Student Alcohol Consumption

Vicki Nomwesigwa

11/8/2021

## DATA WRANGLING

#### Read the data

## school sex age address famsize Pstatus Medu Fedu Mjob Fjob reason  
## 1 GP F 18 U GT3 A 4 4 at\_home teacher course  
## 2 GP F 17 U GT3 T 1 1 at\_home other course  
## 3 GP F 15 U LE3 T 1 1 at\_home other other  
## 4 GP F 15 U GT3 T 4 2 health services home  
## 5 GP F 16 U GT3 T 3 3 other other home  
## 6 GP M 16 U LE3 T 4 3 services other reputation  
## guardian traveltime studytime failures schoolsup famsup paid activities  
## 1 mother 2 2 0 yes no no no  
## 2 father 1 2 0 no yes no no  
## 3 mother 1 2 3 yes no yes no  
## 4 mother 1 3 0 no yes yes yes  
## 5 father 1 2 0 no yes yes no  
## 6 mother 1 2 0 no yes yes yes  
## nursery higher internet romantic famrel freetime goout Dalc Walc health  
## 1 yes yes no no 4 3 4 1 1 3  
## 2 no yes yes no 5 3 3 1 1 3  
## 3 yes yes yes no 4 3 2 2 3 3  
## 4 yes yes yes yes 3 2 2 1 1 5  
## 5 yes yes no no 4 3 2 1 2 5  
## 6 yes yes yes no 5 4 2 1 2 5  
## absences G1 G2 G3  
## 1 6 5 6 6  
## 2 4 5 5 6  
## 3 10 7 8 10  
## 4 2 15 14 15  
## 5 4 6 10 10  
## 6 10 15 15 15

## [1] 395 33

## 'data.frame': 395 obs. of 33 variables:  
## $ school : chr "GP" "GP" "GP" "GP" ...  
## $ sex : chr "F" "F" "F" "F" ...  
## $ age : int 18 17 15 15 16 16 16 17 15 15 ...  
## $ address : chr "U" "U" "U" "U" ...  
## $ famsize : chr "GT3" "GT3" "LE3" "GT3" ...  
## $ Pstatus : chr "A" "T" "T" "T" ...  
## $ Medu : int 4 1 1 4 3 4 2 4 3 3 ...  
## $ Fedu : int 4 1 1 2 3 3 2 4 2 4 ...  
## $ Mjob : chr "at\_home" "at\_home" "at\_home" "health" ...  
## $ Fjob : chr "teacher" "other" "other" "services" ...  
## $ reason : chr "course" "course" "other" "home" ...  
## $ guardian : chr "mother" "father" "mother" "mother" ...  
## $ traveltime: int 2 1 1 1 1 1 1 2 1 1 ...  
## $ studytime : int 2 2 2 3 2 2 2 2 2 2 ...  
## $ failures : int 0 0 3 0 0 0 0 0 0 0 ...  
## $ schoolsup : chr "yes" "no" "yes" "no" ...  
## $ famsup : chr "no" "yes" "no" "yes" ...  
## $ paid : chr "no" "no" "yes" "yes" ...  
## $ activities: chr "no" "no" "no" "yes" ...  
## $ nursery : chr "yes" "no" "yes" "yes" ...  
## $ higher : chr "yes" "yes" "yes" "yes" ...  
## $ internet : chr "no" "yes" "yes" "yes" ...  
## $ romantic : chr "no" "no" "no" "yes" ...  
## $ famrel : int 4 5 4 3 4 5 4 4 4 5 ...  
## $ freetime : int 3 3 3 2 3 4 4 1 2 5 ...  
## $ goout : int 4 3 2 2 2 2 4 4 2 1 ...  
## $ Dalc : int 1 1 2 1 1 1 1 1 1 1 ...  
## $ Walc : int 1 1 3 1 2 2 1 1 1 1 ...  
## $ health : int 3 3 3 5 5 5 3 1 1 5 ...  
## $ absences : int 6 4 10 2 4 10 0 6 0 0 ...  
## $ G1 : int 5 5 7 15 6 15 12 6 16 14 ...  
## $ G2 : int 6 5 8 14 10 15 12 5 18 15 ...  
## $ G3 : int 6 6 10 15 10 15 11 6 19 15 ...

Math: 395 rows, 33 columns

## 'data.frame': 395 obs. of 33 variables:  
## $ school : Factor w/ 2 levels "GP","MS": 1 1 1 1 1 1 1 1 1 1 ...  
## $ sex : Factor w/ 2 levels "F","M": 1 1 1 1 1 2 2 1 2 2 ...  
## $ age : num 18 17 15 15 16 16 16 17 15 15 ...  
## $ address : Factor w/ 2 levels "U","R": 1 1 1 1 1 1 1 1 1 1 ...  
## $ famsize : Factor w/ 2 levels "LE3","GT3": 2 2 1 2 2 1 1 2 1 2 ...  
## $ Pstatus : Factor w/ 2 levels "T","A": 2 1 1 1 1 1 1 2 2 1 ...  
## $ Medu : Factor w/ 5 levels "None","Primary Education (4th grade)",..: 5 2 2 5 4 5 3 5 4 4 ...  
## $ Fedu : Factor w/ 5 levels "None","Primary Education (4th grade)",..: 5 2 2 3 4 4 3 5 3 5 ...  
## $ Mjob : Factor w/ 5 levels "teacher","health",..: 4 4 4 2 5 3 5 5 3 5 ...  
## $ Fjob : Factor w/ 5 levels "teacher","health",..: 1 5 5 3 5 5 5 1 5 5 ...  
## $ reason : Factor w/ 4 levels "home","reputation",..: 3 3 4 1 1 2 1 1 1 1 ...  
## $ guardian : Factor w/ 3 levels "mother","father",..: 1 2 1 1 2 1 1 1 1 1 ...  
## $ traveltime: Factor w/ 4 levels "<15 min.","15 to 30 min.",..: 2 1 1 1 1 1 1 2 1 1 ...  
## $ studytime : Factor w/ 4 levels "<2 hours","2 to 5 hours",..: 2 2 2 3 2 2 2 2 2 2 ...  
## $ failures : num 0 0 3 0 0 0 0 0 0 0 ...  
## $ schoolsup : Factor w/ 2 levels "yes","no": 1 2 1 2 2 2 2 1 2 2 ...  
## $ famsup : Factor w/ 2 levels "yes","no": 2 1 2 1 1 1 2 1 1 1 ...  
## $ paid : Factor w/ 2 levels "yes","no": 2 2 1 1 1 1 2 2 1 1 ...  
## $ activities: Factor w/ 2 levels "yes","no": 2 2 2 1 2 1 2 2 2 1 ...  
## $ nursery : Factor w/ 2 levels "yes","no": 1 2 1 1 1 1 1 1 1 1 ...  
## $ higher : Factor w/ 2 levels "yes","no": 1 1 1 1 1 1 1 1 1 1 ...  
## $ internet : Factor w/ 2 levels "yes","no": 2 1 1 1 2 1 1 2 1 1 ...  
## $ romantic : Factor w/ 2 levels "yes","no": 2 2 2 1 2 2 2 2 2 2 ...  
## $ famrel : Factor w/ 5 levels "very bad","bad",..: 4 5 4 3 4 5 4 4 4 5 ...  
## $ freetime : Factor w/ 5 levels "very low","low",..: 3 3 3 2 3 4 4 1 2 5 ...  
## $ goout : Factor w/ 5 levels "very low","low",..: 4 3 2 2 2 2 4 4 2 1 ...  
## $ Dalc : Factor w/ 5 levels "very low","low",..: 1 1 2 1 1 1 1 1 1 1 ...  
## $ Walc : Factor w/ 5 levels "very low","low",..: 1 1 3 1 2 2 1 1 1 1 ...  
## $ health : Factor w/ 5 levels "very bad","bad",..: 3 3 3 5 5 5 3 1 1 5 ...  
## $ absences : num 6 4 10 2 4 10 0 6 0 0 ...  
## $ G1 : num 5 5 7 15 6 15 12 6 16 14 ...  
## $ G2 : num 6 5 8 14 10 15 12 5 18 15 ...  
## $ G3 : num 6 6 10 15 10 15 11 6 19 15 ...

## school sex age address famsize Pstatus Medu Fedu Mjob Fjob reason  
## 1 GP F 18 U GT3 A 4 4 at\_home teacher course  
## 2 GP F 17 U GT3 T 1 1 at\_home other course  
## 3 GP F 15 U LE3 T 1 1 at\_home other other  
## 4 GP F 15 U GT3 T 4 2 health services home  
## 5 GP F 16 U GT3 T 3 3 other other home  
## 6 GP M 16 U LE3 T 4 3 services other reputation  
## guardian traveltime studytime failures schoolsup famsup paid activities  
## 1 mother 2 2 0 yes no no no  
## 2 father 1 2 0 no yes no no  
## 3 mother 1 2 0 yes no no no  
## 4 mother 1 3 0 no yes no yes  
## 5 father 1 2 0 no yes no no  
## 6 mother 1 2 0 no yes no yes  
## nursery higher internet romantic famrel freetime goout Dalc Walc health  
## 1 yes yes no no 4 3 4 1 1 3  
## 2 no yes yes no 5 3 3 1 1 3  
## 3 yes yes yes no 4 3 2 2 3 3  
## 4 yes yes yes yes 3 2 2 1 1 5  
## 5 yes yes no no 4 3 2 1 2 5  
## 6 yes yes yes no 5 4 2 1 2 5  
## absences G1 G2 G3  
## 1 4 0 11 11  
## 2 2 9 11 11  
## 3 6 12 13 12  
## 4 0 14 14 14  
## 5 0 11 13 13  
## 6 6 12 12 13

## [1] 649 33

## school sex age address   
## Length:649 Length:649 Min. :15.00 Length:649   
## Class :character Class :character 1st Qu.:16.00 Class :character   
## Mode :character Mode :character Median :17.00 Mode :character   
## Mean :16.74   
## 3rd Qu.:18.00   
## Max. :22.00   
## famsize Pstatus Medu Fedu   
## Length:649 Length:649 Min. :0.000 Min. :0.000   
## Class :character Class :character 1st Qu.:2.000 1st Qu.:1.000   
## Mode :character Mode :character Median :2.000 Median :2.000   
## Mean :2.515 Mean :2.307   
## 3rd Qu.:4.000 3rd Qu.:3.000   
## Max. :4.000 Max. :4.000   
## Mjob Fjob reason guardian   
## Length:649 Length:649 Length:649 Length:649   
## 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:649   
## 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 :1.569 Mean :1.931 Mean :0.2219   
## 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:649 Length:649 Length:649 Length:649   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## higher internet romantic famrel   
## Length:649 Length:649 Length:649 Min. :1.000   
## Class :character Class :character Class :character 1st Qu.:4.000   
## Mode :character Mode :character Mode :character Median :4.000   
## Mean :3.931   
## 3rd Qu.:5.000   
## Max. :5.000   
## freetime goout Dalc Walc health   
## Min. :1.00 Min. :1.000 Min. :1.000 Min. :1.00 Min. :1.000   
## 1st Qu.:3.00 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:2.000   
## Median :3.00 Median :3.000 Median :1.000 Median :2.00 Median :4.000   
## Mean :3.18 Mean :3.185 Mean :1.502 Mean :2.28 Mean :3.536   
## 3rd Qu.:4.00 3rd Qu.:4.000 3rd Qu.:2.000 3rd Qu.:3.00 3rd Qu.:5.000   
## Max. :5.00 Max. :5.000 Max. :5.000 Max. :5.00 Max. :5.000   
## absences G1 G2 G3   
## Min. : 0.000 Min. : 0.0 Min. : 0.00 Min. : 0.00   
## 1st Qu.: 0.000 1st Qu.:10.0 1st Qu.:10.00 1st Qu.:10.00   
## Median : 2.000 Median :11.0 Median :11.00 Median :12.00   
## Mean : 3.659 Mean :11.4 Mean :11.57 Mean :11.91   
## 3rd Qu.: 6.000 3rd Qu.:13.0 3rd Qu.:13.00 3rd Qu.:14.00   
## Max. :32.000 Max. :19.0 Max. :19.00 Max. :19.00

