19	U	LE3	T	1	1	at_home	other
## 324	GP	F	17	U	GT3	T	3	1	services	services
## 2	GP	F	17	U	GT3	T	1	1	at_home	other
## 100	GP	F	16	U	GT3	T	4	3	other	at_home
## 372	MS	M	18	R	LE3	T	1	2	at_home	services
## 325	GP	F	17	U	LE3	T	0	2	at_home	at_home
## 367	MS	M	18	U	LE3	T	4	4	teacher	services
## 84	GP	M	15	U	LE3	T	2	2	services	services
## 90	GP	M	16	U	LE3	A	4	4	teacher	health
## 314	GP	F	19	U	LE3	T	3	2	services	other
## 267	GP	M	17	U	GT3	T	3	1	services	other
## 153	GP	F	15	R	GT3	T	3	3	services	services
## 350	MS	M	18	R	GT3	T	3	2	other	other
## 211	GP	F	19	U	GT3	T	3	3	other	other
## 342	GP	M	18	U	GT3	T	4	4	teacher	services
## 217	GP	F	17	U	GT3	T	4	3	other	other

## 370	MS	F	18	R	GT3	T	4	4	other	teacher
## 208	GP	F	16	U	GT3	T	4	3	teacher	other
## 67	GP	M	15	U	GT3	A	4	4	other	services
## 162	GP	M	15	R	GT3	T	3	2	other	other
## 255	GP	M	17	R	GT3	T	2	1	other	other
## 301	GP	F	18	U	LE3	A	4	4	health	other
## 289	GP	M	18	U	GT3	T	2	1	services	services
## 52	GP	F	15	U	LE3	T	4	2	health	other
## 45	GP	F	16	U	LE3	T	2	2	other	at_home
## 333	GP	F	18	U	GT3	T	3	3	services	services
## 345	GP	F	18	U	GT3	T	2	3	at_home	other
## 316	GP	F	19	R	GT3	T	2	3	other	other
## 277	GP	F	18	R	GT3	A	3	2	other	services
## 189	GP	F	17	U	GT3	A	3	3	health	other
## 63	GP	F	16	U	LE3	T	1	2	other	services
## 142	GP	M	16	U	LE3	T	2	2	services	services
## 106	GP	F	15	U	GT3	A	3	3	other	health
## 384	MS	M	19	R	GT3	T	1	1	other	services
## 122	GP	M	15	U	GT3	T	2	2	services	services
## 263	GP	M	18	R	GT3	T	3	2	other	other
## 297	GP	F	19	U	GT3	T	4	4	health	other
## 306	GP	F	18	U	GT3	T	2	4	services	at_home
## 87	GP	F	16	U	LE3	T	2	2	at_home	other
## 121	GP	F	15	U	GT3	T	1	2	at_home	services
## 6	GP	M	16	U	LE3	T	4	3	services	other
## 293	GP	F	18	U	LE3	T	2	1	services	at_home
## 356	MS	F	18	U	GT3	T	3	3	services	services
## 81	GP	M	15	U	GT3	T	2	3	other	services
## 246	GP	M	16	U	GT3	T	2	1	other	other
## 209	GP	F	16	U	GT3	T	1	1	at_home	other
## 39	GP	F	15	R	GT3	T	3	4	services	health
## 147	GP	F	15	U	GT3	T	3	2	health	services
## 219	GP	F	17	U	GT3	T	2	3	at_home	other
## 169	GP	F	16	U	GT3	T	2	2	other	other
## 268	GP	F	18	R	GT3	T	4	4	teacher	other
## 278	GP	M	18	U	GT3	T	4	4	teacher	services
## 210	GP	F	17	R	GT3	T	4	3	teacher	other
## 124	GP	M	16	U	GT3	T	4	4	health	other
## 20	GP	M	16	U	LE3	T	4	3	health	other
## 5	GP	F	16	U	GT3	T	3	3	other	other
## 203	GP	F	17	U	GT3	T	1	1	other	other
## 236	GP	M	16	U	GT3	T	3	2	at_home	other
## 135	GP	M	15	R	GT3	T	3	4	at_home	teacher
## 177	GP	F	16	U	GT3	T	2	2	services	other
## 69	GP	F	15	R	LE3	T	2	2	health	services
## 291	GP	M	18	U	GT3	T	4	2	teacher	other
## 380	MS	F	17	R	GT3	T	3	1	at_home	other
## 299	GP	F	18	U	GT3	T	4	3	other	other
## 311	GP	F	19	U	LE3	T	1	2	services	services
## 202	GP	F	16	U	GT3	T	2	3	other	other
## 146	GP	F	15	U	GT3	T	1	1	other	services
## 223	GP	F	16	U	GT3	T	2	3	services	teacher
## 390	MS	F	18	U	GT3	T	1	1	other	other
## 181	GP	M	16	U	GT3	T	4	3	teacher	other

## 312	GP	F	19	U	GT3	T	2	1	at_home	other
## 242	GP	M	17	R	LE3	A	4	4	teacher	other
## 158	GP	F	18	R	GT3	T	1	1	at_home	other
## 382	MS	M	18	R	GT3	T	2	1	other	other
## 42	GP	M	15	U	LE3	T	4	4	teacher	other
## 24	GP	M	16	U	LE3	T	2	2	other	other
## 173	GP	M	17	U	LE3	T	4	4	teacher	other
## 294	GP	F	17	R	LE3	T	3	1	services	other
## 159	GP	M	16	R	GT3	T	2	2	at_home	other
## 323	GP	F	17	R	LE3	T	2	2	services	services
## 129	GP	M	18	R	GT3	T	2	2	services	other
## 290	GP	M	18	U	LE3	A	4	4	teacher	teacher
## 79	GP	M	17	U	GT3	T	2	1	other	other
## 286	GP	M	17	U	GT3	T	1	1	other	other
## 331	GP	M	18	U	LE3	T	2	2	other	other
## 144	GP	F	16	U	LE3	T	1	1	at_home	at_home
## 41	GP	F	16	U	LE3	T	2	2	other	other
## 118	GP	M	16	U	GT3	T	3	3	other	services
## 51	GP	F	16	U	LE3	T	2	2	services	services
## 344	GP	F	17	U	GT3	A	2	2	at_home	at_home
## 361	MS	F	18	R	LE3	A	1	4	at_home	other
## 349	GP	F	17	U	GT3	T	4	3	health	other
## 387	MS	F	18	R	GT3	T	4	4	teacher	at_home
## 83	GP	F	15	U	LE3	T	3	2	services	other
## 381	MS	M	18	U	GT3	T	4	4	teacher	teacher
## 375	MS	F	18	R	LE3	T	4	4	other	other
## 64	GP	F	16	U	GT3	T	4	3	teacher	health
## 280	GP	M	18	U	LE3	T	4	3	teacher	services
## 270	GP	F	18	R	GT3	T	2	1	other	other
## 23	GP	M	16	U	LE3	T	4	2	teacher	other
## 27	GP	M	15	U	GT3	T	2	2	other	other
## 230	GP	F	17	U	GT3	A	2	1	other	other
## 172	GP	M	16	U	GT3	T	1	0	other	other
## 26	GP	F	16	U	GT3	T	2	2	services	services
## 251	GP	M	18	U	GT3	T	3	2	services	other
## 86	GP	F	15	U	GT3	T	4	4	services	services
## 184	GP	F	17	U	LE3	T	3	3	other	other
## 360	MS	F	18	U	LE3	T	1	1	at_home	services
## 335	GP	F	18	R	GT3	T	2	2	at_home	other
## 307	GP	M	20	U	GT3	A	3	2	services	other
## 248	GP	M	22	U	GT3	T	3	1	services	services
## 357	MS	F	17	R	GT3	T	4	4	teacher	services
## 339	GP	F	18	U	LE3	T	3	3	services	services
## 132	GP	F	15	U	GT3	T	1	1	at_home	other
## 111	GP	M	15	U	LE3	A	4	4	teacher	teacher
## 199	GP	F	17	U	GT3	T	4	4	services	teacher
## 207	GP	F	16	U	GT3	A	3	1	services	other
## 141	GP	M	15	U	GT3	T	4	3	teacher	services
## 298	GP	F	18	U	LE3	T	4	3	other	other
## 77	GP	M	15	U	GT3	T	4	0	teacher	other
## 355	MS	M	17	R	GT3	T	4	3	services	other
## 257	GP	F	17	U	LE3	T	4	2	teacher	services
## 318	GP	F	18	U	GT3	T	4	3	other	other
## 38	GP	M	16	R	GT3	A	4	4	other	teacher

## 216	GP	F	17	U	LE3	T	3	2	other	other
## 92	GP	F	15	U	GT3	T	4	3	services	other
## 379	MS	F	18	U	GT3	T	3	3	other	other
## 190	GP	M	17	R	GT3	T	1	2	at_home	other
## 305	GP	M	19	U	GT3	T	3	3	other	other
## 103	GP	M	15	U	GT3	T	4	4	services	other
## 320	GP	F	18	U	GT3	T	4	4	teacher	other
## 57	GP	F	15	U	GT3	A	4	3	services	services
## 353	MS	M	18	U	LE3	T	1	3	at_home	services
## 61	GP	F	16	R	GT3	T	4	4	health	teacher
## 134	GP	F	16	U	GT3	A	3	4	services	other
## 220	GP	F	17	U	GT3	T	2	2	at_home	at_home
## 287	GP	F	18	U	GT3	T	2	2	at_home	at_home
## 167	GP	M	16	U	GT3	T	2	2	other	other
## 1	GP	F	18	U	GT3	A	4	4	at_home	teacher
## 215	GP	F	17	R	LE3	T	4	4	services	other
## 3	GP	F	15	U	LE3	T	1	1	at_home	other
## 30	GP	M	16	U	GT3	T	4	4	teacher	teacher
## 82	GP	M	15	U	GT3	T	2	3	other	other
## 143	GP	F	15	U	GT3	T	4	4	teacher	services
## 93	GP	F	16	U	LE3	T	3	1	other	other
## 337	GP	F	19	R	GT3	A	3	1	services	at_home
## 221	GP	F	17	R	GT3	T	2	1	at_home	services
## 364	MS	F	17	U	LE3	T	4	4	at_home	at_home
## 96	GP	F	15	R	GT3	T	1	1	at_home	other
## 126	GP	M	15	U	GT3	T	3	4	services	services
## 214	GP	M	18	U	GT3	T	2	2	services	other
## 188	GP	M	16	U	LE3	T	2	1	other	other
## 285	GP	F	17	U	GT3	T	2	2	other	other
## 371	MS	F	19	U	LE3	T	3	2	services	services
## 271	GP	F	19	U	GT3	T	3	3	other	services
## 295	GP	M	18	R	LE3	T	3	2	services	other
## 157	GP	M	17	R	LE3	T	1	2	other	other
## 25	GP	F	15	R	GT3	T	2	4	services	health
## 212	GP	M	17	U	LE3	T	4	4	services	other
## 238	GP	F	16	U	GT3	T	2	1	other	other
## 156	GP	M	15	R	GT3	T	2	3	at_home	services
## 213	GP	F	16	U	GT3	A	2	2	other	other
## 148	GP	F	15	U	GT3	T	1	2	at_home	other
## 234	GP	M	16	U	GT3	T	4	4	health	other
## 56	GP	F	16	U	GT3	A	2	1	other	other
## 107	GP	F	15	U	GT3	T	2	2	other	other
## 154	GP	M	19	U	GT3	T	3	2	services	at_home
## 191	GP	F	16	U	GT3	T	2	3	services	services
## 249	GP	M	18	R	LE3	T	3	3	other	services
## 366	MS	M	18	R	GT3	T	1	3	at_home	other
## 54	GP	F	15	U	GT3	T	4	4	services	services
##	reason	guardian	traveltime	studytime	failures	schoolsup	famsup	paid		
## 261	home	father	1	2	0	no	yes	yes		
## 186	other	mother	1	2	0	no	yes	no		
## 140	course	mother	2	1	0	no	no	no		
## 36	other	father	2	1	0	no	yes	no		
## 394	course	mother	3	1	0	no	no	no		
## 363	other	mother	2	2	0	no	yes	no		



## 276	course	mother	2	2	0	yes	yes	yes
## 330	course	mother	2	3	0	no	yes	yes
## 183	reputation	father	1	2	0	no	yes	no
## 168	home	father	1	2	0	no	no	yes
## 48	reputation	mother	1	4	0	no	no	no
## 104	home	mother	2	2	0	yes	yes	yes
## 136	course	mother	1	3	0	no	yes	no
## 37	home	mother	1	3	0	no	yes	no
## 108	home	father	1	3	0	no	yes	no
## 395	course	father	1	1	0	no	no	no
## 165	course	mother	4	2	3	no	no	no
## 137	course	mother	3	2	0	no	no	no
## 256	course	mother	2	1	1	no	yes	no
## 376	home	mother	4	3	0	no	no	no
## 374	course	mother	1	1	0	no	no	no
## 274	home	mother	1	2	0	no	yes	yes
## 12	reputation	father	3	3	0	no	yes	no
## 40	reputation	mother	1	1	0	yes	yes	yes
## 22	other	father	1	1	0	no	yes	yes
## 70	reputation	father	2	4	0	no	yes	no
## 296	home	mother	1	1	0	no	yes	yes
## 15	home	other	1	3	0	no	yes	no
## 388	course	mother	1	3	1	no	no	no
## 75	home	mother	1	2	0	yes	yes	yes
## 29	home	mother	1	2	0	yes	yes	no
## 343	home	mother	1	2	0	no	no	no
## 131	course	father	2	3	2	no	yes	no
## 62	course	father	4	1	0	yes	yes	no
## 127	home	mother	1	2	0	yes	no	no
## 101	other	mother	1	1	0	yes	yes	yes
## 247	course	father	2	1	0	no	no	no
## 182	home	mother	1	2	0	no	no	yes
## 383	home	father	2	2	0	no	no	no
## 138	course	other	2	1	2	no	yes	no
## 68	course	mother	1	4	0	yes	yes	yes
## 33	course	mother	1	2	0	no	yes	no
## 275	reputation	mother	2	2	0	no	yes	yes
## 166	course	mother	2	1	1	no	yes	no
## 185	reputation	mother	1	2	0	no	yes	yes
## 237	home	father	1	2	0	no	no	yes
## 192	course	mother	1	2	0	no	no	no
## 391	course	other	1	2	2	no	yes	yes
## 65	reputation	father	1	2	0	yes	no	no
## 9	home	mother	1	2	0	no	yes	yes
## 241	course	mother	2	2	0	no	no	no
## 105	course	mother	1	2	0	no	yes	yes
## 197	home	mother	1	1	0	no	no	no
## 128	course	other	1	2	3	no	yes	no
## 198	home	father	3	1	0	no	yes	yes
## 317	course	mother	2	2	0	no	yes	yes
## 171	course	father	3	1	2	no	yes	no
## 392	course	mother	2	1	0	no	no	no
## 334	home	other	1	2	0	no	no	no
## 73	reputation	mother	1	2	2	yes	yes	no

## 195	home	father	2	1	0	no	no	no
## 204	reputation	mother	1	1	0	no	yes	no
## 50	other	father	1	2	1	yes	yes	no
## 145	home	mother	1	1	3	no	yes	no
## 176	course	mother	2	2	0	no	no	yes
## 164	course	father	1	1	0	no	no	no
## 352	course	mother	2	2	0	no	yes	yes
## 225	home	mother	1	3	0	no	yes	no
## 258	reputation	mother	1	2	0	no	yes	no
## 117	reputation	father	2	2	0	no	yes	no
## 88	reputation	mother	1	3	0	no	yes	no
## 359	home	father	2	1	0	no	no	no
## 71	reputation	father	2	4	0	no	yes	yes
## 266	reputation	mother	2	2	0	no	yes	yes
## 102	other	father	1	3	0	no	yes	no
## 302	home	father	2	1	0	no	no	yes
## 19	course	mother	1	1	3	no	yes	no
## 47	home	mother	1	2	0	no	yes	no
## 161	course	mother	2	1	2	no	no	no
## 66	course	mother	3	2	0	no	yes	no
## 245	course	other	2	3	0	no	yes	yes
## 303	reputation	mother	2	3	0	no	yes	yes
## 139	course	mother	1	2	1	no	no	no
## 218	home	father	1	2	1	no	yes	yes
## 362	other	mother	2	2	1	no	no	no
## 4	home	mother	1	3	0	no	yes	yes
## 16	home	mother	1	1	0	no	yes	no
## 326	course	mother	1	3	0	no	no	no
## 262	course	mother	1	2	0	no	yes	yes
## 309	reputation	father	1	2	1	no	no	no
## 99	reputation	mother	1	1	0	no	no	no
## 393	course	other	1	1	3	no	no	no
## 123	course	father	2	2	0	no	yes	yes
## 365	course	father	2	2	0	no	no	no
## 10	home	mother	1	2	0	no	yes	yes
## 28	other	mother	1	1	0	no	no	yes
## 351	home	other	3	2	3	no	no	no
## 91	home	mother	1	3	0	no	yes	yes
## 114	course	mother	1	1	0	no	no	no
## 265	home	mother	1	3	0	no	yes	yes
## 313	course	other	1	2	1	no	no	no
## 368	reputation	mother	3	1	1	no	yes	yes
## 233	reputation	mother	1	2	0	yes	yes	no
## 206	home	mother	1	3	1	no	yes	yes
## 7	home	mother	1	2	0	no	no	no
## 332	course	father	1	3	0	no	yes	no
## 222	reputation	mother	1	3	1	no	yes	no
## 109	home	father	4	4	0	no	yes	yes
## 252	reputation	other	3	2	0	yes	yes	no
## 348	course	mother	1	3	0	no	yes	yes
## 232	course	father	2	2	0	no	yes	yes
## 346	other	mother	1	3	0	no	no	no
## 55	other	mother	1	1	0	no	no	yes
## 259	home	mother	1	2	0	no	no	no

## 373	home	mother	1	3	0	no	no	no
## 193	other	other	2	2	0	no	no	yes
## 112	reputation	father	1	3	1	yes	yes	no
## 44	course	father	1	1	0	yes	yes	no
## 354	home	other	3	1	1	no	yes	no
## 53	other	father	2	1	1	no	no	no
## 13	course	father	1	1	0	no	yes	yes
## 110	other	mother	1	3	0	no	yes	yes
## 14	course	mother	2	2	0	no	yes	yes
## 377	course	other	2	3	2	no	yes	yes
## 85	home	father	1	2	0	no	yes	no
## 179	other	mother	1	1	0	no	yes	no
## 163	course	mother	2	1	1	no	no	no
## 272	reputation	father	1	4	0	no	yes	yes
## 205	reputation	mother	2	4	0	no	yes	yes
## 113	home	mother	1	2	1	yes	no	no
## 228	reputation	father	1	2	0	no	yes	yes
## 170	reputation	mother	1	2	0	no	yes	yes
## 8	home	mother	2	2	0	yes	yes	no
## 327	reputation	mother	1	1	0	no	no	no
## 319	course	father	1	3	0	no	yes	yes
## 201	home	mother	1	2	0	no	yes	no
## 74	reputation	mother	1	1	0	no	no	no
## 49	home	mother	1	2	0	no	yes	yes
## 151	course	mother	1	1	3	no	no	no
## 338	home	mother	1	2	0	no	yes	yes
## 253	other	mother	1	1	1	no	no	no
## 35	home	mother	1	1	0	no	yes	yes
## 308	reputation	other	2	1	1	no	yes	yes
## 321	course	mother	1	2	0	no	yes	yes
## 125	home	mother	1	2	0	no	no	yes
## 310	reputation	other	1	2	1	yes	yes	no
## 324	course	father	1	3	0	no	yes	no
## 2	course	father	1	2	0	no	yes	no
## 100	course	mother	1	3	0	yes	yes	yes
## 372	other	father	3	1	0	no	yes	yes
## 325	home	father	2	3	0	no	no	no
## 367	other	mother	2	3	0	no	no	yes
## 84	home	mother	2	2	0	no	no	yes
## 90	reputation	mother	1	2	0	no	yes	no
## 314	reputation	other	2	2	1	no	yes	yes
## 267	other	mother	1	2	0	no	no	yes
## 153	reputation	other	2	3	2	no	yes	yes
## 350	course	mother	2	1	1	no	yes	no
## 211	reputation	other	1	4	0	no	yes	yes
## 342	home	father	1	2	1	no	yes	no
## 217	reputation	mother	1	2	2	no	no	yes
## 370	other	father	3	2	0	no	yes	yes
## 208	other	mother	1	2	0	no	no	yes
## 67	reputation	mother	1	4	0	no	yes	no
## 162	course	mother	2	2	2	yes	yes	no
## 255	course	mother	1	1	0	no	no	no
## 301	home	mother	1	2	0	no	yes	no
## 289	reputation	mother	1	3	0	no	no	yes

## 52	other	mother	1	2	0	no	yes	yes
## 45	course	father	2	2	1	yes	no	no
## 333	home	mother	1	2	0	no	no	no
## 345	course	mother	1	3	0	no	yes	no
## 316	reputation	other	1	3	1	no	no	no
## 277	home	mother	2	2	0	no	no	no
## 189	reputation	mother	1	2	0	no	yes	no
## 63	reputation	father	1	2	0	yes	no	no
## 142	reputation	father	2	1	2	no	yes	no
## 106	reputation	father	1	4	0	yes	no	no
## 384	other	mother	2	1	1	no	no	no
## 122	home	father	1	4	0	no	yes	yes
## 263	course	mother	1	3	0	no	no	no
## 297	reputation	other	2	2	0	no	yes	yes
## 306	reputation	other	1	2	1	no	yes	yes
## 87	course	mother	1	2	0	no	yes	no
## 121	course	mother	1	2	0	no	no	no
## 6	reputation	mother	1	2	0	no	yes	yes
## 293	reputation	mother	1	2	1	no	no	no
## 356	course	father	1	2	0	no	yes	no
## 81	course	father	1	1	0	yes	yes	yes
## 246	course	mother	3	1	0	no	no	no
## 209	home	mother	2	1	0	no	yes	yes
## 39	course	mother	1	3	0	yes	yes	yes
## 147	home	father	1	2	3	no	yes	no
## 219	home	father	2	1	0	no	yes	yes
## 169	home	mother	1	2	0	no	yes	yes
## 268	reputation	mother	2	2	0	no	no	yes
## 278	home	mother	2	1	0	no	no	yes
## 210	reputation	mother	2	3	0	no	yes	yes
## 124	course	mother	1	1	0	no	yes	no
## 20	home	father	1	1	0	no	no	yes
## 5	home	father	1	2	0	no	yes	yes
## 203	course	mother	1	2	0	no	yes	yes
## 236	reputation	mother	2	3	0	no	no	no
## 135	course	mother	4	2	0	no	yes	no
## 177	reputation	mother	2	2	0	no	no	yes
## 69	reputation	mother	2	2	0	yes	yes	yes
## 291	home	mother	1	2	0	no	yes	yes
## 380	reputation	mother	1	2	0	no	yes	yes
## 299	reputation	father	1	4	0	no	yes	yes
## 311	home	other	1	2	1	no	no	no
## 202	reputation	mother	1	2	0	yes	yes	yes
## 146	course	father	1	2	0	no	yes	yes
## 223	other	mother	1	2	0	yes	no	no
## 390	course	mother	2	2	1	no	no	no
## 181	home	mother	1	2	0	no	yes	yes
## 312	other	other	3	2	0	no	yes	no
## 242	course	mother	2	2	0	no	yes	yes
## 158	course	mother	3	1	3	no	yes	no
## 382	other	mother	2	1	0	no	no	no
## 42	home	other	1	1	0	no	yes	no
## 24	reputation	mother	2	2	0	no	yes	no
## 173	reputation	mother	1	2	0	no	yes	yes

## 294	reputation	mother	2	4	0	no	yes	yes
## 159	course	mother	3	1	0	no	no	no
## 323	course	mother	1	3	0	no	yes	yes
## 129	reputation	mother	1	1	2	no	yes	no
## 290	reputation	mother	1	2	0	no	yes	yes
## 79	home	mother	2	1	3	yes	yes	no
## 286	reputation	father	1	2	0	no	no	yes
## 331	course	mother	1	4	0	no	yes	no
## 144	course	mother	1	1	0	no	no	no
## 41	home	mother	2	2	1	no	yes	no
## 118	home	father	2	1	0	no	no	no
## 51	course	mother	3	2	0	no	yes	yes
## 344	home	father	1	2	1	no	yes	no
## 361	course	mother	3	2	0	no	no	no
## 349	reputation	mother	1	3	0	no	yes	yes
## 387	reputation	mother	3	1	0	no	yes	yes
## 83	reputation	mother	1	2	0	no	yes	yes
## 381	home	father	1	2	0	no	no	yes
## 375	reputation	mother	2	3	0	no	no	no
## 64	home	mother	1	3	0	yes	yes	yes
## 280	course	mother	2	1	0	no	no	yes
## 270	reputation	mother	2	2	0	no	yes	no
## 23	course	mother	1	2	0	no	no	no
## 27	home	mother	1	1	0	no	yes	yes
## 230	course	mother	2	3	0	no	no	no
## 172	reputation	mother	2	2	0	no	yes	yes
## 26	home	mother	1	1	2	no	yes	yes
## 251	course	mother	2	1	1	no	no	no
## 86	reputation	father	2	2	2	no	no	yes
## 184	reputation	mother	1	2	0	no	yes	no
## 360	course	father	2	3	0	no	no	no
## 335	course	mother	2	4	0	no	no	no
## 307	course	other	1	1	0	no	no	no
## 248	other	mother	1	1	3	no	no	no
## 357	other	father	2	2	0	no	yes	yes
## 339	home	mother	1	4	0	no	yes	no
## 132	course	mother	3	1	0	no	yes	no
## 111	course	mother	1	1	0	no	no	no
## 199	home	mother	2	1	1	no	yes	no
## 207	course	mother	1	2	3	no	yes	yes
## 141	course	father	2	4	0	yes	yes	no
## 298	home	other	2	2	0	no	yes	yes
## 77	course	mother	2	4	0	no	no	no
## 355	home	mother	2	2	0	no	yes	yes
## 257	reputation	mother	1	4	0	no	yes	yes
## 318	course	mother	1	3	0	no	yes	yes
## 38	reputation	mother	2	3	0	no	yes	no
## 216	reputation	mother	2	2	0	no	no	yes
## 92	reputation	mother	1	1	0	no	no	yes
## 379	home	mother	1	2	0	no	no	yes
## 190	home	mother	1	2	0	no	no	no
## 305	home	other	1	2	1	no	yes	no
## 103	course	mother	1	1	0	no	yes	no
## 320	course	mother	1	2	0	no	yes	yes

## 57	reputation	mother	1	2	0	no	yes	yes		
## 353	course	mother	1	1	1	no	no	no		
## 61	other	mother	1	2	0	no	yes	no		
## 134	course	father	1	1	0	no	no	no		
## 220	course	mother	1	3	0	no	yes	yes		
## 287	other	mother	1	3	0	no	yes	yes		
## 167	course	father	1	2	0	no	no	no		
## 1	course	mother	2	2	0	yes	no	no		
## 215	other	mother	1	1	0	no	yes	yes		
## 3	other	mother	1	2	3	yes	no	yes		
## 30	home	mother	1	2	0	no	yes	yes		
## 82	home	mother	1	3	0	yes	no	yes		
## 143	course	mother	1	3	0	no	yes	yes		
## 93	home	father	1	2	0	yes	yes	no		
## 337	home	other	1	3	1	no	no	yes		
## 221	reputation	mother	2	2	0	no	yes	no		
## 364	course	mother	1	2	0	no	yes	yes		
## 96	home	mother	2	4	1	yes	yes	yes		
## 126	home	father	1	1	0	yes	no	no		
## 214	home	mother	1	2	1	no	yes	yes		
## 188	course	mother	1	2	0	no	no	yes		
## 285	course	mother	1	2	0	no	yes	no		
## 371	home	other	2	2	2	no	no	no		
## 271	home	other	1	2	2	no	yes	yes		
## 295	reputation	mother	2	3	0	no	yes	yes		
## 157	reputation	mother	1	1	0	no	no	no		
## 25	course	mother	1	3	0	yes	yes	yes		
## 212	home	mother	1	2	0	no	yes	yes		
## 238	home	mother	1	1	0	no	no	no		
## 156	course	mother	1	2	0	yes	no	yes		
## 213	reputation	mother	1	2	0	yes	yes	yes		
## 148	course	mother	1	2	0	no	yes	yes		
## 234	reputation	father	1	2	0	no	yes	yes		
## 56	other	mother	1	2	0	no	no	yes		
## 107	course	mother	1	4	0	yes	yes	yes		
## 154	home	mother	1	1	3	no	yes	no		
## 191	course	mother	1	2	0	no	no	no		
## 249	course	mother	1	2	1	no	yes	no		
## 366	course	mother	2	2	0	no	yes	yes		
## 54	course	mother	1	1	0	yes	yes	yes		
##	activities	nursery	higher	internet	romantic	famrel	freetime	goout	Dalc	Walc
## 261	no	yes	yes	yes	yes	low	low	low	low	low
## 186	yes	yes	yes	yes	yes	high	low	high	low	low
## 140	yes	yes	yes	yes	no	high	low	low	low	low
## 36	yes	yes	yes	no	no	low	high	low	low	low
## 394	no	no	yes	yes	no	high	high	low	low	high
## 363	no	yes	yes	yes	yes	high	low	low	low	low
## 276	no	yes	yes	yes	yes	high	high	high	low	low
## 330	no	no	yes	yes	yes	high	low	low	low	low
## 183	yes	yes	yes	no	no	high	high	low	low	low
## 168	no	yes	yes	yes	yes	high	low	low	low	low
## 48	yes	yes	yes	yes	no	high	low	low	low	low
## 104	no	yes	yes	yes	no	high	low	high	low	low
## 136	yes	yes	yes	yes	yes	high	low	low	low	low

## 37	yes	yes	yes	yes	no	high	high	low	low	low
## 108	yes	yes	yes	yes	no	high	low	low	low	low
## 395	no	yes	yes	yes	no	low	low	low	low	low
## 165	yes	yes	no	no	yes	high	low	high	low	high
## 137	no	yes	yes	no	no	high	high	high	low	high
## 256	yes	yes	yes	yes	no	high	high	high	low	low
## 376	no	yes	yes	yes	no	high	low	low	low	low
## 374	yes	yes	yes	yes	no	low	high	high	low	low
## 274	yes	no	yes	no	yes	low	high	low	low	low
## 12	yes	yes	yes	yes	no	high	low	low	low	low
## 40	yes	yes	yes	no	no	high	low	low	low	low
## 22	no	yes	yes	yes	no	high	high	low	low	low
## 70	no	no	yes	yes	no	high	high	low	low	low
## 296	no	yes	yes	yes	no	high	high	low	low	low
## 15	no	yes	yes	yes	yes	high	high	low	low	low
## 388	yes	no	yes	yes	no	high	high	low	low	low
## 75	yes	yes	yes	yes	no	high	low	low	low	high
## 29	yes	yes	yes	yes	no	high	low	low	low	low
## 343	yes	yes	yes	yes	yes	high	low	low	low	low
## 131	no	yes	yes	yes	yes	high	low	low	low	low
## 62	yes	no	yes	yes	yes	high	high	high	high	high
## 127	yes	yes	yes	yes	yes	high	low	low	low	low
## 101	yes	yes	yes	yes	no	high	high	high	high	high
## 247	no	yes	yes	yes	no	high	low	low	low	low
## 182	yes	yes	yes	yes	yes	high	low	low	low	low
## 383	yes	yes	yes	yes	no	high	high	low	low	low
## 138	yes	no	yes	yes	yes	high	low	low	low	low
## 68	no	yes	yes	yes	no	high	low	low	low	low
## 33	yes	yes	yes	yes	yes	high	high	low	low	low
## 275	no	yes	yes	yes	yes	high	low	low	low	low
## 166	yes	no	no	no	no	high	high	low	low	low
## 185	no	yes	yes	yes	no	low	low	low	low	low
## 237	yes	no	yes	yes	yes	high	high	low	high	high
## 192	yes	yes	yes	yes	no	high	low	low	low	low
## 391	no	yes	yes	no	no	high	high	high	high	high
## 65	yes	yes	yes	yes	yes	high	high	high	low	high
## 9	no	yes	yes	yes	no	high	low	low	low	low
## 241	yes	yes	yes	yes	yes	low	high	high	low	high
## 105	yes	yes	yes	yes	no	high	high	high	low	low
## 197	no	yes	yes	yes	no	high	low	low	low	low
## 128	no	no	no	no	no	low	high	low	low	low
## 198	yes	yes	yes	yes	no	low	low	high	low	high
## 317	yes	yes	yes	yes	no	high	low	low	low	low
## 171	yes	no	yes	yes	no	low	high	high	low	high
## 392	no	no	yes	yes	no	low	high	high	low	high
## 334	yes	no	yes	yes	yes	high	low	low	low	low
## 73	no	no	yes	yes	yes	low	low	high	low	high
## 195	no	yes	yes	yes	no	high	low	low	low	low
## 204	no	yes	yes	yes	no	high	low	low	low	low
## 50	yes	no	yes	yes	no	high	high	high	low	low
## 145	no	yes	yes	yes	no	high	high	high	low	low
## 176	yes	yes	yes	yes	no	high	high	high	high	high
## 164	no	yes	no	yes	no	high	low	low	low	high
## 352	no	yes	yes	yes	no	high	high	high	low	low

## 225	yes	no	yes	yes	no	high	low	low	low	low
## 258	no	yes	yes	yes	no	high	low	low	low	low
## 117	yes	yes	yes	no	no	high	high	low	low	low
## 88	yes	yes	yes	yes	no	high	low	low	low	low
## 359	no	no	yes	yes	yes	low	low	low	low	low
## 71	no	yes	yes	yes	no	high	low	low	low	low
## 266	yes	yes	yes	yes	no	high	low	high	low	high
## 102	yes	yes	yes	yes	yes	high	high	low	low	low
## 302	no	yes	yes	yes	no	high	low	low	low	low
## 19	yes	yes	yes	yes	no	high	high	high	low	high
## 47	no	yes	yes	yes	no	low	low	high	low	high
## 161	yes	yes	no	yes	yes	low	low	low	low	low
## 66	yes	yes	yes	yes	no	high	high	low	low	low
## 245	no	no	yes	yes	yes	high	high	high	low	low
## 303	no	yes	yes	yes	no	high	low	low	low	low
## 139	no	yes	yes	no	yes	high	high	high	low	low
## 218	no	yes	yes	yes	no	low	low	high	low	high
## 362	yes	no	no	no	no	high	high	low	low	low
## 4	yes	yes	yes	yes	yes	low	low	low	low	low
## 16	no	yes	yes	yes	no	high	high	high	low	low
## 326	yes	yes	yes	yes	no	high	low	low	low	low
## 262	no	no	yes	yes	no	high	low	low	low	low
## 309	yes	yes	yes	no	yes	high	high	low	low	low
## 99	yes	no	yes	yes	no	high	low	high	low	low
## 393	no	no	yes	no	no	high	high	low	low	low
## 123	yes	yes	yes	yes	yes	high	low	low	low	low
## 365	no	no	yes	no	no	low	low	low	low	low
## 10	yes	yes	yes	yes	no	high	high	low	low	low
## 28	no	yes	yes	yes	no	low	low	high	low	high
## 351	no	yes	yes	yes	no	high	high	high	low	low
## 91	no	yes	yes	yes	yes	high	low	low	low	low
## 114	no	yes	yes	yes	no	low	high	low	low	low
## 265	yes	yes	yes	yes	yes	high	low	low	low	low
## 313	no	no	yes	yes	no	high	high	low	low	low
## 368	no	yes	yes	yes	yes	high	low	low	low	low
## 233	yes	yes	yes	yes	yes	high	high	high	low	low
## 206	no	yes	yes	yes	yes	high	high	low	low	high
## 7	no	yes	yes	yes	no	high	high	high	low	low
## 332	no	yes	yes	yes	yes	high	high	low	low	low
## 222	yes	yes	yes	no	yes	high	low	high	low	low
## 109	yes	yes	yes	yes	yes	low	low	high	low	high
## 252	no	no	yes	yes	no	high	low	low	low	low
## 348	no	yes	yes	yes	yes	high	high	high	low	low
## 232	yes	yes	yes	yes	no	high	high	low	low	low
## 346	no	yes	yes	yes	yes	high	high	low	low	low
## 55	no	yes	yes	yes	no	high	low	high	high	high
## 259	yes	yes	yes	yes	no	high	low	high	low	low
## 373	yes	yes	yes	no	yes	low	high	low	low	low
## 193	yes	no	yes	yes	no	high	high	high	high	high
## 112	yes	yes	yes	yes	no	high	low	low	low	low
## 44	no	yes	yes	yes	no	high	high	low	low	low
## 354	no	yes	yes	yes	no	high	high	high	low	low
## 53	no	yes	yes	no	no	high	high	high	low	high
## 13	yes	yes	yes	yes	no	high	low	low	low	low



