

# Statistic Model to Analyze Student's Performance - Group 8

Xuchuan Zheng, Sungki Park, Prashant Sharma

29 March, 2024

## 1. *Introduction*

In our final project for Data 603 - Statistical Modelling with Data, we have tried to develop a model to analyze the impact of various demographic and social factors on the performance of students. Academic performance, though it is not the only factor but is one of the crucial factors in shaping a student's future. To get into a good college/university, student must score grades in school, a good college can lead a better future and economic stability. So in order to secure good grades, getting into a great school is enough? Is there something more than a great school that can help a student to perform better? Do the social and demographic factors play any role in student's performance? In our project we are trying to answer these questions.

To answer these questions we are working with a dataset that is collected at 2 Portuguese schools for Mathematics and Portuguese subject. This data is collected by using school reports and questionnaires. The data attributes include students' grades, family size information, education level of parents, free time of student, and many other factors. By working on this project we are hoping to develop more understanding about the factors which can impact the performance of a student.

## 2. *Data*

This data is from [UC Irvine Machine Learning Repository](#). There are 649 rows instances and 30 features in the dataset. Below are details of each feature

1. school - student's school (binary: 'GP' - Gabriel Pereira or 'MS' - Mousinho da Silveira) [Qualitative]
2. sex - student's sex (binary: 'F' - female or 'M' - male) [Qualitative]
3. age - student's age (numeric: from 15 to 22)
4. address - student's home address type (binary: 'U' - urban or 'R' - rural) [Qualitative]
5. famsize - family size (binary: 'LE3' - less or equal to 3 or 'GT3' - greater than 3) [Qualitative]
6. Pstatus - parent's cohabitation status (binary: 'T' - living together or 'A' - apart) [Qualitative]
7. Medu - mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education) [Qualitative]
8. Fedu - father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education) [Qualitative]
9. Mjob - mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at\_home' or 'other') [Qualitative]

10. Fjob - father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at\_home' or 'other') [Qualitative]
11. reason - reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other') [Qualitative]
12. guardian - student's guardian (nominal: 'mother', 'father' or 'other') [Qualitative]
13. traveltime - home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour) [Qualitative]
14. studytime - weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours) [Qualitative]
15. failures - number of past class failures (numeric: n if  $1 \leq n < 3$ , else 4) [Qualitative]
16. schoolsup - extra educational support (binary: yes or no) [Qualitative]
17. famsup - family educational support (binary: yes or no) [Qualitative]
18. paid - extra paid classes within the course subject (Math or Portuguese) (binary: yes or no) [Qualitative]
19. activities - extra-curricular activities (binary: yes or no) [Qualitative]
20. nursery - attended nursery school (binary: yes or no) [Qualitative]
21. higher - wants to take higher education (binary: yes or no) [Qualitative]
22. internet - Internet access at home (binary: yes or no) [Qualitative]
23. romantic - with a romantic relationship (binary: yes or no) [Qualitative]
24. famrel - quality of family relationships (numeric: from 1 - very bad to 5 - excellent) [Qualitative]
25. freetime - free time after school (numeric: from 1 - very low to 5 - very high) [Qualitative]
26. goout - going out with friends (numeric: from 1 - very low to 5 - very high) [Qualitative]
27. Dalc - workday alcohol consumption (numeric: from 1 - very low to 5 - very high) [Qualitative]
28. Walc - weekend alcohol consumption (numeric: from 1 - very low to 5 - very high) [Qualitative]
29. health - current health status (numeric: from 1 - very bad to 5 - very good) [Qualitative]
30. absences - number of school absences (numeric: from 0 to 93) [Quantitative]
31. G1 - first period grade (numeric: from 0 to 20) [Quantitative]
32. G2 - second period grade (numeric: from 0 to 20) [Quantitative]
33. G3 - final grade (numeric: from 0 to 20, output target) [Quantitative]

**G3 (Final grade) is the dependent variable for our model.**

NOTE: We need to convert the qualitative variable from numeric to string

*# Converting the data type of factor variable from int to char*

```
studentDataset$Medu= as.character(studentDataset$Medu)
studentDataset$traveltime= as.character(studentDataset$traveltime)
studentDataset$Fedu= as.character(studentDataset$Fedu)
studentDataset$studytime= as.character(studentDataset$studytime)
studentDataset$famrel= as.character(studentDataset$famrel)
studentDataset$freetime= as.character(studentDataset$freetime)
```

```

studentDataset$goout= as.character(studentDataset$goout)
studentDataset$Walc= as.character(studentDataset$Walc)
studentDataset$Dalc= as.character(studentDataset$Dalc)
studentDataset$health= as.character(studentDataset$health)

head(studentDataset, 3)

```

```

##   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
##   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
##   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
##   absences G1 G2 G3
## 1      4  0 11 11
## 2      2  9 11 11
## 3      6 12 13 12

```

### 3. Methodology

Below is the outline of the steps we are going to perform in our analysis:

1. First Build a full additive model.
2. We will apply some model selection technique to come up with the best additive model.
3. Based on p-value (assuming  $\alpha = 0.05$ ) we will drop variable which are non-significant.
4. Perform partial F-test to verify that dropped variables are indeed non-significant.
5. Provide interpretation for the best additive model to predict our dependent variable (G3 - Final Grades).
6. Based on our best additive model, we will check for interaction between the variables.
7. Using p-value (assuming  $\alpha = 0.05$ ), we will drop the non-significant interaction terms.
8. Use partial F-test and analysis of Variance to verify the usability of our best interaction model.
9. Provide interpretation of our best interaction model.
10. Then we will check if we can include any higher order term in our model (Moving towards Higher order multiple regression model).
11. Verify the significance of higher order terms using p-value (assuming  $\alpha = 0.05$ ).
12. Once we have done all our analysis we will try to define our best regression model (liner or higher order) to predict our dependent variable (G3 - Final Grades).
13. Using our final regression model, we will start checking the regression assumptions.
14. Linearity Assumption.
15. Independence Assumption.
16. Normality Assumption.