Por: 649 rows, 33 Columns

## 'data.frame': 649 obs. of 33 variables:  
## $ school : Factor w/ 2 levels "GP","MS": 1 1 1 1 1 1 1 1 1 1 ...  
## $ sex : Factor w/ 2 levels "F","M": 1 1 1 1 1 2 2 1 2 2 ...  
## $ age : num 18 17 15 15 16 16 16 17 15 15 ...  
## $ address : Factor w/ 2 levels "U","R": 1 1 1 1 1 1 1 1 1 1 ...  
## $ famsize : Factor w/ 2 levels "LE3","GT3": 2 2 1 2 2 1 1 2 1 2 ...  
## $ Pstatus : Factor w/ 2 levels "T","A": 2 1 1 1 1 1 1 2 2 1 ...  
## $ Medu : Factor w/ 5 levels "None","Primary Education (4th grade)",..: 5 2 2 5 4 5 3 5 4 4 ...  
## $ Fedu : Factor w/ 5 levels "None","Primary Education (4th grade)",..: 5 2 2 3 4 4 3 5 3 5 ...  
## $ Mjob : Factor w/ 5 levels "teacher","health",..: 4 4 4 2 5 3 5 5 3 5 ...  
## $ Fjob : Factor w/ 5 levels "teacher","health",..: 1 5 5 3 5 5 5 1 5 5 ...  
## $ reason : Factor w/ 4 levels "home","reputation",..: 3 3 4 1 1 2 1 1 1 1 ...  
## $ guardian : Factor w/ 3 levels "mother","father",..: 1 2 1 1 2 1 1 1 1 1 ...  
## $ traveltime: Factor w/ 4 levels "<15 min.","15 to 30 min.",..: 2 1 1 1 1 1 1 2 1 1 ...  
## $ studytime : Factor w/ 4 levels "<2 hours","2 to 5 hours",..: 2 2 2 3 2 2 2 2 2 2 ...  
## $ failures : num 0 0 0 0 0 0 0 0 0 0 ...  
## $ schoolsup : Factor w/ 2 levels "yes","no": 1 2 1 2 2 2 2 1 2 2 ...  
## $ famsup : Factor w/ 2 levels "yes","no": 2 1 2 1 1 1 2 1 1 1 ...  
## $ paid : Factor w/ 2 levels "yes","no": 2 2 2 2 2 2 2 2 2 2 ...  
## $ activities: Factor w/ 2 levels "yes","no": 2 2 2 1 2 1 2 2 2 1 ...  
## $ nursery : Factor w/ 2 levels "yes","no": 1 2 1 1 1 1 1 1 1 1 ...  
## $ higher : Factor w/ 2 levels "yes","no": 1 1 1 1 1 1 1 1 1 1 ...  
## $ internet : Factor w/ 2 levels "yes","no": 2 1 1 1 2 1 1 2 1 1 ...  
## $ romantic : Factor w/ 2 levels "yes","no": 2 2 2 1 2 2 2 2 2 2 ...  
## $ famrel : Factor w/ 5 levels "very bad","bad",..: 4 5 4 3 4 5 4 4 4 5 ...  
## $ freetime : Factor w/ 5 levels "very low","low",..: 3 3 3 2 3 4 4 1 2 5 ...  
## $ goout : Factor w/ 5 levels "very low","low",..: 4 3 2 2 2 2 4 4 2 1 ...  
## $ Dalc : Factor w/ 5 levels "very low","low",..: 1 1 2 1 1 1 1 1 1 1 ...  
## $ Walc : Factor w/ 5 levels "very low","low",..: 1 1 3 1 2 2 1 1 1 1 ...  
## $ health : Factor w/ 5 levels "very bad","bad",..: 3 3 3 5 5 5 3 1 1 5 ...  
## $ absences : num 4 2 6 0 0 6 0 2 0 0 ...  
## $ G1 : num 0 9 12 14 11 12 13 10 15 12 ...  
## $ G2 : num 11 11 13 14 13 12 12 13 16 12 ...  
## $ G3 : num 11 11 12 14 13 13 13 13 17 13 ...

#### Merge data(get students who take both classes)

merged: 382 rows, 53 columns

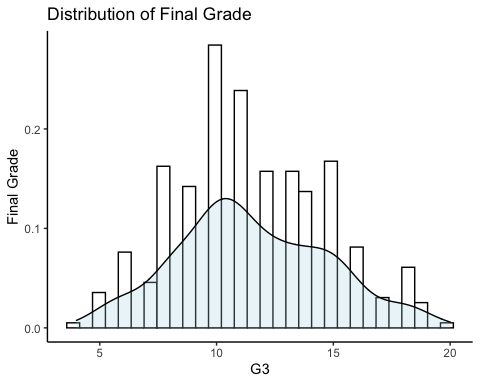
## school sex age address famsize Pstatus Medu  
## 1 GP F 15 R GT3 T 5th to 9th grade  
## 2 GP F 15 R GT3 T 5th to 9th grade  
## 3 GP F 15 R GT3 T Primary Education (4th grade)  
## 4 GP F 15 R GT3 T Primary Education (4th grade)  
## 5 GP F 15 R GT3 T Secondary Education  
## 6 GP F 15 R GT3 T Secondary Education  
## Fedu Mjob Fjob reason nursery internet  
## 1 5th to 9th grade at\_home other reputation yes no  
## 2 Higher Education services health course yes yes  
## 3 Primary Education (4th grade) at\_home other home yes yes  
## 4 Primary Education (4th grade) other other reputation no yes  
## 5 Higher Education services health course yes yes  
## 6 Higher Education services teacher course yes yes  
## guardian\_math traveltime\_math studytime\_math failures\_math schoolsup\_math  
## 1 mother <15 min. <2 hours 0 yes  
## 2 mother <15 min. 5 to 10 hours 0 yes  
## 3 mother 15 to 30 min. >10 hours 1 yes  
## 4 mother <15 min. 2 to 5 hours 2 yes  
## 5 mother <15 min. 5 to 10 hours 0 yes  
## 6 father 15 to 30 min. 5 to 10 hours 2 no  
## famsup\_math paid\_math activities\_math higher\_math romantic\_math famrel\_math  
## 1 yes yes yes yes no good  
## 2 yes yes yes yes no good  
## 3 yes yes yes yes no fair  
## 4 yes no no yes yes fair  
## 5 yes yes yes yes no good  
## 6 yes no no yes yes good  
## freetime\_math goout\_math Dalc\_math Walc\_math health\_math absences\_math  
## 1 neutral very low very low very low bad 8  
## 2 neutral low very low very low very good 2  
## 3 very low low very low very low very bad 2  
## 4 neutral high low high very good 2  
## 5 neutral low very low very low very good 2  
## 6 low low low low very good 0  
## G1\_math G2\_math G3\_math guardian\_por traveltime\_por studytime\_por  
## 1 14 13 13 mother <15 min. <2 hours  
## 2 10 9 8 mother <15 min. 5 to 10 hours  
## 3 7 10 10 mother 15 to 30 min. >10 hours  
## 4 8 6 5 mother <15 min. 2 to 5 hours  
## 5 12 12 11 mother <15 min. 5 to 10 hours  
## 6 12 0 0 father 15 to 30 min. 5 to 10 hours  
## failures\_por schoolsup\_por famsup\_por paid\_por activities\_por higher\_por  
## 1 0 yes yes no yes yes  
## 2 0 yes yes no yes yes  
## 3 0 yes yes yes yes yes  
## 4 0 yes yes no no yes  
## 5 0 yes yes no yes yes  
## 6 0 no yes no no yes  
## romantic\_por famrel\_por freetime\_por goout\_por Dalc\_por Walc\_por health\_por  
## 1 no good neutral very low very low very low bad  
## 2 no good neutral low very low very low very good  
## 3 no fair very low low very low very low very bad  
## 4 yes fair neutral high low high very good  
## 5 no good neutral low very low very low very good  
## 6 yes good low low low low very good  
## absences\_por G1\_por G2\_por G3\_por  
## 1 8 14 13 12  
## 2 2 10 11 10  
## 3 4 13 13 13  
## 4 2 13 11 11  
## 5 2 11 12 12  
## 6 0 10 11 12

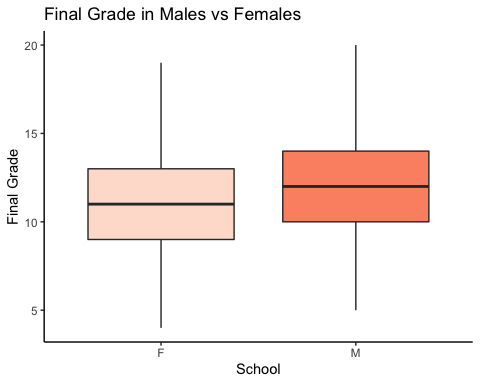
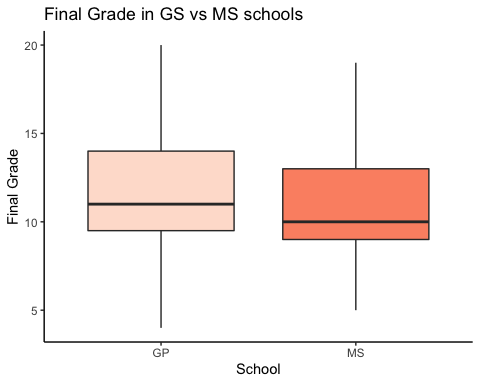
## [1] 382 53