## 110	yes	yes	yes	yes	yes	high	high	high	low	low
## 14	no	yes	yes	yes	no	high	high	low	low	low
## 377	no	no	yes	yes	yes	high	high	low	low	low
## 85	yes	no	yes	yes	no	high	low	low	low	low
## 179	yes	yes	yes	yes	yes	high	low	low	low	high
## 163	yes	yes	yes	no	no	high	high	high	low	high
## 272	yes	yes	yes	yes	yes	high	high	high	low	low
## 205	yes	no	yes	yes	no	high	low	high	low	low
## 113	yes	yes	yes	yes	no	low	low	low	low	low
## 228	no	no	yes	yes	no	high	low	low	low	low
## 170	no	yes	yes	yes	yes	high	high	low	low	low
## 8	no	yes	yes	no	no	high	low	high	low	low
## 327	yes	no	yes	yes	no	high	low	high	low	high
## 319	yes	no	yes	yes	no	high	low	high	low	high
## 201	yes	yes	yes	yes	no	high	low	high	low	high
## 74	yes	yes	yes	no	no	high	low	low	low	low
## 49	no	yes	yes	no	no	high	low	low	low	low
## 151	no	yes	no	yes	yes	low	low	high	low	high
## 338	no	yes	yes	yes	yes	high	low	low	low	low
## 253	no	no	no	yes	no	low	low	high	low	high
## 35	no	no	yes	yes	no	high	high	low	low	low
## 308	no	yes	yes	yes	yes	high	low	high	low	low
## 321	no	yes	yes	yes	yes	high	low	low	low	low
## 125	no	yes	yes	yes	yes	high	high	high	low	low
## 310	yes	no	yes	yes	no	high	high	low	low	low
## 324	no	no	yes	yes	no	low	high	low	low	low
## 2	no	no	yes	yes	no	high	low	low	low	low
## 100	no	yes	yes	yes	no	high	low	high	low	low
## 372	yes	yes	no	yes	yes	high	low	low	low	low
## 325	no	yes	yes	yes	no	low	low	low	low	low
## 367	no	yes	yes	yes	yes	high	low	low	low	low
## 84	yes	yes	yes	yes	no	high	low	low	low	low
## 90	no	yes	yes	no	no	high	low	low	low	high
## 314	no	no	yes	yes	yes	high	low	low	low	low
## 267	yes	yes	yes	yes	yes	high	high	high	low	high
## 153	yes	yes	yes	yes	yes	high	low	low	low	low
## 350	no	no	yes	yes	no	low	high	high	high	high
## 211	yes	yes	yes	yes	no	high	low	low	low	low
## 342	yes	yes	yes	yes	no	high	low	low	low	low
## 217	no	yes	yes	yes	yes	low	high	high	low	high
## 370	no	no	yes	yes	yes	low	low	low	high	low
## 208	yes	yes	yes	yes	yes	low	low	low	low	low
## 67	yes	no	yes	yes	yes	low	low	low	high	high
## 162	no	yes	yes	yes	yes	high	high	high	low	high
## 255	no	no	yes	yes	no	high	high	low	low	high
## 301	no	yes	yes	yes	yes	high	low	high	low	low
## 289	yes	yes	yes	yes	no	high	low	high	low	low
## 52	no	yes	yes	yes	no	high	low	low	low	low
## 45	yes	yes	yes	yes	no	high	low	low	low	low
## 333	yes	yes	yes	yes	no	high	low	high	low	low
## 345	no	yes	yes	yes	no	high	low	low	low	low
## 316	no	yes	yes	yes	yes	high	low	low	low	low
## 277	no	no	no	yes	yes	high	low	low	low	low
## 189	no	no	yes	yes	yes	low	low	low	low	low

## 63	yes	yes	yes	yes	no	high	high	low	low	low
## 142	yes	yes	yes	yes	no	low	low	low	low	low
## 106	no	yes	yes	no	no	high	low	low	low	low
## 384	no	yes	yes	no	no	high	low	low	low	low
## 122	yes	yes	yes	yes	no	high	high	high	low	low
## 263	yes	no	yes	no	no	high	low	low	low	low
## 297	yes	yes	yes	yes	no	low	low	high	low	low
## 306	yes	yes	yes	yes	no	high	high	low	low	low
## 87	no	yes	yes	no	no	high	low	high	low	low
## 121	no	no	yes	yes	no	low	low	low	low	low
## 6	yes	yes	yes	yes	no	high	high	low	low	low
## 293	no	yes	yes	yes	yes	high	high	low	low	low
## 356	no	yes	yes	no	yes	high	low	high	low	low
## 81	yes	no	yes	yes	yes	low	low	low	low	low
## 246	no	yes	yes	yes	no	high	low	low	low	low
## 209	no	yes	yes	no	no	high	low	low	low	high
## 39	yes	yes	yes	yes	no	high	low	low	low	low
## 147	no	yes	yes	yes	no	low	low	low	low	low
## 219	no	yes	yes	no	no	low	low	low	low	high
## 169	no	no	yes	yes	no	high	low	high	low	low
## 268	yes	yes	yes	yes	no	high	low	high	low	low
## 278	yes	yes	yes	yes	no	low	low	high	low	high
## 210	yes	yes	yes	yes	yes	high	high	low	low	low
## 124	yes	yes	yes	yes	no	low	high	high	low	high
## 20	yes	yes	yes	yes	no	low	low	low	low	low
## 5	no	yes	yes	no	no	high	low	low	low	low
## 203	no	no	yes	no	no	high	high	high	low	low
## 236	yes	yes	yes	yes	yes	high	low	low	low	low
## 135	no	yes	yes	no	yes	high	low	low	low	low
## 177	yes	no	yes	yes	no	low	high	high	low	high
## 69	no	yes	yes	yes	no	high	low	low	low	low
## 291	yes	yes	yes	yes	yes	high	low	low	low	high
## 380	yes	no	yes	yes	no	high	high	high	low	low
## 299	no	yes	yes	yes	no	high	low	low	low	low
## 311	yes	no	yes	no	yes	high	low	high	low	low
## 202	yes	yes	yes	no	no	high	high	low	low	low
## 146	no	yes	yes	yes	no	high	high	low	low	low
## 223	no	yes	yes	yes	no	low	low	low	low	low
## 390	yes	yes	yes	no	no	low	low	low	low	low
## 181	yes	yes	yes	yes	no	low	high	low	low	low
## 312	no	yes	no	yes	yes	low	high	low	low	low
## 242	no	yes	yes	yes	no	low	low	low	low	low
## 158	yes	no	yes	no	no	high	low	high	low	high
## 382	yes	no	yes	yes	yes	high	high	low	low	low
## 42	no	no	yes	yes	yes	high	high	low	low	high
## 24	yes	yes	yes	yes	no	high	high	high	low	high
## 173	yes	yes	yes	yes	no	high	high	high	low	low
## 294	no	yes	yes	no	no	low	low	low	low	low
## 159	no	no	yes	no	no	high	low	low	low	low
## 323	yes	yes	yes	yes	no	low	low	low	low	low
## 129	yes	yes	yes	yes	no	low	low	low	low	low
## 290	yes	yes	yes	yes	no	high	high	low	low	low
## 79	yes	yes	no	yes	no	high	high	low	low	low
## 286	no	no	yes	yes	no	high	low	low	low	low

## 331	yes	yes	yes	yes	no	high	high	high	low	high
## 144	no	yes	yes	yes	no	low	high	high	low	low
## 41	yes	no	yes	yes	yes	low	low	low	low	low
## 118	yes	yes	yes	yes	no	high	high	low	low	low
## 51	no	yes	yes	yes	no	high	low	low	low	low
## 344	no	yes	yes	yes	yes	low	low	low	low	low
## 361	no	yes	yes	no	yes	high	low	high	low	high
## 349	yes	yes	yes	yes	yes	high	high	low	low	low
## 387	yes	yes	yes	yes	yes	high	high	low	low	low
## 83	no	yes	yes	yes	no	high	high	high	low	low
## 381	yes	no	yes	yes	no	low	low	high	low	high
## 375	no	yes	yes	yes	no	high	high	high	low	low
## 64	yes	yes	yes	yes	no	low	high	high	low	high
## 280	yes	yes	yes	yes	no	high	low	low	low	low
## 270	no	yes	no	yes	yes	high	low	high	low	low
## 23	yes	yes	yes	yes	no	high	high	low	low	low
## 27	no	yes	yes	yes	no	high	low	low	low	low
## 230	yes	yes	yes	yes	yes	low	low	low	low	low
## 172	yes	yes	yes	yes	yes	high	low	low	low	low
## 26	no	no	yes	yes	no	low	low	low	low	low
## 251	no	yes	no	yes	no	high	high	high	low	high
## 86	no	yes	yes	yes	yes	high	high	high	low	low
## 184	yes	yes	yes	yes	yes	high	low	low	low	low
## 360	no	yes	yes	yes	no	high	low	low	low	low
## 335	yes	yes	yes	no	no	high	high	high	low	low
## 307	yes	yes	yes	no	no	high	high	low	low	low
## 248	no	no	no	yes	yes	high	high	high	high	high
## 357	yes	yes	yes	yes	no	high	low	low	low	low
## 339	no	yes	yes	yes	no	high	low	low	low	low
## 132	yes	no	yes	yes	yes	high	low	low	low	low
## 111	yes	yes	yes	yes	no	high	high	low	low	low
## 199	no	yes	yes	yes	no	high	low	high	low	low
## 207	no	yes	yes	yes	no	low	low	low	low	low
## 141	no	yes	yes	yes	no	low	low	low	low	low
## 298	no	yes	yes	yes	yes	high	high	high	low	low
## 77	yes	yes	yes	yes	no	low	high	low	low	low
## 355	yes	no	yes	yes	yes	high	high	high	low	low
## 257	yes	yes	yes	yes	no	high	low	low	low	low
## 318	yes	yes	yes	yes	yes	high	low	high	low	low
## 38	yes	yes	yes	yes	yes	low	high	low	low	low
## 216	no	yes	yes	yes	no	high	high	high	low	low
## 92	yes	yes	yes	yes	no	high	high	high	low	low
## 379	no	yes	yes	yes	yes	high	low	low	low	low
## 190	no	yes	yes	no	no	low	low	low	low	high
## 305	yes	yes	yes	yes	yes	high	high	high	low	low
## 103	yes	no	yes	yes	no	high	low	low	low	low
## 320	no	yes	yes	yes	no	high	high	high	low	low
## 57	yes	yes	yes	yes	no	high	low	low	low	low
## 353	no	yes	no	yes	yes	high	low	low	low	low
## 61	yes	yes	yes	no	no	low	high	high	low	low
## 134	no	yes	yes	yes	no	low	low	low	low	high
## 220	yes	yes	yes	yes	no	high	low	low	low	low
## 287	no	yes	yes	yes	no	high	low	low	low	low
## 167	no	yes	no	yes	no	high	low	high	low	high

## 1	no	yes	yes	no	no	high	low	high	low	low
## 215	no	yes	yes	no	no	high	low	low	low	low
## 3	no	yes	yes	yes	no	high	low	low	low	low
## 30	yes	yes	yes	yes	yes	high	high	high	high	high
## 82	no	no	yes	yes	no	high	low	low	low	low
## 143	yes	yes	yes	yes	no	high	low	low	low	low
## 93	no	yes	yes	no	no	low	low	low	low	low
## 337	no	yes	yes	no	no	high	high	low	low	low
## 221	yes	yes	yes	yes	no	high	low	high	low	low
## 364	yes	yes	yes	yes	yes	low	low	high	low	low
## 96	yes	yes	yes	yes	no	low	low	low	low	low
## 126	no	yes	yes	yes	no	high	high	high	low	low
## 214	yes	yes	yes	yes	no	high	high	high	low	high
## 188	yes	yes	yes	yes	yes	high	low	low	low	low
## 285	no	no	yes	yes	no	high	high	high	low	low
## 371	yes	yes	yes	no	yes	low	low	low	low	low
## 271	yes	yes	yes	yes	no	high	low	high	low	low
## 295	yes	yes	yes	yes	no	high	high	low	low	low
## 157	no	yes	yes	no	no	low	low	low	low	low
## 25	yes	yes	yes	yes	no	high	low	low	low	low
## 212	no	yes	yes	yes	yes	high	low	high	high	high
## 238	no	yes	yes	yes	yes	high	high	low	low	low
## 156	yes	yes	yes	no	no	high	high	high	low	low
## 213	no	yes	yes	yes	no	low	low	high	low	low
## 148	no	no	yes	yes	no	high	low	low	low	low
## 234	yes	yes	yes	yes	no	high	low	high	low	high
## 56	yes	yes	yes	yes	yes	high	low	high	low	low
## 107	no	yes	yes	yes	no	high	low	low	low	low
## 154	no	yes	no	yes	yes	high	high	high	low	low
## 191	no	yes	yes	yes	no	high	low	low	low	low
## 249	no	yes	yes	yes	yes	high	low	low	low	low
## 366	no	yes	yes	no	no	low	low	high	low	high
## 54	no	yes	yes	yes	no	low	low	high	low	low
##	health	absences	G1	G2	G3	first_gen_college	stable_learning_env			
## 261	low	21	17	18	18	no	yes			
## 186	high	12	12	12	11	yes	yes			
## 140	high	0	16	16	15	no	no			
## 36	high	0	8	7	6	yes	no			
## 394	high	0	11	12	10	yes	no			
## 363	low	0	11	11	10	yes	yes			
## 276	high	6	12	12	12	yes	yes			
## 330	high	4	14	14	14	no	yes			
## 183	high	0	16	17	17	no	no			
## 168	low	0	14	15	16	no	no			
## 48	low	4	19	19	20	no	no			
## 104	low	26	7	6	6	yes	yes			
## 136	high	0	11	0	0	no	yes			
## 37	high	2	15	16	18	no	yes			
## 108	high	2	16	18	18	yes	yes			
## 395	high	5	8	9	9	yes	no			
## 165	high	0	5	8	7	yes	no			
## 137	high	0	10	0	0	no	no			
## 256	high	2	7	9	8	yes	yes			
## 376	high	2	8	8	10	yes	no			

## 374	low	14 6 5 5	yes	no
## 274	low	2 15 14 14	yes	no
## 12	high	4 10 12 12	yes	yes
## 40	low	8 14 13 13	yes	no
## 22	high	0 12 15 15	no	yes
## 70	low	12 16 16 16	yes	yes
## 296	high	4 14 12 11	yes	yes
## 15	low	0 14 16 16	yes	yes
## 388	high	0 7 5 0	yes	no
## 75	high	54 11 12 11	yes	yes
## 29	high	4 11 11 11	no	yes
## 343	high	11 16 15 15	no	no
## 131	high	0 12 0 0	no	yes
## 62	high	6 10 8 11	yes	yes
## 127	low	0 7 10 11	no	no
## 101	high	14 7 7 5	no	yes
## 247	low	4 12 12 13	yes	no
## 182	low	2 12 13 12	yes	no
## 383	low	2 11 11 10	yes	no
## 138	high	0 4 0 0	yes	yes
## 68	high	4 7 7 6	yes	yes
## 33	high	0 17 16 16	no	yes
## 275	low	2 10 10 10	no	yes
## 166	low	16 12 11 12	yes	no
## 185	low	14 12 13 12	yes	yes
## 237	high	4 14 13 13	yes	no
## 192	low	0 8 8 9	yes	no
## 391	high	11 9 9 9	yes	no
## 65	low	0 10 10 10	no	no
## 9	low	0 16 18 19	yes	yes
## 241	high	14 12 12 12	no	no
## 105	low	0 16 18 18	no	yes
## 197	high	4 17 15 16	no	no
## 128	high	2 7 8 9	yes	no
## 198	low	8 9 9 10	yes	yes
## 317	low	0 8 8 0	yes	yes
## 171	low	0 6 5 0	no	yes
## 392	low	3 14 16 16	yes	no
## 334	low	0 8 8 0	yes	no
## 73	high	2 8 6 5	yes	yes
## 195	low	0 13 14 14	yes	no
## 204	low	18 7 6 6	yes	yes
## 50	low	2 7 7 7	no	yes
## 145	high	0 5 0 0	yes	yes
## 176	high	4 10 9 9	no	no
## 164	low	2 10 10 10	yes	no
## 352	low	2 13 13 13	yes	yes
## 225	high	0 13 13 14	no	yes
## 258	low	12 11 11 11	no	yes
## 117	low	2 11 13 14	no	no
## 88	low	4 13 14 14	no	yes
## 359	low	4 10 10 10	yes	no
## 71	high	0 13 15 15	yes	yes
## 266	low	13 17 17 17	no	yes

## 102	high	0 16 17 17	no	yes
## 302	high	0 11 11 10	no	no
## 19	high	16 6 5 5	yes	yes
## 47	low	12 11 12 11	yes	yes
## 161	high	0 7 6 0	yes	no
## 66	low	2 16 15 15	no	yes
## 245	low	0 7 0 0	yes	yes
## 303	low	0 15 12 14	no	yes
## 139	high	0 14 12 12	yes	no
## 218	high	13 6 6 8	yes	yes
## 362	high	2 13 12 12	yes	no
## 4	high	2 15 14 15	no	yes
## 16	low	4 14 14 14	no	yes
## 326	low	3 9 12 11	no	no
## 262	low	2 8 8 8	no	yes
## 309	high	0 15 12 12	yes	no
## 99	low	6 11 14 14	no	no
## 393	low	3 10 8 7	yes	no
## 123	high	2 13 13 13	no	yes
## 365	low	0 12 11 12	yes	no
## 10	high	0 14 15 15	no	yes
## 28	low	4 15 16 15	no	no
## 351	low	8 8 7 8	yes	no
## 91	high	0 7 7 8	yes	yes
## 114	low	10 18 19 19	no	no
## 265	low	0 9 10 0	yes	yes
## 313	high	3 13 11 11	yes	no
## 368	low	0 7 6 0	yes	yes
## 233	low	14 11 9 9	no	yes
## 206	high	28 10 9 9	no	yes
## 7	low	0 12 12 11	yes	no
## 332	high	7 12 14 14	no	yes
## 222	high	0 6 5 0	yes	no
## 109	low	6 10 13 13	no	yes
## 252	low	6 7 10 10	yes	yes
## 348	high	0 10 10 9	no	yes
## 232	low	4 11 11 11	yes	yes
## 346	low	7 13 13 14	yes	no
## 55	low	6 10 13 13	yes	no
## 259	high	8 15 14 14	yes	no
## 373	low	8 13 11 11	yes	no
## 193	high	12 7 8 8	yes	no
## 112	low	0 7 10 10	yes	yes
## 44	low	0 8 8 11	yes	yes
## 354	high	4 8 8 8	yes	yes
## 53	high	6 11 11 10	no	no
## 13	high	2 14 14 14	no	yes
## 110	high	4 14 15 16	no	yes
## 14	low	2 10 10 11	no	yes
## 377	low	4 15 14 15	no	yes
## 85	high	2 9 10 10	yes	yes
## 179	low	10 10 8 9	no	yes
## 163	high	0 7 0 0	yes	no
## 272	low	4 15 14 14	yes	yes

## 205	high	6 10 10 11	yes	yes
## 113	high	6 10 13 13	yes	no
## 228	low	2 12 11 12	yes	yes
## 170	low	0 14 14 14	no	yes
## 8	low	6 6 5 6	no	no
## 327	high	3 14 15 16	yes	no
## 319	high	0 11 11 10	no	yes
## 201	low	2 16 16 16	no	yes
## 74	high	2 12 12 14	yes	no
## 49	high	2 15 15 14	no	no
## 151	high	0 6 5 0	yes	no
## 338	low	0 7 8 0	yes	yes
## 253	high	4 6 9 8	yes	no
## 35	high	0 12 14 15	yes	yes
## 308	high	38 8 9 8	no	yes
## 321	high	23 13 13 13	no	yes
## 125	high	0 8 7 8	yes	no
## 310	low	18 12 10 10	yes	yes
## 324	high	1 12 14 15	yes	yes
## 2	low	4 5 5 6	yes	yes
## 100	low	0 7 9 8	no	yes
## 372	low	3 14 12 12	yes	yes
## 325	low	0 16 15 15	yes	no
## 367	high	0 13 13 13	no	no
## 84	high	4 15 15 15	yes	no
## 90	high	18 8 6 7	no	no
## 314	low	22 13 10 11	yes	yes
## 267	high	2 9 9 10	yes	no
## 153	low	8 10 10 10	yes	yes
## 350	high	10 11 13 13	yes	yes
## 211	low	10 8 8 8	yes	yes
## 342	low	0 10 10 0	no	yes
## 217	low	22 6 6 4	no	no
## 370	high	10 14 12 11	no	yes
## 208	low	10 11 12 13	no	no
## 67	low	4 13 13 12	no	yes
## 162	low	6 5 9 7	yes	yes
## 255	high	0 8 12 12	yes	no
## 301	high	14 12 10 11	no	yes
## 289	low	6 15 14 14	yes	no
## 52	high	2 11 13 13	no	yes
## 45	high	14 10 10 9	yes	no
## 333	high	0 7 0 0	yes	no
## 345	low	4 11 10 10	yes	yes
## 316	low	40 13 11 11	yes	no
## 277	high	75 10 9 9	yes	no
## 189	low	6 8 7 9	yes	yes
## 63	low	4 8 10 9	yes	no
## 142	low	8 9 9 9	yes	yes
## 106	high	10 10 11 11	yes	no
## 384	high	0 6 5 0	yes	no
## 122	high	6 16 14 15	yes	yes
## 263	low	1 13 12 12	yes	no
## 297	low	0 10 9 0	no	yes

## 306	low	8 14 12 12	no	yes
## 87	low	4 8 7 6	yes	no
## 121	low	2 16 15 15	yes	no
## 6	high	10 15 15 15	no	yes
## 293	high	12 12 12 13	yes	no
## 356	high	0 10 9 9	yes	no
## 81	low	2 10 12 12	yes	yes
## 246	high	6 18 18 18	yes	no
## 209	high	6 9 9 10	yes	no
## 39	high	2 12 12 11	no	yes
## 147	low	0 6 7 0	yes	yes
## 219	low	3 7 7 8	yes	no
## 169	high	0 6 7 0	yes	yes
## 268	high	8 12 10 11	no	no
## 278	low	22 9 9 9	no	no
## 210	high	6 7 7 7	no	yes
## 124	high	18 14 11 13	no	yes
## 20	high	4 8 10 10	no	no
## 5	high	4 6 10 10	yes	no
## 203	low	4 9 9 10	yes	no
## 236	low	10 11 9 10	yes	no
## 135	high	0 9 0 0	no	no
## 177	high	2 13 13 11	yes	no
## 69	high	2 8 9 8	yes	yes
## 291	high	11 12 11 11	no	yes
## 380	low	17 10 10 10	yes	yes
## 299	low	0 14 13 14	no	yes
## 311	low	0 9 9 0	yes	no
## 202	high	6 8 10 10	yes	no
## 146	high	0 8 11 11	yes	yes
## 223	low	2 16 16 17	yes	no
## 390	high	0 6 5 0	yes	no
## 181	low	10 9 8 8	no	yes
## 312	low	20 14 12 13	yes	yes
## 242	high	2 10 11 12	no	yes
## 158	high	6 9 8 10	yes	no
## 382	high	5 7 6 7	yes	no
## 42	high	8 12 12 12	no	yes
## 24	high	0 13 13 12	yes	yes
## 173	high	0 13 11 10	no	yes
## 294	low	6 18 18 18	yes	no
## 159	low	2 17 15 15	yes	no
## 323	low	3 11 11 11	yes	yes
## 129	high	0 7 4 0	yes	yes
## 290	low	9 15 13 15	no	yes
## 79	low	2 8 8 10	yes	yes
## 286	high	2 12 10 11	yes	no
## 331	high	2 9 8 8	yes	yes
## 144	low	2 14 14 13	yes	no
## 41	low	25 7 10 11	yes	yes
## 118	high	0 13 14 13	yes	no
## 51	high	2 12 13 13	yes	yes
## 344	high	0 9 8 0	yes	yes
## 361	high	0 13 13 13	no	no



## 349	high	0 13 15 15	no	yes
## 387	high	7 6 5 6	no	yes
## 83	high	10 7 6 6	yes	yes
## 381	low	4 15 14 14	no	no
## 375	low	0 19 18 19	no	no
## 64	high	2 10 9 9	no	yes
## 280	low	8 10 11 10	no	no
## 270	low	0 6 0 0	yes	yes
## 23	high	2 15 15 16	no	no
## 27	high	2 12 12 11	yes	yes
## 230	low	10 12 10 12	yes	no
## 172	low	2 13 15 16	yes	yes
## 26	high	14 6 9 8	yes	yes
## 251	high	0 6 8 8	yes	no
## 86	high	6 7 9 8	no	no
## 184	low	56 9 9 8	yes	yes
## 360	high	0 18 16 16	yes	no
## 335	high	0 10 9 0	yes	no
## 307	high	0 17 18 18	yes	no
## 248	low	16 6 8 8	yes	no
## 357	high	4 12 13 13	no	yes
## 339	low	7 16 15 17	yes	yes
## 132	high	0 8 0 0	yes	yes
## 111	high	6 18 19 19	no	no
## 199	low	24 18 18 18	no	yes
## 207	high	5 7 7 7	yes	yes
## 141	low	0 7 9 0	no	yes
## 298	low	10 10 8 8	no	yes
## 77	low	8 11 11 10	no	no
## 355	low	4 13 11 11	no	yes
## 257	high	6 14 12 13	no	yes
## 318	high	9 9 10 9	no	yes
## 38	high	7 15 16 15	no	yes
## 216	low	2 14 15 15	yes	no
## 92	low	4 16 17 18	no	no
## 379	low	0 15 15 15	yes	no
## 190	low	4 8 9 10	yes	no
## 305	low	20 15 14 13	yes	yes
## 103	high	4 10 13 14	no	yes
## 320	high	2 11 11 11	no	yes
## 57	low	0 14 15 15	no	yes
## 353	low	7 8 7 8	yes	no
## 61	high	6 10 11 11	no	no
## 134	high	16 12 11 11	no	no
## 220	high	4 9 10 10	yes	yes
## 287	low	5 18 18 19	yes	yes
## 167	high	4 10 10 10	yes	no
## 1	low	6 5 6 6	no	no
## 215	low	12 8 10 10	no	no
## 3	low	10 7 8 10	yes	no
## 30	high	16 10 12 11	no	yes
## 82	high	4 11 10 11	yes	no
## 143	high	2 9 11 11	no	yes
## 93	low	4 7 6 6	yes	no

## 337	high	12 14 13 13	yes	no
## 221	high	2 6 6 6	yes	yes
## 364	low	0 16 15 15	no	yes
## 96	low	2 7 10 10	yes	yes
## 126	high	0 13 13 12	no	no
## 214	high	15 6 7 8	yes	yes
## 188	high	0 15 15 15	yes	no
## 285	high	4 10 9 11	yes	yes
## 371	low	4 7 7 9	yes	no
## 271	high	15 9 9 9	yes	yes
## 295	high	8 14 13 14	yes	yes
## 157	high	8 16 12 13	yes	no
## 25	high	2 10 9 8	no	yes
## 212	low	13 12 12 13	no	yes
## 238	high	20 13 12 12	yes	no
## 156	low	2 11 8 8	yes	no
## 213	high	0 12 13 14	yes	yes
## 148	high	2 10 11 11	yes	yes
## 234	low	2 14 13 13	no	yes
## 56	low	8 8 9 10	yes	no
## 107	low	8 7 8 8	yes	yes
## 154	high	0 5 0 0	yes	yes
## 191	low	10 11 12 13	yes	no
## 249	high	8 3 5 5	yes	yes
## 366	low	4 10 10 10	yes	no
## 54	high	0 8 10 11	no	yes
##	high_freq_absent	failed	ord_g3	
## 261	yes	no	18	
## 186	yes	no	11	
## 140	no	no	15	
## 36	no	no	6	
## 394	no	no	10	
## 363	no	no	10	
## 276	no	no	12	
## 330	no	no	14	
## 183	no	no	17	
## 168	no	no	16	
## 48	no	no	20	
## 104	yes	no	6	
## 136	no	no	0	
## 37	no	no	18	
## 108	no	no	18	
## 395	no	no	9	
## 165	no	yes	7	
## 137	no	no	0	
## 256	no	yes	8	
## 376	no	no	10	
## 374	yes	no	5	
## 274	no	no	14	
## 12	no	no	12	
## 40	no	no	13	
## 22	no	no	15	
## 70	yes	no	16	
## 296	no	no	11	

## 15	no	no	16
## 388	no	yes	0
## 75	yes	no	11
## 29	no	no	11
## 343	yes	no	15
## 131	no	yes	0
## 62	no	no	11
## 127	no	no	11
## 101	yes	no	5
## 247	no	no	13
## 182	no	no	12
## 383	no	no	10
## 138	no	yes	0
## 68	no	no	6
## 33	no	no	16
## 275	no	no	10
## 166	yes	yes	12
## 185	yes	no	12
## 237	no	no	13
## 192	no	no	9
## 391	yes	yes	9
## 65	no	no	10
## 9	no	no	19
## 241	yes	no	12
## 105	no	no	18
## 197	no	no	16
## 128	no	yes	9
## 198	no	no	10
## 317	no	no	0
## 171	no	yes	0
## 392	no	no	16
## 334	no	no	0
## 73	no	yes	5
## 195	no	no	14
## 204	yes	no	6
## 50	no	yes	7
## 145	no	yes	0
## 176	no	no	9
## 164	no	no	10
## 352	no	no	13
## 225	no	no	14
## 258	yes	no	11
## 117	no	no	14
## 88	no	no	14
## 359	no	no	10
## 71	no	no	15
## 266	yes	no	17
## 102	no	no	17
## 302	no	no	10
## 19	yes	yes	5
## 47	yes	no	11
## 161	no	yes	0
## 66	no	no	15
## 245	no	no	0

## 303	no	no	14
## 139	no	yes	12
## 218	yes	yes	8
## 362	no	yes	12
## 4	no	no	15
## 16	no	no	14
## 326	no	no	11
## 262	no	no	8
## 309	no	yes	12
## 99	no	no	14
## 393	no	yes	7
## 123	no	no	13
## 365	no	no	12
## 10	no	no	15
## 28	no	no	15
## 351	no	yes	8
## 91	no	no	8
## 114	yes	no	19
## 265	no	no	0
## 313	no	yes	11
## 368	no	yes	0
## 233	yes	no	9
## 206	yes	yes	9
## 7	no	no	11
## 332	no	no	14
## 222	no	yes	0
## 109	no	no	13
## 252	no	no	10
## 348	no	no	9
## 232	no	no	11
## 346	no	no	14
## 55	no	no	13
## 259	no	no	14
## 373	no	no	11
## 193	yes	no	8
## 112	no	yes	10
## 44	no	no	11
## 354	no	yes	8
## 53	no	yes	10
## 13	no	no	14
## 110	no	no	16
## 14	no	no	11
## 377	no	yes	15
## 85	no	no	10
## 179	yes	no	9
## 163	no	yes	0
## 272	no	no	14
## 205	no	no	11
## 113	no	yes	13
## 228	no	no	12
## 170	no	no	14
## 8	no	no	6
## 327	no	no	16
## 319	no	no	10

## 201	no	no	16
## 74	no	no	14
## 49	no	no	14
## 151	no	yes	0
## 338	no	no	0
## 253	no	yes	8
## 35	no	no	15
## 308	yes	yes	8
## 321	yes	no	13
## 125	no	no	8
## 310	yes	yes	10
## 324	no	no	15
## 2	no	no	6
## 100	no	no	8
## 372	no	no	12
## 325	no	no	15
## 367	no	no	13
## 84	no	no	15
## 90	yes	no	7
## 314	yes	yes	11
## 267	no	no	10
## 153	no	yes	10
## 350	yes	yes	13
## 211	yes	no	8
## 342	no	yes	0
## 217	yes	yes	4
## 370	yes	no	11
## 208	yes	no	13
## 67	no	no	12
## 162	no	yes	7
## 255	no	no	12
## 301	yes	no	11
## 289	no	no	14
## 52	no	no	13
## 45	yes	yes	9
## 333	no	no	0
## 345	no	no	10
## 316	yes	yes	11
## 277	yes	no	9
## 189	no	no	9
## 63	no	no	9
## 142	no	yes	9
## 106	yes	no	11
## 384	no	yes	0
## 122	no	no	15
## 263	no	no	12
## 297	no	no	0
## 306	no	yes	12
## 87	no	no	6
## 121	no	no	15
## 6	yes	no	15
## 293	yes	yes	13
## 356	no	no	9
## 81	no	no	12

## 246	no	no	18
## 209	no	no	10
## 39	no	no	11
## 147	no	yes	0
## 219	no	no	8
## 169	no	no	0
## 268	no	no	11
## 278	yes	no	9
## 210	no	no	7
## 124	yes	no	13
## 20	no	no	10
## 5	no	no	10
## 203	no	no	10
## 236	yes	no	10
## 135	no	no	0
## 177	no	no	11
## 69	no	no	8
## 291	yes	no	11
## 380	yes	no	10
## 299	no	no	14
## 311	no	yes	0
## 202	no	no	10
## 146	no	no	11
## 223	no	no	17
## 390	no	yes	0
## 181	yes	no	8
## 312	yes	no	13
## 242	no	no	12
## 158	no	yes	10
## 382	no	no	7
## 42	no	no	12
## 24	no	no	12
## 173	no	no	10
## 294	no	no	18
## 159	no	no	15
## 323	no	no	11
## 129	no	yes	0
## 290	no	no	15
## 79	no	yes	10
## 286	no	no	11
## 331	no	no	8
## 144	no	no	13
## 41	yes	yes	11
## 118	no	no	13
## 51	no	no	13
## 344	no	yes	0
## 361	no	no	13
## 349	no	no	15
## 387	no	no	6
## 83	yes	no	6
## 381	no	no	14
## 375	no	no	19
## 64	no	no	9
## 280	no	no	10

## 270	no	no	0
## 23	no	no	16
## 27	no	no	11
## 230	yes	no	12
## 172	no	no	16
## 26	yes	yes	8
## 251	no	yes	8
## 86	no	yes	8
## 184	yes	no	8
## 360	no	no	16
## 335	no	no	0
## 307	no	no	18
## 248	yes	yes	8
## 357	no	no	13
## 339	no	no	17
## 132	no	no	0
## 111	no	no	19
## 199	yes	yes	18
## 207	no	yes	7
## 141	no	no	0
## 298	yes	no	8
## 77	no	no	10
## 355	no	no	11
## 257	no	no	13
## 318	no	no	9
## 38	no	no	15
## 216	no	no	15
## 92	no	no	18
## 379	no	no	15
## 190	no	no	10
## 305	yes	yes	13
## 103	no	no	14
## 320	no	no	11
## 57	no	no	15
## 353	no	yes	8
## 61	no	no	11
## 134	yes	no	11
## 220	no	no	10
## 287	no	no	19
## 167	no	no	10
## 1	no	no	6
## 215	yes	no	10
## 3	yes	yes	10
## 30	yes	no	11
## 82	no	no	11
## 143	no	no	11
## 93	no	no	6
## 337	yes	yes	13
## 221	no	no	6
## 364	no	no	15
## 96	no	yes	10
## 126	no	no	12
## 214	yes	yes	8
## 188	no	no	15