17. Multi-collinearity.

18. Outliers.

Starting our analysis with building additive model, then we will try to include interaction terms and higher order terms in our model.

### 3.1 Full Additive Model

Creating full additive model:

```
studentPerformance_fm = lm(G3 ~ (school+sex+age+address+famsize+Pstatus+Medu+
                                Fedu+Mjob+Fjob+reason+guardian+traveltime+studytime+
                                failures+schoolsup+famsup+activities+nursery+higher+internet+
                                romantic+famrel+freetime+
                                goout+Dalc+Walc+health+absences),
                           data = studentDataset)
```

```
summary(studentPerformance_fm)
```

```
##
## Call:
## lm(formula = G3 ~ (school + sex + age + address + famsize + Pstatus +
##     Medu + Fedu + Mjob + Fjob + reason + guardian + traveltime +
##     studytime + failures + schoolsup + famsup + activities +
##     nursery + higher + internet + romantic + famrel + freetime +
##     goout + Dalc + Walc + health + absences), data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.2538  -1.4424   0.4217   2.0952   7.7895
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   11.481745    2.696281    4.258 2.26e-05 ***
## schoolMS      -0.309885    0.300992   -1.030 0.303479
## sexM          -0.104478    0.256079   -0.408 0.683368
## age           0.050072    0.104884    0.477 0.633179
## addressU       0.444262    0.278425    1.596 0.110896
## famsizeLE3     0.483175    0.250921    1.926 0.054443 .
## PstatusT      -0.110171    0.364903   -0.302 0.762778
## Medu1         -1.599656    1.214228   -1.317 0.188005
## Medu2         -1.633354    1.218213   -1.341 0.180304
## Medu3         -1.271212    1.233317   -1.031 0.302925
## Medu4         -0.693972    1.270242   -0.546 0.584964
## Fedu1         -1.170353    1.209054   -0.968 0.333289
## Fedu2         -1.083472    1.218187   -0.889 0.374000
## Fedu3         -1.324331    1.231681   -1.075 0.282541
## Fedu4         -1.153623    1.260240   -0.915 0.360208
## Mjobhealth     0.644770    0.566474    1.138 0.255309
## Mjobother     -0.097260    0.329016   -0.296 0.767593
## Mjobservices   0.573808    0.390234    1.470 0.141770
## Mjobteacher   -0.537369    0.535929   -1.003 0.316260
```

```

## Fjobhealth      -0.024084    0.756051   -0.032  0.974594
## Fjobother       0.146518    0.488355    0.300  0.764223
## Fjobservices    -0.169258    0.510364   -0.332  0.740231
## Fjobteacher     1.199449    0.693844    1.729  0.084178 .
## reasonhome      -0.118871    0.286125   -0.415  0.677903
## reasonother     -0.209278    0.384312   -0.545  0.586186
## reasonreputation 0.194450    0.300371    0.647  0.517546
## guardianmother  -0.329956    0.275944   -1.196  0.232089
## guardianother    0.118881    0.525016    0.226  0.820911
## traveltime2     -0.112876    0.256404   -0.440  0.659870
## traveltime3      0.332622    0.457432    0.727  0.467308
## traveltime4     -0.530444    0.776076   -0.683  0.494456
## studytime2       0.345119    0.274679    1.256  0.209255
## studytime3       1.189312    0.378589    3.141  0.001732 **
## studytime4       0.563988    0.520239    1.084  0.278590
## failures        -1.838328    0.185251   -9.923 < 2e-16 ***
## schoolsupyes     -1.114261    0.361618   -3.081  0.002119 **
## famsupyes        -0.296530    0.234051   -1.267  0.205476
## activitiesyes    -0.009298    0.228325   -0.041  0.967526
## nurseryyes       -0.061422    0.281003   -0.219  0.827021
## higheryes        1.190993    0.428916    2.777  0.005595 **
## internetyes      0.363775    0.293088    1.241  0.214836
## romanticyes     -0.648752    0.239575   -2.708  0.006889 **
## famrel2          0.297864    0.847050    0.352  0.725177
## famrel3          0.544597    0.722939    0.753  0.451445
## famrel4          0.956470    0.687110    1.392  0.164234
## famrel5          0.729534    0.701134    1.041  0.298363
## freetime2        0.974211    0.522148    1.866  0.062372 .
## freetime3       -0.021338    0.483631   -0.044  0.964818
## freetime4        0.332933    0.512091    0.650  0.515751
## freetime5        0.968995    0.585733    1.654  0.098382 .
## goout2           1.335736    0.480183    2.782  0.005511 **
## goout3           0.876640    0.473569    1.851  0.064452 .
## goout4           0.278556    0.501895    0.555  0.579016
## goout5          -0.079134    0.538137   -0.147  0.883121
## Dalc2            -0.709801    0.332067   -2.138  0.032803 *
## Dalc3            -0.092302    0.529267   -0.174  0.861590
## Dalc4            -2.000157    0.759479   -2.634  0.008582 **
## Dalc5            -0.617704    0.890539   -0.694  0.488080
## Walc2            -0.304902    0.309526   -0.985  0.324838
## Walc3            0.126708    0.350062    0.362  0.717461
## Walc4            -0.126274    0.442092   -0.286  0.775224
## Walc5            0.944524    0.659359    1.432  0.152324
## health2          -1.037651    0.444244   -2.336  0.019705 *
## health3          -1.300660    0.395549   -3.288  0.001044 **
## health4          -0.853415    0.411859   -2.072  0.038518 *
## health5          -1.249195    0.365081   -3.422  0.000648 ***
## absences         0.005013    0.018571    0.270  0.787265
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.393 on 977 degrees of freedom
## Multiple R-squared:  0.2781, Adjusted R-squared:  0.2293
## F-statistic: 5.703 on 66 and 977 DF,  p-value: < 2.2e-16

```

*Comments:* From the summary of the full model we can see that many of the variables are non-significant. We can apply some techniques for model selection to get significant parameters.