## 'data.frame': 382 obs. of 53 variables:  
## $ school : Factor w/ 2 levels "GP","MS": 1 1 1 1 1 1 1 1 1 1 ...  
## $ sex : Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 1 ...  
## $ age : num 15 15 15 15 15 15 15 15 15 15 ...  
## $ address : Factor w/ 2 levels "U","R": 2 2 2 2 2 2 2 2 2 1 ...  
## $ famsize : Factor w/ 2 levels "LE3","GT3": 2 2 2 2 2 2 2 1 1 2 ...  
## $ Pstatus : Factor w/ 2 levels "T","A": 1 1 1 1 1 1 1 1 1 2 ...  
## $ Medu : Factor w/ 5 levels "None","Primary Education (4th grade)",..: 3 3 2 2 4 4 4 3 4 5 ...  
## $ Fedu : Factor w/ 5 levels "None","Primary Education (4th grade)",..: 3 5 2 2 5 5 4 3 2 4 ...  
## $ Mjob : Factor w/ 5 levels "teacher","health",..: 4 3 4 5 3 3 3 2 5 3 ...  
## $ Fjob : Factor w/ 5 levels "teacher","health",..: 5 2 5 5 2 1 3 3 5 3 ...  
## $ reason : Factor w/ 4 levels "home","reputation",..: 2 3 1 2 3 3 2 2 2 2 ...  
## $ nursery : Factor w/ 2 levels "yes","no": 1 1 1 2 1 1 1 1 2 1 ...  
## $ internet : Factor w/ 2 levels "yes","no": 2 1 1 1 1 1 1 1 1 1 ...  
## $ guardian\_math : Factor w/ 3 levels "mother","father",..: 1 1 1 1 1 2 3 1 2 1 ...  
## $ traveltime\_math: Factor w/ 4 levels "<15 min.","15 to 30 min.",..: 1 1 2 1 1 2 2 2 2 1 ...  
## $ studytime\_math : Factor w/ 4 levels "<2 hours","2 to 5 hours",..: 1 3 4 2 3 3 3 2 4 2 ...  
## $ failures\_math : num 0 0 1 2 0 2 2 0 0 0 ...  
## $ schoolsup\_math : Factor w/ 2 levels "yes","no": 1 1 1 1 1 2 2 1 2 2 ...  
## $ famsup\_math : Factor w/ 2 levels "yes","no": 1 1 1 1 1 1 1 1 1 1 ...  
## $ paid\_math : Factor w/ 2 levels "yes","no": 1 1 1 2 1 2 1 1 2 1 ...  
## $ activities\_math: Factor w/ 2 levels "yes","no": 1 1 1 2 1 2 1 2 2 1 ...  
## $ higher\_math : Factor w/ 2 levels "yes","no": 1 1 1 1 1 1 1 1 1 1 ...  
## $ romantic\_math : Factor w/ 2 levels "yes","no": 2 2 2 1 2 1 1 2 2 2 ...  
## $ famrel\_math : Factor w/ 5 levels "very bad","bad",..: 4 4 3 3 4 4 4 4 4 4 ...  
## $ freetime\_math : Factor w/ 5 levels "very low","low",..: 3 3 1 3 3 2 2 1 4 3 ...  
## $ goout\_math : Factor w/ 5 levels "very low","low",..: 1 2 2 4 2 2 1 3 2 2 ...  
## $ Dalc\_math : Factor w/ 5 levels "very low","low",..: 1 1 1 2 1 2 2 1 2 1 ...  
## $ Walc\_math : Factor w/ 5 levels "very low","low",..: 1 1 1 4 1 2 3 3 3 1 ...  
## $ health\_math : Factor w/ 5 levels "very bad","bad",..: 2 5 1 5 5 5 3 4 3 1 ...  
## $ absences\_math : num 8 2 2 2 2 0 8 2 12 0 ...  
## $ G1\_math : num 14 10 7 8 12 12 10 8 16 14 ...  
## $ G2\_math : num 13 9 10 6 12 0 10 9 16 15 ...  
## $ G3\_math : num 13 8 10 5 11 0 10 8 16 15 ...  
## $ guardian\_por : Factor w/ 3 levels "mother","father",..: 1 1 1 1 1 2 3 1 2 1 ...  
## $ traveltime\_por : Factor w/ 4 levels "<15 min.","15 to 30 min.",..: 1 1 2 1 1 2 2 2 2 1 ...  
## $ studytime\_por : Factor w/ 4 levels "<2 hours","2 to 5 hours",..: 1 3 4 2 3 3 3 2 4 2 ...  
## $ failures\_por : num 0 0 0 0 0 0 0 0 0 0 ...  
## $ schoolsup\_por : Factor w/ 2 levels "yes","no": 1 1 1 1 1 2 2 1 2 2 ...  
## $ famsup\_por : Factor w/ 2 levels "yes","no": 1 1 1 1 1 1 1 1 1 1 ...  
## $ paid\_por : Factor w/ 2 levels "yes","no": 2 2 1 2 2 2 1 2 2 2 ...  
## $ activities\_por : Factor w/ 2 levels "yes","no": 1 1 1 2 1 2 1 2 2 1 ...  
## $ higher\_por : Factor w/ 2 levels "yes","no": 1 1 1 1 1 1 1 1 1 1 ...  
## $ romantic\_por : Factor w/ 2 levels "yes","no": 2 2 2 1 2 1 1 2 2 2 ...  
## $ famrel\_por : Factor w/ 5 levels "very bad","bad",..: 4 4 3 3 4 4 4 4 4 4 ...  
## $ freetime\_por : Factor w/ 5 levels "very low","low",..: 3 3 1 3 3 2 2 1 4 3 ...  
## $ goout\_por : Factor w/ 5 levels "very low","low",..: 1 2 2 4 2 2 1 3 2 2 ...  
## $ Dalc\_por : Factor w/ 5 levels "very low","low",..: 1 1 1 2 1 2 2 1 2 1 ...  
## $ Walc\_por : Factor w/ 5 levels "very low","low",..: 1 1 1 4 1 2 3 3 3 1 ...  
## $ health\_por : Factor w/ 5 levels "very bad","bad",..: 2 5 1 5 5 5 3 4 3 1 ...  
## $ absences\_por : num 8 2 4 2 2 0 2 0 6 0 ...  
## $ G1\_por : num 14 10 13 13 11 10 13 11 15 15 ...  
## $ G2\_por : num 13 11 13 11 12 11 13 10 15 14 ...  
## $ G3\_por : num 12 10 13 11 12 12 13 11 15 15 ...

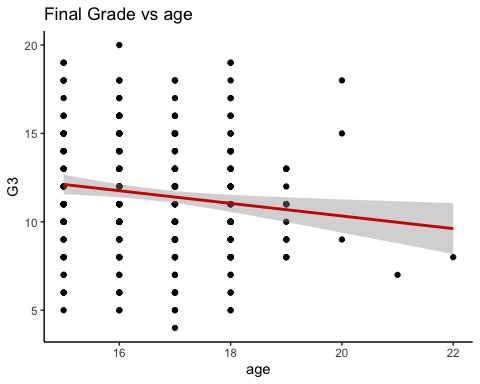
#### EDA MATH

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

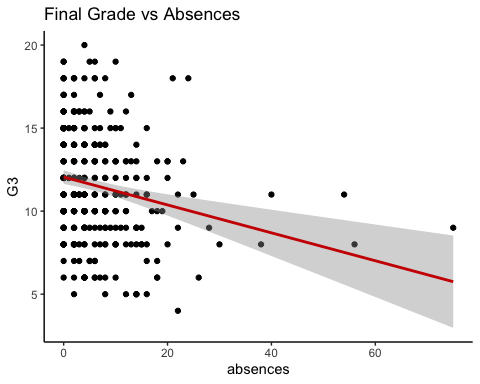




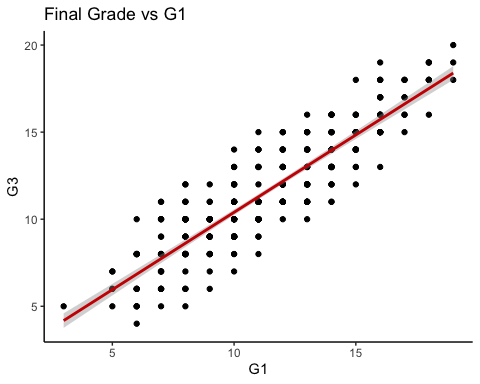
## `geom\_smooth()` using formula 'y ~ x'



## `geom\_smooth()` using formula 'y ~ x'



## `geom\_smooth()` using formula 'y ~ x'



## `geom\_smooth()` using formula 'y ~ x'



## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : at -0.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : radius 0.000225

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : all data on boundary of neighborhood. make span bigger

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : pseudoinverse used at -0.015