## 285		no	no	11
## 371		no	yes	9
## 271		yes	yes	9
## 295		no	no	14
## 157		no	no	13
## 25		no	no	8
## 212		yes	no	13
## 238		yes	no	12
## 156		no	no	8
## 213		no	no	14
## 148		no	no	11
## 234		no	no	13
## 56		no	no	10
## 107		no	no	8
## 154		no	yes	0
## 191		yes	no	13
## 249		no	yes	5
## 366		no	no	10
## 54		no	no	11

testing

##	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob
## 11	GP	F	15	U	GT3	T	4	4	teacher	health
## 17	GP	F	16	U	GT3	T	4	4	services	services
## 18	GP	F	16	U	GT3	T	3	3	other	other
## 21	GP	M	15	U	GT3	T	4	3	teacher	other
## 31	GP	M	15	U	GT3	T	4	4	health	services
## 32	GP	M	15	U	GT3	T	4	4	services	services
## 34	GP	M	15	U	LE3	T	3	3	other	other
## 43	GP	M	15	U	GT3	T	4	4	services	teacher
## 46	GP	F	15	U	LE3	A	4	3	other	other
## 58	GP	M	15	U	GT3	T	4	4	teacher	health
## 59	GP	M	15	U	LE3	T	1	2	other	at_home
## 60	GP	F	16	U	GT3	T	4	2	services	other
## 72	GP	M	15	U	GT3	T	4	2	other	other
## 76	GP	M	15	U	GT3	T	4	3	teacher	other
## 78	GP	F	16	U	GT3	T	2	2	other	other
## 80	GP	F	16	U	GT3	T	3	4	at_home	other
## 89	GP	M	16	U	GT3	T	2	2	services	other
## 94	GP	F	16	U	GT3	T	4	2	teacher	services
## 95	GP	M	15	U	LE3	T	2	2	services	health
## 97	GP	M	16	R	GT3	T	4	3	services	other
## 98	GP	F	16	U	GT3	T	2	1	other	other
## 115	GP	M	15	R	GT3	T	2	1	health	services
## 116	GP	M	16	U	GT3	T	4	4	teacher	teacher
## 119	GP	M	17	R	GT3	T	1	3	other	other
## 120	GP	M	15	U	GT3	T	3	4	other	other
## 130	GP	M	16	R	GT3	T	4	4	teacher	teacher
## 133	GP	F	17	U	LE3	T	2	2	other	other
## 149	GP	M	16	U	GT3	T	4	4	teacher	teacher
## 150	GP	M	15	U	LE3	A	2	1	services	other
## 152	GP	M	16	U	LE3	T	2	1	at_home	other
## 155	GP	F	17	U	GT3	T	4	4	other	teacher
## 160	GP	M	16	U	GT3	T	3	3	other	services



## 174	GP	F	16	U	GT3	T	1	3	at_home	services
## 175	GP	F	16	U	LE3	T	3	3	other	other
## 178	GP	M	17	U	GT3	T	3	3	other	other
## 180	GP	M	17	U	GT3	T	4	3	other	other
## 187	GP	M	16	U	GT3	T	1	2	services	services
## 194	GP	M	16	R	GT3	T	3	3	services	services
## 196	GP	F	17	U	LE3	T	2	4	services	services
## 200	GP	F	16	U	LE3	T	4	4	teacher	teacher
## 224	GP	M	18	U	GT3	T	2	2	other	other
## 226	GP	F	18	R	GT3	T	3	1	other	other
## 227	GP	F	17	U	GT3	T	3	2	other	other
## 229	GP	M	18	U	LE3	T	2	1	at_home	other
## 231	GP	F	17	U	LE3	T	4	3	health	other
## 235	GP	M	16	U	LE3	T	1	1	other	other
## 239	GP	F	17	R	GT3	T	2	1	at_home	services
## 240	GP	M	18	U	GT3	T	2	2	other	services
## 243	GP	M	16	U	LE3	T	4	3	teacher	other
## 244	GP	M	16	U	GT3	T	4	4	services	services
## 250	GP	M	16	U	GT3	T	0	2	other	other
## 254	GP	M	16	R	GT3	T	2	1	other	other
## 260	GP	F	17	U	LE3	T	2	2	services	services
## 264	GP	F	17	U	GT3	T	3	3	other	other
## 269	GP	M	18	U	GT3	T	4	2	health	other
## 273	GP	F	18	U	LE3	T	1	1	other	other
## 279	GP	F	18	U	GT3	T	4	4	health	health
## 281	GP	M	17	U	LE3	A	4	1	services	other
## 282	GP	M	17	U	LE3	A	3	2	teacher	services
## 283	GP	F	18	R	LE3	T	1	1	at_home	other
## 284	GP	F	18	U	GT3	T	1	1	other	other
## 288	GP	F	17	U	GT3	T	1	1	services	teacher
## 292	GP	F	17	U	GT3	T	4	3	health	services
## 300	GP	M	18	U	LE3	T	4	4	teacher	teacher
## 304	GP	F	17	U	GT3	T	3	2	health	health
## 315	GP	F	19	U	GT3	T	1	1	at_home	health
## 322	GP	F	17	U	GT3	T	2	2	other	other
## 328	GP	M	17	R	GT3	T	2	2	services	other
## 329	GP	F	17	U	GT3	T	4	4	teacher	services
## 336	GP	F	17	U	GT3	T	3	4	services	other
## 340	GP	F	17	R	GT3	A	3	2	other	other
## 341	GP	F	19	U	GT3	T	2	1	services	services
## 347	GP	M	18	R	GT3	T	4	3	teacher	services
## 358	MS	F	17	U	LE3	A	3	2	services	other
## 369	MS	F	18	U	GT3	T	2	3	at_home	services
## 378	MS	F	18	R	LE3	T	4	4	teacher	services
## 385	MS	M	18	R	GT3	T	4	2	other	other
## 386	MS	F	18	R	GT3	T	2	2	at_home	other
## 389	MS	F	18	U	LE3	T	3	1	teacher	services
##	reason	guardian	traveltime	studytime	failures	schoolsup	famsup	paid		
## 11	reputation	mother	1	2	0	no	yes	yes		
## 17	reputation	mother	1	3	0	no	yes	yes		
## 18	reputation	mother	3	2	0	yes	yes	no		
## 21	reputation	mother	1	2	0	no	no	no		
## 31	home	mother	1	2	0	no	yes	yes		
## 32	reputation	mother	2	2	0	no	yes	no		

## 34	course	mother	1	2	0	no	no	no
## 43	course	father	1	2	0	no	yes	no
## 46	course	mother	1	2	0	yes	yes	yes
## 58	reputation	mother	1	2	0	no	yes	no
## 59	home	father	1	2	0	yes	yes	no
## 60	course	mother	1	2	0	no	yes	no
## 72	course	mother	1	4	0	no	no	no
## 76	home	mother	1	2	0	no	yes	yes
## 78	reputation	mother	1	4	0	no	no	yes
## 80	course	mother	1	2	0	no	yes	no
## 89	reputation	father	2	2	1	no	no	yes
## 94	home	mother	2	2	0	no	yes	yes
## 95	reputation	mother	1	4	0	no	yes	no
## 97	reputation	mother	2	1	0	yes	yes	no
## 98	course	mother	1	2	0	no	yes	yes
## 115	reputation	mother	1	2	0	no	no	no
## 116	course	father	1	2	0	no	yes	no
## 119	course	father	3	2	1	no	yes	no
## 120	reputation	father	1	1	0	no	no	no
## 130	course	mother	1	1	0	no	no	yes
## 133	course	father	1	1	0	no	yes	no
## 149	course	mother	1	1	0	no	yes	no
## 150	course	mother	4	1	3	no	no	no
## 152	course	mother	1	1	1	no	no	no
## 155	course	mother	1	1	0	yes	yes	no
## 160	course	father	1	2	1	no	yes	yes
## 174	home	mother	1	2	3	no	no	no
## 175	reputation	mother	2	2	0	no	yes	yes
## 178	reputation	father	1	2	0	no	no	no
## 180	course	mother	1	2	0	no	yes	no
## 187	other	mother	1	1	0	no	yes	yes
## 194	reputation	mother	1	1	0	no	yes	no
## 196	course	father	1	2	0	no	no	no
## 200	reputation	mother	1	2	0	no	yes	yes
## 224	home	mother	2	2	0	no	yes	yes
## 226	reputation	mother	1	2	1	no	no	no
## 227	course	mother	1	2	0	no	no	no
## 229	course	mother	4	2	0	yes	yes	yes
## 231	reputation	father	1	2	0	no	no	no
## 235	home	mother	2	2	0	no	yes	yes
## 239	course	mother	3	2	0	no	no	no
## 240	reputation	father	1	2	1	no	no	no
## 243	course	mother	1	1	0	no	no	no
## 244	course	mother	1	1	0	no	no	no
## 250	other	mother	1	1	0	no	no	yes
## 254	course	mother	2	1	0	no	no	no
## 260	course	father	1	4	0	no	no	yes
## 264	home	mother	1	3	0	no	no	no
## 269	reputation	father	1	2	0	no	yes	yes
## 273	home	mother	2	2	0	no	yes	yes
## 279	reputation	father	1	2	1	yes	yes	no
## 281	home	mother	2	1	0	no	no	yes
## 282	home	mother	1	1	1	no	no	no
## 283	reputation	mother	2	4	0	no	yes	yes

## 284	home	mother	2	2	0	yes	no	no		
## 288	reputation	mother	1	3	0	no	yes	yes		
## 292	reputation	mother	1	3	0	no	yes	yes		
## 300	home	mother	1	1	0	no	yes	yes		
## 304	reputation	father	1	4	0	no	yes	yes		
## 315	home	other	1	3	2	no	no	no		
## 322	course	mother	1	2	0	no	yes	no		
## 328	course	mother	4	1	0	no	yes	no		
## 329	course	mother	1	3	0	no	yes	yes		
## 336	course	mother	1	3	0	no	no	no		
## 340	home	mother	1	2	0	no	yes	yes		
## 341	home	other	1	3	1	no	no	yes		
## 347	course	mother	1	3	0	no	no	no		
## 358	reputation	mother	2	2	0	no	no	no		
## 369	course	father	2	1	0	no	yes	yes		
## 378	course	mother	1	2	0	no	no	yes		
## 385	home	father	2	1	1	no	no	yes		
## 386	other	mother	2	3	0	no	no	yes		
## 389	course	mother	1	2	0	no	yes	yes		
##	activities	nursery	higher	internet	romantic	famrel	freetime	goout	Dalc	Walc
## 11	no	yes	yes	yes	no	low	low	low	low	low
## 17	yes	yes	yes	yes	no	low	low	low	low	low
## 18	yes	yes	yes	no	no	high	low	low	low	low
## 21	no	yes	yes	yes	no	high	high	low	low	low
## 31	no	no	yes	yes	no	high	high	low	low	high
## 32	yes	yes	yes	yes	no	high	low	low	low	low
## 34	yes	no	yes	yes	no	high	low	low	low	low
## 43	yes	yes	yes	yes	no	high	low	low	low	low
## 46	yes	yes	yes	yes	yes	high	low	low	low	low
## 58	yes	yes	yes	no	no	low	low	low	low	low
## 59	yes	yes	yes	yes	no	high	low	low	low	low
## 60	no	yes	yes	yes	no	high	low	low	low	low
## 72	no	yes	yes	yes	no	low	low	low	low	low
## 76	yes	yes	yes	yes	no	high	low	low	low	low
## 78	no	yes	yes	yes	yes	high	low	low	low	low
## 80	no	yes	yes	yes	no	low	high	low	low	low
## 89	yes	no	yes	yes	no	high	high	low	low	low
## 94	yes	yes	yes	yes	no	high	low	low	low	low
## 95	yes	yes	yes	yes	no	high	low	high	low	low
## 97	yes	no	yes	yes	no	low	low	low	low	low
## 98	no	yes	yes	no	yes	high	low	high	low	low
## 115	yes	yes	yes	yes	yes	high	high	low	low	low
## 116	yes	yes	yes	yes	no	high	high	high	low	low
## 119	yes	yes	yes	yes	no	high	low	high	low	high
## 120	no	yes	yes	yes	no	low	high	low	low	low
## 130	yes	yes	yes	yes	no	low	high	high	low	high
## 133	no	yes	yes	yes	yes	low	high	high	low	low
## 149	no	yes	no	yes	yes	low	low	low	low	low
## 150	no	yes	yes	yes	no	high	high	high	low	high
## 152	yes	yes	yes	no	yes	high	high	high	low	high
## 155	no	yes	yes	no	yes	high	low	low	low	low
## 160	no	yes	yes	yes	yes	high	high	high	high	high
## 174	yes	no	yes	yes	yes	high	low	high	low	low
## 175	yes	yes	yes	yes	no	high	high	high	low	low

## 178	yes	no	yes	yes	no	high	low	high	low	high
## 180	yes	yes	yes	yes	yes	high	low	low	low	low
## 187	yes	yes	yes	yes	yes	low	low	low	low	low
## 194	yes	yes	yes	yes	no	high	low	low	low	high
## 196	yes	yes	yes	yes	yes	high	low	low	low	low
## 200	no	yes	yes	yes	no	high	high	low	low	low
## 224	no	yes	yes	yes	no	low	low	low	high	high
## 226	yes	yes	yes	yes	yes	high	low	low	low	low
## 227	yes	no	yes	yes	no	high	low	high	low	low
## 229	yes	yes	yes	yes	yes	high	low	low	high	high
## 231	yes	yes	yes	yes	yes	low	low	low	low	low
## 235	no	yes	yes	yes	no	low	high	low	low	low
## 239	yes	yes	yes	no	no	low	low	low	low	low
## 240	no	yes	no	yes	no	high	high	high	low	high
## 243	yes	no	yes	yes	no	high	high	high	low	low
## 244	yes	yes	yes	yes	no	high	low	low	low	low
## 250	no	no	yes	yes	no	high	low	low	low	high
## 254	yes	no	yes	no	no	low	low	low	low	low
## 260	yes	yes	yes	yes	yes	low	high	low	low	low
## 264	yes	no	yes	no	no	low	low	low	low	low
## 269	yes	yes	yes	yes	yes	high	high	high	low	low
## 273	no	no	yes	no	no	high	high	low	low	low
## 279	yes	yes	yes	yes	yes	low	high	high	low	low
## 281	yes	yes	yes	yes	yes	high	high	high	low	high
## 282	no	yes	yes	yes	no	high	high	high	low	high
## 283	yes	yes	yes	no	no	high	low	low	low	low
## 284	yes	yes	yes	yes	no	high	high	high	low	low
## 288	no	yes	yes	yes	no	high	low	low	low	low
## 292	no	yes	yes	yes	no	high	low	low	low	low
## 300	no	yes	yes	yes	yes	low	high	low	low	low
## 304	yes	no	yes	yes	no	high	low	low	low	low
## 315	no	no	yes	yes	yes	high	low	low	low	low
## 322	no	yes	yes	no	yes	high	low	low	low	low
## 328	no	yes	yes	yes	no	high	high	high	high	high
## 329	yes	yes	yes	yes	no	high	high	high	low	low
## 336	no	yes	yes	yes	no	high	high	high	low	low
## 340	no	yes	yes	yes	no	high	low	low	low	low
## 341	yes	yes	yes	yes	yes	high	low	high	low	low
## 347	no	yes	yes	yes	yes	high	low	low	low	low
## 358	no	yes	yes	no	yes	low	low	low	low	low
## 369	no	yes	yes	yes	yes	high	low	low	low	low
## 378	yes	yes	yes	yes	no	high	high	low	low	high
## 385	no	yes	yes	no	no	high	high	low	high	low
## 386	no	yes	yes	no	no	high	low	low	low	low
## 389	no	yes	yes	yes	no	high	low	high	low	low
##	health	absences	G1	G2	G3	first_gen_college	stable_learning_env			
## 11	low		0	10	8	9	no	yes		
## 17	low		6	13	14	14	no	yes		
## 18	high		4	8	10	10	yes	no		
## 21	low		0	13	14	15	no	no		
## 31	high		0	9	11	12	no	yes		
## 32	high		0	17	16	17	no	yes		
## 34	low		0	8	10	12	yes	no		
## 43	high		2	19	18	18	no	yes		

## 46	high	8 8 8 6	no	yes
## 58	high	4 14 15 15	no	no
## 59	high	2 9 10 9	yes	yes
## 60	high	2 15 16 16	no	yes
## 72	low	0 10 10 10	no	no
## 76	high	6 9 9 10	no	yes
## 78	low	0 11 11 11	yes	no
## 80	low	12 5 5 5	no	yes
## 89	low	12 11 10 10	yes	no
## 94	low	0 11 10 10	no	yes
## 95	high	6 11 13 14	yes	yes
## 97	high	2 11 15 15	no	yes
## 98	high	2 8 9 10	yes	no
## 115	high	8 9 9 9	yes	no
## 116	high	2 15 15 16	no	yes
## 119	high	20 9 7 8	yes	yes
## 120	high	6 14 13 13	no	no
## 130	high	8 18 18 18	no	no
## 133	high	12 10 13 12	yes	yes
## 149	high	0 7 6 0	no	yes
## 150	high	0 8 9 10	yes	no
## 152	high	6 12 13 14	yes	no
## 155	high	0 11 11 12	no	no
## 160	high	4 10 12 12	yes	yes
## 174	low	0 8 7 0	yes	no
## 175	high	4 10 11 9	yes	yes
## 178	high	4 6 5 6	yes	no
## 180	low	4 10 10 11	no	yes
## 187	low	2 11 12 11	yes	yes
## 194	high	8 8 9 10	yes	yes
## 196	high	0 14 15 15	no	no
## 200	low	0 9 9 10	no	yes
## 224	high	0 12 13 13	yes	yes
## 226	high	16 9 8 7	yes	no
## 227	low	10 16 15 15	yes	no
## 229	low	14 10 8 9	yes	yes
## 231	low	14 13 13 14	no	no
## 235	high	18 9 7 6	yes	yes
## 239	low	2 13 11 11	yes	no
## 240	low	0 7 7 0	yes	no
## 243	low	0 6 0 0	no	no
## 244	high	0 13 12 12	no	no
## 250	high	0 13 15 15	yes	no
## 254	low	0 8 9 8	yes	no
## 260	low	0 10 9 0	yes	no
## 264	high	4 10 9 9	yes	no
## 269	high	10 10 9 10	no	yes
## 273	low	2 11 11 11	yes	no
## 279	high	15 9 8 8	no	yes
## 281	high	30 8 8 8	no	no
## 282	low	19 11 9 10	yes	no
## 283	low	1 12 12 12	yes	no
## 284	high	4 8 9 10	yes	no
## 288	low	6 13 12 12	yes	yes

## 292	low	0 15 15 15	no	yes
## 300	low	5 16 15 16	no	yes
## 304	high	0 17 17 18	yes	yes
## 315	low	14 15 13 13	yes	no
## 322	low	12 11 9 9	yes	no
## 328	high	8 11 10 10	yes	yes
## 329	high	7 10 9 9	no	yes
## 336	high	16 16 15 15	no	no
## 340	low	4 9 10 10	yes	yes
## 341	low	4 11 12 11	yes	no
## 347	high	9 16 15 16	no	no
## 358	high	2 12 12 11	yes	no
## 369	high	0 11 10 10	yes	yes
## 378	low	4 8 9 10	no	no
## 385	low	14 6 5 5	no	no
## 386	high	2 10 9 10	yes	no
## 389	low	0 7 9 8	yes	yes
##	high_freq_absent	failed	ord_g3	
## 11		no no	9	
## 17		no no	14	
## 18		no no	10	
## 21		no no	15	
## 31		no no	12	
## 32		no no	17	
## 34		no no	12	
## 43		no no	18	
## 46		no no	6	
## 58		no no	15	
## 59		no no	9	
## 60		no no	16	
## 72		no no	10	
## 76		no no	10	
## 78		no no	11	
## 80		yes no	5	
## 89		yes yes	10	
## 94		no no	10	
## 95		no no	14	
## 97		no no	15	
## 98		no no	10	
## 115		no no	9	
## 116		no no	16	
## 119		yes yes	8	
## 120		no no	13	
## 130		no no	18	
## 133		yes no	12	
## 149		no no	0	
## 150		no yes	10	
## 152		no yes	14	
## 155		no no	12	
## 160		no yes	12	
## 174		no yes	0	
## 175		no no	9	
## 178		no no	6	
## 180		no no	11	

## 187	no	no	11
## 194	no	no	10
## 196	no	no	15
## 200	no	no	10
## 224	no	no	13
## 226	yes	yes	7
## 227	yes	no	15
## 229	yes	no	9
## 231	yes	no	14
## 235	yes	no	6
## 239	no	no	11
## 240	no	yes	0
## 243	no	no	0
## 244	no	no	12
## 250	no	no	15
## 254	no	no	8
## 260	no	no	0
## 264	no	no	9
## 269	yes	no	10
## 273	no	no	11
## 279	yes	yes	8
## 281	yes	no	8
## 282	yes	yes	10
## 283	no	no	12
## 284	no	no	10
## 288	no	no	12
## 292	no	no	15
## 300	no	no	16
## 304	no	no	18
## 315	yes	yes	13
## 322	yes	no	9
## 328	no	no	10
## 329	no	no	9
## 336	yes	no	15
## 340	no	no	10
## 341	no	yes	11
## 347	no	no	16
## 358	no	no	11
## 369	no	no	10
## 378	no	no	10
## 385	yes	yes	5
## 386	no	no	10
## 389	no	no	8

## Linear model

Modeling, diagnostics, predictions

```
base_lm <- lm(G3 ~ . -G2 -G1 -ord_g3 -stable_learning_env, data)
summary(base_lm)
```

```
##
## Call:
## lm(formula = G3 ~ . - G2 - G1 - ord_g3 - stable_learning_env,
##     data = data)
```

```

##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.4629  -1.8563   0.3768   2.7129   8.4183
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    14.26449    4.83968   2.947  0.00342 **
## schoolMS         0.71627    0.78940   0.907  0.36483
## sexM             1.24064    0.49743   2.494  0.01309 *
## age            -0.29977    0.21631  -1.386  0.16667
## addressU         0.50527    0.57951   0.872  0.38386
## famsizeLE3       0.70815    0.48336   1.465  0.14380
## PstatusT        -0.33706    0.72697  -0.464  0.64318
## Medu             0.19327    0.35893   0.538  0.59060
## Fedu            -0.24782    0.30312  -0.818  0.41416
## Mjobhealth       0.77671    1.11854   0.694  0.48789
## Mjobother        -0.37068    0.71864  -0.516  0.60631
## Mjobservices     0.77141    0.79944   0.965  0.33524
## Mjobteacher      -1.45523    1.04350  -1.395  0.16402
## Fjobhealth       0.50264    1.43071   0.351  0.72556
## Fjobother        -0.44303    1.01109  -0.438  0.66153
## Fjobservices     -0.28558    1.05170  -0.272  0.78614
## Fjobteacher       1.18454    1.28968   0.918  0.35900
## reasonhome       0.15104    0.55120   0.274  0.78422
## reasonother      0.90615    0.81174   1.116  0.26505
## reasonreputation 0.78970    0.57320   1.378  0.16917
## guardianmother   -0.06654    0.54705  -0.122  0.90326
## guardianother    0.74621    0.99921   0.747  0.45568
## traveltime       -0.33731    0.34177  -0.987  0.32435
## studytime        0.43921    0.28472   1.543  0.12382
## failures         -0.68669    0.59765  -1.149  0.25134
## schoolsupyes     -1.25886    0.66531  -1.892  0.05929 .
## famsupyes        -0.87494    0.47671  -1.835  0.06729 .
## paidyes          0.27979    0.47654   0.587  0.55749
## activitiesyes    -0.27894    0.44015  -0.634  0.52667
## nurseryyes       -0.12389    0.54987  -0.225  0.82188
## higheryes        1.48003    1.08166   1.368  0.17209
## internetyes      0.45356    0.61399   0.739  0.46058
## romanticyes     -1.16064    0.46719  -2.484  0.01345 *
## famrellow        -0.30748    0.51272  -0.600  0.54909
## freetimelow      -0.79441    0.47382  -1.677  0.09450 .
## gooutlow         1.33550    0.51112   2.613  0.00936 **
## Dalclow          -0.21717    1.15359  -0.188  0.85079
## Walclow          -0.02676    0.67033  -0.040  0.96818
## healthlow        0.20864    0.44481   0.469  0.63932
## absences         0.09519    0.04094   2.325  0.02063 *
## first_gen_collegeyes -1.02075    0.79309  -1.287  0.19892
## high_freq_absentyes -0.77063    0.81373  -0.947  0.34427
## failedyes        -2.27412    1.10329  -2.061  0.04002 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.082 on 352 degrees of freedom

```



```
## Multiple R-squared:  0.2907, Adjusted R-squared:  0.2061
## F-statistic: 3.435 on 42 and 352 DF,  p-value: 1.498e-10
```

```
step.model <- stepAIC(base_lm, direction="both")
```

```
## Start:  AIC=1151.72
## G3 ~ (school + sex + age + address + famsize + Pstatus + Medu +
##       Fedu + Mjob + Fjob + reason + guardian + traveltime + studytime +
##       failures + schoolsup + famsup + paid + activities + nursery +
##       higher + internet + romantic + famrel + freetime + goout +
##       Dalc + Walc + health + absences + G1 + G2 + first_gen_college +
##       stable_learning_env + high_freq_absent + failed + ord_g3) -
##       G2 - G1 - ord_g3 - stable_learning_env
```

	Df	Sum of Sq	RSS	AIC
## - Fjob	4	58.108	5924.0	1147.6
## - guardian	2	12.927	5878.8	1148.6
## - reason	3	46.316	5912.2	1148.8
## - Walc	1	0.027	5865.9	1149.7
## - Dalc	1	0.591	5866.5	1149.8
## - nursery	1	0.846	5866.7	1149.8
## - Pstatus	1	3.582	5869.5	1150.0
## - health	1	3.666	5869.6	1150.0
## - Medu	1	4.832	5870.7	1150.0
## - paid	1	5.745	5871.6	1150.1
## - famrel	1	5.993	5871.9	1150.1
## - activities	1	6.693	5872.6	1150.2
## - internet	1	9.094	5875.0	1150.3
## - Fedu	1	11.139	5877.0	1150.5
## - address	1	12.668	5878.6	1150.6
## - school	1	13.720	5879.6	1150.6
## - high_freq_absent	1	14.946	5880.8	1150.7
## - traveltime	1	16.232	5882.1	1150.8
## - failures	1	22.000	5887.9	1151.2
## - first_gen_college	1	27.605	5893.5	1151.6
## <none>			5865.9	1151.7
## - higher	1	31.200	5897.1	1151.8
## - age	1	32.006	5897.9	1151.9
## - famsize	1	35.768	5901.7	1152.1
## - studytime	1	39.656	5905.5	1152.4
## - freetime	1	46.845	5912.7	1152.9
## - famsup	1	56.136	5922.0	1153.5
## - schoolsup	1	59.662	5925.6	1153.7
## - failed	1	70.801	5936.7	1154.5
## - absences	1	90.094	5956.0	1155.7
## - Mjob	4	192.395	6058.3	1156.5
## - romantic	1	102.848	5968.7	1156.6
## - sex	1	103.661	5969.6	1156.6
## - goout	1	113.773	5979.7	1157.3

```
## Step:  AIC=1147.61
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##       Fedu + Mjob + reason + guardian + traveltime + studytime +
##       failures + schoolsup + famsup + paid + activities + nursery +
##       higher + internet + romantic + famrel + freetime + goout +
```

```

##      Dalc + Walc + health + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq    RSS    AIC
## - guardian      2    10.384 5934.4 1144.3
## - reason         3    59.440 5983.4 1145.6
## - Walc           1     0.559 5924.6 1145.7
## - nursery        1     0.970 5925.0 1145.7
## - Dalc           1     1.241 5925.2 1145.7
## - health         1     2.560 5926.6 1145.8
## - Fedu           1     2.588 5926.6 1145.8
## - paid           1     2.969 5927.0 1145.8
## - Medu           1     3.297 5927.3 1145.8
## - famrel         1     4.231 5928.2 1145.9
## - internet       1     5.645 5929.6 1146.0
## - Pstatus        1     5.731 5929.7 1146.0
## - activities     1    10.513 5934.5 1146.3
## - school         1    11.795 5935.8 1146.4
## - address        1    14.753 5938.8 1146.6
## - traveltime     1    14.820 5938.8 1146.6
## - high_freq_absent 1    18.403 5942.4 1146.8
## - failures       1    22.209 5946.2 1147.1
## <none>                                5924.0 1147.6
## - age           1    30.631 5954.6 1147.7
## - higher        1    30.845 5954.8 1147.7
## - famsize       1    32.187 5956.2 1147.8
## - first_gen_college 1    34.350 5958.4 1147.9
## - studytime     1    43.941 5967.9 1148.5
## - schoolsup     1    49.517 5973.5 1148.9
## - freetime      1    50.138 5974.1 1148.9
## - famsup        1    53.020 5977.0 1149.1
## - failed        1    66.563 5990.6 1150.0
## - Mjob          4   180.518 6104.5 1151.5
## + Fjob          4    58.108 5865.9 1151.7
## - absences      1    94.253 6018.3 1151.8
## - romantic      1    94.936 6018.9 1151.9
## - sex           1   105.155 6029.2 1152.6
## - goout         1   110.854 6034.9 1152.9
##
## Step:  AIC=1144.3
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + reason + traveltime + studytime + failures +
##      schoolsup + famsup + paid + activities + nursery + higher +
##      internet + romantic + famrel + freetime + goout + Dalc +
##      Walc + health + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq    RSS    AIC
## - reason         3    60.268 5994.7 1142.3
## - Walc           1     1.015 5935.4 1142.4
## - Dalc           1     1.194 5935.6 1142.4
## - Fedu           1     1.464 5935.8 1142.4
## - Medu           1     2.315 5936.7 1142.5
## - nursery        1     2.354 5936.7 1142.5

```

```

## - paid          1      3.146 5937.5 1142.5
## - health        1      3.681 5938.1 1142.5
## - famrel        1      4.627 5939.0 1142.6
## - internet      1      5.511 5939.9 1142.7
## - Pstatus       1      6.799 5941.2 1142.8
## - school        1     10.510 5944.9 1143.0
## - activities    1     10.605 5945.0 1143.0
## - traveltime    1     12.005 5946.4 1143.1
## - high_freq_absent 1     18.082 5952.5 1143.5
## - address       1     19.026 5953.4 1143.6
## - failures      1     21.806 5956.2 1143.8
## - age           1     23.499 5957.9 1143.9
## <none>          1          5934.4 1144.3
## - famsize       1     30.756 5965.1 1144.3
## - first_gen_college 1    33.861 5968.2 1144.5
## - higher        1     35.225 5969.6 1144.6
## - studytime     1     47.430 5981.8 1145.5
## - schoolsup     1     48.694 5983.1 1145.5
## - famsup        1     51.296 5985.7 1145.7
## - freetime      1     56.659 5991.0 1146.1
## - failed        1     61.246 5995.6 1146.4
## + guardian      2     10.384 5924.0 1147.6
## - Mjob           4    183.657 6118.0 1148.3
## - romantic      1     92.824 6027.2 1148.4
## + Fjob           4     55.565 5878.8 1148.6
## - absences      1     95.558 6029.9 1148.6
## - sex           1    107.385 6041.8 1149.4
## - goout         1    118.595 6053.0 1150.1
##
## Step:  AIC=1142.3
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + traveltime + studytime + failures + schoolsup +
##      famsup + paid + activities + nursery + higher + internet +
##      romantic + famrel + freetime + goout + Dalc + Walc + health +
##      absences + first_gen_college + high_freq_absent + failed
##
##              Df Sum of Sq    RSS    AIC
## - Walc        1      1.317 5996.0 1140.4
## - Fedu        1      1.469 5996.1 1140.4
## - Dalc        1      1.838 5996.5 1140.4
## - nursery     1      1.951 5996.6 1140.4
## - Medu        1      3.918 5998.6 1140.5
## - internet    1      4.868 5999.5 1140.6
## - activities  1      6.230 6000.9 1140.7
## - famrel      1      6.241 6000.9 1140.7
## - Pstatus     1      6.498 6001.2 1140.7
## - health      1      6.870 6001.5 1140.8
## - paid        1      7.053 6001.7 1140.8
## - school      1     10.560 6005.2 1141.0
## - address     1     12.612 6007.3 1141.1
## - traveltime  1     13.824 6008.5 1141.2
## - high_freq_absent 1    14.302 6009.0 1141.2
## - age         1     25.650 6020.3 1142.0
## - failures    1     27.179 6021.8 1142.1

```

```

## - higher          1    28.065 6022.7 1142.1
## - famsize         1    29.828 6024.5 1142.3
## <none>              5994.7 1142.3
## - first_gen_college 1    30.439 6025.1 1142.3
## - schoolsup        1    48.574 6043.2 1143.5
## - studytime        1    50.699 6045.4 1143.6
## - famsup           1    54.155 6048.8 1143.8
## - failed           1    55.751 6050.4 1144.0
## - freetime         1    56.307 6051.0 1144.0
## + reason           3    60.268 5934.4 1144.3
## + guardian         2    11.211 5983.4 1145.6
## + Fjob              4    68.895 5925.8 1145.7
## - romantic         1    91.877 6086.5 1146.3
## - absences         1    96.791 6091.4 1146.6
## - sex              1    99.491 6094.1 1146.8
## - Mjob              4   216.467 6211.1 1148.3
## - goout            1   126.588 6121.2 1148.5
##
## Step:  AIC=1140.38
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + traveltime + studytime + failures + schoolsup +
##      famsup + paid + activities + nursery + higher + internet +
##      romantic + famrel + freetime + goout + Dalc + health + absences +
##      first_gen_college + high_freq_absent + failed
##
##              Df Sum of Sq    RSS    AIC
## - Dalc         1      1.109 5997.1 1138.5
## - Fedu         1      1.644 5997.6 1138.5
## - nursery       1      1.737 5997.7 1138.5
## - Medu         1      4.078 6000.0 1138.7
## - internet      1      4.884 6000.9 1138.7
## - activities    1      6.130 6002.1 1138.8
## - Pstatus       1      6.137 6002.1 1138.8
## - paid          1      6.727 6002.7 1138.8
## - famrel        1      7.078 6003.0 1138.8
## - health        1      7.451 6003.4 1138.9
## - school        1     11.132 6007.1 1139.1
## - address       1     13.102 6009.1 1139.2
## - traveltime    1     14.450 6010.4 1139.3
## - high_freq_absent 1     14.481 6010.5 1139.3
## - age           1     24.951 6020.9 1140.0
## - failures      1     27.554 6023.5 1140.2
## - higher        1     28.788 6024.8 1140.3
## - famsize       1     29.376 6025.3 1140.3
## <none>           5996.0 1140.4
## - first_gen_college 1     30.478 6026.4 1140.4
## - schoolsup      1     47.915 6043.9 1141.5
## - studytime     1     51.565 6047.5 1141.8
## - famsup        1     53.086 6049.1 1141.9
## - failed        1     55.730 6051.7 1142.0
## - freetime      1     57.401 6053.4 1142.2
## + Walc          1      1.317 5994.7 1142.3
## + reason        3     60.569 5935.4 1142.4
## + guardian      2     11.743 5984.2 1143.6