### 3.2 Model Selection

Using Stepwise forward selection procedure to get the significant parameters:

```
studentPerformance_Forward_subsets = ols_step_forward_p(studentPerformance_fm, p_val = 0.05, details = F)
student_performance_forwardMdl = studentPerformance_Forward_subsets$model
summary(student_performance_forwardMdl)
```

```
##
## Call:
## lm(formula = paste(response, "~", paste(preds, collapse = " + ")),
##     data = l)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-13.0125	-1.4391	0.3944	2.1255	8.2166

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	11.70133	1.38394	8.455	< 2e-16 ***
failures	-1.76047	0.17299	-10.177	< 2e-16 ***
higheryes	1.20109	0.41419	2.900	0.00381 **
schoolsupyes	-1.08027	0.34127	-3.165	0.00159 **
schoolMS	-0.52091	0.27020	-1.928	0.05415 .
romanticyes	-0.57404	0.22855	-2.512	0.01217 *
studytime2	0.33184	0.26152	1.269	0.20476
studytime3	1.19724	0.35335	3.388	0.00073 ***
studytime4	0.53768	0.48696	1.104	0.26979
addressU	0.45472	0.25698	1.769	0.07712 .
Medu1	-1.68348	1.17642	-1.431	0.15273
Medu2	-1.64019	1.16935	-1.403	0.16103
Medu3	-1.22194	1.17050	-1.044	0.29676
Medu4	-0.73276	1.16900	-0.627	0.53092
Dalc2	-0.75785	0.28374	-2.671	0.00769 **
Dalc3	-0.05715	0.45502	-0.126	0.90007
Dalc4	-1.74603	0.68863	-2.536	0.01138 *
Dalc5	0.04561	0.70762	0.064	0.94862
famsizeLE3	0.51236	0.23591	2.172	0.03010 *
freetime2	0.84199	0.50684	1.661	0.09697 .
freetime3	-0.15376	0.46711	-0.329	0.74208
freetime4	0.25650	0.48784	0.526	0.59915
freetime5	0.87384	0.56062	1.559	0.11938
goout2	1.20786	0.46546	2.595	0.00960 **
goout3	0.82742	0.45858	1.804	0.07148 .
goout4	0.40653	0.48356	0.841	0.40071
goout5	0.01356	0.50561	0.027	0.97861

```
## health2      -0.71212    0.42745   -1.666    0.09603 .
## health3      -1.11917    0.37841   -2.958    0.00317 **
## health4      -0.62825    0.39712   -1.582    0.11396
## health5      -1.03268    0.34383   -3.003    0.00273 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.392 on 1013 degrees of freedom
## Multiple R-squared:  0.252, Adjusted R-squared:  0.2298
## F-statistic: 11.38 on 30 and 1013 DF, p-value: < 2.2e-16
```

```
#studentPerformance_Foward_subsets$metric
```

Using the result of forward selection process and dropping non-significant variables, the best additive model is:

```
student_performance_addMdl = lm(G3 ~ (failures+higher+studytime+schoolsup+Dalc+health+romantic+famsize+
+goout), data = studentDataset)
summary(student_performance_addMdl)
```

```
##
## Call:
## lm(formula = G3 ~ (failures + higher + studytime + schoolsup +
##      Dalc + health + romantic + famsize + goout), data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.1848  -1.5137   0.3029   2.1295   8.6106
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.68826    0.63642   16.794 < 2e-16 ***
## failures     -1.82887    0.17261  -10.595 < 2e-16 ***
## higheryes     1.54720    0.41456    3.732 0.000200 ***
## studytime2     0.26960    0.26170    1.030 0.303179
## studytime3     1.18603    0.35442    3.346 0.000849 ***
## studytime4     0.62526    0.49164    1.272 0.203738
## schoolsupyes  -1.16127    0.34238   -3.392 0.000721 ***
## Dalc2         -0.73721    0.28608   -2.577 0.010107 *
## Dalc3         -0.14301    0.45454   -0.315 0.753108
## Dalc4        -1.65491    0.69909   -2.367 0.018107 *
## Dalc5          0.11089    0.71654    0.155 0.877041
## health2       -0.66440    0.43152   -1.540 0.123946
## health3       -1.21726    0.38251   -3.182 0.001505 **
## health4       -0.74820    0.40008   -1.870 0.061751 .
## health5       -0.99991    0.34699   -2.882 0.004038 **
## romanticyes   -0.59306    0.23040   -2.574 0.010190 *
## famsizeLE3     0.48561    0.23806    2.040 0.041625 *
## goout2         1.15597    0.46996    2.460 0.014070 *
## goout3         0.67978    0.45855    1.482 0.138527
## goout4         0.38442    0.48055    0.800 0.423919
## goout5         0.08698    0.50253    0.173 0.862617
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 3.449 on 1023 degrees of freedom
## Multiple R-squared:  0.2189, Adjusted R-squared:  0.2036
## F-statistic: 14.33 on 20 and 1023 DF,  p-value: < 2.2e-16
```

**Our best additive model:**

$$\widehat{G3} = 10.68826 - 1.82887(\text{failures}) + 1.54720(\text{higheryes}) + 0.26960(\text{studytime2}) + 1.18603(\text{studytime3}) \\ + 0.62526(\text{studytime4}) - 1.16127(\text{schoolsupyes}) - 0.73721(\text{Dalc2}) - 0.14301(\text{Dalc3}) \\ - 1.65491(\text{Dalc4}) + 0.11089(\text{Dalc5}) - 0.66440(\text{health2}) - 1.21726(\text{health3}) - 0.74820(\text{health4}) \\ - 0.99991(\text{health5}) - 0.59306(\text{romanticyes}) + 0.48561(\text{famsizeLE3}) + 1.15597(\text{goout2}) + 0.67978(\text{goout3}) + 0.38442(\text{goout4}) \\ + 0.08698(\text{goout5})$$

**Add some comment**

### 3.3 Interaction Model

Using our best additive model, we can check for interaction term:

```
student_performance_intMdl1 = lm(G3 ~ (failures+higher+studytime+schoolsup+Dalc+health+romantic+famsize
+goout)^2, data = studentDataset)
summary(student_performance_intMdl1)
```

```
##
## Call:
## lm(formula = G3 ~ (failures + higher + studytime + schoolsup +
##       Dalc + health + romantic + famsize + goout)^2, data = studentDataset)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-12.6059	-1.3914	0.1717	1.9203	6.6996

```
##
## Coefficients: (4 not defined because of singularities)
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.19551    2.81649   2.200  0.02809 *
## failures        -2.75207    1.23301  -2.232  0.02587 *
## higheryes         4.44810    2.38735   1.863  0.06278 .
## studytime2        4.39372    1.64797   2.666  0.00782 **
## studytime3        7.01518    3.50387   2.002  0.04559 *
## studytime4        3.97662    5.20509   0.764  0.44509
## schoolsupyes      0.10943    4.34254   0.025  0.97990
## Dalc2            -2.26256    2.33965  -0.967  0.33379
## Dalc3            -0.80477    3.91410  -0.206  0.83715
## Dalc4           -9.95939    8.68152  -1.147  0.25162
## Dalc5           -2.63007    5.98232  -0.440  0.66031
## health2         -0.35899    2.71301  -0.132  0.89476
## health3          2.56397    2.58846   0.991  0.32219
## health4          3.97467    2.79449   1.422  0.15530
## health5          2.52100    2.41385   1.044  0.29660
## romanticyes     -2.61645    1.58806  -1.648  0.09981 .
## famsizeLE3       0.10877    1.69830   0.064  0.94895
```



## goout2	3.90749	2.56623	1.523	0.12821
## goout3	3.76392	2.43585	1.545	0.12266
## goout4	1.21308	2.45073	0.495	0.62074
## goout5	1.70105	2.51234	0.677	0.49854
## failures:higheryes	-0.82183	0.52111	-1.577	0.11515
## failures:studytime2	-0.34101	0.45599	-0.748	0.45476
## failures:studytime3	0.06996	0.83190	0.084	0.93300
## failures:studytime4	1.31009	2.96755	0.441	0.65898
## failures:schoolsupyes	1.45452	0.73978	1.966	0.04960 *
## failures:Dalc2	-0.39955	0.47394	-0.843	0.39945
## failures:Dalc3	0.40039	0.70669	0.567	0.57116
## failures:Dalc4	0.13264	1.56996	0.084	0.93269
## failures:Dalc5	1.51596	1.20670	1.256	0.20935
## failures:health2	0.48151	1.04401	0.461	0.64477
## failures:health3	1.07117	0.83315	1.286	0.19890
## failures:health4	-0.46263	0.86941	-0.532	0.59478
## failures:health5	0.59094	0.79540	0.743	0.45772
## failures:romanticyes	-0.67921	0.48000	-1.415	0.15742
## failures:famsizeLE3	0.45925	0.47859	0.960	0.33753
## failures:goout2	0.31027	0.99617	0.311	0.75552
## failures:goout3	1.83641	0.99821	1.840	0.06616 .
## failures:goout4	1.46403	1.07575	1.361	0.17389
## failures:goout5	1.28334	0.99637	1.288	0.19809
## higheryes:studytime2	-2.03805	1.09505	-1.861	0.06307 .
## higheryes:studytime3	1.63302	2.59693	0.629	0.52963
## higheryes:studytime4	-8.78982	5.35875	-1.640	0.10131
## higheryes:schoolsupyes	0.71248	4.37144	0.163	0.87057
## higheryes:Dalc2	1.88685	1.19508	1.579	0.11474
## higheryes:Dalc3	2.64174	2.20154	1.200	0.23049
## higheryes:Dalc4	4.25797	5.00007	0.852	0.39468
## higheryes:Dalc5	4.75483	4.24757	1.119	0.26327
## higheryes:health2	-1.00878	2.15258	-0.469	0.63945
## higheryes:health3	-1.39330	2.04161	-0.682	0.49514
## higheryes:health4	-2.61984	2.12349	-1.234	0.21764
## higheryes:health5	-2.32303	1.90791	-1.218	0.22372
## higheryes:romanticyes	0.31145	1.02087	0.305	0.76038
## higheryes:famsizeLE3	-0.27038	1.17042	-0.231	0.81736
## higheryes:goout2	-1.19710	1.93327	-0.619	0.53594
## higheryes:goout3	-3.37541	1.79200	-1.884	0.05996 .
## higheryes:goout4	1.16620	1.90917	0.611	0.54147
## higheryes:goout5	0.51551	1.74393	0.296	0.76761
## studytime2:schoolsupyes	-0.27602	1.37089	-0.201	0.84048
## studytime3:schoolsupyes	-0.78356	1.66954	-0.469	0.63896
## studytime4:schoolsupyes	-2.24127	2.01097	-1.115	0.26537
## studytime2:Dalc2	-1.03569	0.78399	-1.321	0.18684
## studytime3:Dalc2	-0.83481	1.03042	-0.810	0.41807
## studytime4:Dalc2	-0.37492	2.42822	-0.154	0.87733
## studytime2:Dalc3	-0.35496	1.20886	-0.294	0.76911
## studytime3:Dalc3	-3.06285	3.23996	-0.945	0.34475
## studytime4:Dalc3	-1.18069	2.98535	-0.395	0.69258
## studytime2:Dalc4	1.69346	2.78591	0.608	0.54344
## studytime3:Dalc4	7.96165	7.45785	1.068	0.28602
## studytime4:Dalc4	NA	NA	NA	NA
## studytime2:Dalc5	-4.55563	4.39151	-1.037	0.29985