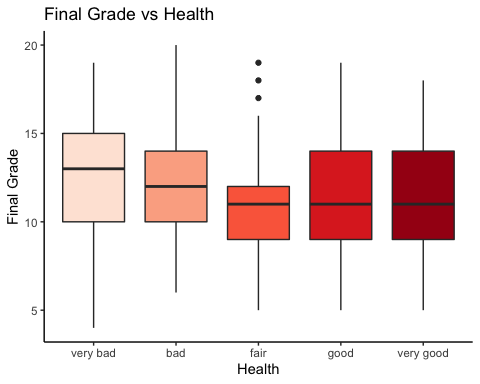
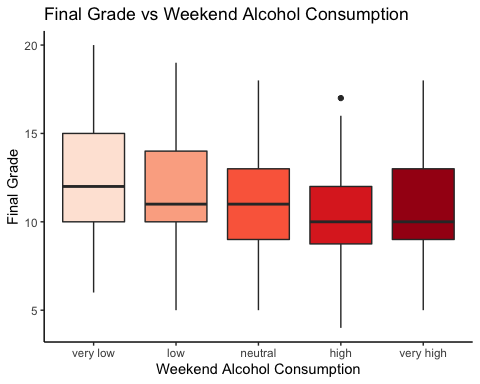
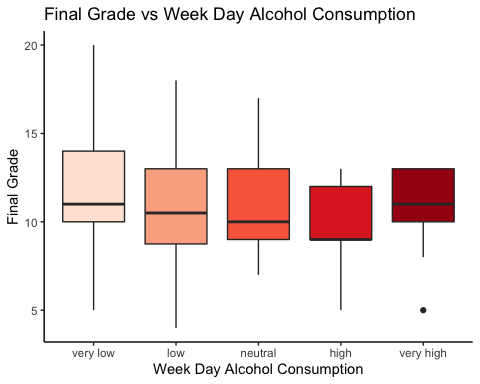
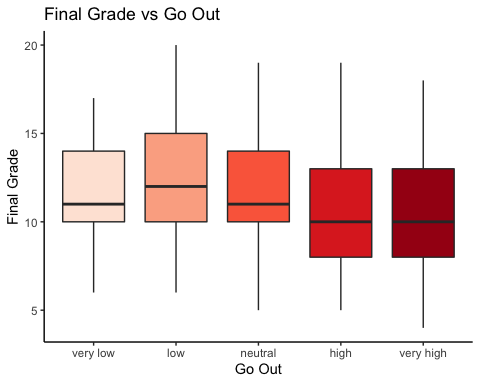
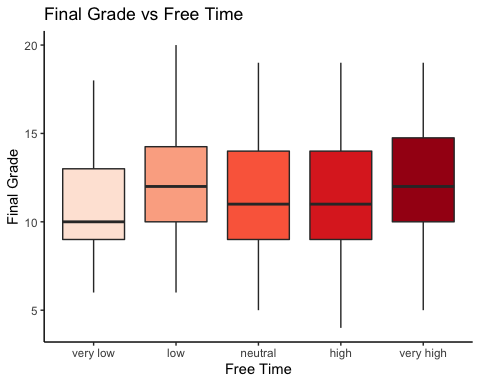
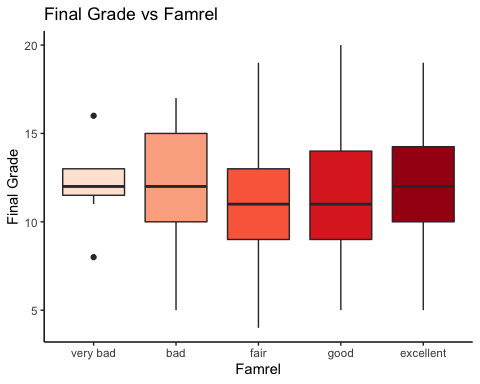
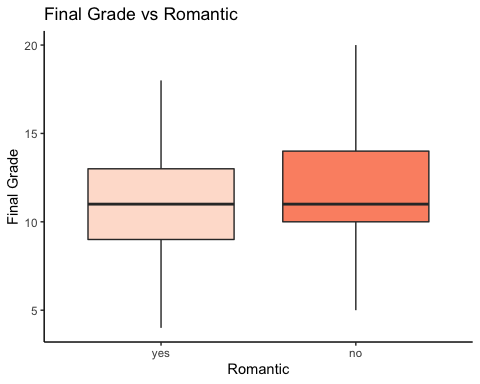
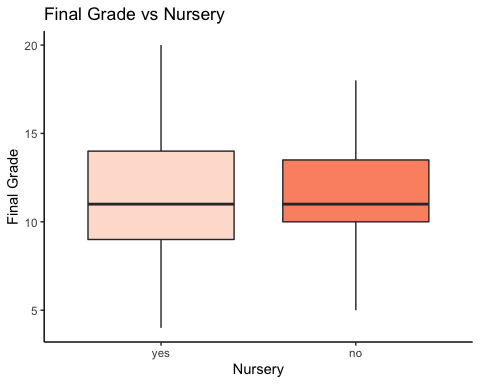
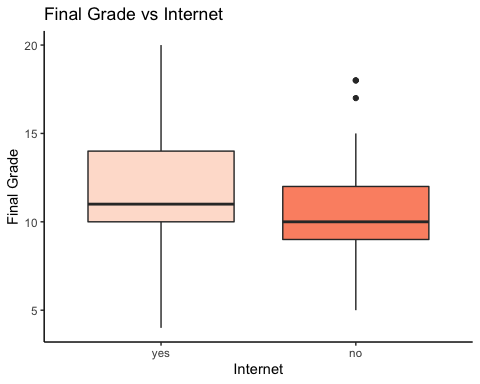
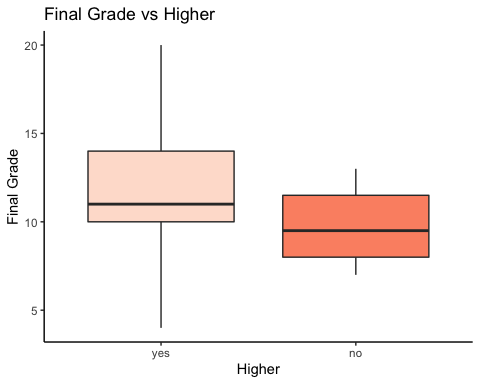
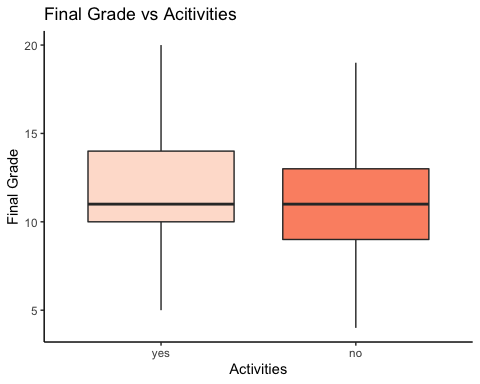
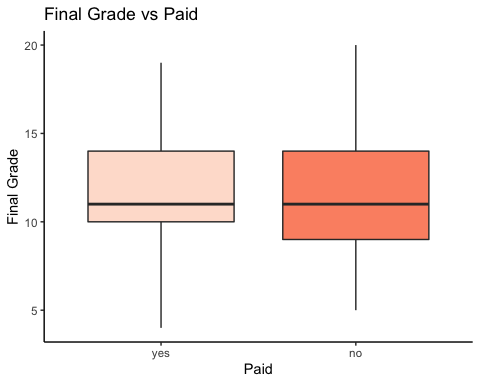
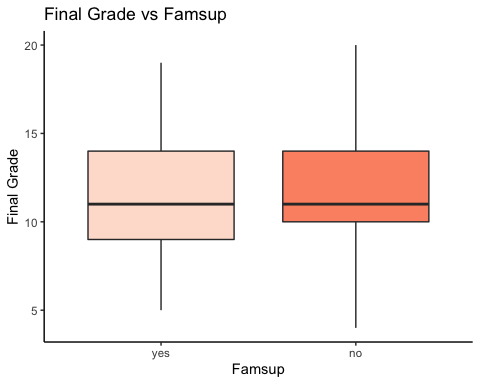
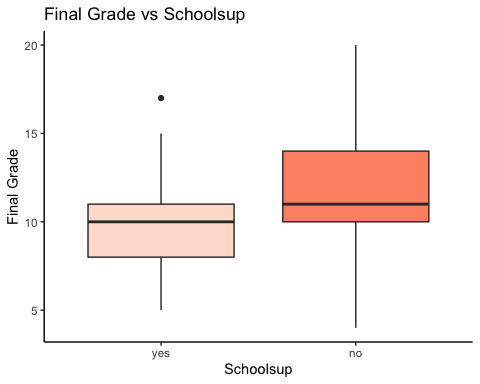
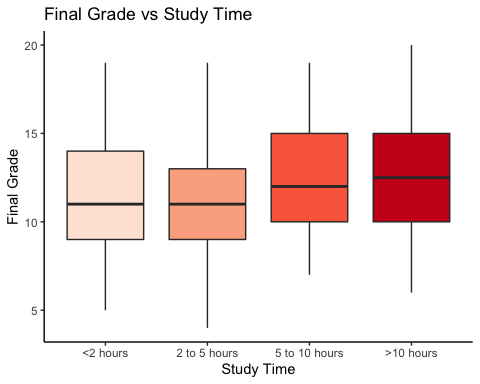
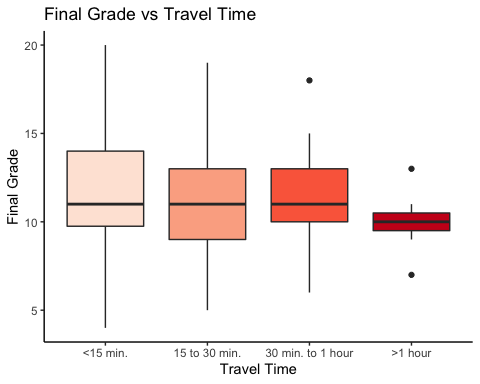
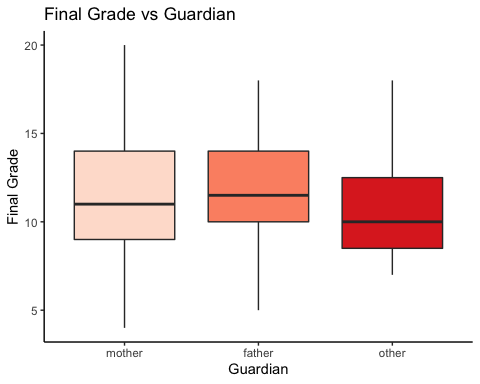
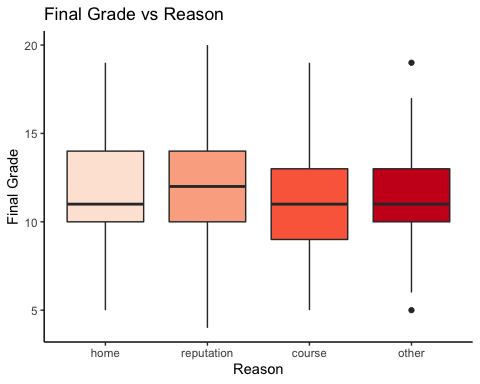
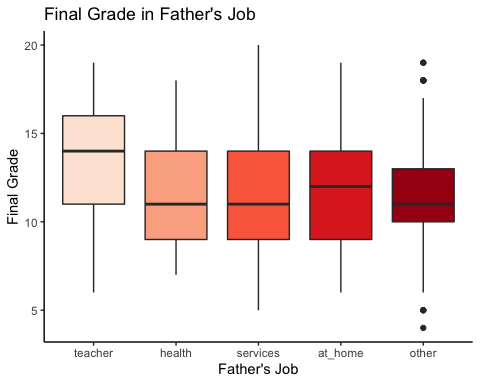
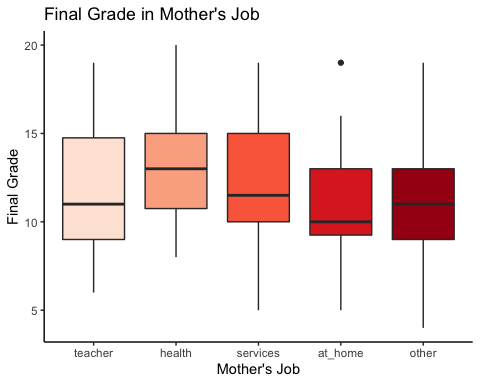
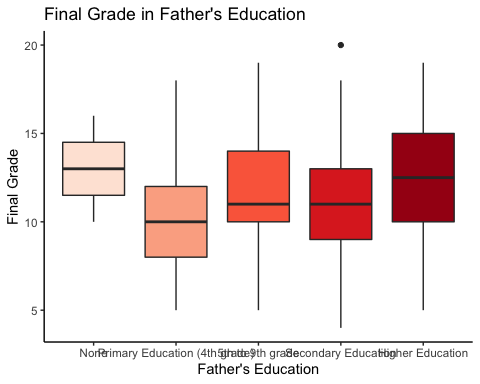
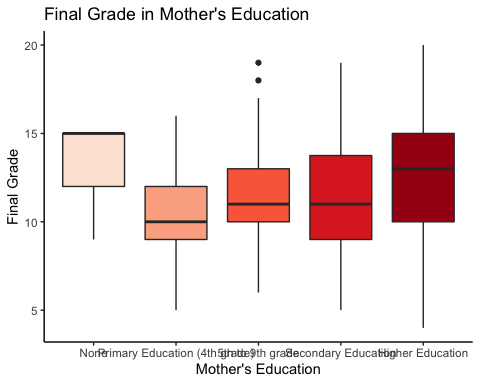
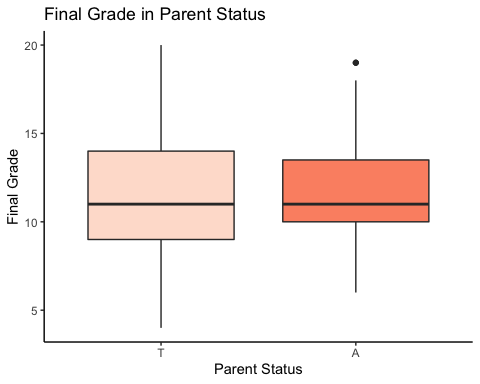
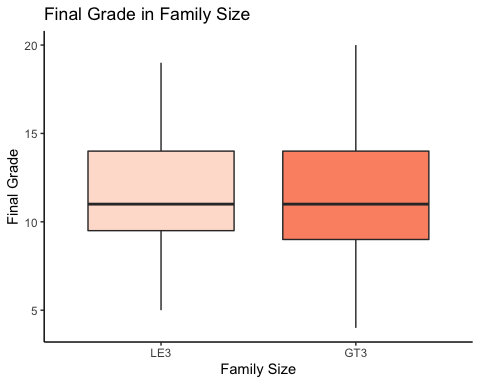
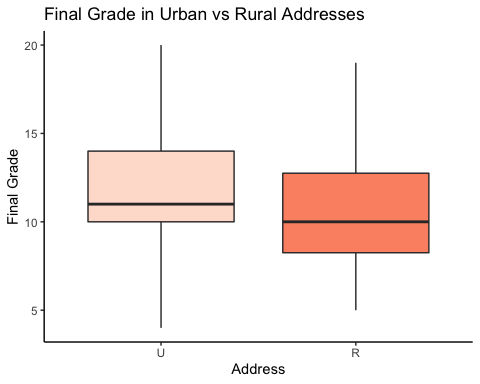
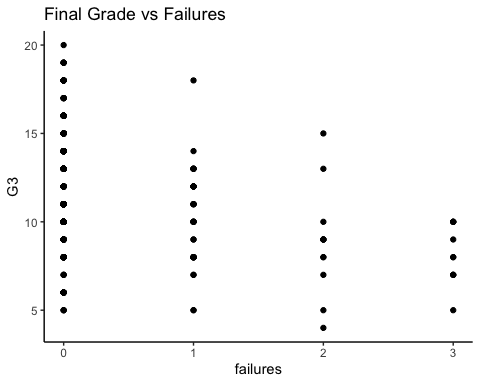
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : neighborhood radius 0.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : reciprocal condition number 1