```

```

## + Fjob          4    70.097 5925.9 1143.7
## - romantic      1    90.982 6087.0 1144.3
## - absences      1    95.639 6091.6 1144.6
## - sex           1    99.432 6095.4 1144.9
## - Mjob          4   217.116 6213.1 1146.4
## - goout         1   159.887 6155.9 1148.8
##
## Step:  AIC=1138.46
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + traveltime + studytime + failures + schoolsup +
##      famsup + paid + activities + nursery + higher + internet +
##      romantic + famrel + freetime + goout + health + absences +
##      first_gen_college + high_freq_absent + failed
##
##              Df Sum of Sq   RSS   AIC
## - Fedu        1     1.590 5998.7 1136.6
## - nursery     1     2.133 5999.2 1136.6
## - Medu        1     4.304 6001.4 1136.7
## - internet    1     4.928 6002.0 1136.8
## - activities  1     6.292 6003.4 1136.9
## - Pstatus     1     6.388 6003.5 1136.9
## - famrel      1     6.962 6004.0 1136.9
## - health      1     7.081 6004.2 1136.9
## - paid        1     7.849 6004.9 1137.0
## - school      1    11.411 6008.5 1137.2
## - address     1    13.515 6010.6 1137.3
## - high_freq_absent 1    13.555 6010.6 1137.3
## - traveltime  1    13.601 6010.7 1137.3
## - age         1    24.411 6021.5 1138.1
## - failures    1    27.503 6024.6 1138.3
## - higher      1    29.028 6026.1 1138.4
## - famsize     1    29.269 6026.3 1138.4
## - first_gen_college 1    30.243 6027.3 1138.4
## <none>                5997.1 1138.5
## - schoolsup    1    47.036 6044.1 1139.5
## - studytime    1    51.478 6048.6 1139.8
## - famsup      1    52.900 6050.0 1139.9
## - failed      1    56.066 6053.1 1140.1
## - freetime    1    58.687 6055.8 1140.3
## + Dalc        1     1.109 5996.0 1140.4
## + Walc        1     0.588 5996.5 1140.4
## + reason      3    60.996 5936.1 1140.4
## + guardian    2    11.494 5985.6 1141.7
## + Fjob        4    70.556 5926.5 1141.8
## - romantic    1    89.994 6087.1 1142.3
## - absences    1    94.651 6091.7 1142.6
## - sex         1   106.450 6103.5 1143.4
## - Mjob        4   217.084 6214.2 1144.5
## - goout       1   158.780 6155.9 1146.8
##
## Step:  AIC=1136.56
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Mjob + traveltime + studytime + failures + schoolsup + famsup +
##      paid + activities + nursery + higher + internet + romantic +

```

```

##      famrel + freetime + goout + health + absences + first_gen_college +
##      high_freq_absent + failed
##
##              Df Sum of Sq      RSS      AIC
## - nursery      1      2.229 6000.9 1134.7
## - Medu          1      3.471 6002.1 1134.8
## - internet      1      4.751 6003.4 1134.9
## - Pstatus       1      6.265 6004.9 1135.0
## - activities    1      6.509 6005.2 1135.0
## - famrel        1      6.807 6005.5 1135.0
## - health        1      7.276 6005.9 1135.0
## - paid          1      8.368 6007.0 1135.1
## - school        1     10.964 6009.6 1135.3
## - traveltime    1     12.621 6011.3 1135.4
## - high_freq_absent 1     12.860 6011.5 1135.4
## - address       1     13.704 6012.4 1135.5
## - age           1     24.490 6023.2 1136.2
## - failures      1     26.412 6025.1 1136.3
## - higher        1     28.308 6027.0 1136.4
## - famsize       1     29.765 6028.4 1136.5
## - first_gen_college 1     30.436 6029.1 1136.6
## <none>                      5998.7 1136.6
## - schoolsup     1     47.826 6046.5 1137.7
## - studytime     1     54.191 6052.9 1138.1
## - famsup        1     56.277 6054.9 1138.2
## - failed        1     57.385 6056.1 1138.3
## + Fedu          1      1.590 5997.1 1138.5
## + Dalc          1      1.055 5997.6 1138.5
## + Walc          1      0.713 5998.0 1138.5
## + reason        3     61.019 5937.6 1138.5
## - freetime      1     60.685 6059.4 1138.5
## + guardian      2     10.228 5988.4 1139.9
## - romantic      1     89.702 6088.4 1140.4
## + Fjob          4     61.291 5937.4 1140.5
## - absences      1     93.526 6092.2 1140.7
## - sex           1    106.050 6104.7 1141.5
## - Mjob          4    216.236 6214.9 1142.5
## - goout         1    161.916 6160.6 1145.1
##
## Step:  AIC=1134.71
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Mjob + traveltime + studytime + failures + schoolsup + famsup +
##      paid + activities + higher + internet + romantic + famrel +
##      freetime + goout + health + absences + first_gen_college +
##      high_freq_absent + failed
##
##              Df Sum of Sq      RSS      AIC
## - Medu          1      3.280 6004.2 1132.9
## - internet      1      5.167 6006.1 1133.0
## - Pstatus       1      5.949 6006.8 1133.1
## - activities    1      6.240 6007.1 1133.1
## - famrel        1      6.524 6007.4 1133.1
## - health        1      7.491 6008.4 1133.2
## - paid          1      7.883 6008.8 1133.2

```

```

## - school          1    11.399 6012.3 1133.5
## - traveltime      1    12.798 6013.7 1133.5
## - high_freq_absent 1    13.147 6014.0 1133.6
## - address         1    13.559 6014.5 1133.6
## - age             1    24.023 6024.9 1134.3
## - failures        1    26.018 6026.9 1134.4
## - famsize         1    28.442 6029.3 1134.6
## - higher          1    28.613 6029.5 1134.6
## - first_gen_college 1    29.504 6030.4 1134.6
## <none>              6000.9 1134.7
## - schoolsup       1    48.791 6049.7 1135.9
## - studytime       1    52.898 6053.8 1136.2
## - famsup          1    56.130 6057.0 1136.4
## - failed          1    57.614 6058.5 1136.5
## + nursery         1     2.229 5998.7 1136.6
## + Fedu            1     1.686 5999.2 1136.6
## + Dalc            1     1.451 5999.4 1136.6
## + Walc            1     0.466 6000.4 1136.7
## + reason          3    60.661 5940.2 1136.7
## - freetime        1    61.848 6062.7 1136.8
## + guardian        2    11.467 5989.4 1138.0
## - romantic        1    90.641 6091.5 1138.6
## + Fjob            4    61.040 5939.9 1138.7
## - absences        1    93.825 6094.7 1138.8
## - sex             1   105.719 6106.6 1139.6
## - Mjob            4   215.495 6216.4 1140.6
## - goout           1   163.450 6164.3 1143.3
##
## Step:  AIC=1132.92
## G3 ~ school + sex + age + address + famsize + Pstatus + Mjob +
##       traveltime + studytime + failures + schoolsup + famsup +
##       paid + activities + higher + internet + romantic + famrel +
##       freetime + goout + health + absences + first_gen_college +
##       high_freq_absent + failed
##
##              Df Sum of Sq   RSS   AIC
## - internet      1     4.974 6009.2 1131.2
## - activities     1     6.204 6010.4 1131.3
## - Pstatus        1     6.921 6011.1 1131.4
## - famrel         1     7.053 6011.2 1131.4
## - paid           1     7.813 6012.0 1131.4
## - health         1     8.650 6012.8 1131.5
## - school         1    11.485 6015.7 1131.7
## - high_freq_absent 1    12.445 6016.6 1131.7
## - traveltime     1    13.540 6017.7 1131.8
## - address        1    13.940 6018.1 1131.8
## - age            1    24.508 6028.7 1132.5
## - failures       1    26.253 6030.4 1132.7
## - famsize        1    26.629 6030.8 1132.7
## - higher         1    28.529 6032.7 1132.8
## <none>              6004.2 1132.9
## - schoolsup      1    49.290 6053.5 1134.2
## - famsup         1    54.249 6058.4 1134.5
## - studytime      1    55.376 6059.6 1134.5

```

```

## + Medu          1      3.280 6000.9 1134.7
## + nursery       1      2.038 6002.1 1134.8
## - failed        1     59.069 6063.2 1134.8
## + reason        3     62.168 5942.0 1134.8
## + Dalc          1      1.674 6002.5 1134.8
## + Fedu          1      0.838 6003.3 1134.9
## - freetime      1     60.465 6064.6 1134.9
## + Walc          1      0.486 6003.7 1134.9
## - first_gen_college 1     72.647 6076.8 1135.7
## + guardian      2     10.588 5993.6 1136.2
## - romantic      1     89.697 6093.9 1136.8
## + Fjob          4     60.982 5943.2 1136.9
## - absences      1     97.391 6101.6 1137.3
## - sex           1    108.378 6112.6 1138.0
## - Mjob          4    218.135 6222.3 1139.0
## - goout         1    161.253 6165.4 1141.4
##
## Step:  AIC=1131.25
## G3 ~ school + sex + age + address + famsize + Pstatus + Mjob +
##      traveltime + studytime + failures + schoolsup + famsup +
##      paid + activities + higher + romantic + famrel + freetime +
##      goout + health + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq   RSS   AIC
## - Pstatus      1      5.758 6014.9 1129.6
## - activities    1      6.067 6015.2 1129.7
## - famrel        1      7.719 6016.9 1129.8
## - paid          1      9.030 6018.2 1129.8
## - health        1     9.542 6018.7 1129.9
## - school        1    11.340 6020.5 1130.0
## - high_freq_absent 1    11.969 6021.1 1130.0
## - traveltime    1    13.664 6022.8 1130.2
## - address       1    16.917 6026.1 1130.4
## - failures      1    25.245 6034.4 1130.9
## - famsize       1    26.447 6035.6 1131.0
## - age           1    26.726 6035.9 1131.0
## - higher        1    27.484 6036.6 1131.0
## <none>                      6009.2 1131.2
## - schoolsup     1    49.694 6058.8 1132.5
## - famsup        1    53.344 6062.5 1132.7
## + internet      1      4.974 6004.2 1132.9
## - studytime     1    57.740 6066.9 1133.0
## + Medu          1      3.088 6006.1 1133.0
## + nursery       1      2.435 6006.7 1133.1
## + Dalc          1      1.768 6007.4 1133.1
## - freetime      1    59.458 6068.6 1133.1
## + reason        3    61.415 5947.7 1133.2
## + Fedu          1      0.735 6008.4 1133.2
## + Walc          1      0.461 6008.7 1133.2
## - failed        1    61.239 6070.4 1133.3
## - first_gen_college 1    73.519 6082.7 1134.0
## + guardian      2    10.612 5998.5 1134.5
## - romantic      1    86.466 6095.6 1134.9

```



```

## + Fjob          4    57.783 5951.4 1135.4
## - absences      1   100.838 6110.0 1135.8
## - sex           1   111.435 6120.6 1136.5
## - Mjob          4   221.815 6231.0 1137.6
## - goout         1   159.583 6168.7 1139.6
##
## Step:  AIC=1129.63
## G3 ~ school + sex + age + address + famsize + Mjob + traveltime +
##      studytime + failures + schoolsup + famsup + paid + activities +
##      higher + romantic + famrel + freetime + goout + health +
##      absences + first_gen_college + high_freq_absent + failed
##
##              Df Sum of Sq    RSS    AIC
## - famrel      1     7.528 6022.4 1128.1
## - activities  1     7.590 6022.5 1128.1
## - paid        1     8.219 6023.1 1128.2
## - health      1     9.145 6024.1 1128.2
## - school      1    10.994 6025.9 1128.3
## - high_freq_absent 1    12.063 6027.0 1128.4
## - traveltime  1    14.038 6028.9 1128.5
## - address     1    16.902 6031.8 1128.7
## - failures    1    24.055 6039.0 1129.2
## - age         1    27.173 6042.1 1129.4
## - higher      1    28.849 6043.8 1129.5
## <none>                6014.9 1129.6
## - famsize     1    31.277 6046.2 1129.7
## - schoolsup    1    48.815 6063.7 1130.8
## - famsup      1    53.992 6068.9 1131.2
## + Pstatus     1     5.758 6009.2 1131.2
## - freetime    1    56.915 6071.8 1131.3
## + Medu        1     3.970 6010.9 1131.4
## + internet    1     3.812 6011.1 1131.4
## - studytime   1    57.652 6072.6 1131.4
## + Dalc        1     2.052 6012.9 1131.5
## + nursery     1     2.034 6012.9 1131.5
## + reason      3    61.467 5953.4 1131.6
## + Fedu        1     0.587 6014.3 1131.6
## + Walc        1     0.251 6014.7 1131.6
## - failed      1    63.224 6078.1 1131.8
## - first_gen_college 1    79.217 6094.1 1132.8
## + guardian    2    11.389 6003.5 1132.9
## - romantic    1    85.462 6100.4 1133.2
## + Fjob        4    59.633 5955.3 1133.7
## - absences    1   106.295 6121.2 1134.5
## - sex         1   111.160 6126.1 1134.9
## - Mjob        4   225.315 6240.2 1136.2
## - goout       1   158.110 6173.0 1137.9
##
## Step:  AIC=1128.12
## G3 ~ school + sex + age + address + famsize + Mjob + traveltime +
##      studytime + failures + schoolsup + famsup + paid + activities +
##      higher + romantic + freetime + goout + health + absences +
##      first_gen_college + high_freq_absent + failed
##

```

```

##          Df Sum of Sq    RSS    AIC
## - activities      1      7.435 6029.9 1126.6
## - health          1      7.584 6030.0 1126.6
## - paid            1      9.149 6031.6 1126.7
## - school          1      9.664 6032.1 1126.8
## - traveltime      1     13.537 6036.0 1127.0
## - high_freq_absent 1     14.005 6036.4 1127.0
## - address         1     17.905 6040.3 1127.3
## - age             1     24.406 6046.8 1127.7
## - failures        1     25.641 6048.1 1127.8
## - higher          1     29.910 6052.3 1128.1
## <none>                        6022.4 1128.1
## - famsize         1     31.152 6053.6 1128.2
## - schoolsup        1     48.344 6070.8 1129.3
## - famsup          1     53.199 6075.6 1129.6
## + famrel          1      7.528 6014.9 1129.6
## + Pstatus         1      5.567 6016.9 1129.8
## + Medu            1      4.536 6017.9 1129.8
## + internet        1      4.405 6018.0 1129.8
## + reason          3     63.460 5959.0 1129.9
## + Dalc            1      1.879 6020.6 1130.0
## + nursery         1      1.761 6020.7 1130.0
## - studytime       1     60.428 6082.9 1130.1
## + Walc           1      0.714 6021.7 1130.1
## + Fedu           1      0.445 6022.0 1130.1
## - freetime        1     61.557 6084.0 1130.1
## - failed          1     61.850 6084.3 1130.2
## + guardian        2     12.164 6010.3 1131.3
## - first_gen_college 1     80.105 6102.5 1131.3
## - romantic        1     85.144 6107.6 1131.7
## + Fjob            4     57.663 5964.8 1132.3
## - absences        1    108.680 6131.1 1133.2
## - sex             1    113.609 6136.0 1133.5
## - Mjob            4    225.936 6248.4 1134.7
## - goout           1    156.890 6179.3 1136.3
##
## Step:  AIC=1126.61
## G3 ~ school + sex + age + address + famsize + Mjob + traveltime +
##      studytime + failures + schoolsup + famsup + paid + higher +
##      romantic + freetime + goout + health + absences + first_gen_college +
##      high_freq_absent + failed
##
##          Df Sum of Sq    RSS    AIC
## - health          1      7.995 6037.9 1125.1
## - paid            1      9.999 6039.9 1125.3
## - school          1     11.483 6041.4 1125.4
## - high_freq_absent 1     13.124 6043.0 1125.5
## - traveltime      1     13.727 6043.6 1125.5
## - address         1     20.017 6049.9 1125.9
## - age             1     23.452 6053.3 1126.1
## - failures        1     24.299 6054.2 1126.2
## - higher          1     27.659 6057.5 1126.4
## <none>                        6029.9 1126.6
## - famsize         1     31.219 6061.1 1126.7

```

```

## - schoolsup      1    50.051 6079.9 1127.9
## - famsup        1    51.843 6081.7 1128.0
## + activities    1     7.435 6022.4 1128.1
## + famrel        1     7.372 6022.5 1128.1
## + Pstatus       1     7.052 6022.8 1128.2
## - studytime     1    56.694 6086.6 1128.3
## + Medu          1     4.622 6025.3 1128.3
## + internet      1     4.114 6025.8 1128.3
## + Dalc          1     2.101 6027.8 1128.5
## + nursery       1     1.458 6028.4 1128.5
## + Walc          1     0.583 6029.3 1128.6
## + Fedu          1     0.555 6029.3 1128.6
## - freetime      1    61.065 6090.9 1128.6
## + reason        3    58.641 5971.2 1128.8
## - failed        1    63.924 6093.8 1128.8
## - first_gen_college 1    78.223 6108.1 1129.7
## + guardian      2    12.095 6017.8 1129.8
## - romantic      1    88.630 6118.5 1130.4
## + Fjob          4    60.550 5969.3 1130.6
## - absences      1   107.390 6137.3 1131.6
## - sex           1   108.674 6138.5 1131.7
## - Mjob          4   225.207 6255.1 1133.1
## - goout         1   161.707 6191.6 1135.1
##
## Step:  AIC=1125.13
## G3 ~ school + sex + age + address + famsize + Mjob + traveltime +
##      studytime + failures + schoolsup + famsup + paid + higher +
##      romantic + freetime + goout + absences + first_gen_college +
##      high_freq_absent + failed
##
##              Df Sum of Sq    RSS    AIC
## - paid        1    10.716 6048.6 1123.8
## - high_freq_absent 1    11.598 6049.5 1123.9
## - school      1    12.514 6050.4 1124.0
## - traveltime  1    13.912 6051.8 1124.0
## - address     1    21.600 6059.5 1124.5
## - age         1    22.034 6059.9 1124.6
## - failures    1    23.340 6061.2 1124.7
## - higher      1    27.393 6065.3 1124.9
## <none>                6037.9 1125.1
## - famsize     1    31.772 6069.6 1125.2
## - schoolsup    1    48.738 6086.6 1126.3
## + health      1     7.995 6029.9 1126.6
## + activities  1     7.846 6030.0 1126.6
## - famsup      1    53.732 6091.6 1126.6
## + Pstatus     1     6.708 6031.2 1126.7
## + Medu        1     5.799 6032.1 1126.8
## + famrel      1     5.787 6032.1 1126.8
## + internet    1     4.816 6033.1 1126.8
## - freetime    1    58.585 6096.5 1127.0
## - studytime   1    59.536 6097.4 1127.0
## + nursery     1     1.671 6036.2 1127.0
## + Dalc        1     1.623 6036.2 1127.0
## + reason      3    61.887 5976.0 1127.1

```

```

## + Walc          1      1.035 6036.8 1127.1
## + Fedu          1      0.606 6037.3 1127.1
## - failed        1     66.411 6104.3 1127.5
## + guardian      2     14.111 6023.8 1128.2
## - first_gen_college 1     78.383 6116.3 1128.2
## - romantic      1     90.039 6127.9 1129.0
## + Fjob          4     58.303 5979.6 1129.3
## - sex           1    105.710 6143.6 1130.0
## - absences      1    105.891 6143.8 1130.0
## - Mjob          4    220.425 6258.3 1131.3
## - goout         1    165.403 6203.3 1133.8
##
## Step:  AIC=1123.83
## G3 ~ school + sex + age + address + famsize + Mjob + traveltime +
##      studytime + failures + schoolsup + famsup + higher + romantic +
##      freetime + goout + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq   RSS   AIC
## - high_freq_absent  1    11.893 6060.5 1122.6
## - school            1    13.648 6062.2 1122.7
## - traveltime        1    14.707 6063.3 1122.8
## - age               1    20.993 6069.6 1123.2
## - address           1    22.351 6070.9 1123.3
## - failures          1    23.406 6072.0 1123.4
## <none>                6048.6 1123.8
## - higher            1    31.507 6080.1 1123.9
## - famsize           1    32.108 6080.7 1123.9
## - famsup            1    45.157 6093.7 1124.8
## - schoolsup         1    50.481 6099.1 1125.1
## + paid              1    10.716 6037.9 1125.1
## + activities        1     8.770 6039.8 1125.3
## + health            1     8.713 6039.9 1125.3
## + famrel            1     6.604 6042.0 1125.4
## + reason            3    67.181 5981.4 1125.4
## + internet          1     6.288 6042.3 1125.4
## + Pstatus           1     5.780 6042.8 1125.5
## + Medu              1     5.668 6042.9 1125.5
## - freetime          1    57.147 6105.7 1125.5
## + Dalc              1     3.067 6045.5 1125.6
## + nursery           1     1.236 6047.3 1125.8
## + Fedu              1     0.985 6047.6 1125.8
## + Walc              1     0.555 6048.0 1125.8
## - studytime         1    64.142 6112.7 1126.0
## - failed            1    70.758 6119.3 1126.4
## - first_gen_college 1    77.297 6125.9 1126.8
## + guardian          2    14.405 6034.2 1126.9
## - romantic          1    89.248 6137.8 1127.6
## + Fjob              4    52.599 5996.0 1128.4
## - sex               1   102.339 6150.9 1128.5
## - absences          1   107.768 6156.4 1128.8
## - Mjob              4   219.664 6268.2 1129.9
## - goout             1   160.633 6209.2 1132.2
##

```

```

## Step: AIC=1122.61
## G3 ~ school + sex + age + address + famsize + Mjob + traveltime +
##      studytime + failures + schoolsup + famsup + higher + romantic +
##      freetime + goout + absences + first_gen_college + failed
##
##      Df Sum of Sq    RSS    AIC
## - traveltime      1    13.259 6073.7 1121.5
## - school           1    14.199 6074.7 1121.5
## - address          1    19.683 6080.2 1121.9
## - failures         1    23.090 6083.6 1122.1
## - age              1    24.292 6084.8 1122.2
## - higher           1    27.693 6088.2 1122.4
## - famsize          1    29.953 6090.4 1122.6
## <none>                                6060.5 1122.6
## - famsup           1    46.840 6107.3 1123.7
## - schoolsup         1    48.944 6109.4 1123.8
## + high_freq_absent  1    11.893 6048.6 1123.8
## + paid             1    11.012 6049.5 1123.9
## + famrel           1     8.447 6052.0 1124.1
## + activities       1     7.823 6052.7 1124.1
## - freetime         1    54.175 6114.7 1124.1
## + health           1     7.106 6053.4 1124.2
## + internet         1     5.790 6054.7 1124.2
## + Pstatus          1     5.784 6054.7 1124.2
## + Medu             1     4.718 6055.8 1124.3
## + reason           3    63.156 5997.3 1124.5
## + nursery          1     1.403 6059.1 1124.5
## + Dalc             1     1.390 6059.1 1124.5
## + Walc             1     0.944 6059.5 1124.5
## + Fedu             1     0.522 6060.0 1124.6
## - studytime        1    67.327 6127.8 1125.0
## - first_gen_college 1    73.000 6133.5 1125.3
## - failed           1    74.711 6135.2 1125.5
## + guardian         2    14.839 6045.6 1125.6
## - romantic         1    90.329 6150.8 1126.5
## + Fjob             4    54.936 6005.5 1127.0
## - sex              1   101.693 6162.2 1127.2
## - Mjob             4   214.393 6274.9 1128.3
## - absences         1   128.452 6188.9 1128.9
## - goout            1   162.672 6223.2 1131.1
##
## Step: AIC=1121.47
## G3 ~ school + sex + age + address + famsize + Mjob + studytime +
##      failures + schoolsup + famsup + higher + romantic + freetime +
##      goout + absences + first_gen_college + failed
##
##      Df Sum of Sq    RSS    AIC
## - school           1    10.285 6084.0 1120.1
## - age              1    22.293 6096.0 1120.9
## - failures         1    25.256 6099.0 1121.1
## - famsize          1    26.818 6100.6 1121.2
## - higher           1    28.844 6102.6 1121.3
## <none>                                6073.7 1121.5
## - address          1    31.581 6105.3 1121.5

```

```

## + traveltime      1    13.259 6060.5 1122.6
## - schoolsup       1    49.034 6122.8 1122.7
## + paid            1    11.759 6062.0 1122.7
## - famsup          1    50.943 6124.7 1122.8
## + high_freq_absent 1    10.446 6063.3 1122.8
## + activities      1     8.106 6065.6 1123.0
## + famrel          1     7.816 6065.9 1123.0
## + health          1     7.390 6066.3 1123.0
## - freetime        1    55.528 6129.3 1123.1
## + Pstatus         1     6.136 6067.6 1123.1
## + internet        1     5.946 6067.8 1123.1
## + Medu            1     5.657 6068.1 1123.1
## + reason          3    65.406 6008.3 1123.2
## + Walc            1     1.643 6072.1 1123.4
## + nursery         1     1.504 6072.2 1123.4
## + Dalc            1     0.543 6073.2 1123.4
## + Fedu            1     0.055 6073.7 1123.5
## - studytime       1    71.853 6145.6 1124.1
## - failed          1    71.972 6145.7 1124.1
## - first_gen_college 1    74.364 6148.1 1124.3
## + guardian        2    12.487 6061.3 1124.7
## - romantic        1    92.108 6165.8 1125.4
## - sex             1    98.165 6171.9 1125.8
## + Fjob            4    55.338 6018.4 1125.9
## - Mjob            4   221.092 6294.8 1127.6
## - absences        1   128.487 6202.2 1127.7
## - goout           1   169.495 6243.2 1130.3
##
## Step:  AIC=1120.14
## G3 ~ sex + age + address + famsize + Mjob + studytime + failures +
##       schoolsup + famsup + higher + romantic + freetime + goout +
##       absences + first_gen_college + failed
##
##              Df Sum of Sq    RSS    AIC
## - age          1    14.853 6098.9 1119.1
## - address       1    24.577 6108.6 1119.7
## - failures      1    26.417 6110.4 1119.8
## - famsize       1    29.156 6113.2 1120.0
## <none>                      6084.0 1120.1
## - higher        1    31.483 6115.5 1120.2
## + paid          1    12.688 6071.3 1121.3
## - schoolsup     1    50.582 6134.6 1121.4
## + high_freq_absent 1    11.091 6072.9 1121.4
## + school        1    10.285 6073.7 1121.5
## + activities    1     9.919 6074.1 1121.5
## + traveltime    1     9.345 6074.7 1121.5
## + health        1     8.255 6075.8 1121.6
## + famrel        1     6.463 6077.6 1121.7
## + Pstatus       1     5.867 6078.2 1121.8
## + internet      1     5.813 6078.2 1121.8
## - famsup        1    56.178 6140.2 1121.8
## + Medu          1     5.582 6078.4 1121.8
## - freetime      1    57.740 6141.8 1121.9
## + reason        3    65.002 6019.0 1121.9

```

```

## + Walc          1      1.832 6082.2 1122.0
## + nursery       1      1.823 6082.2 1122.0
## + Dalc          1      0.903 6083.1 1122.1
## + Fedu          1      0.023 6084.0 1122.1
## - studytime     1     65.829 6149.9 1122.4
## - failed        1     71.247 6155.3 1122.7
## - first_gen_college 1     73.351 6157.4 1122.9
## + guardian      2     11.176 6072.8 1123.4
## - romantic      1     89.829 6173.9 1123.9
## - sex           1     93.837 6177.9 1124.2
## + Fjob          4     56.297 6027.7 1124.5
## - absences      1    119.451 6203.5 1125.8
## - Mjob          4    218.306 6302.3 1126.1
## - goout         1    173.743 6257.8 1129.3
##
## Step:  AIC=1119.1
## G3 ~ sex + address + famsize + Mjob + studytime + failures +
##      schoolsup + famsup + higher + romantic + freetime + goout +
##      absences + first_gen_college + failed
##
##              Df Sum of Sq   RSS   AIC
## - failures      1    25.369 6124.2 1118.7
## - famsize       1    28.020 6126.9 1118.9
## - address       1    29.428 6128.3 1119.0
## <none>                      6098.9 1119.1
## - higher        1    37.542 6136.4 1119.5
## - schoolsup      1    39.918 6138.8 1119.7
## + age           1    14.853 6084.0 1120.1
## + high_freq_absent 1    13.491 6085.4 1120.2
## + paid          1    11.222 6087.7 1120.4
## - famsup        1    51.416 6150.3 1120.4
## + traveltime    1     9.314 6089.6 1120.5
## + activities     1     7.916 6091.0 1120.6
## + internet      1     7.349 6091.5 1120.6
## + health        1     6.386 6092.5 1120.7
## + Pstatus       1     6.330 6092.5 1120.7
## + Medu          1     5.734 6093.1 1120.7
## - freetime      1    56.506 6155.4 1120.8
## + famrel        1     5.089 6093.8 1120.8
## + reason        3    65.484 6033.4 1120.8
## + school        1     2.845 6096.0 1120.9
## + nursery       1     1.437 6097.4 1121.0
## + Walc          1     1.141 6097.7 1121.0
## + Dalc          1     0.330 6098.5 1121.1
## + Fedu          1     0.032 6098.8 1121.1
## - studytime     1    62.167 6161.0 1121.1
## - first_gen_college 1    80.992 6179.9 1122.3
## - failed        1    81.064 6179.9 1122.3
## + guardian      2     3.431 6095.4 1122.9
## + Fjob          4    56.965 6041.9 1123.4
## - romantic      1    98.243 6197.1 1123.4
## - sex           1   100.196 6199.1 1123.5
## - absences      1   110.458 6209.3 1124.2
## - Mjob          4   224.920 6323.8 1125.4

```

```

## - goout          1    186.393 6285.3 1129.0
##
## Step:  AIC=1118.74
## G3 ~ sex + address + famsize + Mjob + studytime + schoolsup +
##      famsup + higher + romantic + freetime + goout + absences +
##      first_gen_college + failed
##
##              Df Sum of Sq    RSS    AIC
## - address      1      28.92 6153.2 1118.6
## - famsize       1      29.58 6153.8 1118.7
## <none>                      6124.2 1118.7
## + failures     1      25.37 6098.9 1119.1
## - schoolsup     1      39.39 6163.6 1119.3
## + age          1      13.81 6110.4 1119.8
## + high_freq_absent 1      12.96 6111.3 1119.9
## - higher        1      50.12 6174.4 1120.0
## + paid          1      11.45 6112.8 1120.0
## + traveltime    1      10.94 6113.3 1120.0
## + reason        3      71.38 6052.9 1120.1
## - famsup        1      54.90 6179.1 1120.3
## + activities    1       6.67 6117.6 1120.3
## + famrel        1       6.41 6117.8 1120.3
## + internet      1       6.29 6118.0 1120.3
## + Medu          1       5.99 6118.3 1120.4
## + health        1       5.64 6118.6 1120.4
## + Pstatus       1       4.91 6119.3 1120.4
## - freetime      1      58.15 6182.4 1120.5
## + school        1       3.60 6120.6 1120.5
## + Walc          1       1.63 6122.6 1120.6
## + nursery       1       1.15 6123.1 1120.7
## + Dalc          1       0.28 6124.0 1120.7
## + Fedu          1       0.16 6124.1 1120.7
## - studytime     1      68.48 6192.7 1121.1
## - first_gen_college 1      83.32 6207.6 1122.1
## + guardian      2       3.27 6121.0 1122.5
## + Fjob          4      60.07 6064.2 1122.8
## - romantic      1      98.67 6222.9 1123.1
## - sex           1     103.72 6228.0 1123.4
## - absences      1     127.06 6251.3 1124.9
## - Mjob          4     223.16 6347.4 1124.9
## - goout         1     194.44 6318.7 1129.1
## - failed        1     640.58 6764.8 1156.0
##
## Step:  AIC=1118.61
## G3 ~ sex + famsize + Mjob + studytime + schoolsup + famsup +
##      higher + romantic + freetime + goout + absences + first_gen_college +
##      failed
##
##              Df Sum of Sq    RSS    AIC
## <none>                      6153.2 1118.6
## + address      1      28.92 6124.2 1118.7
## - famsize       1      33.82 6187.0 1118.8
## - schoolsup     1      37.38 6190.5 1119.0
## + failures     1      24.86 6128.3 1119.0

```



```
## + traveltime      1      23.30 6129.9 1119.1
## + age             1      18.47 6134.7 1119.4
## - higher          1      49.47 6202.6 1119.8
## + paid            1      12.04 6141.1 1119.8
## + internet        1      11.88 6141.3 1119.8
## + high_freq_absent 1       8.90 6144.3 1120.0
## + activities       1       8.59 6144.6 1120.0
## - famsup          1     54.87 6208.0 1120.1
## + Medu            1       7.64 6145.5 1120.1
## + famrel          1       7.32 6145.8 1120.1
## + health          1       7.07 6146.1 1120.2
## + Pstatus         1       5.39 6147.8 1120.3
## - freetime        1     57.29 6210.5 1120.3
## + Walc            1       2.41 6150.8 1120.5
## + reason           3     63.85 6089.3 1120.5
## + nursery         1       0.88 6152.3 1120.5
## + Dalc            1       0.33 6152.8 1120.6
## + Fedu            1       0.20 6153.0 1120.6
## + school          1       0.10 6153.1 1120.6
## - studytime       1     64.86 6218.0 1120.8
## - first_gen_college 1     84.51 6237.7 1122.0
## + guardian        2       5.72 6147.4 1122.2
## + Fjob            4     62.03 6091.1 1122.6
## - romantic        1     96.69 6249.9 1122.8
## - sex             1     97.75 6250.9 1122.8
## - absences        1    123.11 6276.3 1124.4
## - Mjob            4    242.27 6395.4 1125.9
## - goout           1    189.15 6342.3 1128.6
## - failed          1    670.10 6823.3 1157.4
```

```
summary(step.model)
```

```
##
## Call:
## lm(formula = G3 ~ sex + famsize + Mjob + studytime + schoolsup +
##     famsup + higher + romantic + freetime + goout + absences +
##     first_gen_college + failed, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.2041  -2.0732   0.3544   2.6786   9.6991
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    8.11450    1.39218   5.829 1.20e-08 ***
## sexM           1.12016    0.45712   2.450 0.014719 *
## famsizeLE3     0.65890    0.45713   1.441 0.150309
## Mjobhealth     1.68430    0.93895   1.794 0.073641 .
## Mjobother      0.05209    0.63950   0.081 0.935123
## Mjobservices   1.36030    0.68184   1.995 0.046757 *
## Mjobteacher    -0.71854    0.87864  -0.818 0.413992
## studytime      0.52329    0.26215   1.996 0.046634 *
## schoolsupyes   -0.94383    0.62288  -1.515 0.130542
## famsupyes      -0.80258    0.43716  -1.836 0.067160 .
## higheryes      1.73643    0.99609   1.743 0.082102 .
```

```
## romanticyes          -1.09643    0.44989   -2.437 0.015265 *
## freetimelow          -0.84135    0.44848   -1.876 0.061429 .
## gooutlow             1.53330    0.44981    3.409 0.000723 ***
## absences              0.07208    0.02621    2.750 0.006245 **
## first_gen_collegeyes -1.19608    0.52493   -2.279 0.023250 *
## failedyes            -3.51954    0.54855   -6.416 4.19e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.035 on 378 degrees of freedom
## Multiple R-squared:  0.256, Adjusted R-squared:  0.2245
## F-statistic: 8.127 on 16 and 378 DF,  p-value: < 2.2e-16
```

From fitting the base variables in a linear model, we can see that the variables sex, schoolsup, famsup, romantic, freetime, goout, absences, and failed are active.

Based on the backwards stepwise regression model, it seems that the variables sex, Mjob, studytime, higher, romantic, freetime, goout, absences, first\_gen\_college, famsup, failed.