## studytime3:Dalc5	-16.46327	7.80422	-2.110	0.03519	*
## studytime4:Dalc5	-14.57926	6.00065	-2.430	0.01532	*
## studytime2:health2	-1.68591	1.19312	-1.413	0.15801	
## studytime3:health2	-1.68791	1.83688	-0.919	0.35841	
## studytime4:health2	3.11430	2.72881	1.141	0.25408	
## studytime2:health3	-0.93001	1.07452	-0.866	0.38700	
## studytime3:health3	-2.61377	1.63669	-1.597	0.11064	
## studytime4:health3	-0.86750	1.78248	-0.487	0.62661	
## studytime2:health4	-2.73272	1.12655	-2.426	0.01548	*
## studytime3:health4	-3.82429	1.65255	-2.314	0.02089	*
## studytime4:health4	-1.99454	2.37549	-0.840	0.40135	
## studytime2:health5	-0.38549	0.90176	-0.427	0.66913	
## studytime3:health5	-2.31601	1.52067	-1.523	0.12812	
## studytime4:health5	3.05423	1.87299	1.631	0.10333	
## studytime2:romanticyes	0.95350	0.63427	1.503	0.13313	
## studytime3:romanticyes	0.67832	0.82952	0.818	0.41374	
## studytime4:romanticyes	1.76484	2.12974	0.829	0.40753	
## studytime2:famsizeLE3	0.12089	0.66683	0.181	0.85619	
## studytime3:famsizeLE3	0.51127	0.96248	0.531	0.59542	
## studytime4:famsizeLE3	2.35353	1.54717	1.521	0.12858	
## studytime2:goout2	-1.46852	1.23334	-1.191	0.23410	
## studytime3:goout2	-5.21845	2.22681	-2.343	0.01933	*
## studytime4:goout2	6.10169	3.11451	1.959	0.05042	.
## studytime2:goout3	-0.85596	1.22958	-0.696	0.48653	
## studytime3:goout3	-4.20232	2.24467	-1.872	0.06153	.
## studytime4:goout3	5.90070	3.23892	1.822	0.06883	.
## studytime2:goout4	-1.96383	1.29053	-1.522	0.12845	
## studytime3:goout4	-7.51204	2.39313	-3.139	0.00175	**
## studytime4:goout4	1.87345	3.74103	0.501	0.61665	
## studytime2:goout5	-2.35473	1.34618	-1.749	0.08062	.
## studytime3:goout5	-4.89583	2.48231	-1.972	0.04890	*
## studytime4:goout5	1.24967	3.16468	0.395	0.69303	
## schoolsupyes:Dalc2	1.82186	1.24149	1.467	0.14261	
## schoolsupyes:Dalc3	0.10663	2.67987	0.040	0.96827	
## schoolsupyes:Dalc4	5.99897	4.34381	1.381	0.16763	
## schoolsupyes:Dalc5	-1.24564	3.72701	-0.334	0.73830	
## schoolsupyes:health2	2.09829	1.62032	1.295	0.19567	
## schoolsupyes:health3	2.14663	1.44409	1.486	0.13752	
## schoolsupyes:health4	3.67709	1.43966	2.554	0.01082	*
## schoolsupyes:health5	1.83856	1.38313	1.329	0.18411	
## schoolsupyes:romanticyes	0.40506	1.03780	0.390	0.69641	
## schoolsupyes:famsizeLE3	-0.06095	1.12973	-0.054	0.95699	
## schoolsupyes:goout2	-4.66081	1.77450	-2.627	0.00878	**
## schoolsupyes:goout3	-4.98956	1.84686	-2.702	0.00704	**
## schoolsupyes:goout4	-4.56433	1.86739	-2.444	0.01472	*
## schoolsupyes:goout5	-5.27996	2.10258	-2.511	0.01222	*
## Dalc2:health2	-1.28246	1.26426	-1.014	0.31068	
## Dalc3:health2	-0.14983	2.21054	-0.068	0.94598	
## Dalc4:health2	11.53218	6.58596	1.751	0.08030	.
## Dalc5:health2	NA	NA	NA	NA	
## Dalc2:health3	-0.69766	1.30187	-0.536	0.59217	
## Dalc3:health3	0.27260	2.31822	0.118	0.90642	
## Dalc4:health3	-7.49135	4.35025	-1.722	0.08542	.
## Dalc5:health3	-0.49121	3.62666	-0.135	0.89229	