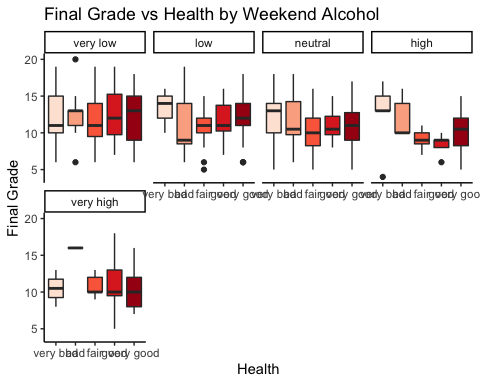
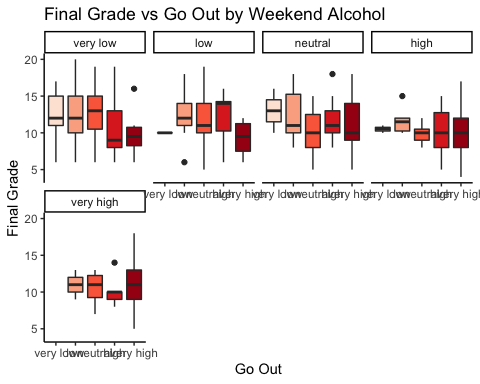
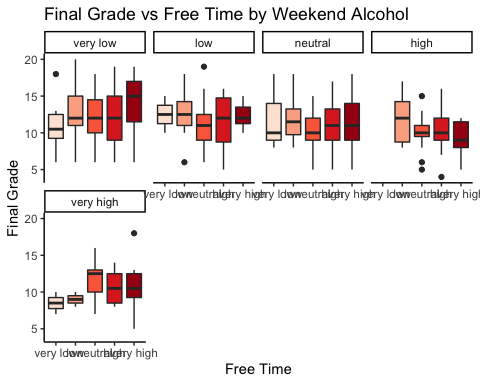
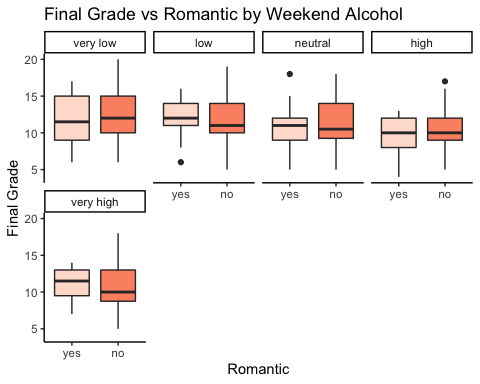
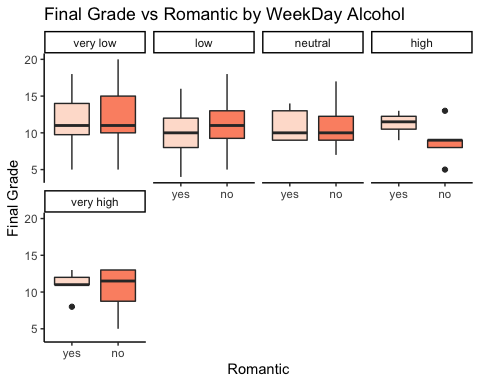
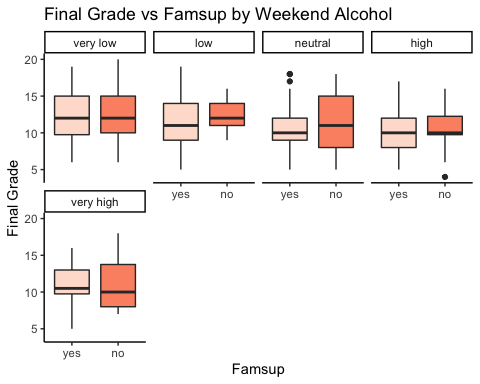
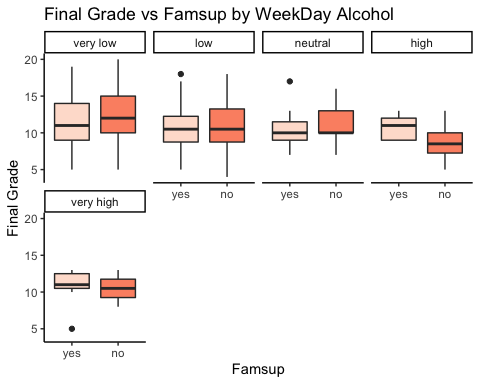
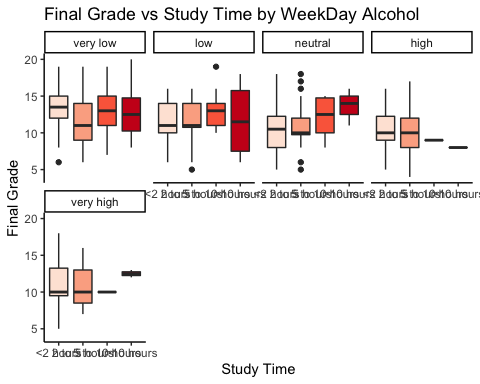
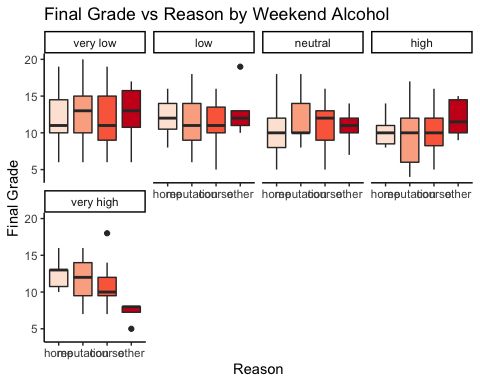
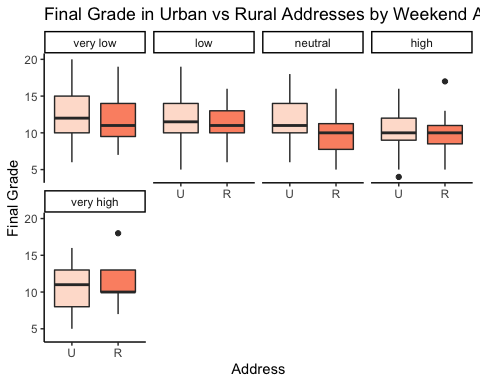
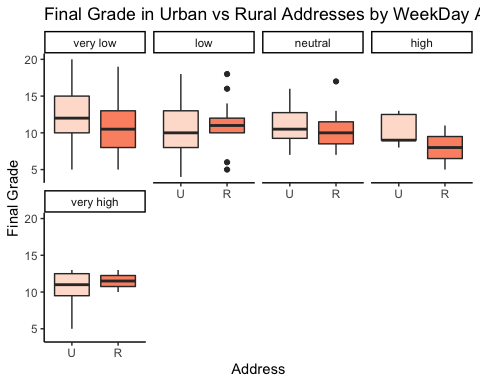
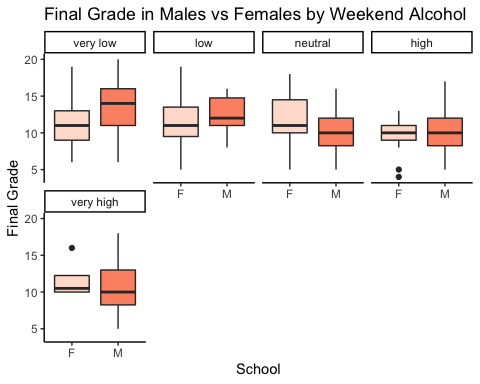
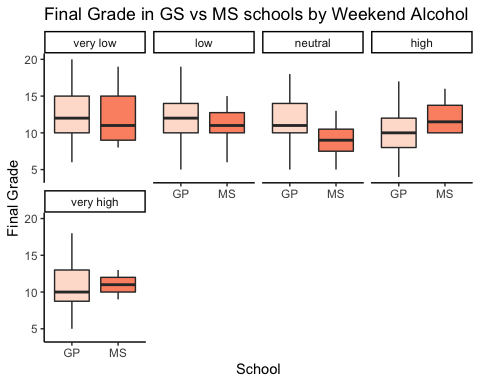
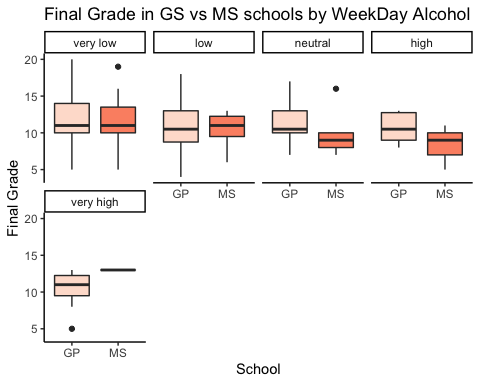
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : zero-width neighborhood. make span bigger

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =  
## parametric, : There are other near singularities as well. 1

## Warning: Computation failed in `stat\_smooth()`:  
## NA/NaN/Inf in foreign function call (arg 5)



## Interactions



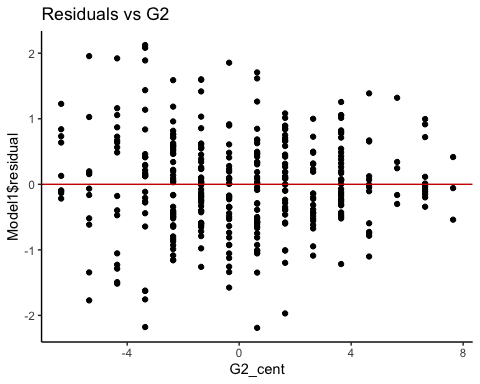
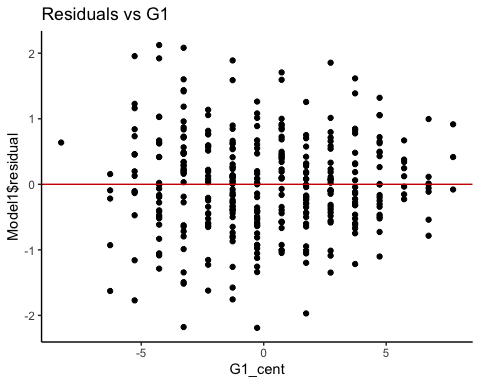
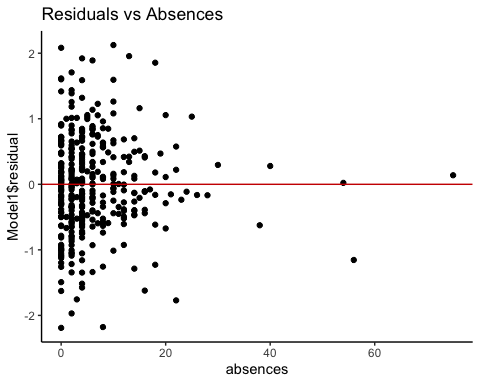
# Fit linear model