Based on these active variables, some interactions that we think could be significant are: schoolsup*failed*, famsupfirst\_gen\_college, higher\*first\_gen\_college.

```
activelm <- lm(G3 ~ (sex + schoolsup + romantic + freetime + goout + absences + failed + Mjob + studytime + higher + first_gen_college + famsup)^2)
summary(activelm)
```

```
##
## Call:
## lm(formula = G3 ~ (sex + schoolsup + romantic + freetime + goout +
##      absences + failed + Mjob + studytime + higher + first_gen_college +
##      famsup)^2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.6429  -1.7534   0.2846   2.2075   7.3177
##
## Coefficients: (2 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1.506826   13.118897  -0.115 0.908638
## sexM             0.892073    5.622189   0.159 0.874042
## schoolsupyes    13.761155    9.218122   1.493 0.136598
## romanticyes    -1.449784    4.692335  -0.309 0.757574
## freetimelow     2.630885    4.451459   0.591 0.554983
## gooutlow       -1.031107   10.156300  -0.102 0.919207
## absences        0.148759    0.316984   0.469 0.639221
## failedyes      -8.234032    4.894235  -1.682 0.093599 .
## Mjobhealth     -1.259093    5.227716  -0.241 0.809847
## Mjobother      -5.143290    9.033916  -0.569 0.569585
## Mjobservices    1.528706   11.532458   0.133 0.894638
## Mjobteacher    -9.988699    6.801967  -1.469 0.143083
## studytime       5.357552    4.658888   1.150 0.251133
## higheryes      15.703850   12.568400   1.249 0.212529
## first_gen_collegeyes -2.042557    2.769453  -0.738 0.461413
## famsupyes      -1.124951    5.956632  -0.189 0.850341
## sexM:schoolsupyes -3.306582    1.597411  -2.070 0.039366 *
## sexM:romanticyes  1.516326    1.076859   1.408 0.160203
```

## sexM:freetimelow	0.440768	1.021810	0.431	0.666536	
## sexM:gooutlow	0.140986	1.100729	0.128	0.898174	
## sexM:absences	-0.168685	0.101499	-1.662	0.097636	.
## sexM:failedyes	-0.812131	1.739573	-0.467	0.640963	
## sexM:Mjobhealth	4.072742	2.744648	1.484	0.138956	
## sexM:Mjobother	2.715976	1.700434	1.597	0.111336	
## sexM:Mjobservices	3.912599	1.832894	2.135	0.033651	*
## sexM:Mjobteacher	4.222718	2.177167	1.940	0.053432	.
## sexM:studytime	-0.342144	0.612314	-0.559	0.576761	
## sexM:higheryes	-4.325565	5.196762	-0.832	0.405911	
## sexM:first_gen_collegeyes	3.448751	1.258091	2.741	0.006511	**
## sexM:famsupyes	1.181025	1.031439	1.145	0.253169	
## schoolsupyes:romanticyes	1.976315	1.776126	1.113	0.266780	
## schoolsupyes:freetimelow	1.219474	1.826300	0.668	0.504853	
## schoolsupyes:gooutlow	-0.008335	1.789930	-0.005	0.996288	
## schoolsupyes:absences	-0.181005	0.097024	-1.866	0.063140	.
## schoolsupyes:failedyes	2.056166	1.959551	1.049	0.294936	
## schoolsupyes:Mjobhealth	-2.452853	4.078269	-0.601	0.548027	
## schoolsupyes:Mjobother	-0.957822	2.166765	-0.442	0.658791	
## schoolsupyes:Mjobservices	0.062923	2.350522	0.027	0.978662	
## schoolsupyes:Mjobteacher	3.535611	3.934416	0.899	0.369613	
## schoolsupyes:studytime	-2.615114	0.957261	-2.732	0.006695	**
## schoolsupyes:higheryes	-11.448554	8.737679	-1.310	0.191177	
## schoolsupyes:first_gen_collegeyes	5.330960	1.715583	3.107	0.002080	**
## schoolsupyes:famsupyes	-1.051195	1.608959	-0.653	0.514070	
## romanticyes:freetimelow	-0.104303	1.110758	-0.094	0.925253	
## romanticyes:gooutlow	-0.993412	1.145600	-0.867	0.386595	
## romanticyes:absences	-0.067529	0.077543	-0.871	0.384569	
## romanticyes:failedyes	1.610302	1.391467	1.157	0.248142	
## romanticyes:Mjobhealth	-0.122641	2.593575	-0.047	0.962318	
## romanticyes:Mjobother	-0.060667	1.643023	-0.037	0.970572	
## romanticyes:Mjobservices	-1.237220	1.700337	-0.728	0.467443	
## romanticyes:Mjobteacher	0.882332	2.215259	0.398	0.690712	
## romanticyes:studytime	-0.687214	0.691971	-0.993	0.321500	
## romanticyes:higheryes	2.788889	3.990571	0.699	0.485209	
## romanticyes:first_gen_collegeyes	0.458349	1.253016	0.366	0.714791	
## romanticyes:famsupyes	-0.577383	1.077682	-0.536	0.592544	
## freetimelow:gooutlow	-0.941444	1.057541	-0.890	0.374107	
## freetimelow:absences	0.111240	0.091142	1.221	0.223287	
## freetimelow:failedyes	-2.274094	1.541957	-1.475	0.141379	
## freetimelow:Mjobhealth	-1.322847	2.588477	-0.511	0.609714	
## freetimelow:Mjobother	-1.531530	1.864238	-0.822	0.412038	
## freetimelow:Mjobservices	-0.462435	1.972455	-0.234	0.814808	
## freetimelow:Mjobteacher	-1.231910	2.244147	-0.549	0.583479	
## freetimelow:studytime	0.604095	0.669282	0.903	0.367508	
## freetimelow:higheryes	-2.919575	3.673405	-0.795	0.427406	
## freetimelow:first_gen_collegeyes	0.854057	1.330632	0.642	0.521496	
## freetimelow:famsupyes	-1.330642	1.077589	-1.235	0.217921	
## gooutlow:absences	-0.049939	0.089046	-0.561	0.575366	
## gooutlow:failedyes	-0.174341	1.370430	-0.127	0.898860	
## gooutlow:Mjobhealth	1.457008	2.348034	0.621	0.535415	
## gooutlow:Mjobother	-0.418606	1.709665	-0.245	0.806753	
## gooutlow:Mjobservices	0.912677	1.784840	0.511	0.609506	
## gooutlow:Mjobteacher	2.080526	2.307532	0.902	0.368026	

```

## gooutlow:studytime      1.127611    0.699657    1.612 0.108154
## gooutlow:higheryes     -0.372584    9.920550   -0.038 0.970068
## gooutlow:first_gen_collegeyes  1.024237    1.296276    0.790 0.430111
## gooutlow:famsupyes     -0.093945    1.093377   -0.086 0.931590
## absences:failedyes      0.289277    0.073870    3.916 0.000113 ***
## absences:Mjobhealth     -0.044829    0.244754   -0.183 0.854806
## absences:Mjobother      -0.139807    0.139956   -0.999 0.318683
## absences:Mjobservices   -0.080993    0.153141   -0.529 0.597307
## absences:Mjobteacher    -0.105356    0.196984   -0.535 0.593178
## absences:studytime      0.017934    0.057633    0.311 0.755901
## absences:higheryes     -0.102366    0.248758   -0.412 0.681012
## absences:first_gen_collegeyes  0.066378    0.098061    0.677 0.499021
## absences:famsupyes      0.068059    0.073031    0.932 0.352175
## failedyes:Mjobhealth     3.888284    2.901794    1.340 0.181337
## failedyes:Mjobother     -1.029702    2.006996   -0.513 0.608314
## failedyes:Mjobservices   1.625498    2.056012    0.791 0.429838
## failedyes:Mjobteacher    -0.667717    4.128115   -0.162 0.871620
## failedyes:studytime      0.274374    0.971061    0.283 0.777729
## failedyes:higheryes     -1.349987    4.152803   -0.325 0.745363
## failedyes:first_gen_collegeyes  4.246604    1.838110    2.310 0.021593 *
## failedyes:famsupyes      0.068774    1.322558    0.052 0.958565
## Mjobhealth:studytime     2.929895    1.371083    2.137 0.033465 *
## Mjobother:studytime      0.867948    0.899516    0.965 0.335419
## Mjobservices:studytime    0.977495    0.967555    1.010 0.313230
## Mjobteacher:studytime     0.783722    1.376670    0.569 0.569614
## Mjobhealth:higheryes      NA          NA          NA          NA
## Mjobother:higheryes      8.361265    8.588696    0.974 0.331129
## Mjobservices:higheryes    1.376403   11.159423    0.123 0.901926
## Mjobteacher:higheryes    10.776012    5.658085    1.905 0.057858 .
## Mjobhealth:first_gen_collegeyes -5.333834    2.483592   -2.148 0.032595 *
## Mjobother:first_gen_collegeyes -3.997660    1.895764   -2.109 0.035849 *
## Mjobservices:first_gen_collegeyes -5.602517    1.941616   -2.885 0.004210 **
## Mjobteacher:first_gen_collegeyes -1.139043    3.603667   -0.316 0.752176
## Mjobhealth:famsupyes     -0.971463    2.603701   -0.373 0.709348
## Mjobother:famsupyes      0.228631    1.502886    0.152 0.879195
## Mjobservices:famsupyes   -0.129235    1.643959   -0.079 0.937397
## Mjobteacher:famsupyes    -1.875211    2.166962   -0.865 0.387575
## studytime:higheryes     -7.929807    4.479400   -1.770 0.077760 .
## studytime:first_gen_collegeyes  0.748786    0.744561    1.006 0.315434
## studytime:famsupyes      1.418357    0.614634    2.308 0.021743 *
## higheryes:first_gen_collegeyes      NA          NA          NA          NA
## higheryes:famsupyes     -1.524942    5.664120   -0.269 0.787951
## first_gen_collegeyes:famsupyes -1.101570    1.255553   -0.877 0.381039
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.84 on 282 degrees of freedom
## Multiple R-squared:  0.4973, Adjusted R-squared:  0.2976
## F-statistic: 2.491 on 112 and 282 DF,  p-value: 5.445e-10

```

From this, we can see that there seem to be significant interaction effects between sex and schoolsup, sex and first\_gen\_college, sex and Mjob, schoolsup and absences, schoolsup and studytime, schoolsup and first\_gen\_college, absences and failed, failed and first\_gen\_college, Mjob and studytime, Mjob and first\_gen\_college, studytime and famsup.

## Linear model on the training data set

```
base_lm_tr <- lm(G3 ~ . -G2 -G1 -ord_g3 -stable_learning_env, data = training)
step.model_tr <- stepAIC(base_lm_tr, direction="both")
```

```
## Start: AIC=927.72
## G3 ~ (school + sex + age + address + famsize + Pstatus + Medu +
##       Fedu + Mjob + Fjob + reason + guardian + traveltime + studytime +
##       failures + schoolsup + famsup + paid + activities + nursery +
##       higher + internet + romantic + famrel + freetime + goout +
##       Dalc + Walc + health + absences + G1 + G2 + first_gen_college +
##       stable_learning_env + high_freq_absent + failed + ord_g3) -
##       G2 - G1 - ord_g3 - stable_learning_env
##
##           Df Sum of Sq    RSS    AIC
## - Fjob      4    25.052 4559.3 921.46
## - guardian  2     3.373 4537.6 923.96
## - reason    3    50.313 4584.6 925.21
## - higher    1     0.036 4534.3 925.72
## - Walc      1     0.204 4534.5 925.73
## - Dalc      1     0.358 4534.6 925.75
## - famrel    1     0.371 4534.6 925.75
## - Medu      1     0.921 4535.2 925.78
## - Fedu      1     5.124 4539.4 926.08
## - nursery   1     5.172 4539.4 926.08
## - high_freq_absent 1     7.005 4541.3 926.21
## - activities 1     7.286 4541.5 926.23
## - address   1    11.194 4545.4 926.50
## - internet  1    12.933 4547.2 926.62
## - traveltime 1    17.745 4552.0 926.95
## - failures  1    18.130 4552.4 926.98
## - Pstatus   1    19.946 4554.2 927.11
## - health    1    26.319 4560.6 927.55
## <none>                4534.3 927.72
## - school    1    30.003 4564.3 927.80
## - paid      1    34.856 4569.1 928.14
## - first_gen_college 1    37.337 4571.6 928.31
## - studytime 1    55.699 4589.9 929.58
## - failed    1    56.947 4591.2 929.66
## - Mjob      4   145.059 4679.3 929.67
## - schoolsup  1    67.383 4601.6 930.38
## - famsize   1    68.749 4603.0 930.48
## - age       1    69.869 4604.1 930.55
## - absences  1    81.239 4615.5 931.33
## - famsup    1    81.674 4615.9 931.36
## - sex       1    89.555 4623.8 931.90
## - romantic  1   110.934 4645.2 933.36
## - goout     1   127.344 4661.6 934.47
## - freetime  1   135.158 4669.4 935.00
##
## Step: AIC=921.46
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##       Fedu + Mjob + reason + guardian + traveltime + studytime +
##       failures + schoolsup + famsup + paid + activities + nursery +
```

```
## higher + internet + romantic + famrel + freetime + goout +
## Dalc + Walc + health + absences + first_gen_college + high_freq_absent +
## failed
```

```
##
##          Df Sum of Sq    RSS    AIC
## - guardian      2      2.456 4561.8 917.63
## - reason         3     54.677 4614.0 919.23
## - higher         1      0.001 4559.3 919.46
## - Dalc           1      0.003 4559.3 919.46
## - famrel         1      0.064 4559.4 919.47
## - Medu           1      1.476 4560.8 919.56
## - Fedu           1      2.289 4561.6 919.62
## - Walc           1      2.340 4561.6 919.62
## - nursery        1      5.992 4565.3 919.88
## - activities     1      7.690 4567.0 919.99
## - high_freq_absent 1      9.336 4568.6 920.11
## - internet       1     11.909 4571.2 920.29
## - address        1     12.789 4572.1 920.35
## - traveltime     1     14.191 4573.5 920.44
## - failures       1     17.216 4576.5 920.65
## - Pstatus        1     22.362 4581.7 921.01
## - paid           1     26.525 4585.8 921.29
## - health         1     27.034 4586.3 921.33
## <none>                                4559.3 921.46
## - school         1     30.178 4589.5 921.55
## - first_gen_college 1     41.045 4600.3 922.29
## - Mjob           4    142.087 4701.4 923.16
## - studytime      1     55.029 4614.3 923.25
## - failed         1     56.172 4615.5 923.33
## - famsize        1     62.299 4621.6 923.75
## - schoolsup       1     65.516 4624.8 923.97
## - age            1     67.063 4626.4 924.08
## - famsup         1     85.588 4644.9 925.34
## - absences       1     88.683 4648.0 925.55
## - sex            1     92.448 4651.8 925.81
## - romantic       1    109.267 4668.6 926.95
## + Fjob           4     25.052 4534.3 927.72
## - goout          1    122.960 4682.3 927.87
## - freetime       1    140.161 4699.5 929.03
```

```
## Step: AIC=917.63
```

```
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
## Fedu + Mjob + reason + traveltime + studytime + failures +
## schoolsup + famsup + paid + activities + nursery + higher +
## internet + romantic + famrel + freetime + goout + Dalc +
## Walc + health + absences + first_gen_college + high_freq_absent +
## failed
```

```
##
##          Df Sum of Sq    RSS    AIC
## - reason      3     56.461 4618.2 915.52
## - Dalc         1      0.005 4561.8 915.63
## - higher       1      0.035 4561.8 915.63
## - famrel       1      0.062 4561.8 915.64
## - Medu         1      1.075 4562.8 915.71
```

```

## - Fedu          1      1.599 4563.4 915.74
## - Walc          1      2.545 4564.3 915.81
## - nursery       1      7.566 4569.3 916.16
## - activities    1      7.927 4569.7 916.18
## - high_freq_absent 1      9.108 4570.9 916.26
## - internet      1     11.546 4573.3 916.43
## - traveltime    1     12.455 4574.2 916.49
## - address       1     15.272 4577.0 916.69
## - failures      1     17.008 4578.8 916.81
## - Pstatus       1     23.387 4585.1 917.25
## - paid          1     27.112 4588.9 917.50
## <none>          1     4561.8 917.63
## - health        1     29.203 4591.0 917.65
## - school        1     29.267 4591.0 917.65
## - first_gen_college 1     40.635 4602.4 918.43
## - failed        1     54.354 4616.1 919.37
## - Mjob          4    143.807 4705.6 919.44
## - studytime     1     57.318 4619.1 919.58
## - famsize       1     61.877 4623.6 919.89
## - schoolsup     1     64.734 4626.5 920.08
## - age           1     68.830 4630.6 920.36
## - famsup        1     84.566 4646.3 921.44
## + guardian      2      2.456 4559.3 921.46
## - absences      1     88.161 4649.9 921.68
## - sex           1     93.058 4654.8 922.01
## - romantic      1    110.186 4671.9 923.17
## + Fjob          4     24.135 4537.6 923.96
## - goout         1    132.764 4694.5 924.70
## - freetime      1    149.181 4710.9 925.80
##
## Step:  AIC=915.52
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + traveltime + studytime + failures + schoolsup +
##      famsup + paid + activities + nursery + higher + internet +
##      romantic + famrel + freetime + goout + Dalc + Walc + health +
##      absences + first_gen_college + high_freq_absent + failed
##
##              Df Sum of Sq   RSS   AIC
## - higher      1      0.000 4618.2 913.52
## - Dalc        1      0.197 4618.4 913.53
## - famrel      1      0.440 4618.7 913.55
## - Fedu        1      2.425 4620.6 913.68
## - Walc        1      3.203 4621.4 913.74
## - Medu        1      3.626 4621.8 913.77
## - activities  1      3.649 4621.9 913.77
## - nursery     1      5.593 4623.8 913.90
## - high_freq_absent 1      6.845 4625.1 913.99
## - address     1      9.786 4628.0 914.19
## - internet    1     11.003 4629.2 914.27
## - traveltime  1     15.053 4633.3 914.55
## - Pstatus     1     22.545 4640.8 915.06
## - failures    1     22.829 4641.0 915.08
## - school      1     26.310 4644.5 915.31
## <none>        1     4618.2 915.52

```

```

## - first_gen_college 1 36.895 4655.1 916.03
## - paid 1 37.204 4655.4 916.05
## - health 1 41.384 4659.6 916.34
## - failed 1 46.596 4664.8 916.69
## - studytime 1 57.716 4675.9 917.44
## + reason 3 56.461 4561.8 917.63
## - schoolsup 1 64.672 4682.9 917.91
## - famsize 1 64.838 4683.1 917.92
## - age 1 67.842 4686.1 918.13
## - Mjob 4 158.447 4776.7 918.18
## + guardian 2 4.240 4614.0 919.23
## - famsup 1 84.960 4703.2 919.28
## - sex 1 85.096 4703.3 919.29
## - absences 1 93.470 4711.7 919.85
## - romantic 1 105.080 4723.3 920.63
## + Fjob 4 28.845 4589.4 921.54
## - goout 1 137.376 4755.6 922.78
## - freetime 1 146.968 4765.2 923.42
##
## Step: AIC=913.52
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
## Fedu + Mjob + traveltime + studytime + failures + schoolsup +
## famsup + paid + activities + nursery + internet + romantic +
## famrel + freetime + goout + Dalc + Walc + health + absences +
## first_gen_college + high_freq_absent + failed
##
## Df Sum of Sq RSS AIC
## - Dalc 1 0.196 4618.4 911.53
## - famrel 1 0.439 4618.7 911.55
## - Fedu 1 2.451 4620.7 911.69
## - Walc 1 3.207 4621.4 911.74
## - Medu 1 3.628 4621.8 911.77
## - activities 1 3.662 4621.9 911.77
## - nursery 1 5.593 4623.8 911.90
## - high_freq_absent 1 6.942 4625.2 911.99
## - address 1 9.786 4628.0 912.19
## - internet 1 11.037 4629.3 912.27
## - traveltime 1 15.075 4633.3 912.55
## - Pstatus 1 22.556 4640.8 913.06
## - failures 1 23.561 4641.8 913.13
## - school 1 26.483 4644.7 913.33
## <none> 4618.2 913.52
## - first_gen_college 1 36.895 4655.1 914.03
## - paid 1 37.615 4655.8 914.08
## - health 1 41.386 4659.6 914.34
## - failed 1 46.981 4665.2 914.72
## - studytime 1 58.310 4676.5 915.48
## + higher 1 0.000 4618.2 915.52
## + reason 3 56.425 4561.8 915.63
## - schoolsup 1 64.687 4682.9 915.91
## - famsize 1 64.840 4683.1 915.92
## - Mjob 4 158.597 4776.8 916.19
## - age 1 69.331 4687.6 916.23
## + guardian 2 4.201 4614.0 917.23

```



```

## - famsup          1      85.007 4703.2 917.28
## - sex             1      86.601 4704.8 917.39
## - absences        1      94.813 4713.0 917.94
## - romantic         1     106.608 4724.8 918.73
## + Fjob            4      28.844 4589.4 919.54
## - goout           1     137.383 4755.6 920.78
## - freetime        1     146.990 4765.2 921.42
##
## Step: AIC=911.53
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + traveltime + studytime + failures + schoolsup +
##      famsup + paid + activities + nursery + internet + romantic +
##      famrel + freetime + goout + Walc + health + absences + first_gen_college +
##      high_freq_absent + failed
##
##              Df Sum of Sq    RSS    AIC
## - famrel      1      0.467 4618.9 909.56
## - Fedu         1      2.410 4620.8 909.70
## - Walc         1      3.018 4621.4 909.74
## - activities   1      3.619 4622.0 909.78
## - Medu         1      3.662 4622.1 909.78
## - nursery      1      6.145 4624.6 909.95
## - high_freq_absent 1      6.746 4625.2 909.99
## - address      1     10.060 4628.5 910.22
## - internet     1     11.105 4629.5 910.29
## - traveltime   1     14.931 4633.3 910.55
## - Pstatus      1     23.051 4641.5 911.11
## - failures     1     23.421 4641.8 911.13
## - school       1     26.986 4645.4 911.37
## <none>                                4618.4 911.53
## - first_gen_college 1     36.800 4655.2 912.04
## - paid         1     38.982 4657.4 912.19
## - health       1     41.199 4659.6 912.34
## - failed       1     47.590 4666.0 912.77
## - studytime    1     58.437 4676.9 913.51
## + Dalc         1      0.196 4618.2 913.52
## + higher       1      0.000 4618.4 913.53
## + reason       3     56.616 4561.8 913.63
## - schoolsup     1     64.538 4683.0 913.92
## - famsize      1     65.013 4683.4 913.95
## - Mjob         4    158.749 4777.2 914.21
## - age          1     69.138 4687.6 914.23
## + guardian     2      4.233 4614.2 915.24
## - famsup       1     85.027 4703.4 915.30
## - sex          1     87.210 4705.6 915.44
## - absences     1     94.843 4713.3 915.96
## - romantic     1    106.605 4725.0 916.74
## + Fjob         4     28.899 4589.5 917.55
## - goout        1    137.482 4755.9 918.80
## - freetime     1    148.808 4767.2 919.55
##
## Step: AIC=909.56
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + traveltime + studytime + failures + schoolsup +

```

```

##      famsup + paid + activities + nursery + internet + romantic +
##      freetime + goout + Walc + health + absences + first_gen_college +
##      high_freq_absent + failed
##
##              Df Sum of Sq    RSS    AIC
## - Fedu      1      2.288 4621.2 907.72
## - Walc      1      3.540 4622.4 907.81
## - activities 1      3.632 4622.5 907.81
## - Medu      1      3.758 4622.6 907.82
## - nursery   1      6.014 4624.9 907.98
## - high_freq_absent 1      7.094 4626.0 908.05
## - address   1     10.211 4629.1 908.26
## - internet  1     11.223 4630.1 908.33
## - traveltime 1     14.719 4633.6 908.57
## - Pstatus   1     22.801 4641.7 909.12
## - failures  1     23.858 4642.7 909.19
## - school    1     26.528 4645.4 909.37
## <none>                                4618.9 909.56
## - first_gen_college 1     36.728 4655.6 910.07
## - paid      1     39.415 4658.3 910.25
## - health    1     40.797 4659.7 910.34
## - failed    1     47.519 4666.4 910.80
## + famrel    1      0.467 4618.4 911.53
## + Dalc      1      0.224 4618.7 911.55
## - studytime 1     58.712 4677.6 911.56
## + higher    1      0.000 4618.9 911.56
## + reason    3     57.019 4561.9 911.64
## - schoolsup  1     64.366 4683.2 911.94
## - famsize   1     65.123 4684.0 911.99
## - age       1     68.765 4687.6 912.23
## - Mjob      4    159.308 4778.2 912.28
## + guardian  2      4.154 4614.7 913.28
## - famsup    1     84.826 4703.7 913.32
## - sex       1     88.733 4707.6 913.58
## - absences  1     95.626 4714.5 914.04
## - romantic  1    106.423 4725.3 914.76
## + Fjob      4     28.327 4590.6 915.62
## - goout     1    137.094 4756.0 916.81
## - freetime  1    153.451 4772.3 917.89
##
## Step:  AIC=907.72
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Mjob + traveltime + studytime + failures + schoolsup + famsup +
##      paid + activities + nursery + internet + romantic + freetime +
##      goout + Walc + health + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq    RSS    AIC
## - Medu      1      2.716 4623.9 905.91
## - Walc      1      3.732 4624.9 905.98
## - activities 1      3.937 4625.1 905.99
## - high_freq_absent 1      6.461 4627.6 906.16
## - nursery   1      6.563 4627.7 906.17
## - address   1     10.636 4631.8 906.45

```

```

## - internet          1    11.233 4632.4 906.49
## - traveltime        1    13.638 4634.8 906.65
## - Pstatus           1    22.105 4643.3 907.23
## - failures          1    22.136 4643.3 907.23
## - school            1    25.743 4646.9 907.48
## <none>               4621.2 907.72
## - first_gen_college 1    36.511 4657.7 908.21
## - paid              1    40.059 4661.2 908.45
## - health            1    41.096 4662.3 908.52
## - failed            1    49.573 4670.7 909.09
## + Fedu              1     2.288 4618.9 909.56
## + famrel            1     0.345 4620.8 909.70
## + Dalc              1     0.178 4621.0 909.71
## + higher            1     0.022 4621.1 909.72
## + reason            3    57.777 4563.4 909.75
## - studytime         1    62.725 4683.9 909.98
## - famsize           1    65.935 4687.1 910.20
## - schoolsup         1    66.605 4687.8 910.24
## - Mjob              4   157.259 4778.4 910.30
## - age               1    69.879 4691.1 910.46
## + guardian          2     2.871 4618.3 911.52
## - sex               1    87.997 4709.2 911.68
## - famsup            1    89.835 4711.0 911.81
## - absences          1    94.024 4715.2 912.09
## - romantic          1   106.822 4728.0 912.94
## + Fjob              4    24.289 4596.9 914.06
## - goout             1   138.827 4760.0 915.07
## - freetime          1   158.481 4779.7 916.38
##
## Step:  AIC=905.91
## G3 ~ school + sex + age + address + famsize + Pstatus + Mjob +
##      traveltime + studytime + failures + schoolsup + famsup +
##      paid + activities + nursery + internet + romantic + freetime +
##      goout + Walc + health + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq   RSS   AIC
## - Walc        1     3.904 4627.8 904.17
## - activities  1     4.142 4628.0 904.19
## - high_freq_absent 1     5.900 4629.8 904.31
## - nursery     1     6.240 4630.1 904.33
## - address     1    11.026 4634.9 904.66
## - internet    1    11.360 4635.2 904.68
## - traveltime  1    14.136 4638.0 904.87
## - failures    1    22.039 4645.9 905.41
## - Pstatus     1    23.280 4647.2 905.49
## - school      1    25.026 4648.9 905.61
## <none>         4623.9 905.91
## - paid        1    40.304 4664.2 906.65
## - health      1    44.173 4668.1 906.91
## - failed      1    51.557 4675.4 907.41
## + Medu        1     2.716 4621.2 907.72
## + reason      3    59.868 4564.0 907.79
## + Fedu        1     1.247 4622.6 907.82

```

```

## + famrel          1      0.446 4623.4 907.88
## + Dalc            1      0.221 4623.7 907.89
## + higher          1      0.020 4623.9 907.91
## - famsize         1     63.719 4687.6 908.23
## - studytime       1     66.557 4690.4 908.42
## - schoolsup        1     67.735 4691.6 908.50
## - age             1     70.228 4694.1 908.67
## - Mjob            4    160.854 4784.7 908.71
## - first_gen_college 1     80.781 4704.7 909.38
## + guardian        2      2.473 4621.4 909.74
## - famsup          1     87.926 4711.8 909.86
## - sex             1     91.639 4715.5 910.11
## - absences        1     98.034 4721.9 910.54
## - romantic        1    105.730 4729.6 911.05
## + Fjob            4     25.623 4598.3 912.15
## - goout           1    136.934 4760.8 913.13
## - freetime        1    157.319 4781.2 914.48
##
## Step: AIC=904.17
## G3 ~ school + sex + age + address + famsize + Pstatus + Mjob +
##      traveltime + studytime + failures + schoolsup + famsup +
##      paid + activities + nursery + internet + romantic + freetime +
##      goout + health + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq   RSS   AIC
## - activities      1      4.178 4632.0 902.46
## - nursery         1      5.015 4632.8 902.52
## - high_freq_absent 1      7.291 4635.1 902.67
## - address         1     11.211 4639.0 902.94
## - internet        1     11.897 4639.7 902.98
## - traveltime      1     15.286 4643.1 903.22
## - Pstatus         1     22.412 4650.2 903.70
## - failures        1     23.632 4651.4 903.78
## - school          1     25.623 4653.4 903.92
## <none>                                4627.8 904.17
## - paid           1     39.528 4667.3 904.86
## - health         1     47.178 4675.0 905.38
## - failed         1     50.269 4678.1 905.59
## + Walc           1      3.904 4623.9 905.91
## + Medu           1      2.888 4624.9 905.98
## + reason         3     60.583 4567.2 906.01
## + Fedu           1      1.364 4626.4 906.08
## + famrel         1      0.987 4626.8 906.11
## + Dalc           1      0.015 4627.8 906.17
## + higher         1      0.003 4627.8 906.17
## - famsize        1     62.702 4690.5 906.43
## - schoolsup       1     67.153 4694.9 906.73
## - studytime      1     67.748 4695.5 906.77
## - age            1     67.910 4695.7 906.78
## - Mjob           4    162.533 4790.3 907.08
## - first_gen_college 1     79.320 4707.1 907.54
## - famsup         1     85.618 4713.4 907.97
## + guardian       2      2.682 4625.1 907.99

```

```

## - sex          1      87.930 4715.7 908.12
## - absences     1      96.725 4724.5 908.71
## - romantic     1     105.538 4733.3 909.30
## + Fjob         4      27.949 4599.8 910.26
## - freetime     1     161.494 4789.3 913.01
## - goout        1     189.734 4817.5 914.87
##
## Step: AIC=902.46
## G3 ~ school + sex + age + address + famsize + Pstatus + Mjob +
##      traveltime + studytime + failures + schoolsup + famsup +
##      paid + nursery + internet + romantic + freetime + goout +
##      health + absences + first_gen_college + high_freq_absent +
##      failed
##
##              Df Sum of Sq    RSS    AIC
## - nursery      1      5.067 4637.0 900.80
## - high_freq_absent 1      6.862 4638.8 900.93
## - internet     1     11.689 4643.7 901.25
## - address      1     12.273 4644.2 901.29
## - traveltime   1     14.998 4647.0 901.48
## - failures     1     22.502 4654.5 901.99
## - Pstatus      1     24.454 4656.4 902.12
## - school       1     27.167 4659.1 902.31
## <none>                4632.0 902.46
## - paid         1     40.599 4672.6 903.22
## - health       1     47.606 4679.6 903.69
## - failed       1     52.022 4684.0 903.99
## + activities   1      4.178 4627.8 904.17
## + Walc         1      3.940 4628.0 904.19
## + Medu         1      3.101 4628.9 904.25
## + Fedu         1      1.572 4630.4 904.35
## + famrel       1      1.003 4631.0 904.39
## + Dalc         1      0.032 4631.9 904.46
## + higher       1      0.028 4631.9 904.46
## + reason       3     56.135 4575.8 904.61
## - famsize      1     63.140 4695.1 904.74
## - studytime    1     65.022 4697.0 904.86
## - age          1     66.050 4698.0 904.93
## - schoolsup     1     67.088 4699.1 905.00
## - Mjob         4    164.263 4796.2 905.47
## - first_gen_college 1    77.150 4709.1 905.68
## - famsup       1     85.006 4717.0 906.21
## - sex          1     85.140 4717.1 906.21
## + guardian     2      2.670 4629.3 906.28
## - absences     1     95.919 4727.9 906.94
## - romantic     1    107.519 4739.5 907.71
## + Fjob         4     28.113 4603.9 908.53
## - freetime     1    159.222 4791.2 911.14
## - goout        1    192.427 4824.4 913.32
##
## Step: AIC=900.8
## G3 ~ school + sex + age + address + famsize + Pstatus + Mjob +
##      traveltime + studytime + failures + schoolsup + famsup +
##      paid + internet + romantic + freetime + goout + health +

```

```

##      absences + first_gen_college + high_freq_absent + failed
##
##      Df Sum of Sq    RSS    AIC
## - high_freq_absent  1      7.525 4644.6 899.32
## - address          1     11.795 4648.8 899.61
## - internet         1     12.829 4649.9 899.68
## - traveltime       1     15.140 4652.2 899.83
## - failures         1     22.776 4659.8 900.35
## - Pstatus          1     23.686 4660.7 900.41
## - school           1     28.953 4666.0 900.77
## <none>
##                  4637.0 900.80
## - paid             1     38.795 4675.8 901.44
## - health           1     47.660 4684.7 902.04
## - failed           1     51.399 4688.4 902.29
## + nursery          1      5.067 4632.0 902.46
## + activities        1      4.230 4632.8 902.52
## + Medu              1      2.757 4634.3 902.62
## + Walc              1      2.704 4634.3 902.62
## + Fedu              1      2.000 4635.0 902.67
## + famrel            1      0.699 4636.3 902.76
## + Dalc              1      0.077 4637.0 902.80
## + higher            1      0.025 4637.0 902.80
## - famsize           1     60.239 4697.3 902.88
## - studytime         1     62.151 4699.2 903.01
## + reason            3     54.246 4582.8 903.09
## - age               1     64.954 4702.0 903.20
## - schoolsup         1     68.628 4705.7 903.45
## - Mjob              4    162.543 4799.6 903.69
## - first_gen_college 1     74.036 4711.1 903.81
## - famsup            1     83.040 4720.1 904.41
## - sex               1     84.148 4721.2 904.49
## + guardian          2      3.830 4633.2 904.54
## - absences          1     97.372 4734.4 905.37
## - romantic          1    107.431 4744.5 906.04
## + Fjob              4     28.374 4608.7 906.86
## - freetime          1    163.247 4800.3 909.74
## - goout             1    194.154 4831.2 911.77
##
## Step:  AIC=899.32
## G3 ~ school + sex + age + address + famsize + Pstatus + Mjob +
##      traveltime + studytime + failures + schoolsup + famsup +
##      paid + internet + romantic + freetime + goout + health +
##      absences + first_gen_college + failed
##
##      Df Sum of Sq    RSS    AIC
## - address          1     10.888 4655.4 898.06
## - internet         1     12.243 4656.8 898.15
## - traveltime       1     13.276 4657.8 898.22
## - failures         1     22.376 4666.9 898.84
## - Pstatus          1     22.751 4667.3 898.86
## - school           1     29.054 4673.6 899.29
## <none>
##                  4644.6 899.32
## - paid             1     38.496 4683.1 899.92
## - health           1     45.625 4690.2 900.41

```

```

## + high_freq_absent 1 7.525 4637.0 900.80
## + nursery 1 5.730 4638.8 900.93
## - failed 1 53.561 4698.1 900.94
## + Walc 1 3.825 4640.7 901.06
## + activities 1 3.782 4640.8 901.06
## + Medu 1 2.117 4642.4 901.17
## - famsize 1 57.573 4702.1 901.21
## + Fedu 1 1.525 4643.0 901.21
## + famrel 1 1.236 4643.3 901.23
## + higher 1 0.191 4644.4 901.30
## + Dalc 1 0.053 4644.5 901.31
## - studytime 1 64.806 4709.4 901.70
## + reason 3 51.488 4593.1 901.79
## - schoolsup 1 66.391 4711.0 901.80
## - age 1 67.062 4711.6 901.85
## - Mjob 4 159.553 4804.1 901.99
## - first_gen_college 1 70.839 4715.4 902.10
## - sex 1 83.878 4728.4 902.97
## + guardian 2 3.663 4640.9 903.07
## - famsup 1 85.915 4730.5 903.11
## - romantic 1 107.706 4752.3 904.56
## + Fjob 4 30.169 4614.4 905.26
## - absences 1 122.998 4767.6 905.58
## - freetime 1 159.558 4804.1 907.99
## - goout 1 198.114 4842.7 910.52
##
## Step: AIC=898.06
## G3 ~ school + sex + age + famsize + Pstatus + Mjob + traveltime +
## studytime + failures + schoolsup + famsup + paid + internet +
## romantic + freetime + goout + health + absences + first_gen_college +
## failed
##
## Df Sum of Sq RSS AIC
## - internet 1 17.076 4672.5 897.21
## - traveltime 1 22.097 4677.5 897.55
## - failures 1 22.963 4678.4 897.61
## - school 1 23.121 4678.6 897.62
## - Pstatus 1 23.721 4679.2 897.66
## <none> 4655.4 898.06
## - paid 1 38.359 4693.8 898.65
## + address 1 10.888 4644.6 899.32
## - health 1 48.864 4704.3 899.36
## + high_freq_absent 1 6.618 4648.8 899.61
## + nursery 1 5.198 4650.3 899.70
## - failed 1 54.289 4709.7 899.72
## + activities 1 4.759 4650.7 899.73
## + Walc 1 3.988 4651.5 899.79
## + Medu 1 2.539 4652.9 899.88
## + Fedu 1 1.844 4653.6 899.93
## + famrel 1 1.469 4654.0 899.96
## + higher 1 0.188 4655.3 900.04
## + Dalc 1 0.000 4655.4 900.06
## - studytime 1 59.649 4715.1 900.08
## - famsize 1 60.883 4716.3 900.16