## Dalc2:health4	1.08093	1.18367	0.913	0.36139
## Dalc3:health4	2.77572	4.11811	0.674	0.50048
## Dalc4:health4	-2.47874	3.45081	-0.718	0.47276
## Dalc5:health4	-7.00544	5.08326	-1.378	0.16852
## Dalc2:health5	-0.07002	1.02722	-0.068	0.94567
## Dalc3:health5	-0.13611	1.84256	-0.074	0.94113
## Dalc4:health5	1.82151	4.05185	0.450	0.65315
## Dalc5:health5	-0.94226	4.35674	-0.216	0.82882
## Dalc2:romanticyes	-0.18658	0.69807	-0.267	0.78931
## Dalc3:romanticyes	2.56848	1.34287	1.913	0.05612 .
## Dalc4:romanticyes	3.06578	3.47672	0.882	0.37813
## Dalc5:romanticyes	4.21448	4.09434	1.029	0.30361
## Dalc2:famsizeLE3	-0.92768	0.69149	-1.342	0.18009
## Dalc3:famsizeLE3	0.17557	1.18086	0.149	0.88184
## Dalc4:famsizeLE3	7.40395	3.82153	1.937	0.05302 .
## Dalc5:famsizeLE3	-1.58703	2.65124	-0.599	0.54960
## Dalc2:goout2	0.98252	1.89981	0.517	0.60517
## Dalc3:goout2	-2.36548	2.68578	-0.881	0.37870
## Dalc4:goout2	-7.90805	3.48983	-2.266	0.02370 *
## Dalc5:goout2	8.17369	5.93387	1.377	0.16873
## Dalc2:goout3	1.44074	1.88254	0.765	0.44429
## Dalc3:goout3	-3.95701	2.69239	-1.470	0.14201
## Dalc4:goout3	6.94833	7.71471	0.901	0.36802
## Dalc5:goout3	9.46036	4.46007	2.121	0.03420 *
## Dalc2:goout4	0.33667	1.89862	0.177	0.85930
## Dalc3:goout4	-3.56439	2.57833	-1.382	0.16720
## Dalc4:goout4	-1.71477	3.94362	-0.435	0.66380
## Dalc5:goout4	-3.84918	3.93697	-0.978	0.32850
## Dalc2:goout5	-0.31882	1.97374	-0.162	0.87171
## Dalc3:goout5	-0.89956	2.62620	-0.343	0.73203
## Dalc4:goout5	NA	NA	NA	NA
## Dalc5:goout5	NA	NA	NA	NA
## health2:romanticyes	0.21476	1.04665	0.205	0.83747
## health3:romanticyes	0.70414	0.93910	0.750	0.45358
## health4:romanticyes	-0.70717	0.99705	-0.709	0.47835
## health5:romanticyes	-0.53777	0.83714	-0.642	0.52080
## health2:famsizeLE3	-1.88495	1.04453	-1.805	0.07149 .
## health3:famsizeLE3	-0.60730	0.93707	-0.648	0.51710
## health4:famsizeLE3	-0.32377	0.97530	-0.332	0.74000
## health5:famsizeLE3	-1.19048	0.83867	-1.419	0.15612
## health2:goout2	1.39577	2.08052	0.671	0.50248
## health3:goout2	-2.21354	1.85301	-1.195	0.23259
## health4:goout2	-1.23126	2.32097	-0.530	0.59591
## health5:goout2	-1.69178	1.66702	-1.015	0.31046
## health2:goout3	3.10719	1.96421	1.582	0.11404
## health3:goout3	-1.38610	1.81554	-0.763	0.44540
## health4:goout3	0.44052	2.18944	0.201	0.84059
## health5:goout3	0.84908	1.63247	0.520	0.60312
## health2:goout4	3.08032	2.00483	1.536	0.12480
## health3:goout4	-2.21220	1.88143	-1.176	0.24000
## health4:goout4	-0.12331	2.20087	-0.056	0.95533
## health5:goout4	-0.52818	1.64521	-0.321	0.74826
## health2:goout5	1.70747	2.11700	0.807	0.42015
## health3:goout5	-3.81315	2.07869	-1.834	0.06694 .

```
## health4:goout5          0.17455    2.38932    0.073    0.94178
## health5:goout5         -1.16655    1.78371   -0.654    0.51329
## romanticyes:famsizeLE3  0.19945    0.57899    0.344    0.73057
## romanticyes:goout2      1.71975    1.16526    1.476    0.14035
## romanticyes:goout3      0.56134    1.15334    0.487    0.62659
## romanticyes:goout4      1.52420    1.19973    1.270    0.20427
## romanticyes:goout5      1.53200    1.24633    1.229    0.21933
## famsizeLE3:goout2       2.00716    1.20129    1.671    0.09512
## famsizeLE3:goout3       1.33000    1.17687    1.130    0.25874
## famsizeLE3:goout4       0.68444    1.22336    0.559    0.57598
## famsizeLE3:goout5       1.39645    1.30170    1.073    0.28367
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.357 on 858 degrees of freedom
## Multiple R-squared:  0.3792, Adjusted R-squared:  0.2454
## F-statistic: 2.833 on 185 and 858 DF,  p-value: < 2.2e-16
```

In the summary of our full interaction model the interaction term *Dalc:goout and studytime:Dalc* has some NA entries, it simply mean that there is not enough data to generate values for those interaction. So we will drop these interaction completely.