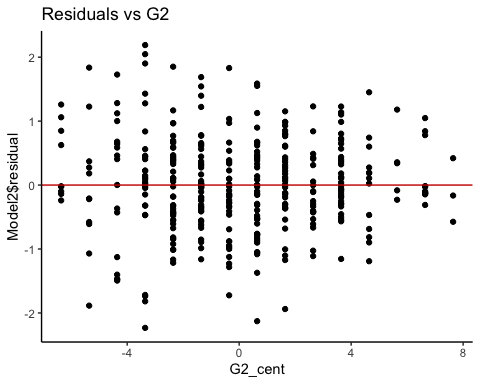
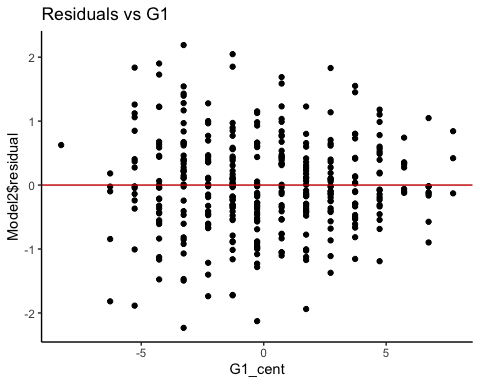
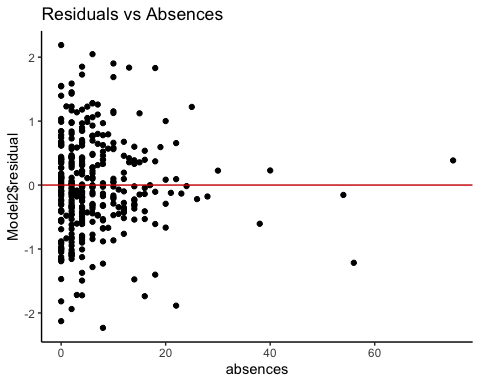
##   
## Call:  
## lm(formula = G3 ~ ., data = math)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.1902 -0.4719 -0.0007 0.4809 2.1244   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 11.548043 1.010482 11.428 < 2e-16 \*\*\*  
## schoolMS -0.009186 0.167169 -0.055 0.95621   
## sexM -0.055561 0.110682 -0.502 0.61606   
## addressR -0.163192 0.129723 -1.258 0.20941   
## famsizeGT3 0.072178 0.107443 0.672 0.50226   
## PstatusA 0.134431 0.159605 0.842 0.40033   
## MeduPrimary Education (4th grade) -0.165904 0.525529 -0.316 0.75247   
## Medu5th to 9th grade -0.179731 0.525792 -0.342 0.73273   
## MeduSecondary Education -0.307018 0.533259 -0.576 0.56524   
## MeduHigher Education -0.302011 0.546766 -0.552 0.58113   
## FeduPrimary Education (4th grade) 0.233562 0.628938 0.371 0.71064   
## Fedu5th to 9th grade 0.172070 0.629855 0.273 0.78490   
## FeduSecondary Education 0.145956 0.630716 0.231 0.81716   
## FeduHigher Education 0.149426 0.641579 0.233 0.81600   
## Mjobhealth 0.019930 0.209567 0.095 0.92430   
## Mjobservices -0.159567 0.181422 -0.880 0.37984   
## Mjobat\_home -0.215241 0.235681 -0.913 0.36186   
## Mjobother -0.383583 0.191343 -2.005 0.04593 \*   
## Fjobhealth -0.141397 0.282199 -0.501 0.61672   
## Fjobservices -0.132479 0.222338 -0.596 0.55175   
## Fjobat\_home -0.328752 0.299758 -1.097 0.27367   
## Fjobother -0.002116 0.215947 -0.010 0.99219   
## reasonreputation -0.059705 0.129542 -0.461 0.64522   
## reasoncourse -0.185114 0.123747 -1.496 0.13577   
## reasonother -0.169408 0.182243 -0.930 0.35337   
## guardianfather -0.001935 0.117756 -0.016 0.98690   
## guardianother -0.241294 0.189527 -1.273 0.20399   
## traveltime15 to 30 min. -0.094050 0.112432 -0.836 0.40356   
## traveltime30 min. to 1 hour -0.129129 0.223634 -0.577 0.56411   
## traveltime>1 hour 0.785733 0.381879 2.058 0.04053 \*   
## studytime2 to 5 hours 0.002820 0.123055 0.023 0.98173   
## studytime5 to 10 hours 0.014223 0.170919 0.083 0.93374   
## studytime>10 hours 0.129816 0.222321 0.584 0.55974   
## schoolsupno 0.197680 0.144395 1.369 0.17205   
## famsupno -0.097530 0.104885 -0.930 0.35321   
## paidno 0.126844 0.105299 1.205 0.22934   
## activitiesno 0.049903 0.099327 0.502 0.61576   
## nurseryno 0.199223 0.123372 1.615 0.10744   
## higherno 0.106109 0.265256 0.400 0.68943   
## internetno -0.021443 0.135924 -0.158 0.87476   
## romanticno -0.059356 0.106884 -0.555 0.57910   
## famrelbad 0.507241 0.420996 1.205 0.22924   
## famrelfair 0.617554 0.361713 1.707 0.08884 .   
## famrelgood 0.556036 0.350291 1.587 0.11353   
## famrelexcellent 0.938903 0.355245 2.643 0.00867 \*\*   
## freetimelow 0.141127 0.253314 0.557 0.57787   
## freetimeneutral 0.143913 0.242355 0.594 0.55310   
## freetimehigh 0.063608 0.251049 0.253 0.80016   
## freetimevery high 0.114819 0.286086 0.401 0.68846   
## gooutlow -0.248612 0.228417 -1.088 0.27732   
## gooutneutral -0.290941 0.226831 -1.283 0.20065   
## goouthigh -0.449221 0.238263 -1.885 0.06038 .   
## gooutvery high -0.535220 0.259153 -2.065 0.03979 \*   
## Dalclow 0.027719 0.144053 0.192 0.84755   
## Dalcneutral 0.310145 0.224070 1.384 0.16738   
## Dalchigh 0.069997 0.338287 0.207 0.83622   
## Dalcvery high -0.469389 0.393583 -1.193 0.23400   
## Walclow -0.119590 0.137130 -0.872 0.38388   
## Walcneutral -0.136103 0.149402 -0.911 0.36306   
## Walchigh -0.139986 0.192718 -0.726 0.46819   
## Walcvery high 0.131175 0.305371 0.430 0.66784   
## healthbad 0.143647 0.198522 0.724 0.46991   
## healthfair -0.121120 0.173899 -0.696 0.48668   
## healthgood -0.017801 0.184389 -0.097 0.92316   
## healthvery good -0.261000 0.161444 -1.617 0.10704   
## absences -0.008877 0.006427 -1.381 0.16825   
## G1\_cent 0.101960 0.035818 2.847 0.00473 \*\*   
## G2\_cent 0.877459 0.037072 23.669 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.819 on 289 degrees of freedom  
## Multiple R-squared: 0.9477, Adjusted R-squared: 0.9356   
## F-statistic: 78.22 on 67 and 289 DF, p-value: < 2.2e-16

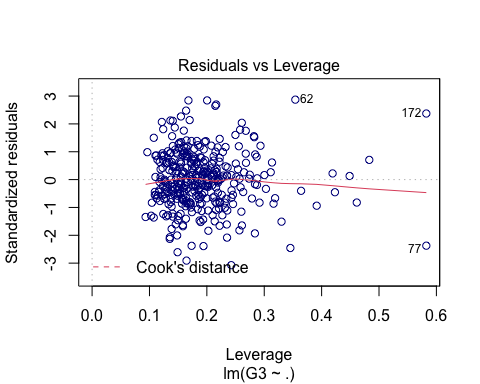
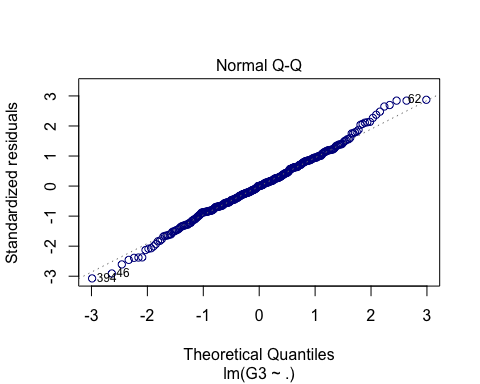
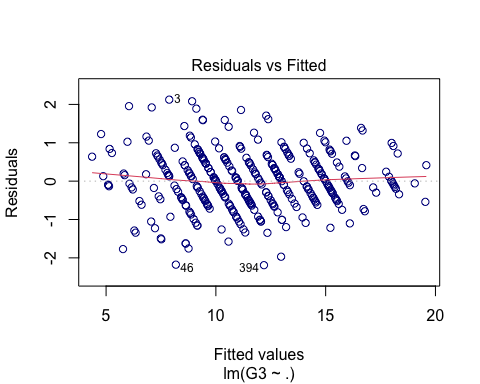
# Fit EDA model and significant values

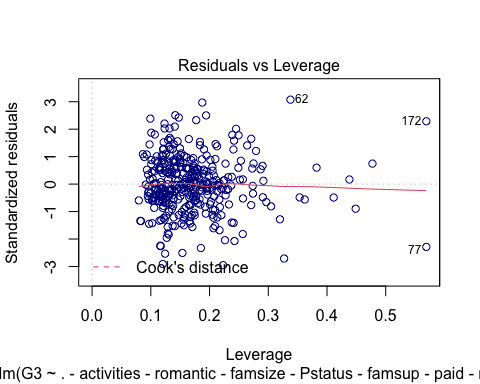
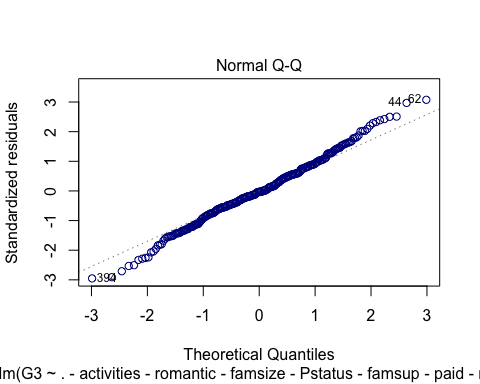
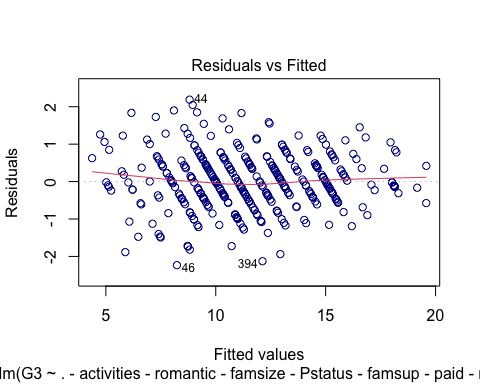
##   
## Call:  
## lm(formula = G3 ~ . - activities - romantic - famsize - Pstatus -   
## famsup - paid - nursery, data = math)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.23331 -0.42823 -0.02331 0.44154 2.18961   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 11.6782317 0.9830753 11.879 < 2e-16 \*\*\*  
## schoolMS -0.0026040 0.1628145 -0.016 0.98725   
## sexM -0.0684995 0.1082971 -0.633 0.52754   
## addressR -0.1475076 0.1269605 -1.162 0.24624   
## MeduPrimary Education (4th grade) -0.2066674 0.5223748 -0.396 0.69266   
## Medu5th to 9th grade -0.2386351 0.5196867 -0.459 0.64643   
## MeduSecondary Education -0.3178001 0.5285647 -0.601 0.54813   
## MeduHigher Education -0.3275643 0.5398733 -0.607 0.54449   
## FeduPrimary Education (4th grade) 0.2551283 0.6199093 0.412 0.68096   
## Fedu5th to 9th grade 0.1832651 0.6188120 0.296 0.76732   
## FeduSecondary Education 0.1446380 0.6230539 0.232 0.81659   
## FeduHigher Education 0.1554386 0.6326916 0.246 0.80610   
## Mjobhealth 0.0809120 0.2070668 0.391 0.69626   
## Mjobservices -0.1002446 0.1792705 -0.559 0.57646   
## Mjobat\_home -0.1462585 0.2315430 -0.632 0.52809   
## Mjobother -0.2983124 0.1846984 -1.615 0.10735   
## Fjobhealth -0.1894034 0.2801874 -0.676 0.49958   
## Fjobservices -0.2044744 0.2194724 -0.932 0.35227   
## Fjobat\_home -0.4392081 0.2942699 -1.493 0.13662   
## Fjobother -0.0705843 0.2134844 -0.331 0.74116   
## reasonreputation -0.0455313 0.1282560 -0.355 0.72284   
## reasoncourse -0.1488524 0.1212019 -1.228 0.22037   
## reasonother -0.1705699 0.1802886 -0.946 0.34487   
## guardianfather -0.0001776 0.1167316 -0.002 0.99879   
## guardianother -0.1698164 0.1862707 -0.912 0.36269   
## traveltime15 to 30 min. -0.1146632 0.1104367 -1.038 0.29999   
## traveltime30 min. to 1 hour -0.1078792 0.2206323 -0.489 0.62524   
## traveltime>1 hour 0.7428506 0.3761563 1.975 0.04921 \*   
## studytime2 to 5 hours 0.0038697 0.1187264 0.033 0.97402   
## studytime5 to 10 hours 0.0105958 0.1672208 0.063 0.94952   
## studytime>10 hours 0.0953779 0.2174388 0.439 0.66124   
## schoolsupno 0.1982128 0.1433578 1.383 0.16782   
## higherno 0.1587213 0.2633670 0.603 0.54720   
## internetno -0.0193997 0.1337474 -0.145 0.88477   
## famrelbad 0.4498798 0.4146261 1.085 0.27879   
## famrelfair 0.5702248 0.3572038 1.596 0.11148   
## famrelgood 0.4765309 0.3452764 1.380 0.16858   
## famrelexcellent 0.8798309 0.3501926 2.512 0.01252 \*   
## freetimelow 0.1707929 0.2487470 0.687 0.49286   
## freetimeneutral 0.1914127 0.2370136 0.808 0.41997   
## freetimehigh 0.1036012 0.2454950 0.422 0.67332   
## freetimevery high 0.1715568 0.2769890 0.619 0.53615   
## gooutlow -0.2319888 0.2258162 -1.027 0.30510   
## gooutneutral -0.2843619 0.2251032 -1.263 0.20749   
## goouthigh -0.4644570 0.2359889 -1.968 0.04999 \*   
## gooutvery high -0.5500973 0.2574906 -2.136 0.03347 \*   
## Dalclow 0.0075849 0.1424646 0.053 0.95758   
## Dalcneutral 0.2810625 0.2204502 1.275 0.20333   
## Dalchigh -0.0218432 0.3288080 -0.066 0.94708   
## Dalcvery high -0.3814729 0.3885934 -0.982 0.32706   
## Walclow -0.1347396 0.1341568 -1.004 0.31603   
## Walcneutral -0.1459727 0.1464278 -0.997 0.31963   
## Walchigh -0.1312703 0.1897246 -0.692 0.48954   
## Walcvery high 0.1648774 0.2997012 0.550 0.58264   
## healthbad 0.1653923 0.1965699 0.841 0.40081   
## healthfair -0.0898301 0.1722185 -0.522 0.60234   
## healthgood -0.0100769 0.1837796 -0.055 0.95631   
## healthvery good -0.2317663 0.1593652 -1.454 0.14692   
## absences -0.0076671 0.0062974 -1.218 0.22438   
## G1\_cent 0.1018777 0.0354172 2.877 0.00431 \*\*   
## G2\_cent 0.8755647 0.0368167 23.782 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.8181 on 296 degrees of freedom  
## Multiple R-squared: 0.9466, Adjusted R-squared: 0.9358   
## F-statistic: 87.44 on 60 and 296 DF, p-value: < 2.2e-16