```

```

## - age          1      67.556 4723.0 900.61
## - schoolsup    1      67.754 4723.2 900.62
## - first_gen_college 1      69.505 4725.0 900.74
## + reason       3      46.296 4609.2 900.90
## - Mjob         4     163.441 4818.9 900.96
## - sex          1      80.414 4735.9 901.47
## + guardian     2       5.868 4649.6 901.66
## - famsup       1      89.836 4745.3 902.10
## - romantic     1     107.007 4762.5 903.24
## + Fjob         4      31.716 4623.7 903.90
## - absences     1     118.305 4773.8 903.99
## - freetime     1     160.073 4815.5 906.74
## - goout        1     195.465 4850.9 909.05
##
## Step:  AIC=897.21
## G3 ~ school + sex + age + famsize + Pstatus + Mjob + traveltime +
##      studytime + failures + schoolsup + famsup + paid + romantic +
##      freetime + goout + health + absences + first_gen_college +
##      failed
##
##              Df Sum of Sq   RSS   AIC
## - Pstatus      1    19.015 4691.5 896.50
## - school        1    21.569 4694.1 896.67
## - failures      1    21.788 4694.3 896.68
## - traveltime    1    23.853 4696.4 896.82
## <none>                      4672.5 897.21
## + internet      1    17.076 4655.4 898.06
## - paid          1    43.019 4715.5 898.11
## + address       1    15.721 4656.8 898.15
## - health        1    52.668 4725.2 898.76
## + nursery       1     6.408 4666.1 898.78
## + high_freq_absent 1     5.782 4666.7 898.82
## + activities    1     4.747 4667.8 898.89
## + Walc          1     4.446 4668.1 898.91
## + Medu          1     2.790 4669.7 899.02
## + Fedu          1     1.987 4670.5 899.08
## + famrel        1     1.765 4670.8 899.09
## - failed        1    58.055 4730.6 899.12
## + higher        1     0.379 4672.1 899.19
## + Dalc          1     0.031 4672.5 899.21
## - famsize       1    61.777 4734.3 899.36
## - studytime     1    61.937 4734.5 899.37
## - schoolsup     1    68.343 4740.9 899.80
## - first_gen_college 1    70.528 4743.1 899.95
## - age          1    72.763 4745.3 900.10
## + reason       3    45.034 4627.5 900.15
## - Mjob         4   168.611 4841.1 900.42
## - sex          1    82.535 4755.1 900.75
## + guardian     2     5.820 4666.7 900.82
## - famsup       1    86.871 4759.4 901.03
## - romantic     1    98.108 4770.6 901.78
## + Fjob         4    30.735 4641.8 903.13
## - absences     1   128.687 4801.2 903.80
## - freetime     1   155.787 4828.3 905.58

```



```

## - goout          1    194.404 4866.9 908.09
##
## Step: AIC=896.5
## G3 ~ school + sex + age + famsize + Mjob + traveltime + studytime +
##      failures + schoolsup + famsup + paid + romantic + freetime +
##      goout + health + absences + first_gen_college + failed
##
##              Df Sum of Sq    RSS    AIC
## - school      1     20.534 4712.1 895.88
## - failures     1     20.551 4712.1 895.88
## - traveltime   1     26.808 4718.3 896.30
## <none>                    4691.5 896.50
## - paid         1     37.875 4729.4 897.04
## + Pstatus      1     19.015 4672.5 897.21
## + address      1     15.946 4675.6 897.42
## + internet     1     12.371 4679.2 897.66
## - health       1     50.727 4742.3 897.90
## + activities   1      6.716 4684.8 898.04
## + nursery      1      5.386 4686.2 898.13
## + high_freq_absent 1      5.106 4686.4 898.15
## + Medu         1      3.943 4687.6 898.23
## + Walc         1      3.552 4688.0 898.26
## + Fedu         1      1.258 4690.3 898.41
## + famrel       1      1.224 4690.3 898.41
## + Dalc         1      0.296 4691.2 898.48
## + higher       1      0.216 4691.3 898.48
## - failed       1     61.500 4753.0 898.61
## - studytime    1     62.139 4753.7 898.65
## - schoolsup    1     66.578 4758.1 898.95
## - famsize      1     71.246 4762.8 899.26
## - age          1     72.054 4763.6 899.31
## + reason       3     44.134 4647.4 899.51
## - sex          1     81.989 4773.5 899.97
## + guardian     2      6.841 4684.7 900.04
## - first_gen_college 1     83.323 4774.9 900.06
## - Mjob         4    175.278 4866.8 900.09
## - famsup       1     85.654 4777.2 900.21
## - romantic     1     97.402 4788.9 900.99
## + Fjob         4     32.692 4658.8 902.29
## - absences     1    140.967 4832.5 903.85
## - freetime     1    146.356 4837.9 904.20
## - goout        1    193.877 4885.4 907.29
##
## Step: AIC=895.88
## G3 ~ sex + age + famsize + Mjob + traveltime + studytime + failures +
##      schoolsup + famsup + paid + romantic + freetime + goout +
##      health + absences + first_gen_college + failed
##
##              Df Sum of Sq    RSS    AIC
## - traveltime   1     17.052 4729.1 895.02
## - failures     1     23.556 4735.6 895.45
## <none>                    4712.1 895.88
## + school      1     20.534 4691.5 896.50
## - paid         1     39.965 4752.0 896.55

```

```

## + Pstatus      1    17.980 4694.1 896.67
## + internet     1    11.208 4700.9 897.12
## + address      1     8.647 4703.4 897.30
## + activities   1     7.951 4704.1 897.34
## + nursery      1     7.080 4705.0 897.40
## - age          1    54.142 4766.2 897.49
## + high_freq_absent 1     5.453 4706.6 897.51
## - health       1    54.765 4766.8 897.53
## - studytime    1    56.040 4768.1 897.61
## + Walc         1     3.855 4708.2 897.62
## + Medu         1     2.933 4709.1 897.68
## - failed       1    58.468 4770.5 897.77
## + Fedu         1     0.852 4711.2 897.82
## + Dalc         1     0.610 4711.5 897.84
## + famrel       1     0.320 4711.8 897.86
## + higher       1     0.003 4712.1 897.88
## - schoolsup     1    66.967 4779.0 898.34
## - famsize      1    72.499 4784.6 898.70
## - sex          1    77.262 4789.3 899.02
## + reason       3    41.679 4670.4 899.07
## - first_gen_college 1    81.934 4794.0 899.32
## - Mjob         4   174.514 4886.6 899.37
## + guardian     2     4.868 4707.2 899.55
## - romantic     1    94.621 4806.7 900.16
## - famsup       1    96.177 4808.3 900.26
## + Fjob         4    35.159 4676.9 901.51
## - absences     1   126.801 4838.9 902.27
## - freetime     1   154.525 4866.6 904.07
## - goout        1   202.366 4914.4 907.16
##
## Step:  AIC=895.02
## G3 ~ sex + age + famsize + Mjob + studytime + failures + schoolsup +
##      famsup + paid + romantic + freetime + goout + health + absences +
##      first_gen_college + failed
##
##              Df Sum of Sq   RSS   AIC
## - failures      1    24.582 4753.7 894.66
## <none>              4729.1 895.02
## + Pstatus      1    20.570 4708.6 895.64
## + traveltime    1    17.052 4712.1 895.88
## + address      1    17.020 4712.1 895.88
## - paid         1    43.139 4772.3 895.89
## + internet     1    12.480 4716.6 896.18
## + school       1    10.778 4718.3 896.30
## + activities   1     7.906 4721.2 896.49
## + nursery      1     6.473 4722.7 896.59
## - age          1    54.411 4783.5 896.63
## - health       1    54.994 4784.1 896.67
## + Walc         1     4.849 4724.3 896.69
## + Medu         1     4.223 4724.9 896.74
## + high_freq_absent 1     3.219 4725.9 896.80
## - failed       1    58.351 4787.5 896.89
## + Dalc         1     0.382 4728.7 896.99
## + famrel       1     0.323 4728.8 897.00

```

```

## + Fedu          1      0.267 4728.9 897.00
## + higher        1      0.021 4729.1 897.02
## - studytime     1     60.228 4789.4 897.02
## - schoolsup     1     63.756 4792.9 897.25
## - famsize       1     68.825 4798.0 897.58
## - sex           1     75.532 4804.7 898.03
## + reason        3     43.987 4685.1 898.07
## - first_gen_college 1     82.183 4811.3 898.46
## + guardian      2       3.243 4725.9 898.80
## - famsup        1     99.775 4828.9 899.62
## - romantic      1    101.172 4830.3 899.71
## - Mjob          4    194.583 4923.7 899.76
## + Fjob          4     32.399 4696.7 900.85
## - absences      1    132.613 4861.7 901.76
## - freetime      1    158.505 4887.6 903.44
## - goout         1    210.307 4939.4 906.77
##
## Step:  AIC=894.66
## G3 ~ sex + age + famsize + Mjob + studytime + schoolsup + famsup +
##      paid + romantic + freetime + goout + health + absences +
##      first_gen_college + failed
##
##              Df Sum of Sq    RSS    AIC
## <none>                4753.7 894.66
## + failures           1     24.58 4729.1 895.02
## + Pstatus            1     19.17 4734.5 895.38
## + traveltime         1     18.08 4735.6 895.45
## + address            1     17.27 4736.4 895.51
## - paid              1     43.30 4797.0 895.52
## + school            1     12.75 4741.0 895.81
## + internet          1     11.52 4742.2 895.89
## - health            1     52.68 4806.4 896.14
## + nursery           1      6.86 4746.8 896.20
## + Walc              1      6.61 4747.1 896.22
## + activities         1      6.32 4747.4 896.24
## - age               1     55.75 4809.5 896.34
## + Medu              1      4.03 4749.7 896.39
## + high_freq_absent  1      2.94 4750.8 896.46
## + higher            1      1.39 4752.3 896.56
## + famrel            1      0.81 4752.9 896.60
## + Dalc              1      0.15 4753.6 896.65
## + Fedu              1      0.06 4753.6 896.65
## - schoolsup         1     63.93 4817.6 896.88
## - studytime         1     68.57 4822.3 897.18
## + reason            3     49.10 4704.6 897.38
## - famsize          1     73.95 4827.7 897.53
## - sex              1     74.40 4828.1 897.56
## - first_gen_college 1     86.27 4840.0 898.34
## + guardian          2      3.68 4750.0 898.41
## - Mjob              4    195.19 4948.9 899.37
## - romantic          1    103.42 4857.1 899.46
## - famsup            1    105.75 4859.5 899.61
## + Fjob              4     33.61 4720.1 900.41
## - absences          1    147.05 4900.8 902.28

```

```
## - freetime          1    160.34 4914.0 903.14
## - goout             1    219.35 4973.1 906.91
## - failed            1    497.54 5251.3 924.11
```

```
AIC(step.model_tr)
```

```
## [1] 1793.426
```

```
summary(step.model_tr)
```

```
##
## Call:
## lm(formula = G3 ~ sex + age + famsize + Mjob + studytime + schoolsup +
##     famsup + paid + romantic + freetime + goout + health + absences +
##     first_gen_college + failed, data = training)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.5966  -1.9143   0.4216   2.6383  10.1119
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    15.61947     3.50460   4.457 1.18e-05 ***
## sexM             1.09117     0.50610   2.156 0.031885 *
## age            -0.37192     0.19928  -1.866 0.062990 .
## famsizeLE3       1.09276     0.50838   2.149 0.032403 *
## Mjobhealth       1.51975     1.04176   1.459 0.145669
## Mjobother        0.20870     0.69149   0.302 0.763006
## Mjobservices     1.41891     0.74064   1.916 0.056352 .
## Mjobteacher     -0.95496     0.99932  -0.956 0.340045
## studytime        0.60425     0.29194   2.070 0.039340 *
## schoolsupyes    -1.43855     0.71982  -1.998 0.046575 *
## famsupyes       -1.31666     0.51223  -2.570 0.010644 *
## paidyes          0.81591     0.49608   1.645 0.101084
## romanticyes     -1.27220     0.50050  -2.542 0.011534 *
## freetimelow    -1.56347     0.49398  -3.165 0.001712 **
## gooutlow        1.84589     0.49863   3.702 0.000255 ***
## healthlow       0.84074     0.46341   1.814 0.070648 .
## absences         0.08500     0.02804   3.031 0.002652 **
## first_gen_collegeyes -1.37377   0.59173  -2.322 0.020930 *
## failedyes       -3.43265     0.61568  -5.575 5.55e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.001 on 297 degrees of freedom
## Multiple R-squared:  0.3157, Adjusted R-squared:  0.2742
## F-statistic: 7.612 on 18 and 297 DF,  p-value: 2.832e-16
```

AIC of 1793.426 and adjusted R-squared of 0.2742.

Active variables: goout, absences, first\_gen\_college, failed, freetime, romantic, famsup, schoolsup, studytime, famsize, sex. Let's explore these with interaction terms.

```
activelm_tr <- lm(G3 ~ (sex + schoolsup + romantic + freetime + goout + absences + failed + studytime +
summary(activelm_tr)
```

```
##
```

```
## Call:
## lm(formula = G3 ~ (sex + schoolsup + romantic + freetime + goout +
##     absences + failed + studytime + first_gen_college + famsup +
##     famsize), data = training)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.5751  -2.1790   0.4765   2.7999  11.0657
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    10.60588    0.99491   10.660 < 2e-16 ***
## sexM             0.84292    0.50654    1.664 0.097129 .
## schoolsupyes    -0.92949    0.68924   -1.349 0.178478
## romanticyes     -1.35942    0.50766   -2.678 0.007813 **
## freetimelow    -1.42836    0.50007   -2.856 0.004581 **
## gooutlow        1.70615    0.50120    3.404 0.000753 ***
## absences         0.07599    0.02813    2.701 0.007297 **
## failedyes      -3.75325    0.59064   -6.355 7.63e-10 ***
## studytime        0.56316    0.29587    1.903 0.057930 .
## first_gen_collegeyes -1.34272    0.49968   -2.687 0.007603 **
## famsupyes       -0.87684    0.49917   -1.757 0.079996 .
## famsizeLE3       1.14706    0.51719    2.218 0.027304 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.097 on 304 degrees of freedom
## Multiple R-squared:  0.2655, Adjusted R-squared:  0.2389
## F-statistic: 9.987 on 11 and 304 DF,  p-value: 1.635e-15
```

```
AIC(activelm_tr)
```

```
## [1] 1801.82
```

AIC of 1801.82, Adjusted R-squared of 0.2389.

Adding all interaction terms:

```
interlm_tr <- lm(G3 ~ (sex + schoolsup + romantic + freetime + goout + absences + failed + studytime +
summary(interlm_tr)
```

```
##
## Call:
## lm(formula = G3 ~ (sex + schoolsup + romantic + freetime + goout +
##     absences + failed + studytime + first_gen_college + famsup +
##     famsize)^2, data = training)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.1133  -1.9546   0.3402   2.2558   8.0897
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    17.828089    2.724408    6.544 3.39e-10 ***
## sexM           -2.236645    1.989617   -1.124 0.262028
## schoolsupyes    2.491733    2.900969    0.859 0.391205
```

## romanticyes	-2.924267	2.589279	-1.129	0.259826	
## freetimelow	-2.019097	2.075748	-0.973	0.331642	
## gooutlow	1.696368	2.184918	0.776	0.438251	
## absences	-0.123817	0.172425	-0.718	0.473376	
## failedyes	-5.471058	2.886823	-1.895	0.059226	.
## studytime	-2.406763	1.037396	-2.320	0.021150	*
## first_gen_collegeyes	-7.000153	1.957673	-3.576	0.000419	***
## famsupyes	-3.708275	1.989413	-1.864	0.063498	.
## famsizeLE3	1.443291	2.318042	0.623	0.534095	
## sexM:schoolsupyes	-2.993395	1.677001	-1.785	0.075483	.
## sexM:romanticyes	1.287244	1.071327	1.202	0.230682	
## sexM:freetimelow	1.086732	1.060899	1.024	0.306664	
## sexM:gooutlow	-0.667762	1.122071	-0.595	0.552307	
## sexM:absences	-0.009894	0.099224	-0.100	0.920654	
## sexM:failedyes	-3.771541	1.544833	-2.441	0.015328	*
## sexM:studytime	0.167683	0.614759	0.273	0.785263	
## sexM:first_gen_collegeyes	4.029339	1.103393	3.652	0.000317	***
## sexM:famsupyes	1.168833	1.082460	1.080	0.281279	
## sexM:famsizeLE3	-0.415000	1.199880	-0.346	0.729734	
## schoolsupyes:romanticyes	2.301138	1.966379	1.170	0.243023	
## schoolsupyes:freetimelow	0.798298	1.701002	0.469	0.639258	
## schoolsupyes:gooutlow	0.312639	1.893847	0.165	0.869014	
## schoolsupyes:absences	-0.244795	0.104775	-2.336	0.020265	*
## schoolsupyes:failedyes	2.951505	1.786648	1.652	0.099799	.
## schoolsupyes:studytime	-2.227824	0.914124	-2.437	0.015506	*
## schoolsupyes:first_gen_collegeyes	3.809002	1.644812	2.316	0.021384	*
## schoolsupyes:famsupyes	-0.150740	1.589188	-0.095	0.924508	
## schoolsupyes:famsizeLE3	-3.380800	1.893210	-1.786	0.075357	.
## romanticyes:freetimelow	-1.775277	1.097661	-1.617	0.107074	
## romanticyes:gooutlow	-0.260153	1.149870	-0.226	0.821196	
## romanticyes:absences	-0.018206	0.079839	-0.228	0.819808	
## romanticyes:failedyes	-0.209183	1.298112	-0.161	0.872111	
## romanticyes:studytime	1.104782	0.738687	1.496	0.136023	
## romanticyes:first_gen_collegeyes	0.487793	1.144907	0.426	0.670436	
## romanticyes:famsupyes	-0.478762	1.171035	-0.409	0.683011	
## romanticyes:famsizeLE3	0.419434	1.265768	0.331	0.740646	
## freetimelow:gooutlow	-1.544626	0.993952	-1.554	0.121449	
## freetimelow:absences	0.133237	0.079625	1.673	0.095522	.
## freetimelow:failedyes	-2.576119	1.337191	-1.927	0.055178	.
## freetimelow:studytime	0.592657	0.655277	0.904	0.366638	
## freetimelow:first_gen_collegeyes	0.974338	1.060793	0.918	0.359246	
## freetimelow:famsupyes	-0.383726	1.107122	-0.347	0.729187	
## freetimelow:famsizeLE3	-0.642029	1.131025	-0.568	0.570783	
## gooutlow:absences	-0.067710	0.086300	-0.785	0.433442	
## gooutlow:failedyes	0.526401	1.319646	0.399	0.690312	
## gooutlow:studytime	1.208907	0.697376	1.734	0.084243	.
## gooutlow:first_gen_collegeyes	0.114464	1.128380	0.101	0.919282	
## gooutlow:famsupyes	-0.982709	1.089033	-0.902	0.367733	
## gooutlow:famsizeLE3	-0.573774	1.175861	-0.488	0.626007	
## absences:failedyes	0.215436	0.070183	3.070	0.002381	**
## absences:studytime	-0.006741	0.056951	-0.118	0.905869	
## absences:first_gen_collegeyes	0.116287	0.081528	1.426	0.155018	
## absences:famsupyes	0.171607	0.072160	2.378	0.018156	*
## absences:famsizeLE3	-0.152720	0.081077	-1.884	0.060779	.

```
## failedyes:studytime          -0.340481    0.950143   -0.358  0.720387
## failedyes:first_gen_collegeyes  2.933576    1.635308    1.794  0.074043 .
## failedyes:famsupyes           0.694334    1.308619    0.531  0.596179
## failedyes:famsizeLE3          1.667873    1.465060    1.138  0.256034
## studytime:first_gen_collegeyes  0.989623    0.642846    1.539  0.124967
## studytime:famsupyes           1.177579    0.656217    1.794  0.073947 .
## studytime:famsizeLE3          0.913739    0.752968    1.214  0.226083
## first_gen_collegeyes:famsupyes  0.202676    1.124631    0.180  0.857130
## first_gen_collegeyes:famsizeLE3 -0.920305    1.184496   -0.777  0.437920
## famsupyes:famsizeLE3          0.116438    1.163129    0.100  0.920339
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.79 on 249 degrees of freedom
## Multiple R-squared:  0.4852, Adjusted R-squared:  0.3488
## F-statistic: 3.556 on 66 and 249 DF,  p-value: 3.206e-13
```

```
AIC(interlm_tr)
```

```
## [1] 1799.471
```

Significant interactions: *absencesfamsup*, *absencesfailed* *schoolsupfirst\_gen\_college*, *schoolsupabsences*, *schoolsupstudytime*, *sexfailed*, *sex\*first\_gen\_college*

```
interlm_tr1 <- lm(G3 ~ (sex + schoolsup + romantic + freetime + goout + absences + failed + studytime +
summary(interlm_tr1)
```

```
##
## Call:
## lm(formula = G3 ~ (sex + schoolsup + romantic + freetime + goout +
##   absences + failed + studytime + first_gen_college + famsup +
##   famsize + absences * famsup + absences * failed + schoolsup *
##   first_gen_college + schoolsup * absences + schoolsup * studytime +
##   sex * failed + sex * first_gen_college), data = training)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.4178  -2.2623   0.5875   2.6372   8.9101
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    11.693293    1.020762   11.455 < 2e-16 ***
## sexM           -0.887111    0.727008   -1.220  0.223349
## schoolsupyes    0.352447    1.999087    0.176  0.860175
## romanticyes    -1.643156    0.483126   -3.401  0.000763 ***
## freetimelow   -1.265993    0.472756   -2.678  0.007821 **
## gooutlow       1.556515    0.482777    3.224  0.001405 **
## absences        0.006135    0.044127    0.139  0.889517
## failedyes      -4.414005    0.906859   -4.867  1.84e-06 ***
## studytime       0.805798    0.304397    2.647  0.008549 **
## first_gen_collegeyes -3.405193    0.672109   -5.066  7.13e-07 ***
## famsupyes      -1.267101    0.578374   -2.191  0.029243 *
## famsizeLE3      1.136219    0.492371    2.308  0.021706 *
## absences:famsupyes  0.066071    0.056944    1.160  0.246872
## absences:failedyes  0.202897    0.061362    3.307  0.001061 **
## schoolsupyes:first_gen_collegeyes  4.459123    1.373647    3.246  0.001303 **
```

```
## schoolsupyes:absences          -0.099735    0.077493   -1.287  0.199090
## schoolsupyes:studytime         -1.634406    0.749457   -2.181  0.029984 *
## sexM:failedyes                 -2.147021    1.123717   -1.911  0.057013 .
## sexM:first_gen_collegeyes      3.729817    0.948321    3.933  0.000104 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.84 on 297 degrees of freedom
## Multiple R-squared:  0.3696, Adjusted R-squared:  0.3314
## F-statistic: 9.673 on 18 and 297 DF,  p-value: < 2.2e-16
```

```
AIC(interlm_tr1)
```

```
## [1] 1767.509
```

Schoolsup seems to have a high correlation with first\_gen\_college and studytime. Absences has a high correlation with failed. Both of these make sense intuitively. Sex and first\_Gen\_college seems to have a high correlation, which does not make much sense intuitively and should be explored further.

Including only active variables and paring down the model:

```
interlm_tr2 <- lm(G3 ~ (romantic + freetime + goout + failed + studytime + first_gen_college + famsup +
summary(interlm_tr2)
```

```
##
## Call:
## lm(formula = G3 ~ (romantic + freetime + goout + failed + studytime +
##   first_gen_college + famsup + famsize + absences * failed +
##   schoolsup * first_gen_college + schoolsup * studytime + sex *
##   first_gen_college), data = training)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.1510  -2.1029   0.4616   2.4331   8.6819
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    11.56650    1.00954   11.457 < 2e-16 ***
## romanticyes     -1.57486    0.48280   -3.262  0.001234 **
## freetimelow     -1.20937    0.47412   -2.551  0.011246 *
## gooutlow        1.67484    0.47822    3.502  0.000532 ***
## failedyes       -5.57431    0.70675   -7.887  5.80e-14 ***
## studytime        0.82586    0.30416    2.715  0.007008 **
## first_gen_collegeyes -3.32647    0.67292   -4.943  1.28e-06 ***
## famsupyes       -0.96589    0.47204   -2.046  0.041608 *
## famsizeLE3       1.19057    0.48812    2.439  0.015304 *
## absences         0.02018    0.03122    0.646  0.518453
## schoolsupyes     0.03990    1.96359    0.020  0.983804
## sexM            -1.06718    0.72433   -1.473  0.141707
## failedyes:absences  0.23837    0.05954    4.004  7.87e-05 ***
## first_gen_collegeyes:schoolsupyes  4.18746    1.35602    3.088  0.002203 **
## studytime:schoolsupyes -1.62659    0.75000   -2.169  0.030884 *
## first_gen_collegeyes:sexM  3.34855    0.92366    3.625  0.000339 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```



```
## Residual standard error: 3.857 on 300 degrees of freedom
## Multiple R-squared:  0.3576, Adjusted R-squared:  0.3255
## F-statistic: 11.14 on 15 and 300 DF,  p-value: < 2.2e-16
```

```
AIC(interlm_tr2)
```

```
## [1] 1767.442
```

AIC of 1767.442, Adjusted R-squared of 0.3255.

Using the model on the testing set:

```
pred.lm <- predict(interlm_tr2, testing)
pred.lm
```

```
##      11      17      18      21      31      32      34      43
## 12.717789 13.664748 10.446235 13.825867 12.859976 11.650605 13.829146 11.690972
##      46      58      59      60      72      76      78      80
##  9.281686 11.731340 13.877800 12.758157 14.268208 11.771707 10.434069 14.169365
##      89      94      95      97      98      115      116      119
## 11.376307 12.717789 12.961230  9.278423  6.181993 12.434563 11.225503  9.594648
##     120     130     133     149     150     152     155     160
## 13.121112 11.486639  7.957911  9.249891  6.963512  6.939987  8.730383  5.092277
##     174     175     178     180     187     194     196     200
##  1.533208 10.197155 11.044474 10.156482  9.312343 11.008302 13.299388 15.117727
##     224     226     227     229     231     235     239     240
## 11.672689  7.344933  8.884210 12.545147 13.581960 14.435933 10.397581  6.598803
##     243     244     250     254     260     264     269     273
## 12.515736 11.790639 11.812723 11.812723 12.834007 11.263805  9.812116 11.831628
##     279     281     282     283     284     288     292     300
##  5.768204 11.546389 11.876063 12.253786 10.946656 10.338282 13.543646 11.750747
##     304     315     322     328     329     336     340     341
## 11.043036  7.653679  8.058670 10.542833 13.219463 14.367006  9.472058  3.393286
##     347     358     369     378     385     386     389
## 12.049147 10.013290  6.990610 16.164352 11.045477 11.223437  8.907049
```

```
mse_test <- mean((pred.lm - testing$G3)^2)
testing$G3
```

```
## [1]  9 14 10 15 12 17 12 18  6 15  9 16 10 10 11  5 10 10 14 15 10  9 16  8 13
## [26] 18 12  0 10 14 12 12  0  9  6 11 11 10 15 10 13  7 15  9 14  6 11  0  0 12
## [51] 15  8  0  9 10 11  8  8 10 12 10 12 15 16 18 13  9 10  9 15 10 11 16 11 10
## [76] 10  5 10  8
```

```
mse_test
```

```
## [1] 18.86165
```

## Multicategory ordinal logit model

Due to the way grades are assigned as values between 0 and 20, we would like to consider G3 as an ordered categorical variable with 21 levels. This would allow us to fit a multicategory ordinal logistic model to the data.

We examine the EDA and active variables in the linear model to choose the predictors in our base model.

Fitting the base model:

```
require(foreign)
```

```

## Loading required package: foreign
require(nnet)
require(ggplot2)
require(reshape2)

## Loading required package: reshape2
##
## Attaching package: 'reshape2'
## The following object is masked from 'package:tidyr':
##
##     smiths
## The following object is masked from 'package:openintro':
##
##     tips
require(MASS)
require(Hmisc)

## Loading required package: Hmisc
## Loading required package: lattice
##
## Attaching package: 'lattice'
## The following objects are masked from 'package:openintro':
##
##     ethanol, lsegments
## Loading required package: survival
##
## Attaching package: 'survival'
## The following object is masked from 'package:openintro':
##
##     transplant
## Loading required package: Formula
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:dplyr':
##
##     src, summarize
## The following object is masked from 'package:rvest':
##
##     html
## The following objects are masked from 'package:base':
##
##     format.pval, units
mod <- polr(ord_g3 ~ . -G1 -G2 -G3, data = training)
summary(mod)

```

```
##
## Re-fitting to get Hessian

## Call:
## polr(formula = ord_g3 ~ . - G1 - G2 - G3, data = training)
##
## Coefficients:
##              Value Std. Error t value
## schoolMS      0.32605    0.38236  0.8527
## sexM          0.56589    0.24600  2.3004
## age          -0.28037    0.10765 -2.6045
## addressU      0.18757    0.28015  0.6695
## famsizeLE3    0.50097    0.23647  2.1185
## PstatusT     -0.32472    0.34653 -0.9371
## Medu         0.06932    0.18259  0.3796
## Fedu        -0.08955    0.15325 -0.5844
## Mjobhealth    0.60023    0.54994  1.0915
## Mjobother    -0.05752    0.35774 -0.1608
## Mjobservices  0.54672    0.39896  1.3704
## Mjobteacher  -0.61096    0.52074 -1.1732
## Fjobhealth   -0.40733    0.69420 -0.5868
## Fjobother    -0.17855    0.48037 -0.3717
## Fjobservices -0.05255    0.49757 -0.1056
## Fjobteacher  0.46618    0.66467  0.7014
## reasonhome   0.23177    0.26518  0.8740
## reasonother  0.25966    0.38080  0.6819
## reasonreputation 0.39833    0.28845  1.3809
## guardianmother 0.04580    0.26486  0.1729
## guardianother 0.46437    0.48249  0.9624
## traveltime   -0.14252    0.17642 -0.8079
## studytime    0.37461    0.14709  2.5467
## failures     -0.40709    0.28580 -1.4244
## schoolsupyes -1.07299    0.33629 -3.1907
## famsupyes    -0.49298    0.54604 -0.9028
## paidyes      0.20052    0.23526  0.8523
## activitiesyes -0.11540    0.21736 -0.5309
## nurseryyes   -0.17096    0.26950 -0.6343
## higheryes    -0.15919    0.49945 -0.3187
## internetyes  0.39537    0.44970  0.8792
## romanticyes  -0.50355    0.23510 -2.1419
## famrellow    0.05791    0.26647  0.2173
## freetimelow  -0.67456    0.23496 -2.8709
## gooutlow     0.77564    0.24852  3.1210
## Dalclow      0.07991    0.59539  0.1342
## Walclow      0.25638    0.32121  0.7982
## healthlow    0.29549    0.21720  1.3604
## absences     0.02002    0.02009  0.9966
## first_gen_collegeyes -0.70199    0.38869 -1.8060
## stable_learning_envyes -0.16982    0.59977 -0.2831
## high_freq_absentyes -0.16074    0.38422 -0.4184
## failedyes    -0.91443    0.54086 -1.6907
##
## Intercepts:
##              Value Std. Error t value
```

```
## 0|4 -7.1685 2.3463 -3.0552
## 4|5 -7.1297 2.3461 -3.0390
## 5|6 -6.9426 2.3448 -2.9609
## 6|7 -6.5614 2.3424 -2.8012
## 7|8 -6.3439 2.3413 -2.7096
## 8|9 -5.7195 2.3376 -2.4467
## 9|10 -5.3221 2.3345 -2.2798
## 10|11 -4.6194 2.3278 -1.9844
## 11|12 -3.9471 2.3222 -1.6997
## 12|13 -3.5499 2.3203 -1.5299
## 13|14 -3.0007 2.3201 -1.2933
## 14|15 -2.4425 2.3202 -1.0527
## 15|16 -1.6383 2.3195 -0.7063
## 16|17 -1.0749 2.3222 -0.4629
## 17|18 -0.7489 2.3251 -0.3221
## 18|19 0.2523 2.3411 0.1078
## 19|20 2.1350 2.5097 0.8507
##
## Residual Deviance: 1513.116
## AIC: 1633.116
```

```
acc.ord <- predict(mod, training)
ctable <- table(training$G3, acc.ord)
round((sum(diag(ctable))/sum(ctable))*100,2)
```

```
## [1] 20.89
```

```
ctable
```

```
##      acc.ord
##      0  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
## 0  18  0  0  0  0  0  0 10  5  0  0  0  0  0  0  0  0
## 4   0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0
## 5   3  0  0  0  0  0  0  2  0  0  0  0  0  0  0  0  0
## 6   1  0  0  0  0  0  0  9  1  0  0  0  1  0  0  0  0
## 7   4  0  0  0  0  0  0  1  3  0  0  0  0  0  0  0  0
## 8  11  0  0  0  0  0  0  8  7  0  0  0  1  0  0  0  0
## 9   6  0  0  0  0  0  0  8  6  0  0  0  0  0  0  0  0
## 10  8  0  0  0  0  0  0 14 12  0  0  0  5  0  0  0  0
## 11  2  0  0  0  0  0  0 12 20  0  1  0  5  0  0  0  0
## 12  3  0  0  0  0  0  0  3 12  0  0  0  5  0  0  0  0
## 13  5  0  0  0  0  0  0  2 17  0  0  0  4  0  0  0  0
## 14  1  0  0  0  0  0  0  2 12  0  0  0  8  0  0  0  0
## 15  0  0  0  0  0  0  0  3  7  0  1  0 14  0  0  0  0
## 16  0  0  0  0  0  0  0  1  4  0  0  0  7  0  0  0  0
## 17  0  0  0  0  0  0  0  0  3  0  0  0  2  0  0  0  0
## 18  0  0  0  0  0  0  0  1  4  0  0  0  4  0  0  0  0
## 19  0  0  0  0  0  0  0  0  1  0  0  0  4  0  0  0  0
## 20  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0
```

```
mod1 <- polr(ord_g3 ~ failed + high_freq_absent + romantic + internet + goout + first_gen_college + Wal
summary(mod1)
```

```
##
## Re-fitting to get Hessian
## Call:
```

```
## polr(formula = ord_g3 ~ failed + high_freq_absent + romantic +
##       internet + goout + first_gen_college + Walc + sex + schoolsup +
##       famsup + absences + studytime + higher, data = training)
##
```

```
## Coefficients:
```

	Value	Std. Error	t value
failedyes	-1.34171	0.27018	-4.9659
high_freq_absentyes	0.03565	0.34970	0.1019
romanticyes	-0.45766	0.22322	-2.0502
internetyes	0.36012	0.27651	1.3024
gooutlow	0.58461	0.22879	2.5552
first_gen_collegeyes	-0.66042	0.22335	-2.9568
Walclow	0.33622	0.27525	1.2215
sexM	0.55216	0.22206	2.4865
schoolsupyes	-0.69882	0.28905	-2.4177
famsupyes	-0.48594	0.21594	-2.2503
absences	0.01048	0.01726	0.6071
studytime	0.25616	0.13427	1.9078
higheryes	0.37412	0.45201	0.8277