Our interaction model will be:

```
student_performance_intMdl2 = lm(G3 ~ (failures+higher+studytime+schoolsup+Dalc+health+romantic+famsize
                                     +goout+failures:schoolsup+studytime:health+studytime:goout
                                     +schoolsup:health+schoolsup:goout),
                                data = studentDataset)
summary(student_performance_intMdl2)
```

```
##
## Call:
## lm(formula = G3 ~ (failures + higher + studytime + schoolsup +
##   Dalc + health + romantic + famsize + +goout + failures:schoolsup +
##   studytime:health + studytime:goout + schoolsup:health + schoolsup:goout),
##   data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.5009  -1.5919   0.2916   2.1990   7.9895
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    8.498798   0.981759   8.657 < 2e-16 ***
## failures       -2.046312   0.183899 -11.127 < 2e-16 ***
## higheryes       1.571610   0.413108   3.804 0.000151 ***
## studytime2      2.629073   1.128137   2.330 0.019982 *
## studytime3      7.110347   2.120459   3.353 0.000829 ***
## studytime4     -1.775067   2.219330  -0.800 0.424006
## schoolsupyes     0.776582   1.496920   0.519 0.604026
## Dalc2          -0.818135   0.290378  -2.817 0.004937 **
## Dalc3           0.008952   0.458562   0.020 0.984428
## Dalc4          -1.642254   0.698916  -2.350 0.018983 *
```

```

## Dalc5          0.318271    0.745456    0.427 0.669510
## health2        0.533736    0.855741    0.624 0.532960
## health3       -0.862240    0.756791   -1.139 0.254838
## health4        0.376706    0.772522    0.488 0.625919
## health5       -0.468239    0.654243   -0.716 0.474348
## romanticyes   -0.601559    0.234131   -2.569 0.010335 *
## famsizeLE3     0.611245    0.240307    2.544 0.011123 *
## goout2         2.701405    0.826061    3.270 0.001112 **
## goout3         2.120678    0.806759    2.629 0.008705 **
## goout4         2.270820    0.841135    2.700 0.007058 **
## goout5         1.910583    0.848629    2.251 0.024580 *
## failures:schoolsupyes 1.470464    0.531499    2.767 0.005769 **
## studytime2:health2 -1.995736    1.020678   -1.955 0.050828 .
## studytime3:health2 -1.342237    1.636566   -0.820 0.412326
## studytime4:health2  0.633159    2.065026    0.307 0.759204
## studytime2:health3 -0.720734    0.915681   -0.787 0.431411
## studytime3:health3 -2.050107    1.438783   -1.425 0.154504
## studytime4:health3 -1.209162    1.562797   -0.774 0.439283
## studytime2:health4 -1.586649    0.944165   -1.680 0.093180 .
## studytime3:health4 -2.868793    1.453510   -1.974 0.048694 *
## studytime4:health4 -2.212189    2.067158   -1.070 0.284808
## studytime2:health5 -0.677095    0.802073   -0.844 0.398772
## studytime3:health5 -2.416723    1.359039   -1.778 0.075668 .
## studytime4:health5 -0.130447    1.579813   -0.083 0.934209
## studytime2:goout2  -1.117962    1.061877   -1.053 0.292682
## studytime3:goout2  -3.996336    1.818851   -2.197 0.028239 *
## studytime4:goout2   4.912056    2.193727    2.239 0.025368 *
## studytime2:goout3  -1.319228    1.033860   -1.276 0.202247
## studytime3:goout3  -3.495151    1.791390   -1.951 0.051329 .
## studytime4:goout3   3.472643    2.242616    1.548 0.121827
## studytime2:goout4  -1.842089    1.061063   -1.736 0.082861 .
## studytime3:goout4  -4.748364    1.921606   -2.471 0.013640 *
## studytime4:goout4   1.986642    2.727053    0.728 0.466483
## studytime2:goout5  -1.770224    1.095119   -1.616 0.106312
## studytime3:goout5  -2.911700    2.013425   -1.446 0.148454
## studytime4:goout5   2.181123    2.207602    0.988 0.323391
## schoolsupyes:health2 2.214681    1.424101    1.555 0.120232
## schoolsupyes:health3 2.450585    1.269637    1.930 0.053874 .
## schoolsupyes:health4 2.768818    1.283108    2.158 0.031176 *
## schoolsupyes:health5 2.849506    1.192475    2.390 0.017054 *
## schoolsupyes:goout2 -5.901575    1.351639   -4.366 1.4e-05 ***
## schoolsupyes:goout3 -4.808884    1.366416   -3.519 0.000452 ***
## schoolsupyes:goout4 -4.135247    1.368517   -3.022 0.002578 **
## schoolsupyes:goout5 -5.866249    1.562109   -3.755 0.000183 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.398 on 990 degrees of freedom
## Multiple R-squared:  0.2664, Adjusted R-squared:  0.2271
## F-statistic: 6.783 on 53 and 990 DF,  p-value: < 2.2e-16

```

### 3.4 Comparing Additive and Interaction model

Comparing our additive and interaction models to see which one is better. To verify this we can use partial F-test using below hypothesis:

$H_0$  : All the interaction coefficients are not significant  $H_a$  : At least one of the interaction coefficient is significant

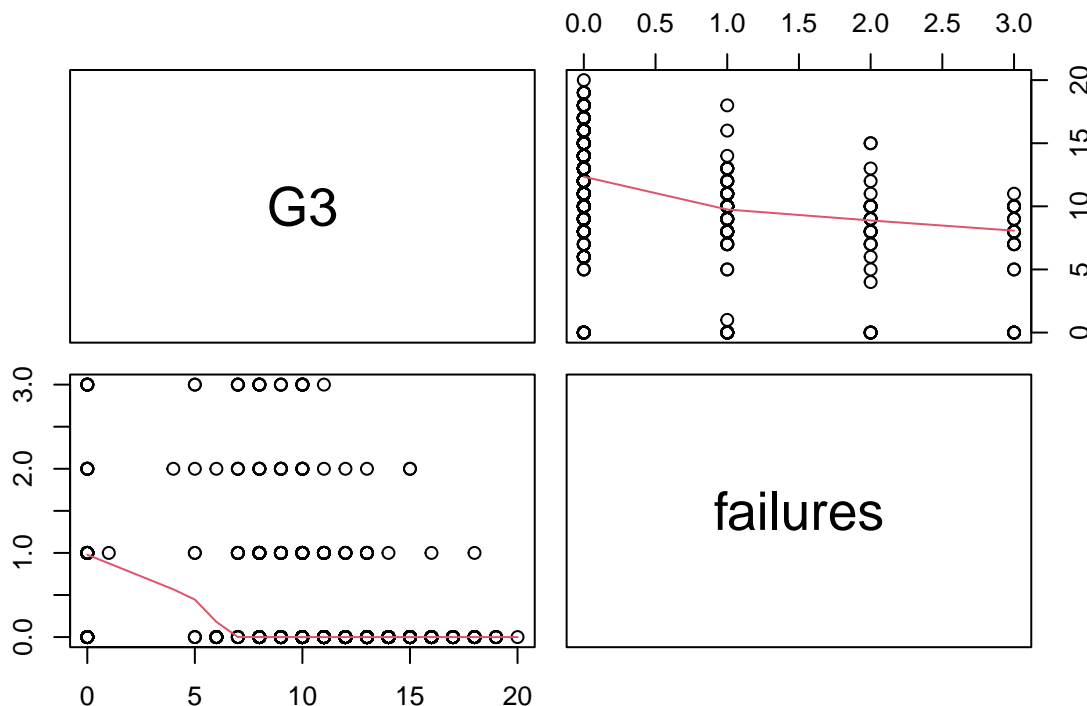
Source of variation	DF	Sum of Square	Mean Square	F-Statistic
Regression	33	740.32	22.43394	1.9433
Residual	990	11429	11.54444	
Total	1023	12169		

The p-value from our ANOVA test is **0.001228** which is less than our assumed  $\alpha = 0.05$  hence we can reject our  $H_0$  and conclude that at least one of our interaction coefficient is significant, and