# Model Assessment



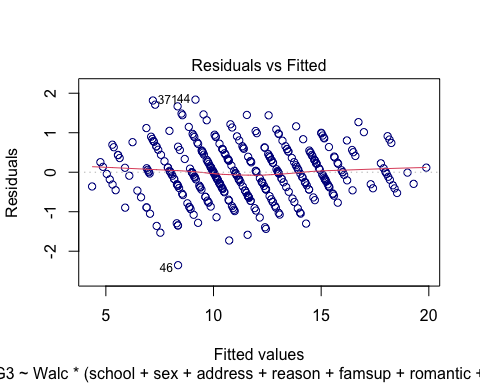




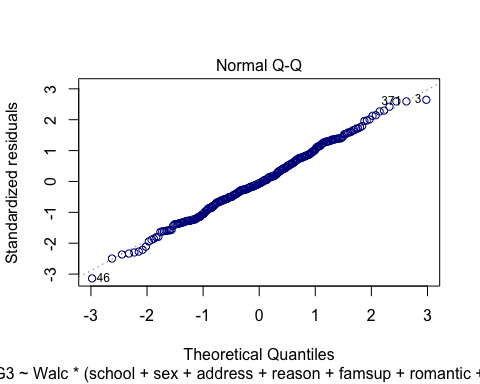


# Model with interactions

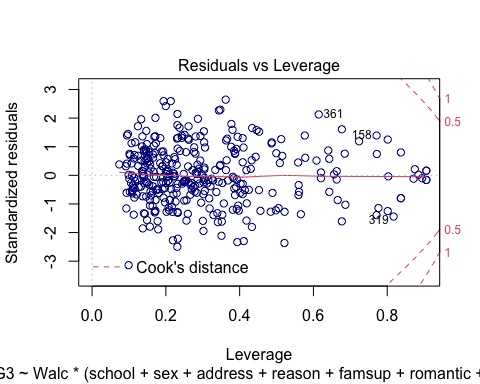
##   
## Call:  
## lm(formula = G3 ~ Walc \* (school + sex + address + reason + famsup +   
## romantic + goout + health) + Dalc \* (school + studytime +   
## romantic) + famrel + traveltime + G1\_cent + G2\_cent, data = math)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.35305 -0.39476 -0.01764 0.41917 1.83778   
##   
## Coefficients: (4 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 12.08444 0.54106 22.335 < 2e-16 \*\*\*  
## Walclow -1.89498 0.81791 -2.317 0.02135 \*   
## Walcneutral -0.03735 0.87814 -0.043 0.96611   
## Walchigh -2.14343 1.06124 -2.020 0.04452 \*   
## Walcvery high -1.26744 3.45864 -0.366 0.71435   
## schoolMS -0.18544 0.30945 -0.599 0.54957   
## sexM 0.01601 0.16092 0.099 0.92085   
## addressR -0.43406 0.20584 -2.109 0.03601 \*   
## reasonreputation -0.04535 0.20301 -0.223 0.82343   
## reasoncourse -0.14270 0.18575 -0.768 0.44310   
## reasonother -0.10158 0.30565 -0.332 0.73993   
## famsupno -0.12129 0.15517 -0.782 0.43516   
## romanticno -0.10433 0.16449 -0.634 0.52651   
## gooutlow -0.47639 0.26990 -1.765 0.07883 .   
## gooutneutral -0.42055 0.27513 -1.529 0.12770   
## goouthigh -0.61249 0.30375 -2.016 0.04487 \*   
## gooutvery high -0.77823 0.42305 -1.840 0.06707 .   
## healthbad 0.43445 0.29745 1.461 0.14545   
## healthfair -0.24832 0.23438 -1.059 0.29045   
## healthgood -0.20962 0.26369 -0.795 0.42744   
## healthvery good -0.50281 0.23036 -2.183 0.03003 \*   
## Dalclow 0.31195 0.38098 0.819 0.41370   
## Dalcneutral 2.03336 0.90945 2.236 0.02628 \*   
## Dalchigh 2.46031 1.60275 1.535 0.12609   
## Dalcvery high 1.07888 1.18796 0.908 0.36469   
## studytime2 to 5 hours 0.08170 0.15729 0.519 0.60394   
## studytime5 to 10 hours 0.17049 0.19988 0.853 0.39452   
## studytime>10 hours 0.22763 0.24617 0.925 0.35606   
## famrelbad 0.25582 0.48059 0.532 0.59501   
## famrelfair 0.46457 0.41889 1.109 0.26852   
## famrelgood 0.22478 0.41261 0.545 0.58641   
## famrelexcellent 0.65473 0.41836 1.565 0.11891   
## traveltime15 to 30 min. -0.07033 0.11700 -0.601 0.54832   
## traveltime30 min. to 1 hour -0.07173 0.24077 -0.298 0.76603   
## traveltime>1 hour 2.37781 0.79208 3.002 0.00297 \*\*   
## G1\_cent 0.10602 0.03769 2.813 0.00532 \*\*   
## G2\_cent 0.89305 0.03979 22.442 < 2e-16 \*\*\*  
## Walclow:schoolMS -0.03656 0.48597 -0.075 0.94010   
## Walcneutral:schoolMS -0.10391 0.53180 -0.195 0.84525   
## Walchigh:schoolMS -0.20686 0.68171 -0.303 0.76182   
## Walcvery high:schoolMS 0.77242 2.09553 0.369 0.71275   
## Walclow:sexM -0.28973 0.28970 -1.000 0.31826   
## Walcneutral:sexM -0.11150 0.27966 -0.399 0.69046   
## Walchigh:sexM 0.18143 0.35515 0.511 0.60993   
## Walcvery high:sexM -2.48869 1.03324 -2.409 0.01677 \*   
## Walclow:addressR 0.39817 0.33981 1.172 0.24247   
## Walcneutral:addressR 0.40666 0.33834 1.202 0.23058   
## Walchigh:addressR 0.50961 0.51513 0.989 0.32352   
## Walcvery high:addressR -0.65564 0.73140 -0.896 0.37093   
## Walclow:reasonreputation 0.27133 0.35267 0.769 0.44244   
## Walcneutral:reasonreputation 0.10105 0.34407 0.294 0.76925   
## Walchigh:reasonreputation -0.54462 0.53931 -1.010 0.31358   
## Walcvery high:reasonreputation 3.98741 2.70920 1.472 0.14238   
## Walclow:reasoncourse 0.17747 0.33126 0.536 0.59263   
## Walcneutral:reasoncourse 0.24886 0.33795 0.736 0.46221   
## Walchigh:reasoncourse -0.03506 0.50156 -0.070 0.94433   
## Walcvery high:reasoncourse 0.97641 2.02946 0.481 0.63087   
## Walclow:reasonother 0.51905 0.55131 0.941 0.34740   
## Walcneutral:reasonother 0.61905 0.50361 1.229 0.22019   
## Walchigh:reasonother -1.13088 0.70199 -1.611 0.10851   
## Walcvery high:reasonother 1.27567 2.37103 0.538 0.59106   
## Walclow:famsupno 0.20342 0.28404 0.716 0.47459   
## Walcneutral:famsupno 0.20923 0.26465 0.791 0.42997   
## Walchigh:famsupno -0.07519 0.37566 -0.200 0.84153   
## Walcvery high:famsupno 1.20545 0.82319 1.464 0.14440   
## Walclow:romanticno -0.12199 0.29451 -0.414 0.67909   
## Walcneutral:romanticno 0.13052 0.30032 0.435 0.66424   
## Walchigh:romanticno 0.43691 0.41579 1.051 0.29441   
## Walcvery high:romanticno 0.60250 0.85170 0.707 0.48000   
## Walclow:gooutlow 1.51169 0.72138 2.096 0.03717 \*   
## Walcneutral:gooutlow -0.47251 0.73125 -0.646 0.51879   
## Walchigh:gooutlow 1.70632 0.84283 2.025 0.04403 \*   
## Walcvery high:gooutlow -3.32531 1.82686 -1.820 0.06997 .   
## Walclow:gooutneutral 1.18397 0.71838 1.648 0.10064   
## Walcneutral:gooutneutral -0.43762 0.68890 -0.635 0.52588   
## Walchigh:gooutneutral 0.85473 0.87170 0.981 0.32781   
## Walcvery high:gooutneutral -0.12126 0.86505 -0.140 0.88864   
## Walclow:goouthigh 1.60958 0.78062 2.062 0.04029 \*   
## Walcneutral:goouthigh -0.67386 0.71755 -0.939 0.34861   
## Walchigh:goouthigh 1.27840 0.78692 1.625 0.10557   
## Walcvery high:goouthigh 1.30003 0.94959 1.369 0.17226   
## Walclow:gooutvery high 1.79958 0.87861 2.048 0.04163 \*   
## Walcneutral:gooutvery high -0.21778 0.80296 -0.271 0.78645   
## Walchigh:gooutvery high 1.13416 0.83751 1.354 0.17694   
## Walcvery high:gooutvery high NA NA NA NA   
## Walclow:healthbad -0.44578 0.53815 -0.828 0.40830   
## Walcneutral:healthbad -0.50265 0.51896 -0.969 0.33374   
## Walchigh:healthbad 0.42359 0.79950 0.530 0.59673   
## Walcvery high:healthbad 0.07549 3.65735 0.021 0.98355   
## Walclow:healthfair 0.53592 0.42147 1.272 0.20477   
## Walcneutral:healthfair -0.01345 0.45119 -0.030 0.97624   
## Walchigh:healthfair 0.37327 0.71450 0.522 0.60186   
## Walcvery high:healthfair 3.82072 3.17856 1.202 0.23054   
## Walclow:healthgood 0.38672 0.47857 0.808 0.41985   
## Walcneutral:healthgood 0.07155 0.48363 0.148 0.88251   
## Walchigh:healthgood 1.09278 0.71520 1.528 0.12785   
## Walcvery high:healthgood 2.05987 1.94016 1.062 0.28944   
## Walclow:healthvery good 0.52464 0.40637 1.291 0.19793   
## Walcneutral:healthvery good 0.44345 0.43767 1.013 0.31198   
## Walchigh:healthvery good 0.65960 0.58139 1.135 0.25771   
## Walcvery high:healthvery good -0.30817 1.41618 -0.218 0.82792   
## schoolMS:Dalclow 0.39701 0.48067 0.826 0.40965   
## schoolMS:Dalcneutral 0.13358 0.59292 0.225 0.82194   
## schoolMS:Dalchigh -1.98921 0.99388 -2.001 0.04647 \*   
## schoolMS:Dalcvery high 5.30625 3.25308 1.631 0.10417   
## Dalclow:studytime2 to 5 hours -0.66411 0.33337 -1.992 0.04749 \*   
## Dalcneutral:studytime2 to 5 hours 0.43961 0.46659 0.942 0.34705   
## Dalchigh:studytime2 to 5 hours -1.95944 1.18851 -1.649 0.10053   
## Dalcvery high:studytime2 to 5 hours 2.38215 2.58183 0.923 0.35711   
## Dalclow:studytime5 to 10 hours -1.16931 0.44767 -2.612 0.00957 \*\*   
## Dalcneutral:studytime5 to 10 hours -1.71979 1.23879 -1.388 0.16634   
## Dalchigh:studytime5 to 10 hours NA NA NA NA   
## Dalcvery high:studytime5 to 10 hours NA NA NA NA   
## Dalclow:studytime>10 hours -0.64270 0.72637 -0.885 0.37714   
## Dalcneutral:studytime>10 hours 0.55342 3.80396 0.145 0.88445   
## Dalchigh:studytime>10 hours NA NA NA NA   
## Dalcvery high:studytime>10 hours -6.27293 3.17508 -1.976 0.04934 \*   
## romanticno:Dalclow 0.36169 0.37066 0.976 0.33014   
## romanticno:Dalcneutral -2.07919 0.88529 -2.349 0.01966 \*   
## romanticno:Dalchigh -0.70220 1.19130 -0.589 0.55612   
## romanticno:Dalcvery high -2.59229 1.57938 -1.641 0.10204   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7896 on 240 degrees of freedom  
## Multiple R-squared: 0.9597, Adjusted R-squared: 0.9402   
## F-statistic: 49.21 on 116 and 240 DF, p-value: < 2.2e-16