```
##
```

```
## Intercepts:
```

	Value	Std. Error	t value
0 4	-1.6658	0.6572	-2.5347
4 5	-1.6293	0.6566	-2.4816
5 6	-1.4556	0.6534	-2.2276
6 7	-1.1027	0.6485	-1.7004
7 8	-0.9014	0.6467	-1.3939
8 9	-0.3219	0.6437	-0.5000
9 10	0.0449	0.6427	0.0698
10 11	0.6847	0.6438	1.0635
11 12	1.2954	0.6464	2.0040
12 13	1.6591	0.6479	2.5610
13 14	2.1640	0.6516	3.3212
14 15	2.6641	0.6579	4.0496
15 16	3.3940	0.6711	5.0577
16 17	3.9280	0.6858	5.7276
17 18	4.2419	0.6982	6.0754
18 19	5.2132	0.7670	6.7971
19 20	7.0544	1.1934	5.9112

```
##
```

```
## Residual Deviance: 1562.392
```

```
## AIC: 1622.392
```

```
(ctable <- coef(summary(mod1)))
```

```
##
```

```
## Re-fitting to get Hessian
```

	Value	Std. Error	t value
failedyes	-1.34171397	0.27018479	-4.96591237
high_freq_absentyes	0.03565031	0.34969709	0.10194625
romanticyes	-0.45766071	0.22322374	-2.05023314
internetyes	0.36011512	0.27651008	1.30235803
gooutlow	0.58461188	0.22879147	2.55521711
first_gen_collegeyes	-0.66042320	0.22335452	-2.95683831

## Walcflow	0.33622440	0.27525303	1.22151026
## sexM	0.55215988	0.22206496	2.48647905
## schoolsupyes	-0.69881781	0.28904608	-2.41766920
## famsupyes	-0.48593653	0.21594151	-2.25031553
## absences	0.01047719	0.01725748	0.60710999
## studytime	0.25615867	0.13426713	1.90782854
## higheryes	0.37411910	0.45201352	0.82767238
## 0 4	-1.66581605	0.65721302	-2.53466683
## 4 5	-1.62930927	0.65655494	-2.48160386
## 5 6	-1.45558081	0.65342826	-2.22760613
## 6 7	-1.10268571	0.64846971	-1.70044290
## 7 8	-0.90144383	0.64670148	-1.39391027
## 8 9	-0.32186448	0.64369588	-0.50002569
## 9 10	0.04485165	0.64272846	0.06978319
## 10 11	0.68472806	0.64384467	1.06349884
## 11 12	1.29538659	0.64638677	2.00404255
## 12 13	1.65914607	0.64785012	2.56100295
## 13 14	2.16402771	0.65158608	3.32116935
## 14 15	2.66414639	0.65787988	4.04959395
## 15 16	3.39401724	0.67106437	5.05766269
## 16 17	3.92801161	0.68580175	5.72761970
## 17 18	4.24194365	0.69821069	6.07544927
## 18 19	5.21317802	0.76696900	6.79711702
## 19 20	7.05442354	1.19339185	5.91123825

Calculate and store p-values:

```
p1 <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
(ctable <- cbind(ctable, "p value" = p1))
```

##	Value	Std. Error	t value	p value
## failedyes	-1.34171397	0.27018479	-4.96591237	6.837883e-07
## high_freq_absentyes	0.03565031	0.34969709	0.10194625	9.187993e-01
## romanticyes	-0.45766071	0.22322374	-2.05023314	4.034169e-02
## internetyes	0.36011512	0.27651008	1.30235803	1.927940e-01
## gooutlow	0.58461188	0.22879147	2.55521711	1.061216e-02
## first_gen_collegeyes	-0.66042320	0.22335452	-2.95683831	3.108111e-03
## Walcflow	0.33622440	0.27525303	1.22151026	2.218929e-01
## sexM	0.55215988	0.22206496	2.48647905	1.290142e-02
## schoolsupyes	-0.69881781	0.28904608	-2.41766920	1.562027e-02
## famsupyes	-0.48593653	0.21594151	-2.25031553	2.442892e-02
## absences	0.01047719	0.01725748	0.60710999	5.437779e-01
## studytime	0.25615867	0.13426713	1.90782854	5.641338e-02
## higheryes	0.37411910	0.45201352	0.82767238	4.078561e-01
## 0 4	-1.66581605	0.65721302	-2.53466683	1.125543e-02
## 4 5	-1.62930927	0.65655494	-2.48160386	1.307926e-02
## 5 6	-1.45558081	0.65342826	-2.22760613	2.590679e-02
## 6 7	-1.10268571	0.64846971	-1.70044290	8.904765e-02
## 7 8	-0.90144383	0.64670148	-1.39391027	1.633447e-01
## 8 9	-0.32186448	0.64369588	-0.50002569	6.170570e-01
## 9 10	0.04485165	0.64272846	0.06978319	9.443662e-01
## 10 11	0.68472806	0.64384467	1.06349884	2.875558e-01
## 11 12	1.29538659	0.64638677	2.00404255	4.506550e-02
## 12 13	1.65914607	0.64785012	2.56100295	1.043705e-02
## 13 14	2.16402771	0.65158608	3.32116935	8.964113e-04

```
## 14|15          2.66414639 0.65787988 4.04959395 5.130658e-05
## 15|16          3.39401724 0.67106437 5.05766269 4.244263e-07
## 16|17          3.92801161 0.68580175 5.72761970 1.018496e-08
## 17|18          4.24194365 0.69821069 6.07544927 1.236411e-09
## 18|19          5.21317802 0.76696900 6.79711702 1.067334e-11
## 19|20          7.05442354 1.19339185 5.91123825 3.395455e-09
```

Confidence intervals for parameter estimates:

```
(ci1 <- confint(mod1))
```

```
## Waiting for profiling to be done...
```

```
##
```

```
## Re-fitting to get Hessian
```

```
##              2.5 %      97.5 %
## failedyes      -1.876197497 -0.81556618
## high_freq_absentyes -0.656411452 0.71764477
## romanticyes     -0.897663611 -0.02188466
## internetyes     -0.180649012 0.90497042
## gooutlow        0.137337761 1.03514595
## first_gen_collegeyes -1.100400976 -0.22416437
## Walclow         -0.203641426 0.87718770
## sexM            0.118264255 0.98945878
## schoolsupyes    -1.268171885 -0.13267874
## famsupyes       -0.910849856 -0.06369438
## absences        -0.021851056 0.04656485
## studytime       -0.006577242 0.52033547
## higheryes       -0.505211121 1.27490555
```

Analyzing the p-values and confidence intervals allows us to determine whether the coefficient estimates are significant. Based on these, failed, romantic, goout, first\_gen\_college, sex, schoolsup, famsup, studytime seem to be active. (Studytime is dubious, but we will include it in the next model)

Refitting a model with these predictors:

```
mod2 <- polr(ord_g3 ~ failed + romantic + goout + first_gen_college + studytime + sex + schoolsup + famsup,
summary(mod2))
```

```
##
```

```
## Re-fitting to get Hessian
```

```
## Call:
```

```
## polr(formula = ord_g3 ~ failed + romantic + goout + first_gen_college +
##       studytime + sex + schoolsup + famsup, data = training)
```

```
##
```

```
## Coefficients:
```

```
##              Value Std. Error t value
## failedyes      -1.3882    0.2606  -5.327
## romanticyes     -0.4157    0.2160  -1.924
## gooutlow        0.6874    0.2119   3.245
## first_gen_collegeyes -0.6953    0.2189  -3.177
## studytime       0.2749    0.1333   2.062
## sexM            0.4689    0.2127   2.205
## schoolsupyes    -0.6642    0.2881  -2.305
## famsupyes       -0.4203    0.2133  -1.970
##
```

```
## Intercepts:
##      Value Std. Error t value
## 0|4 -2.5429 0.4453 -5.7101
## 4|5 -2.5068 0.4443 -5.6425
## 5|6 -2.3344 0.4396 -5.3103
## 6|7 -1.9834 0.4324 -4.5866
## 7|8 -1.7842 0.4294 -4.1546
## 8|9 -1.2081 0.4228 -2.8573
## 9|10 -0.8412 0.4200 -2.0027
## 10|11 -0.2072 0.4175 -0.4963
## 11|12 0.3981 0.4182 0.9519
## 12|13 0.7604 0.4192 1.8142
## 13|14 1.2610 0.4225 2.9847
## 14|15 1.7560 0.4295 4.0886
## 15|16 2.4801 0.4464 5.5558
## 16|17 3.0130 0.4678 6.4414
## 17|18 3.3261 0.4857 6.8484
## 18|19 4.2924 0.5793 7.4098
## 19|20 6.1276 1.0817 5.6650
##
## Residual Deviance: 1567.543
## AIC: 1617.543
```

```
(ctable <- coef(summary(mod2)))
```

```
##
## Re-fitting to get Hessian

##      Value Std. Error t value
## failedyes -1.3881639 0.2605997 -5.3268047
## romanticyes -0.4156752 0.2160358 -1.9241039
## gooutlow 0.6873737 0.2118547 3.2445526
## first_gen_collegeyes -0.6953188 0.2188570 -3.1770460
## studytime 0.2748992 0.1332997 2.0622646
## sexM 0.4689357 0.2126871 2.2048153
## schoolsupyes -0.6642084 0.2881418 -2.3051443
## famsupyes -0.4203058 0.2133383 -1.9701374
## 0|4 -2.5429074 0.4453387 -5.7100521
## 4|5 -2.5067847 0.4442706 -5.6424726
## 5|6 -2.3344286 0.4396061 -5.3102731
## 6|7 -1.9834475 0.4324423 -4.5866175
## 7|8 -1.7841536 0.4294407 -4.1545983
## 8|9 -1.2081500 0.4228275 -2.8573116
## 9|10 -0.8412191 0.4200487 -2.0026706
## 10|11 -0.2072168 0.4175067 -0.4963196
## 11|12 0.3980517 0.4181620 0.9519078
## 12|13 0.7604341 0.4191666 1.8141570
## 13|14 1.2609616 0.4224815 2.9846550
## 14|15 1.7559528 0.4294754 4.0885992
## 15|16 2.4801341 0.4464006 5.5558481
## 16|17 3.0129643 0.4677514 6.4413788
## 17|18 3.3260960 0.4856779 6.8483581
## 18|19 4.2923566 0.5792787 7.4098294
## 19|20 6.1275584 1.0816527 5.6649962
```



```
p2 <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
(ctable <- cbind(ctable, "p value" = p2))
```

	Value	Std. Error	t value	p value
## failedyes	-1.3881639	0.2605997	-5.3268047	9.995554e-08
## romanticyes	-0.4156752	0.2160358	-1.9241039	5.434156e-02
## gooutlow	0.6873737	0.2118547	3.2445526	1.176353e-03
## first_gen_collegeyes	-0.6953188	0.2188570	-3.1770460	1.487835e-03
## studytime	0.2748992	0.1332997	2.0622646	3.918255e-02
## sexM	0.4689357	0.2126871	2.2048153	2.746706e-02
## schoolsupyes	-0.6642084	0.2881418	-2.3051443	2.115849e-02
## famsupyes	-0.4203058	0.2133383	-1.9701374	4.882263e-02
## 0 4	-2.5429074	0.4453387	-5.7100521	1.129416e-08
## 4 5	-2.5067847	0.4442706	-5.6424726	1.676252e-08
## 5 6	-2.3344286	0.4396061	-5.3102731	1.094611e-07
## 6 7	-1.9834475	0.4324423	-4.5866175	4.504849e-06
## 7 8	-1.7841536	0.4294407	-4.1545983	3.258595e-05
## 8 9	-1.2081500	0.4228275	-2.8573116	4.272462e-03
## 9 10	-0.8412191	0.4200487	-2.0026706	4.521266e-02
## 10 11	-0.2072168	0.4175067	-0.4963196	6.196689e-01
## 11 12	0.3980517	0.4181620	0.9519078	3.411438e-01
## 12 13	0.7604341	0.4191666	1.8141570	6.965355e-02
## 13 14	1.2609616	0.4224815	2.9846550	2.838983e-03
## 14 15	1.7559528	0.4294754	4.0885992	4.339860e-05
## 15 16	2.4801341	0.4464006	5.5558481	2.762670e-08
## 16 17	3.0129643	0.4677514	6.4413788	1.183930e-10
## 17 18	3.3260960	0.4856779	6.8483581	7.470236e-12
## 18 19	4.2923566	0.5792787	7.4098294	1.264620e-13
## 19 20	6.1275584	1.0816527	5.6649962	1.470278e-08

```
(ci2 <- confint(mod2))
```

```
## Waiting for profiling to be done...
```

```
##
```

```
## Re-fitting to get Hessian
```

	2.5 %	97.5 %
## failedyes	-1.90462685	-0.881708189
## romanticyes	-0.84097272	0.006566793
## gooutlow	0.27365841	1.104747592
## first_gen_collegeyes	-1.12660028	-0.268012697
## studytime	0.01411035	0.537164395
## sexM	0.05297154	0.887306542
## schoolsupyes	-1.23199282	-0.100060789
## famsupyes	-0.83978194	-0.002875816

AIC has decreased.

Based on the p-values and confidence intervals, romantic does not seem to be significant. Let's try excluding it.

Pared-down model again:

```
mod3 <- polr(ord_g3 ~ failed + goout + first_gen_college + sex + schoolsup + studytime, data = training)
summary(mod3)
```

```
## Call:
```

```
## polr(formula = ord_g3 ~ failed + goout + first_gen_college +
##       sex + schoolsup + studytime, data = training, Hess = TRUE)
##
## Coefficients:
##               Value Std. Error t value
## failedyes      -1.4470    0.2594  -5.577
## gooutlow        0.6862    0.2115   3.244
## first_gen_collegeyes -0.5623    0.2119  -2.654
## sexM            0.5365    0.2106   2.547
## schoolsupyes    -0.6138    0.2822  -2.175
## studytime       0.2189    0.1311   1.670
##
## Intercepts:
##      Value Std. Error t value
## 0|4    -2.1140  0.4116   -5.1354
## 4|5    -2.0782  0.4106   -5.0620
## 5|6    -1.9083  0.4060   -4.7005
## 6|7    -1.5631  0.3995   -3.9125
## 7|8    -1.3668  0.3969   -3.4435
## 8|9    -0.7982  0.3913   -2.0401
## 9|10   -0.4358  0.3892   -1.1199
## 10|11   0.1921  0.3879    0.4952
## 11|12   0.7943  0.3898    2.0379
## 12|13   1.1538  0.3918    2.9448
## 13|14   1.6485  0.3967    4.1551
## 14|15   2.1387  0.4054    5.2759
## 15|16   2.8588  0.4247    6.7305
## 16|17   3.3897  0.4478    7.5696
## 17|18   3.7015  0.4667    7.9310
## 18|19   4.6606  0.5642    8.2598
## 19|20   6.4828  1.0738    6.0371
##
## Residual Deviance: 1574.549
## AIC: 1620.549
```

```
(ctable <- coef(summary(mod3)))
```

```
##               Value Std. Error    t value
## failedyes      -1.4470044  0.2594410 -5.5773921
## gooutlow        0.6862095  0.2115274  3.2440692
## first_gen_collegeyes -0.5623425  0.2119141 -2.6536341
## sexM            0.5364682  0.2106266  2.5470106
## schoolsupyes    -0.6138372  0.2821635 -2.1754659
## studytime       0.2188984  0.1310851  1.6698957
## 0|4             -2.1139669  0.4116495 -5.1353568
## 4|5             -2.0782289  0.4105540 -5.0620109
## 5|6             -1.9083008  0.4059759 -4.7005269
## 6|7             -1.5630672  0.3995011 -3.9125480
## 7|8             -1.3667896  0.3969204 -3.4434854
## 8|9             -0.7982483  0.3912773 -2.0401089
## 9|10            -0.4358495  0.3891718 -1.1199410
## 10|11            0.1920664  0.3878675  0.4951856
## 11|12            0.7942916  0.3897682  2.0378563
## 12|13            1.1537589  0.3917941  2.9448095
## 13|14            1.6485090  0.3967448  4.1550867
```

```
## 14|15          2.1387107  0.4053722  5.2759186
## 15|16          2.8587993  0.4247500  6.7305461
## 16|17          3.3897018  0.4478022  7.5696401
## 17|18          3.7014831  0.4667133  7.9309568
## 18|19          4.6605957  0.5642498  8.2598092
## 19|20          6.4827907  1.0738176  6.0371431
```

```
p3 <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
(ctable <- cbind(ctable, "p value" = p3))
```

```
##              Value Std. Error   t value    p value
## failedyes      -1.4470044  0.2594410 -5.5773921 2.441512e-08
## gooutlow         0.6862095  0.2115274  3.2440692 1.178351e-03
## first_gen_collegeyes -0.5623425  0.2119141 -2.6536341 7.963013e-03
## sexM             0.5364682  0.2106266  2.5470106 1.086501e-02
## schoolsupyes     -0.6138372  0.2821635 -2.1754659 2.959522e-02
## studytime        0.2188984  0.1310851  1.6698957 9.494000e-02
## 0|4             -2.1139669  0.4116495 -5.1353568 2.816093e-07
## 4|5             -2.0782289  0.4105540 -5.0620109 4.148574e-07
## 5|6             -1.9083008  0.4059759 -4.7005269 2.594911e-06
## 6|7             -1.5630672  0.3995011 -3.9125480 9.132736e-05
## 7|8             -1.3667896  0.3969204 -3.4434854 5.742675e-04
## 8|9             -0.7982483  0.3912773 -2.0401089 4.133948e-02
## 9|10            -0.4358495  0.3891718 -1.1199410 2.627389e-01
## 10|11           0.1920664  0.3878675  0.4951856 6.204691e-01
## 11|12           0.7942916  0.3897682  2.0378563 4.156431e-02
## 12|13           1.1537589  0.3917941  2.9448095 3.231536e-03
## 13|14           1.6485090  0.3967448  4.1550867 3.251642e-05
## 14|15           2.1387107  0.4053722  5.2759186 1.320927e-07
## 15|16           2.8587993  0.4247500  6.7305461 1.690275e-11
## 16|17           3.3897018  0.4478022  7.5696401 3.742596e-14
## 17|18           3.7014831  0.4667133  7.9309568 2.174638e-15
## 18|19           4.6605957  0.5642498  8.2598092 1.459055e-16
## 19|20           6.4827907  1.0738176  6.0371431 1.568666e-09
```

```
(ci3 <- confint(mod3))
```

```
## Waiting for profiling to be done...
```

```
##              2.5 %      97.5 %
## failedyes      -1.96148845 -0.94318819
## gooutlow         0.27307924  1.10288470
## first_gen_collegeyes -0.97950820 -0.14817187
## sexM             0.12485281  0.95095558
## schoolsupyes     -1.16913473 -0.06070744
## studytime       -0.03766773  0.47670960
```

All predictors are significant, but AIC has increased compared to mod2.

Evaluating accuracy of the model for the training set:

```
acc.ord3 <- predict(mod3, training)
ctable <- table(training$G3, acc.ord3)
round((sum(diag(ctable))/sum(ctable))*100,2)
```

```
## [1] 17.09
```

ctable

```
##      acc.ord3
##      0  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
## 0 16 0 0 0 0 0 0 7 7 0 0 0 3 0 0 0 0
## 4 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## 5 2 0 0 0 0 0 0 3 0 0 0 0 0 0 0 0 0
## 6 1 0 0 0 0 0 0 6 5 0 0 0 0 0 0 0 0
## 7 5 0 0 0 0 0 0 0 1 0 0 0 2 0 0 0 0
## 8 12 0 0 0 0 0 0 4 9 0 0 0 2 0 0 0 0
## 9 6 0 0 0 0 0 0 4 9 0 0 0 1 0 0 0 0
## 10 8 0 0 0 0 0 0 6 20 0 0 0 5 0 0 0 0
## 11 4 0 0 0 0 0 0 8 23 0 0 0 5 0 0 0 0
## 12 3 0 0 0 0 0 0 5 11 0 0 0 4 0 0 0 0
## 13 5 0 0 0 0 0 0 3 17 0 0 0 3 0 0 0 0
## 14 1 0 0 0 0 0 0 3 8 0 0 0 11 0 0 0 0
## 15 0 0 0 0 0 0 0 2 14 0 0 0 9 0 0 0 0
## 16 0 0 0 0 0 0 0 2 6 0 0 0 4 0 0 0 0
## 17 0 0 0 0 0 0 0 1 3 0 0 0 1 0 0 0 0
## 18 1 0 0 0 0 0 0 1 5 0 0 0 2 0 0 0 0
## 19 0 0 0 0 0 0 0 0 3 0 0 0 2 0 0 0 0
## 20 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
```

Very terrible accuracy even for the training set.

What if we add interaction terms?

Let's base our interaction terms on the discussion for the linear model.

```
mod4 <- polr(ord_g3 ~ failed + goout + romantic + first_gen_college + sex + schoolsup + sex*schoolsup +
summary(mod4)
```

```
##
## Re-fitting to get Hessian

## Call:
## polr(formula = ord_g3 ~ failed + goout + romantic + first_gen_college +
##      sex + schoolsup + sex * schoolsup + sex * first_gen_college +
##      schoolsup * failed + schoolsup * studytime + schoolsup *
##      first_gen_college + studytime * famsup, data = training)
##
## Coefficients:
##                                     Value Std. Error t value
## failedyes                        -1.72945     0.2981 -5.80124
## gooutlow                         0.67401     0.2169  3.10806
## romanticyes                     -0.48910     0.2204 -2.21930
## first_gen_collegeyes            -1.43664     0.3180 -4.51738
## sexM                           -0.01446     0.3391 -0.04265
## schoolsupyes                     0.73242     0.9305  0.78712
## studytime                       0.36175     0.2487  1.45475
## famsupyes                      -0.76421     0.5894 -1.29664
## sexM:schoolsupyes               -1.06747     0.6217 -1.71701
## first_gen_collegeyes:sexM       1.06194     0.4232  2.50904
## failedyes:schoolsupyes          1.15774     0.6593  1.75607
## schoolsupyes:studytime          -1.11414     0.3374 -3.30178
## first_gen_collegeyes:schoolsupyes 1.59232     0.6093  2.61341
## studytime:famsupyes            0.14783     0.2850  0.51876
```

```
##
## Intercepts:
##      Value   Std. Error t value
## 0|4   -3.0101   0.5949   -5.0599
## 4|5   -2.9717   0.5939   -5.0034
## 5|6   -2.7903   0.5898   -4.7309
## 6|7   -2.4230   0.5833   -4.1537
## 7|8   -2.2159   0.5806   -3.8163
## 8|9   -1.6123   0.5757   -2.8006
## 9|10  -1.2173   0.5741   -2.1202
## 10|11 -0.5423   0.5738   -0.9451
## 11|12  0.0965   0.5744    0.1680
## 12|13  0.4842   0.5739    0.8438
## 13|14  1.0192   0.5741    1.7752
## 14|15  1.5345   0.5779    2.6551
## 15|16  2.2702   0.5892    3.8531
## 16|17  2.8082   0.6045    4.6455
## 17|18  3.1241   0.6180    5.0554
## 18|19  4.0937   0.6937    5.9014
## 19|20  5.9292   1.1498    5.1569
##
## Residual Deviance: 1537.762
## AIC: 1599.762
```

```
(cetable <- coef(summary(mod4)))
```

```
##
## Re-fitting to get Hessian
##
##      Value Std. Error   t value
## failedyes -1.72945368  0.2981178 -5.80124295
## gooutlow  0.67400726  0.2168580  3.10805786
## romanticyes -0.48909643  0.2203829 -2.21930313
## first_gen_collegeyes -1.43663982  0.3180252 -4.51737778
## sexM -0.01446239  0.3390652 -0.04265371
## schoolsupyes 0.73241720  0.9305008  0.78712151
## studytime 0.36174519  0.2486653  1.45474731
## famsupyes -0.76421191  0.5893770 -1.29664348
## sexM:schoolsupyes -1.06746616  0.6217004 -1.71701055
## first_gen_collegeyes:sexM 1.06194139  0.4232461  2.50904014
## failedyes:schoolsupyes 1.15773926  0.6592797  1.75606684
## schoolsupyes:studytime -1.11413949  0.3374358 -3.30178248
## first_gen_collegeyes:schoolsupyes 1.59232262  0.6092889  2.61341153
## studytime:famsupyes 0.14782789  0.2849627  0.51876227
## 0|4 -3.01008158  0.5948850 -5.05993887
## 4|5 -2.97166463  0.5939322 -5.00337322
## 5|6 -2.79029263  0.5898075 -4.73085276
## 6|7 -2.42295680  0.5833207 -4.15373034
## 7|8 -2.21585268  0.5806252 -3.81632165
## 8|9 -1.61234628  0.5757246 -2.80055115
## 9|10 -1.21732081  0.5741468 -2.12022571
## 10|11 -0.54231022  0.5738069 -0.94510930
## 11|12 0.09649365  0.5743725  0.16799839
## 12|13 0.48423331  0.5738879  0.84377688
## 13|14 1.01921806  0.5741418  1.77520259
```

```
## 14|15          1.53449674  0.5779426  2.65510251
## 15|16          2.27018077  0.5891862  3.85307839
## 16|17          2.80819116  0.6045010  4.64546978
## 17|18          3.12405658  0.6179583  5.05544881
## 18|19          4.09365582  0.6936810  5.90135194
## 19|20          5.92920227  1.1497705  5.15685736
```

```
p4 <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
(ctable <- cbind(ctable, "p value" = p4))
```

```
##              Value Std. Error    t value
## failedyes      -1.72945368  0.2981178 -5.80124295
## gooutlow        0.67400726  0.2168580  3.10805786
## romanticyes     -0.48909643  0.2203829 -2.21930313
## first_gen_collegeyes -1.43663982  0.3180252 -4.51737778
## sexM            -0.01446239  0.3390652 -0.04265371
## schoolsupyes      0.73241720  0.9305008  0.78712151
## studytime        0.36174519  0.2486653  1.45474731
## famsupyes       -0.76421191  0.5893770 -1.29664348
## sexM:schoolsupyes -1.06746616  0.6217004 -1.71701055
## first_gen_collegeyes:sexM  1.06194139  0.4232461  2.50904014
## failedyes:schoolsupyes  1.15773926  0.6592797  1.75606684
## schoolsupyes:studytime -1.11413949  0.3374358 -3.30178248
## first_gen_collegeyes:schoolsupyes  1.59232262  0.6092889  2.61341153
## studytime:famsupyes    0.14782789  0.2849627  0.51876227
## 0|4             -3.01008158  0.5948850 -5.05993887
## 4|5             -2.97166463  0.5939322 -5.00337322
## 5|6             -2.79029263  0.5898075 -4.73085276
## 6|7             -2.42295680  0.5833207 -4.15373034
## 7|8             -2.21585268  0.5806252 -3.81632165
## 8|9             -1.61234628  0.5757246 -2.80055115
## 9|10            -1.21732081  0.5741468 -2.12022571
## 10|11           -0.54231022  0.5738069 -0.94510930
## 11|12           0.09649365  0.5743725  0.16799839
## 12|13           0.48423331  0.5738879  0.84377688
## 13|14           1.01921806  0.5741418  1.77520259
## 14|15           1.53449674  0.5779426  2.65510251
## 15|16           2.27018077  0.5891862  3.85307839
## 16|17           2.80819116  0.6045010  4.64546978
## 17|18           3.12405658  0.6179583  5.05544881
## 18|19           4.09365582  0.6936810  5.90135194
## 19|20           5.92920227  1.1497705  5.15685736
##              p value
## failedyes      6.582515e-09
## gooutlow       1.883212e-03
## romanticyes    2.646611e-02
## first_gen_collegeyes  6.261014e-06
## sexM           9.659776e-01
## schoolsupyes    4.312107e-01
## studytime      1.457392e-01
## famsupyes      1.947539e-01
## sexM:schoolsupyes  8.597724e-02
## first_gen_collegeyes:sexM  1.210597e-02
## failedyes:schoolsupyes  7.907700e-02
## schoolsupyes:studytime  9.607254e-04
```

```
## first_gen_collegeyes:schoolsupyes 8.964329e-03
## studytime:famsupyes 6.039265e-01
## 0|4 4.193909e-07
## 4|5 5.633572e-07
## 5|6 2.235787e-06
## 6|7 3.270986e-05
## 7|8 1.354559e-04
## 8|9 5.101542e-03
## 9|10 3.398702e-02
## 10|11 3.446031e-01
## 11|12 8.665845e-01
## 12|13 3.987941e-01
## 13|14 7.586444e-02
## 14|15 7.928431e-03
## 15|16 1.166420e-04
## 16|17 3.393034e-06
## 17|18 4.293798e-07
## 18|19 3.605349e-09
## 19|20 2.511290e-07
```

```
(ci4 <- confint(mod4))
```

```
## Waiting for profiling to be done...
```

```
##
```

```
## Re-fitting to get Hessian
```

```
##                2.5 %      97.5 %
## failedyes      -2.3214364 -1.15123396
## gooutlow        0.2503920  1.10114465
## romanticyes     -0.9230151 -0.05837766
## first_gen_collegeyes -2.0646870 -0.81670972
## sexM            -0.6799620  0.65071516
## schoolsupyes    -1.0998497  2.56413384
## studytime       -0.1227550  0.85203523
## famsupyes       -1.9233145  0.38932033
## sexM:schoolsupyes -2.2962304  0.15377624
## first_gen_collegeyes:sexM 0.2363092  1.89633281
## failedyes:schoolsupyes -0.1337302  2.46425196
## schoolsupyes:studytime -1.7820611 -0.45214925
## first_gen_collegeyes:schoolsupyes 0.4009686  2.79629741
## studytime:famsupyes -0.4093914  0.70866271
```

AIC has decreased significantly compared to the previous models without interaction terms, by nearly 20. However, in this model, sex, its interaction with schoolsup, and its interaction with first\_gen\_college all seem to be insignificant. The interaction between studytime and famsup and failed and schoolsup do not seem significant either, so let us remove it to pare down the model:

```
mod5 <- polr(ord_g3 ~ failed + goout + romantic + schoolsup + first_gen_college + schoolsup * studytime
summary(mod5))
```

```
##
```

```
## Re-fitting to get Hessian
```

```
## Call:
```

```
## polr(formula = ord_g3 ~ failed + goout + romantic + schoolsup +
##       first_gen_college + schoolsup * studytime + schoolsup * first_gen_college,
##       data = training)
```

```
##
## Coefficients:
##
## Value Std. Error t value
## failedyes -1.39146 0.2622 -5.30667
## gooutlow 0.59649 0.2134 2.79476
## romanticyes -0.50391 0.2163 -2.32943
## schoolsupyes 0.03364 0.8300 0.04052
## first_gen_collegeyes -0.85007 0.2317 -3.66947
## studytime 0.29305 0.1393 2.10388
## schoolsupyes:studytime -0.85318 0.3244 -2.62985
## schoolsupyes:first_gen_collegeyes 1.52426 0.5720 2.66499
##
## Intercepts:
## Value Std. Error t value
## 0|4 -2.6921 0.4042 -6.6607
## 4|5 -2.6558 0.4029 -6.5914
## 5|6 -2.4839 0.3975 -6.2485
## 6|7 -2.1376 0.3891 -5.4942
## 7|8 -1.9409 0.3852 -5.0386
## 8|9 -1.3590 0.3772 -3.6032
## 9|10 -0.9824 0.3739 -2.6277
## 10|11 -0.3366 0.3707 -0.9081
## 11|12 0.2744 0.3708 0.7400
## 12|13 0.6418 0.3719 1.7259
## 13|14 1.1514 0.3752 3.0688
## 14|15 1.6489 0.3824 4.3124
## 15|16 2.3707 0.4009 5.9140
## 16|17 2.9048 0.4245 6.8436
## 17|18 3.2188 0.4440 7.2496
## 18|19 4.1810 0.5450 7.6720
## 19|20 6.0069 1.0636 5.6479
##
## Residual Deviance: 1563.245
## AIC: 1613.245

(ctable <- coef(summary(mod5)))

##
## Re-fitting to get Hessian
##
## Value Std. Error t value
## failedyes -1.39146287 0.2622101 -5.30667226
## gooutlow 0.59648735 0.2134306 2.79475999
## romanticyes -0.50390805 0.2163229 -2.32942552
## schoolsupyes 0.03363511 0.8300358 0.04052249
## first_gen_collegeyes -0.85006782 0.2316595 -3.66947101
## studytime 0.29304592 0.1392885 2.10387705
## schoolsupyes:studytime -0.85318089 0.3244225 -2.62984500
## schoolsupyes:first_gen_collegeyes 1.52425650 0.5719558 2.66498990
## 0|4 -2.69208825 0.4041720 -6.66074948
## 4|5 -2.65581637 0.4029201 -6.59142141
## 5|6 -2.48392676 0.3975209 -6.24854363
## 6|7 -2.13758424 0.3890606 -5.49421935
## 7|8 -1.94093740 0.3852119 -5.03862266
## 8|9 -1.35895482 0.3771508 -3.60321358
```



```
## 9|10 -0.98243182 0.3738798 -2.62766733
## 10|11 -0.33664864 0.3707238 -0.90808488
## 11|12 0.27441039 0.3708471 0.73995555
## 12|13 0.64181542 0.3718808 1.72586325
## 13|14 1.15140976 0.3751961 3.06882090
## 14|15 1.64894914 0.3823753 4.31238440
## 15|16 2.37070007 0.4008641 5.91397458
## 16|17 2.90477643 0.4244524 6.84358578
## 17|18 3.21877485 0.4439939 7.24959282
## 18|19 4.18096474 0.5449623 7.67202626
## 19|20 6.00685744 1.0635640 5.64785684
```

```
p5 <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
(ctable <- cbind(ctable, "p value" = p5))
```

```
## Value Std. Error t value
## failedyes -1.39146287 0.2622101 -5.30667226
## gooutlow 0.59648735 0.2134306 2.79475999
## romanticyes -0.50390805 0.2163229 -2.32942552
## schoolsupyes 0.03363511 0.8300358 0.04052249
## first_gen_collegeyes -0.85006782 0.2316595 -3.66947101
## studytime 0.29304592 0.1392885 2.10387705
## schoolsupyes:studytime -0.85318089 0.3244225 -2.62984500
## schoolsupyes:first_gen_collegeyes 1.52425650 0.5719558 2.66498990
## 0|4 -2.69208825 0.4041720 -6.66074948
## 4|5 -2.65581637 0.4029201 -6.59142141
## 5|6 -2.48392676 0.3975209 -6.24854363
## 6|7 -2.13758424 0.3890606 -5.49421935
## 7|8 -1.94093740 0.3852119 -5.03862266
## 8|9 -1.35895482 0.3771508 -3.60321358
## 9|10 -0.98243182 0.3738798 -2.62766733
## 10|11 -0.33664864 0.3707238 -0.90808488
## 11|12 0.27441039 0.3708471 0.73995555
## 12|13 0.64181542 0.3718808 1.72586325
## 13|14 1.15140976 0.3751961 3.06882090
## 14|15 1.64894914 0.3823753 4.31238440
## 15|16 2.37070007 0.4008641 5.91397458
## 16|17 2.90477643 0.4244524 6.84358578
## 17|18 3.21877485 0.4439939 7.24959282
## 18|19 4.18096474 0.5449623 7.67202626
## 19|20 6.00685744 1.0635640 5.64785684
## p value
## failedyes 1.116447e-07
## gooutlow 5.193826e-03
## romanticyes 1.983653e-02
## schoolsupyes 9.676766e-01
## first_gen_collegeyes 2.430529e-04
## studytime 3.538917e-02
## schoolsupyes:studytime 8.542381e-03
## schoolsupyes:first_gen_collegeyes 7.699064e-03
## 0|4 2.724347e-11
## 4|5 4.356354e-11
## 5|6 4.142975e-10
## 6|7 3.924425e-08
## 7|8 4.688939e-07
```

```
## 8|9          3.143071e-04
## 9|10         8.597255e-03
## 10|11        3.638334e-01
## 11|12        4.593270e-01
## 12|13        8.437202e-02
## 13|14        2.149054e-03
## 14|15        1.615033e-05
## 15|16        3.339494e-09
## 16|17        7.723504e-12
## 17|18        4.180264e-13
## 18|19        1.693003e-14
## 19|20        1.624604e-08
```

```
(ci5 <- confint(mod5))
```

```
## Waiting for profiling to be done...
```

```
##
```

```
## Re-fitting to get Hessian
```

```
##              2.5 %      97.5 %
## failedyes    -1.91099915 -0.88183866
## gooutlow      0.17940546  1.01668340
## romanticyes  -0.92995496 -0.08125812
## schoolsupyes  -1.60642853  1.65956106
## first_gen_collegeyes -1.30698473 -0.39813983
## studytime     0.02039631  0.56697874
## schoolsupyes:studytime -1.49158352 -0.21490640
## schoolsupyes:first_gen_collegeyes 0.40814640  2.65646321
```

This has resulted in an increase in the AIC, which is still lower than the first three models.

Let's check the accuracy of this model with interaction terms:

```
acc.ord4 <- predict(mod4, training)
ctable <- table(training$G3, acc.ord4)
round((sum(diag(ctable))/sum(ctable))*100,2)
```

```
## [1] 19.94
```

```
ctable
```

```
##      acc.ord4
##      0  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
## 0 18  0  0  0  0  0  0 10  3  0  1  0  1  0  0  0  0  0
## 4  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## 5  4  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0
## 6  1  0  0  0  0  0  0 10  1  0  0  0  0  0  0  0  0  0
## 7  4  0  0  0  0  0  0  1  1  0  0  0  2  0  0  0  0  0
## 8 14  0  0  0  0  0  0  7  3  0  0  0  3  0  0  0  0  0
## 9  8  0  0  0  0  0  0  4  6  0  2  0  0  0  0  0  0  0
## 10 5  0  0  0  0  0  0 14 10  0  2  0  8  0  0  0  0  0
## 11 5  0  0  0  0  0  0 13 15  0  1  0  6  0  0  0  0  0
## 12 3  0  0  0  0  0  0  7  9  0  1  0  3  0  0  0  0  0
## 13 4  0  0  0  0  0  0  5 10  0  3  0  6  0  0  0  0  0
## 14 0  0  0  0  0  0  0  2  8  0  4  0  9  0  0  0  0  0
## 15 0  0  0  0  0  0  0  3  8  0  1  0 13  0  0  0  0  0
## 16 0  0  0  0  0  0  0  0  9  0  0  0  3  0  0  0  0  0
## 17 0  0  0  0  0  0  0  0  3  0  0  0  2  0  0  0  0  0
```

```
## 18 1 0 0 0 0 0 0 0 0 4 0 2 0 2 0 0 0 0 0
## 19 0 0 0 0 0 0 0 0 0 2 0 0 0 3 0 0 0 0 0
## 20 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0
```

The accuracy is even lower than mod3, at only 19.94% for the training set.

Checking on testing set:

```
pred.ord3 <- predict(mod3, testing)
ctable <- table(testing$G3, pred.ord3)
round((sum(diag(ctable))/sum(ctable))*100,2)
```

```
## [1] 8.86
```

```
pred.ord4 <- predict(mod4, testing)
ctable <- table(testing$G3, pred.ord4)
round((sum(diag(ctable))/sum(ctable))*100,2)
```

```
## [1] 11.39
```

```
pred.ord5 <- predict(mod5, testing)
ctable <- table(testing$G3, pred.ord5)
round((sum(diag(ctable))/sum(ctable))*100,2)
```

```
## [1] 10.13
```

Accuracy rates are even lower, at 8.86%, 11.39%, and 10.13%.

Highly inaccurate model, not a good fit for the data.

## Recoding g3 with larger categories

```
library(car)
```

```
## Loading required package: carData
```

```
##
```

```
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
```

```
##
```

```
## recode
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
## some
```

```
## The following object is masked from 'package:openintro':
```

```
##
```

```
## densityPlot
```

```
data <- data %>%
  mutate(cat_g3 = case_when(
    G3 == 0 ~ "Poor",
    G3 <= 9 ~ "Weak",
    G3 <= 13 ~ "Sufficient",
    G3 <= 15 ~ "Good",
    G3 <= 17 ~ "Very Good",
    G3 <= 20 ~ "Excellent"
  ))
```

```
data <- data %>%
  mutate(cat_g3 = factor(cat_g3, levels=c("Poor", "Weak", "Sufficient", "Good", "Very Good", "Excellent"))

set.seed(3)
train_ind <- sample(x = nrow(data), size = 0.8 * nrow(data))
test_ind_neg <- -train_ind
ftrain <- data[train_ind, ]
ftest <- data[test_ind_neg, ]
```

Trying out a multicat ordinal logit on this:

```
mod6 <- polr(cat_g3 ~ failed + goout + romantic + schoolsup + first_gen_college + schoolsup * studytime
summary(mod6)
```

```
##
## Re-fitting to get Hessian
## Call:
## polr(formula = cat_g3 ~ failed + goout + romantic + schoolsup +
##       first_gen_college + schoolsup * studytime + schoolsup * first_gen_college,
##       data = ftrain)
##
## Coefficients:
##
##              Value Std. Error t value
## failedyes      -1.55916    0.2803 -5.56248
## gooutlow        0.58202    0.2260  2.57501
## romanticyes    -0.68049    0.2310 -2.94618
## schoolsupyes     0.01527    0.8809  0.01734
## first_gen_collegeyes -0.87441    0.2449 -3.56977
## studytime       0.31024    0.1444  2.14898
## schoolsupyes:studytime -0.89744    0.3472 -2.58483
## schoolsupyes:first_gen_collegeyes 1.78767    0.6142  2.91072
##
## Intercepts:
##              Value Std. Error t value
## Poor|Weak      -2.8407    0.4234 -6.7090
## Weak|Sufficient -1.0846    0.3904 -2.7781
## Sufficient|Good  1.1195    0.3899  2.8716
## Good|Very Good  2.3538    0.4154  5.6661
## Very Good|Excellent 3.2057    0.4575  7.0075
##
## Residual Deviance: 873.2587
## AIC: 899.2587
```

```
(ctable <- coef(summary(mod6)))
```

```
##
## Re-fitting to get Hessian
##
##              Value Std. Error t value
## failedyes      -1.55915962    0.2802994 -5.56247945
## gooutlow        0.58201779    0.2260254  2.57501100
## romanticyes    -0.68048602    0.2309721 -2.94618281
## schoolsupyes     0.01527404    0.8808503  0.01734011
## first_gen_collegeyes -0.87441246    0.2449495 -3.56976600
```

```
## studytime                0.31023730  0.1443649  2.14898050
## schoolsupyes:studytime   -0.89744352  0.3471959 -2.58483360
## schoolsupyes:first_gen_collegeyes  1.78767301  0.6141683  2.91072192
## Poor|Weak                -2.84071658  0.4234218 -6.70895259
## Weak|Sufficient          -1.08462029  0.3904224 -2.77806907
## Sufficient|Good          1.11949859  0.3898567  2.87156458
## Good|Very Good           2.35375515  0.4154121  5.66607292
## Very Good|Excellent      3.20565589  0.4574631  7.00746349
```

```
p6 <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
(ctable <- cbind(ctable, "p value" = p6))
```

```
##                               Value Std. Error    t value
## failedyes                    -1.55915962  0.2802994 -5.56247945
## gooutlow                      0.58201779  0.2260254  2.57501100
## romanticyes                  -0.68048602  0.2309721 -2.94618281
## schoolsupyes                   0.01527404  0.8808503  0.01734011
## first_gen_collegeyes         -0.87441246  0.2449495 -3.56976600
## studytime                     0.31023730  0.1443649  2.14898050
## schoolsupyes:studytime        -0.89744352  0.3471959 -2.58483360
## schoolsupyes:first_gen_collegeyes  1.78767301  0.6141683  2.91072192
## Poor|Weak                    -2.84071658  0.4234218 -6.70895259
## Weak|Sufficient              -1.08462029  0.3904224 -2.77806907
## Sufficient|Good              1.11949859  0.3898567  2.87156458
## Good|Very Good               2.35375515  0.4154121  5.66607292
## Very Good|Excellent          3.20565589  0.4574631  7.00746349
```

```
##                               p value
## failedyes                    2.659684e-08
## gooutlow                     1.002369e-02
## romanticyes                  3.217222e-03
## schoolsupyes                  9.861653e-01
## first_gen_collegeyes         3.573003e-04
## studytime                    3.163595e-02
## schoolsupyes:studytime        9.742600e-03
## schoolsupyes:first_gen_collegeyes 3.605948e-03
## Poor|Weak                    1.960263e-11
## Weak|Sufficient              5.468299e-03
## Sufficient|Good              4.084453e-03
## Good|Very Good               1.461074e-08
## Very Good|Excellent          2.426773e-12
```

```
(ci5 <- confint(mod6))
```

```
## Waiting for profiling to be done...
```

```
##
```

```
## Re-fitting to get Hessian
```

```
##                               2.5 %    97.5 %
## failedyes                    -2.11603937 -1.0156926
## gooutlow                      0.14100166  1.0278318
## romanticyes                  -1.13625714 -0.2299976
## schoolsupyes                  -1.71834716  1.7492857
## first_gen_collegeyes         -1.35840403 -0.3972182
## studytime                     0.02754372  0.5941023
## schoolsupyes:studytime        -1.58076549 -0.2136889
## schoolsupyes:first_gen_collegeyes 0.58713044  3.0005349
```

```
acc.ord6 <- predict(mod6, ftrain)
ctable <- table(ftrain$cat_g3, acc.ord6)
ctable
```

```
##          acc.ord6
##          Poor Weak Sufficient Good Very Good Excellent
## Poor          3  10          20   0          0          0
## Weak          5  22          46   0          0          0
## Sufficient    2  13         113   2          0          0
## Good          0   0          47   1          0          0
## Very Good     0   0          17   0          0          0
## Excellent     0   0          14   1          0          0
```

Still not very accurate for the training

Random forest:

```
library(randomForest)
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
```

```
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:dplyr':
```

```
##
```

```
##      combine
```

```
## The following object is masked from 'package:ggplot2':
```

```
##
```

```
##      margin
```

```
rf.cat<-randomForest(cat_g3~. -G1 -G2 -G3 -ord_g3,data = ftrain,ntree=50, importance=TRUE)
print(rf.cat)
```

```
##
```

```
## Call:
```

```
## randomForest(formula = cat_g3 ~ . - G1 - G2 - G3 - ord_g3, data = ftrain,      ntree = 50, importan
```

```
##          Type of random forest: classification
```

```
##          Number of trees: 50
```

```
## No. of variables tried at each split: 5
```

```
##
```

```
##          OOB estimate of  error rate: 57.59%
```

```
## Confusion matrix:
```

```
##          Poor Weak Sufficient Good Very Good Excellent class.error
## Poor          19   4           8   2          0          0  0.4242424
## Weak          4  22          44   2          1          0  0.6986301
## Sufficient    7  27          82  11          1          2  0.3692308
## Good          3   5          28  11          1          0  0.7708333
## Very Good     2   1           6   7          0          1  1.0000000
## Excellent     1   2           7   5          0          0  1.0000000
```

```
importance(rf.cat)
```

```
##          Poor      Weak Sufficient      Good  Very Good
## school    0.226003250 -0.9187087 -0.85677217 -1.43591632  0.00000000
## sex       1.494651032  1.7762605  0.03821383  1.95443887 -0.07433352
```

## age	1.569527762	-0.9790961	-0.09353067	-0.45194443	0.02905101
## address	0.593232865	1.4805049	-0.32024614	1.81297046	-1.76776695
## famsize	0.105561155	-1.4014503	1.32354745	-1.79028456	-0.79912145
## Pstatus	0.858834610	-0.5609630	-1.07505128	-0.81690009	-0.18738292
## Medu	-0.090460439	-0.5727420	1.46337515	0.77784603	-1.57129772
## Fedu	0.776365954	-0.5223751	-0.61663892	0.02339555	0.16785072
## Mjob	-0.253061442	-0.8243416	1.36113089	1.00045915	-0.81485832
## Fjob	-0.665838466	-0.7241551	-0.13657280	1.12241368	0.43424751
## reason	0.523038396	-0.6379817	-0.51204545	1.58959151	-0.54960331
## guardian	1.026051352	-0.2905394	-0.74376536	2.18978869	0.76367171
## traveltime	0.972437546	-0.4315403	-0.35130163	-0.87522193	0.00000000
## studytime	1.314317442	0.9320422	-0.10289467	-2.21185947	-1.01015254
## failures	2.505922776	2.1480062	0.69661141	3.16401561	0.00000000
## schoolsup	1.544393831	3.2865116	-0.35391285	2.95599795	-1.01015254
## famsup	1.153420048	-1.0686139	0.82078718	1.06578304	-1.73147871
## paid	0.327652042	0.8863156	0.93433356	0.05888746	0.46743264
## activities	0.163392380	0.1570002	-0.56349980	0.20756785	0.00000000
## nursery	-0.528334760	-0.5361549	0.49062376	1.39576068	-0.15621187
## higher	-1.467975974	-1.4230766	-0.73578067	1.01015254	0.00000000
## internet	0.490256317	-0.7623360	0.88494504	0.15631749	-1.33630621
## romantic	0.799040036	-0.2618230	1.18552600	-0.56383214	-1.22152572
## famrel	1.428274026	0.8199781	-0.93253131	-0.51093471	0.69441933
## freetime	-0.003392297	-1.8553676	0.20285302	0.47927194	-1.43642070
## goout	-1.124683098	1.0160149	1.10162389	1.74429690	-1.01015254
## Dalc	0.000000000	-1.1623014	1.43019388	0.00000000	0.00000000
## Walc	1.251666237	-0.7444515	-1.27743971	2.46725776	-0.89209537
## health	-1.532156175	-1.1851722	0.56639789	0.14562287	-0.57928445
## absences	9.378811659	2.0115634	2.07432098	0.15645826	0.52191292
## first_gen_college	1.142965213	-1.0057378	-1.17232382	0.83071549	-0.98441222
## stable_learning_env	-0.178572533	-0.3979485	-0.75368498	0.94471707	-1.11170005
## high_freq_absent	1.970691144	0.4405862	0.25233073	2.72942494	-1.01015254
## failed	0.446574184	1.8520112	1.58493067	2.83378101	1.41421356
##	Excellent	MeanDecreaseAccuracy	MeanDecreaseGini		
## school	0.00000000	-1.18834240	2.5443404		
## sex	0.73822821	1.85978465	5.6312466		
## age	0.15139145	0.08631188	12.2447927		
## address	-1.01015254	0.68231901	5.2743769		
## famsize	-1.01015254	-0.53717450	6.0839526		
## Pstatus	1.01015254	-0.83682978	2.8450952		
## Medu	0.06287195	0.65002473	10.4685866		
## Fedu	-0.19914544	-0.58802493	11.1382283		
## Mjob	1.16956482	1.22904054	11.4021937		
## Fjob	-2.06148198	-0.40018725	10.4249888		
## reason	0.00000000	0.00717601	10.2113621		
## guardian	0.00000000	0.23559066	6.8649196		
## traveltime	1.01015254	-0.28121281	6.7112174		
## studytime	0.00000000	-0.18011625	10.1012299		
## failures	-1.01015254	3.12501468	7.6902074		
## schoolsup	1.01015254	2.88811852	4.8028085		
## famsup	1.44004608	1.14144635	4.7869583		
## paid	1.01015254	1.04674191	6.4322666		
## activities	-0.74611773	-0.01498062	5.3091696		
## nursery	0.00000000	0.43920166	3.7866851		
## higher	0.00000000	-1.65455646	1.0367975		

```
varImpPlot(rf.cat)
```

The figure displays two plots side-by-side, showing the importance of various variables for predicting the 'absences' variable. The left plot shows the importance measured by MeanDecreaseAccuracy, and the right plot shows the importance measured by MeanDecreaseGini. Both plots use a grid of circles to represent the importance of each variable, with the 'absences' variable being the most important in both plots.

**Variables (from top to bottom):**

- absences
- failures
- schoolsup
- failed
- high\_freq\_absent
- Sex
- Dropout
- failop
- failsup
- romantic
- address
- Medu
- nursery
- guardian
- internet
- age
- failrel
- Season
- Daic
- activities
- travelttime
- Walc
- stable\_learning\_env
- partsize
- readin
- Pstatus

**MeanDecreaseAccuracy (Left Plot):**

The x-axis represents the importance score, ranging from 0 to 6. The 'absences' variable has the highest importance score, followed by 'failures', 'schoolsup', 'failed', 'high\_freq\_absent', 'Sex', 'Dropout', 'failop', 'failsup', 'romantic', 'address', 'Medu', 'nursery', 'guardian', 'internet', 'age', 'failrel', 'Season', 'Daic', 'activities', 'travelttime', 'Walc', 'stable\_learning\_env', 'partsize', 'readin', and 'Pstatus'.

**MeanDecreaseGini (Right Plot):**

The x-axis represents the importance score, ranging from 0 to 20. The 'absences' variable has the highest importance score, followed by 'failures', 'schoolsup', 'failed', 'high\_freq\_absent', 'Sex', 'Dropout', 'failop', 'failsup', 'romantic', 'address', 'Medu', 'nursery', 'guardian', 'internet', 'age', 'failrel', 'Season', 'Daic', 'activities', 'travelttime', 'Walc', 'stable\_learning\_env', 'partsize', 'readin', and 'Pstatus'.

##		actual						
##	predictions	Poor	Weak	Sufficient	Good	Very	Good	Excellent
##	Poor	33	0	0	0		0	0
##	Weak	0	73	0	0		0	0
##	Sufficient	0	0	130	0		0	0
##	Good	0	0	0	48		0	0
##	Very Good	0	0	0	0		17	0
##	Excellent	0	0	0	0		0	15

120



```
## [1] 1
```

Very fitted model with accuracy for training data >99%.

Let's see what the accuracy rate for the testing set is:

```
rf.pred<- predict(rf.cat, ftest, type = 'class')
t<-table(predictions=rf.pred, actual=ftest$cat_g3)
t
```

```
##           actual
## predictions  Poor Weak Sufficient Good Very Good Excellent
## Poor         2    0         2    0         0         0
## Weak         0    1         6    0         0         0
## Sufficient    2   16        21    8         2         1
## Good         1    2         6    4         2         2
## Very Good    0    0         0    0         1         0
## Excellent    0    0         0    0         0         0
```

```
sum(diag(t))/sum(t)
```

```
## [1] 0.3670886
```

39.2% accuracy, which is an improvement, but a low value.

Let's choose the most important variables, as well as interaction effects we believe to be important based on previous exploration:

```
rf.cat1<-randomForest(cat_g3~failed + absences+ Mjob + studytime + goout + romantic + schoolsup + first,
print(rf.cat1)
```

```
##
## Call:
## randomForest(formula = cat_g3 ~ failed + absences + Mjob + studytime + goout + romantic + scho
##           Type of random forest: classification
##           Number of trees: 50
## No. of variables tried at each split: 2
##
## OOB estimate of error rate: 56.33%
## Confusion matrix:
##           Poor Weak Sufficient Good Very Good Excellent class.error
## Poor         18   4         10    0         1         0  0.4545455
## Weak         4   20         47    2         0         0  0.7260274
## Sufficient    4   21         93   12         0         0  0.2846154
## Good         3    2         33    7         3         0  0.8541667
## Very Good    2    2          7    5         0         1  1.0000000
## Excellent    0    2         11    2         0         0  1.0000000
```

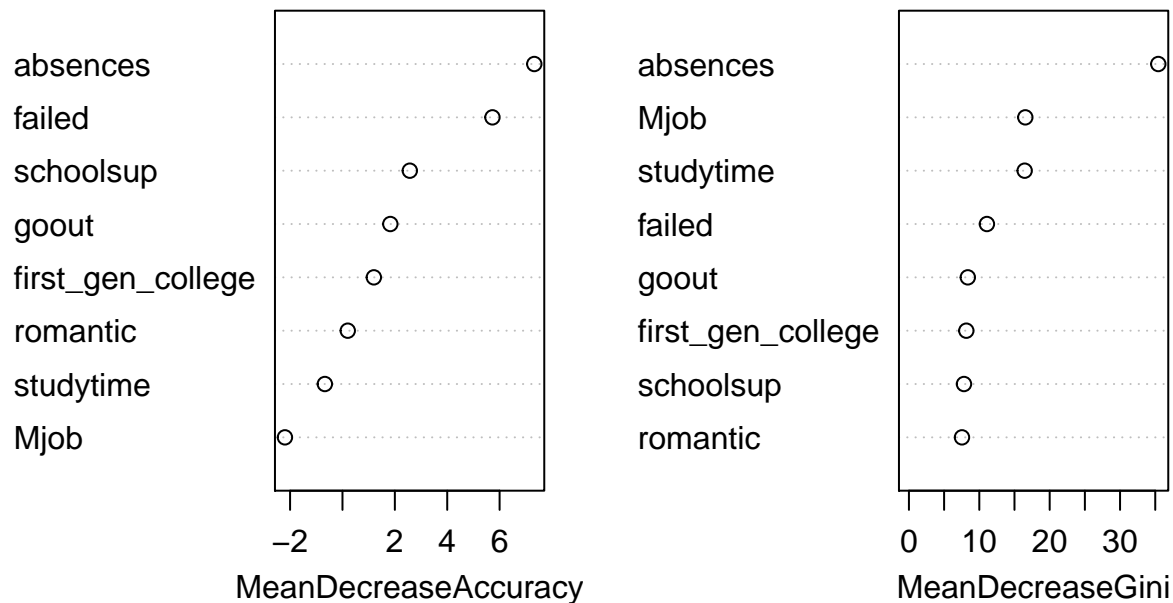
```
importance(rf.cat1)
```

```
##           Poor      Weak Sufficient      Good Very Good
## failed      3.74080536  7.061153083  1.17764779  0.6023557  1.0101525
## absences    9.62023705  2.748465022  2.70089303  1.3877461 -0.6850789
## Mjob        0.96975297 -0.007131597 -1.25574679 -2.1882684 -2.0177331
## studytime  -0.63971257  1.094720741 -0.01208343 -0.1590809 -0.1597932
## goout       -0.07414337  4.723603083 -1.18528474  2.8224822 -1.2856487
## romantic   -0.10727438  0.580291270  0.58831813 -0.8518423 -1.9580067
## schoolsup   3.84798132  3.063717643  0.19984540  1.8072086  0.0000000
## first_gen_college 0.57975639  1.757160423 -1.39942112  2.6458567 -1.3988963
```

```
##           Excellent MeanDecreaseAccuracy MeanDecreaseGini
## failed      -1.7216382           5.726920      11.082086
## absences     -1.1166562           7.324336      35.433386
## Mjob         -0.8172404          -2.197027      16.535533
## studytime    -3.5628692          -0.672507      16.447634
## goout        -1.4229789           1.827199       8.366949
## romantic      1.7126031           0.202445       7.536230
## schoolsup      0.0000000           2.570983       7.822760
## first_gen_college -0.5261332       1.197858       8.153775
```

```
varImpPlot(rf.cat1)
```

rf.cat1



```
rf.acc<- predict(rf.cat1, ftrain, type = 'class')
t<-table(predictions=rf.acc, actual=ftrain$cat_g3)
t
```

```
##           actual
## predictions Poor Weak Sufficient Good Very Good Excellent
## Poor         28   2      2      2      2      0
## Weak          0  45      2      0      0      0
## Sufficient     4  25     125    24     10     11
## Good           0   1      1    22      2      0
## Very Good      1   0      0     0      3      0
## Excellent      0   0      0     0      0      4
```

```
sum(diag(t))/sum(t)
```

```
## [1] 0.7183544
```

Much worse fit for the data, 50.32% OOB estimate of error rate and 70.25% accuracy rate for the training data.

```
rf.pred1<- predict(rf.cat1, ftest, type = 'class')
t<-table(predictions=rf.pred1, actual=ftest$cat_g3)
t
```

```
##           actual
## predictions  Poor Weak Sufficient Good Very Good Excellent
## Poor         3    0         2    0         0         0
## Weak         0    3         6    0         0         0
## Sufficient    2   15        22    9         4         2
## Good          0    1         5    3         1         1
## Very Good    0    0         0    0         0         0
## Excellent    0    0         0    0         0         0
```

```
sum(diag(t))/sum(t)
```

```
## [1] 0.3924051
```

38% Accuracy, which is less than the previous RF model.

## Recoding binary response variable for pass-fail

Considering final grades as a continuous variable and ordinal categorical variable gave poor results. Therefore, we'd like to model a binary variable that indicates whether the student passes (grade  $\geq 10$ ) or fails ( $<10$ ).

```
data <- data %>%
  mutate(pf = case_when(
    G3 >= 10 ~ "pass",
    G3 < 10 ~ "fail"
  ))
data <- data %>%
  mutate(pf = factor(pf, levels=c("pass", "fail"), ordered = FALSE))

set.seed(3)
train_ind <- sample(x = nrow(data), size = 0.8 * nrow(data))
test_ind_neg <- -train_ind
ftrain <- data[train_ind, ]
ftest <- data[test_ind_neg, ]
```

## Fitting a decision tree on pass-fail

```
library(tree)
```

```
## Registered S3 method overwritten by 'tree':
##   method      from
##   print.tree cli
```

```
library(rpart)
```

## Fitting random forest on pass-fail

Fitting with ALL predictors:

```
rf.bin<-randomForest(pf~. -G1 -G2 -G3 -ord_g3 - cat_g3 -failures -reason -health -age -nursery,data = f
print(rf.bin)
```

```
##
```

```
## Call:
## randomForest(formula = pf ~ . - G1 - G2 - G3 - ord_g3 - cat_g3 - failures - reason - health - ,
##               Type of random forest: classification
##               Number of trees: 50
## No. of variables tried at each split: 3
##
## OOB estimate of error rate: 33.23%
## Confusion matrix:
##      pass fail class.error
## pass  179   31  0.1476190
## fail   74   32  0.6981132
```

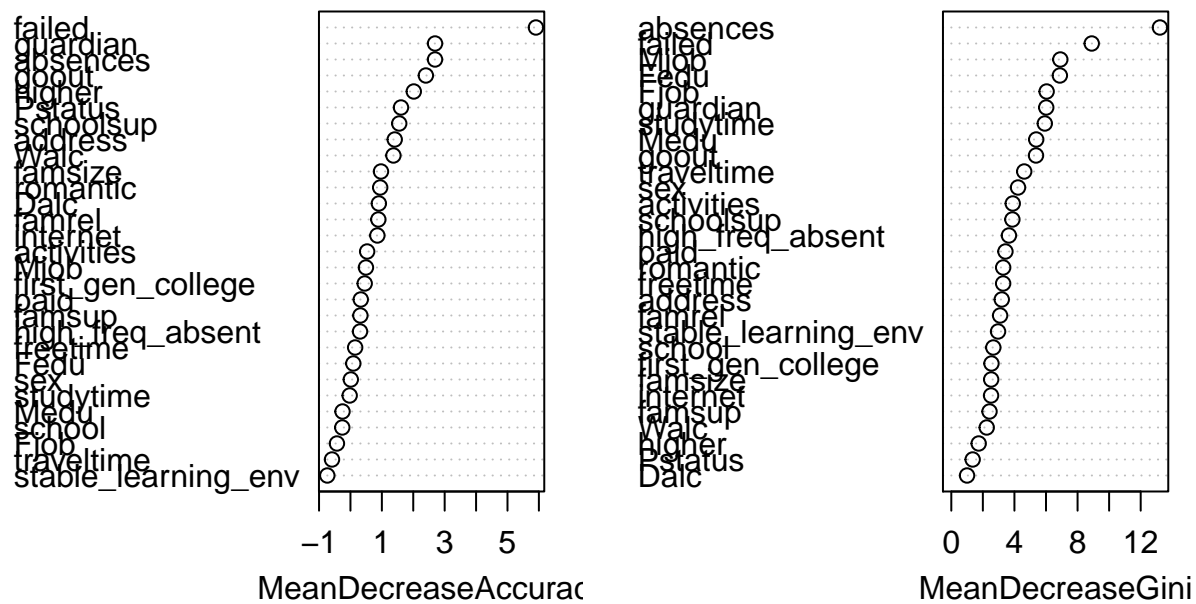
```
importance(rf.bin)
```

	pass	fail	MeanDecreaseAccuracy
## school	-1.2365897	1.21225073	-0.26101214
## sex	-0.1808442	0.29503671	0.01100947
## address	1.5357229	0.08941542	1.40961253
## famsize	1.4233602	-0.03747495	0.97203613
## Pstatus	1.0802031	0.97424218	1.60939637
## Medu	-1.5700298	1.59476857	-0.25288939
## Fedu	0.4273086	-0.35045203	0.09063057
## Mjob	-0.5513467	1.42875815	0.49358129
## Fjob	-0.6371998	0.31467729	-0.43035898
## guardian	2.5160369	0.57529012	2.69334691
## traveltime	-0.5073801	-0.37990679	-0.58765439
## studytime	0.2496549	-0.39105764	-0.02488387
## schoolsup	1.1417140	1.31683439	1.55332807
## famsup	-0.4604047	1.73933103	0.31650754
## paid	-0.6588598	0.89574576	0.32894723
## activities	0.7555376	-0.05140775	0.53108214
## higher	1.5303278	1.69345789	2.01647086
## internet	0.2415090	1.00530975	0.85680874
## romantic	1.3887979	-0.05233816	0.94629305
## famrel	1.2685077	-0.20095748	0.88799032
## freetime	-0.1076713	0.49580824	0.14923493
## goout	0.6555519	2.29575407	2.40084292
## Dalc	0.7749492	0.29415566	0.90390611
## Walc	1.7584529	0.06565175	1.36829804
## absences	2.0164737	1.62817570	2.69250297
## first_gen_college	-0.6244840	1.82670724	0.45339567
## stable_learning_env	-0.5556304	-0.59333857	-0.73730499
## high_freq_absent	0.9232732	-0.62379109	0.30460159
## failed	5.3277863	4.14046663	5.90559225
##	MeanDecreaseGini		
## school	2.6613627		
## sex	4.2278218		
## address	3.1964253		
## famsize	2.5261012		
## Pstatus	1.3541068		
## Medu	5.3808360		
## Fedu	6.8794297		
## Mjob	6.9075638		
## Fjob	6.0401937		
## guardian	6.0147528		

```
## traveltime 4.6198568
## studytime 5.9234659
## schoolsup 3.8701226
## famsup 2.4223571
## paid 3.4304067
## activities 3.8938695
## higher 1.7362249
## internet 2.5199383
## romantic 3.2872393
## famrel 3.0916236
## freetime 3.2846767
## goout 5.3667204
## Dalc 0.9977066
## Walc 2.2450620
## absences 13.2184840
## first_gen_college 2.5447637
## stable_learning_env 2.9624366
## high_freq_absent 3.6521725
## failed 8.8980095
```

```
varImpPlot(rf.bin)
```

rf.bin



```
rf.acc<- predict(rf.bin, ftrain, type = 'class')
t<-table(predictions=rf.acc, actual=ftrain$cat_g3)
t
```

```
##          actual
## predictions Poor Weak Sufficient Good Very Good Excellent
##          pass   0    1         130  48         17         15
##          fail   33   72           0   0           0           0
```

```
sum(diag(t))/sum(t)
```

```
## [1] 0.2278481
```

Predictions on testing set:

```
rf.pred2<- predict(rf.bin, ftest, type = 'class')
t<-table(predictions=rf.pred2, actual=ftest$pf)
t
```

```
##           actual
## predictions pass fail
##           pass  53  21
##           fail   2   3
```

```
sum(diag(t))/sum(t)
```

```
## [1] 0.7088608
```

72.15% accuracy rate.

Finding the best random forest model by including important predictors:

```
rf.bin<-randomForest(pf~failed + absences+ guardian + studytime + goout + schoolsup + first_gen_college
print(rf.bin)
```

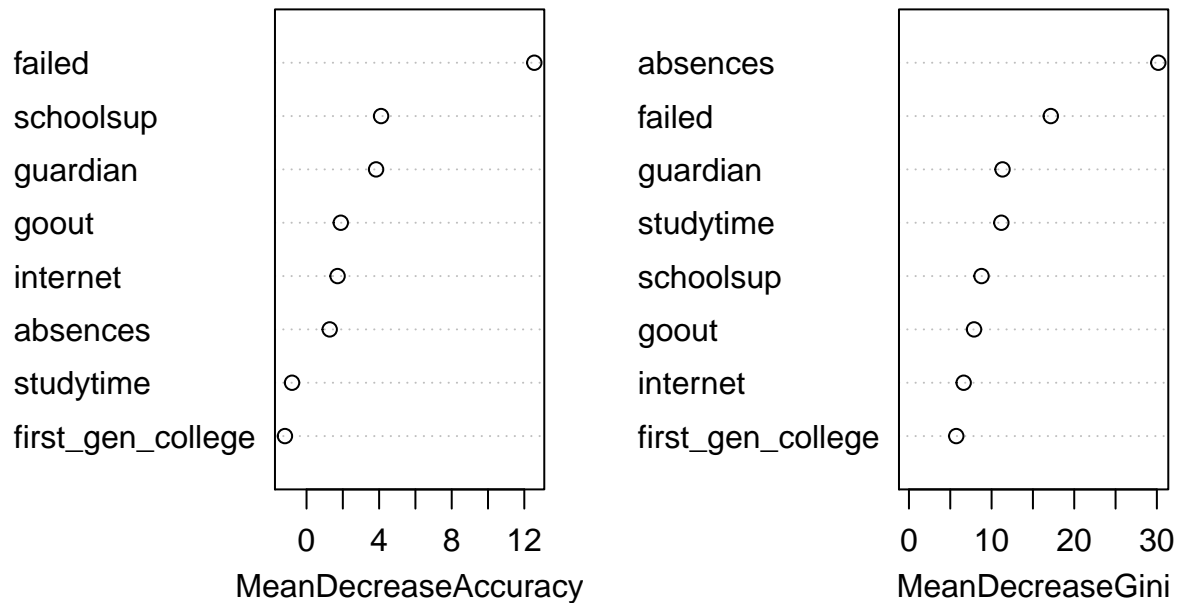
```
##
## Call:
## randomForest(formula = pf ~ failed + absences + guardian + studytime +      goout + schoolsup + fir
##           Type of random forest: classification
##           Number of trees: 50
## No. of variables tried at each split: 3
##
##           OOB estimate of  error rate: 23.42%
## Confusion matrix:
##           pass fail class.error
## pass  185   25  0.1190476
## fail   49   57  0.4622642
```

```
importance(rf.bin)
```

```
##           pass      fail MeanDecreaseAccuracy MeanDecreaseGini
## failed      10.5031127  7.7729302           12.5514428          17.164473
## absences      0.7776225  0.9887959           1.2768025          30.173546
## guardian      3.0322670  3.2505062           3.8326289          11.324691
## studytime     -2.3599001  2.1225447          -0.7993082          11.182709
## goout         1.9456591  0.7599509           1.8876845           7.875670
## schoolsup      2.3393935  4.4977156           4.1079438           8.779473
## first_gen_college -1.7292285  0.0998300          -1.1898968           5.723457
## internet      2.2696157 -0.2422729           1.7107156           6.620175
```

```
varImpPlot(rf.bin)
```

## rf.bin



```
rf.acc<- predict(rf.bin, ftrain, type = 'class')
t<-table(predictions=rf.acc, actual=ftrain$cat_g3)
t
```

```
##          actual
## predictions Poor Weak Sufficient Good Very Good Excellent
##      pass    5   22      124   48        17        15
##      fail   28   51         6    0          0          0
```

```
sum(diag(t))/sum(t)
```

```
## [1] 0.1772152
```

The pared-down model has an OOB estimate of error rate of 25.95% and a training set prediction accuracy rate of 89.6%.

Predictions on testing set:

```
rf.pred3<- predict(rf.bin, ftest, type = 'class')
t<-table(predictions=rf.pred3, actual=ftest$pf)
t
```

```
##          actual
## predictions pass fail
##      pass    50   18
##      fail     5    6
```

```
sum(diag(t))/sum(t)
```

```
## [1] 0.7088608
```

72.15% prediction accuracy rate.

Overall the random-forests for pass-fail indicate that the most important factors affecting whether the student passes/fails are failed, absences, guardian, studytime, goout, schoolsup, first\_gen\_college, internet.