## Warning: not plotting observations with leverage one:  
## 67, 90, 109, 183, 188, 210, 218, 226, 299, 316, 353



## Warning: not plotting observations with leverage one:  
## 67, 90, 109, 183, 188, 210, 218, 226, 299, 316, 353



# Stepwise AIC

## lm(formula = G3 ~ G2\_cent + G1\_cent + famrel + traveltime + absences +   
## health + nursery, data = math)

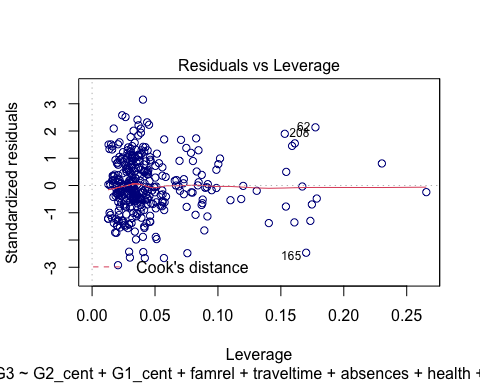
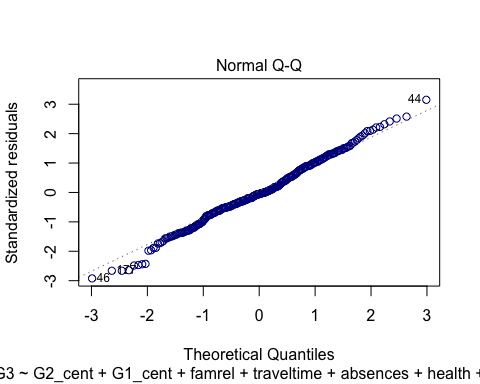
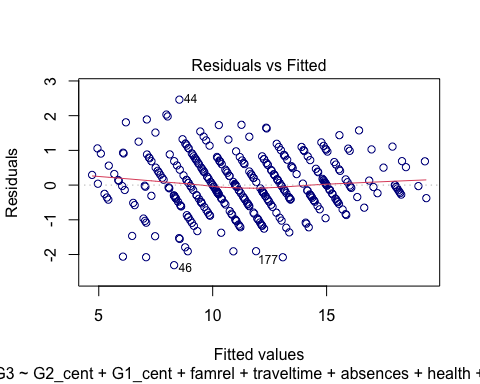
# Stepwise BIC

## lm(formula = G3 ~ G2\_cent + G1\_cent, data = math)

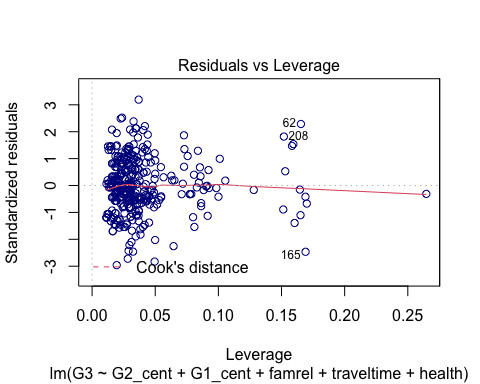
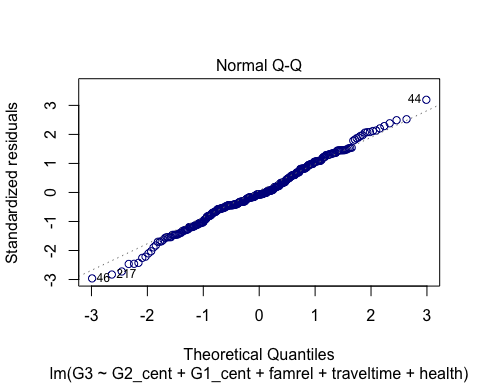
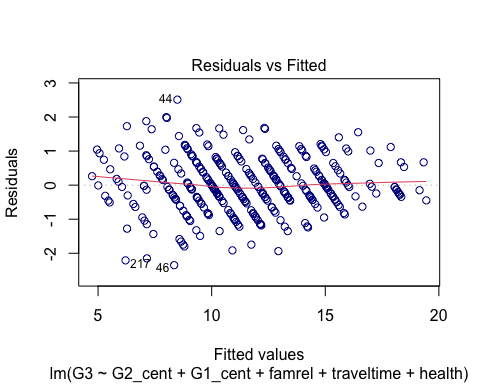
## lm(formula = G3 ~ G2\_cent + G1\_cent + famrel + traveltime + health,   
## data = math)

## lm(formula = G3 ~ G2\_cent + G1\_cent, data = math)

##   
## Call:  
## lm(formula = G3 ~ G2\_cent + G1\_cent + famrel + traveltime + absences +   
## health + nursery, data = math)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.30670 -0.44169 -0.04326 0.51731 2.46160   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 11.163853 0.318174 35.087 < 2e-16 \*\*\*  
## G2\_cent 0.882506 0.032528 27.131 < 2e-16 \*\*\*  
## G1\_cent 0.113831 0.030985 3.674 0.000277 \*\*\*  
## famrelbad 0.214518 0.374586 0.573 0.567239   
## famrelfair 0.428864 0.328878 1.304 0.193105   
## famrelgood 0.327852 0.317759 1.032 0.302915   
## famrelexcellent 0.710589 0.321745 2.209 0.027871 \*   
## traveltime15 to 30 min. -0.132920 0.097950 -1.357 0.175673   
## traveltime30 min. to 1 hour -0.114905 0.192035 -0.598 0.549999   
## traveltime>1 hour 0.753300 0.311806 2.416 0.016220 \*   
## absences -0.009209 0.005355 -1.720 0.086399 .   
## healthbad 0.236410 0.178874 1.322 0.187167   
## healthfair 0.018544 0.152437 0.122 0.903249   
## healthgood 0.047006 0.162209 0.290 0.772157   
## healthvery good -0.158036 0.140378 -1.126 0.261046   
## nurseryno 0.149641 0.107234 1.395 0.163785   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7975 on 341 degrees of freedom  
## Multiple R-squared: 0.9415, Adjusted R-squared: 0.939   
## F-statistic: 366 on 15 and 341 DF, p-value: < 2.2e-16



##   
## Call:  
## lm(formula = G3 ~ G2\_cent + G1\_cent + famrel + traveltime + health,   
## data = math)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.34876 -0.43483 -0.05419 0.53296 2.50952   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 11.11191 0.31387 35.403 < 2e-16 \*\*\*  
## G2\_cent 0.89475 0.03187 28.074 < 2e-16 \*\*\*  
## G1\_cent 0.10544 0.03079 3.425 0.000689 \*\*\*  
## famrelbad 0.20250 0.37558 0.539 0.590124   
## famrelfair 0.44667 0.32994 1.354 0.176693   
## famrelgood 0.32772 0.31840 1.029 0.304069   
## famrelexcellent 0.72828 0.32243 2.259 0.024531 \*   
## traveltime15 to 30 min. -0.13072 0.09828 -1.330 0.184383   
## traveltime30 min. to 1 hour -0.09670 0.19264 -0.502 0.616009   
## traveltime>1 hour 0.76666 0.31286 2.451 0.014763 \*   
## healthbad 0.23855 0.17964 1.328 0.185097   
## healthfair 0.04445 0.15261 0.291 0.771047   
## healthgood 0.06031 0.16274 0.371 0.711162   
## healthvery good -0.14172 0.14076 -1.007 0.314739   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.801 on 343 degrees of freedom  
## Multiple R-squared: 0.9407, Adjusted R-squared: 0.9384   
## F-statistic: 418.3 on 13 and 343 DF, p-value: < 2.2e-16



## Analysis of Variance Table  
##   
## Model 1: G3 ~ G2\_cent + G1\_cent + famrel + traveltime + absences + health +   
## nursery  
## Model 2: G3 ~ G2\_cent + G1\_cent + famrel + traveltime + health  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 341 216.89   
## 2 343 220.05 -2 -3.1606 2.4846 0.08487 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## G2\_cent G1\_cent   
## 5.865895 5.642656   
## famrelbad famrelfair   
## 3.170094 8.600848   
## famrelgood famrelexcellent   
## 14.168388 11.423163   
## traveltime15 to 30 min. traveltime30 min. to 1 hour   
## 1.051686 1.042995   
## traveltime>1 hour absences   
## 1.049029 1.075920   
## healthbad healthfair   
## 1.708121 2.327334   
## healthgood healthvery good   
## 2.009547 2.585512   
## nurseryno   
## 1.028351

## G2\_cent G1\_cent   
## 5.582840 5.522313   
## famrelbad famrelfair   
## 3.159591 8.582092   
## famrelgood famrelexcellent   
## 14.103301 11.373741   
## traveltime15 to 30 min. traveltime30 min. to 1 hour   
## 1.049754 1.040538   
## traveltime>1 hour healthbad   
## 1.047037 1.708070   
## healthfair healthgood   
## 2.312597 2.005416   
## healthvery good   
## 2.577399