### 3.5. Higher Order Model

Since we have finalized our interaction model, we can check if there are any variable (quantitative variable) for which we need to add higher order terms. For this analysis we will use pairs plot. In our final interaction model we have only one quantitative variable (i.e. failures).

```
pairs(~G3 + failures, data = studentDataset, panel = panel.smooth)
```



Though there no clear visual indication that we should add higher order term for our variable, but we can still try by adding the higher order term and check the significance.

```
student_performance_intMdl3 = lm(G3 ~ (failures+I(failures^2)+higher+studytime+schoolsup+Dalc+health+ron
+goout+failures:schoolsup+studytime:health+studytime:goout
+schoolsup:health+schoolsup:goout),
data = studentDataset)
summary(student_performance_intMdl3)
```

```
##
## Call:
## lm(formula = G3 ~ (failures + I(failures^2) + higher + studytime +
##      schoolsup + Dalc + health + romantic + famsize + goout +
##      failures:schoolsup + studytime:health + studytime:goout +
##      schoolsup:health + schoolsup:goout), data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.5025  -1.5588   0.3212   2.1742   8.6013
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      8.6707     0.9755   8.888 < 2e-16 ***
## failures        -3.8453     0.4898  -7.851 1.07e-14 ***
## I(failures^2)     0.7576     0.1914   3.958 8.09e-05 ***
## higheryes        1.4940     0.4105   3.639 0.000288 ***
## studytime2        2.5596     1.1200   2.285 0.022504 *
## studytime3        7.4010     2.1062   3.514 0.000462 ***
## studytime4       -1.7682     2.2031  -0.803 0.422394
## schoolsupyes       0.4547     1.4882   0.306 0.760012
## Dalc2            -0.7863     0.2884  -2.727 0.006511 **
## Dalc3             0.1068     0.4559   0.234 0.814862
## Dalc4            -1.3768     0.6970  -1.975 0.048525 *
## Dalc5             0.2097     0.7405   0.283 0.777071
## health2           0.7798     0.8517   0.916 0.360150
## health3          -0.9041     0.7513  -1.203 0.229152
## health4           0.2467     0.7676   0.321 0.747981
## health5          -0.4384     0.6495  -0.675 0.499822
## romanticyes      -0.5402     0.2329  -2.319 0.020589 *
## famsizeLE3        0.5953     0.2386   2.495 0.012744 *
## goout2            2.6289     0.8202   3.205 0.001393 **
## goout3            2.0995     0.8009   2.622 0.008886 **
## goout4            2.3283     0.8351   2.788 0.005404 **
## goout5            1.7670     0.8432   2.096 0.036370 *
## failures:schoolsupyes 1.4629     0.5276   2.773 0.005665 **
## studytime2:health2 -2.2138     1.0147  -2.182 0.029361 *
## studytime3:health2 -1.4987     1.6251  -0.922 0.356618
## studytime4:health2  0.2277     2.0525   0.111 0.911696
## studytime2:health3 -0.5703     0.9098  -0.627 0.530916
## studytime3:health3 -1.9426     1.4285  -1.360 0.174180
## studytime4:health3 -1.1523     1.5514  -0.743 0.457822
## studytime2:health4 -1.4005     0.9384  -1.492 0.135926
## studytime3:health4 -2.7377     1.4432  -1.897 0.058132 .
## studytime4:health4 -2.1322     2.0521  -1.039 0.299055
## studytime2:health5 -0.6203     0.7963  -0.779 0.436179
## studytime3:health5 -2.3509     1.3492  -1.742 0.081736 .
## studytime4:health5 -0.2144     1.5684  -0.137 0.891296
## studytime2:goout2  -1.1239     1.0541  -1.066 0.286582
## studytime3:goout2  -4.2356     1.8065  -2.345 0.019244 *
## studytime4:goout2   4.9803     2.1777   2.287 0.022411 *
## studytime2:goout3  -1.3213     1.0263  -1.287 0.198229
## studytime3:goout3  -3.9000     1.7812  -2.190 0.028790 *
## studytime4:goout3   3.3769     2.2263   1.517 0.129630
```

```

## studytime2:goout4      -1.8888      1.0534   -1.793  0.073252 .
## studytime3:goout4      -5.1949      1.9109   -2.719  0.006670 **
## studytime4:goout4       1.8847      2.7072    0.696  0.486475
## studytime2:goout5      -1.7584      1.0871   -1.618  0.106088
## studytime3:goout5      -3.2727      2.0008   -1.636  0.102214
## studytime4:goout5       2.3971      2.1921    1.094  0.274432
## schoolsupyes:health2    2.3211      1.4139    1.642  0.100989
## schoolsupyes:health3    2.3842      1.2604    1.892  0.058844 .
## schoolsupyes:health4    2.9360      1.2744    2.304  0.021438 *
## schoolsupyes:health5    2.9259      1.1839    2.471  0.013625 *
## schoolsupyes:goout2     -5.5980      1.3439   -4.165  3.38e-05 ***
## schoolsupyes:goout3     -4.5044      1.3586   -3.316  0.000948 ***
## schoolsupyes:goout4     -3.8287      1.3607   -2.814  0.004993 **
## schoolsupyes:goout5     -5.5483      1.5527   -3.573  0.000370 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.373 on 989 degrees of freedom
## Multiple R-squared:  0.2778, Adjusted R-squared:  0.2384
## F-statistic: 7.047 on 54 and 989 DF,  p-value: < 2.2e-16

```

$$\begin{aligned}
\widehat{G3} = & 8.67068((\text{Intercept})) - 3.84533(\text{failures}) + 0.757609(I(\text{failures}^2)) + 1.49398(\text{higheryes}) \\
& + 2.5596(\text{studytime2}) + 7.40096(\text{studytime3}) - 1.7682(\text{studytime4}) + 0.454711(\text{schoolsupyes}) \\
& - 0.786275(\text{Dalc2}) + 0.106775(\text{Dalc3}) - 1.37676(\text{Dalc4}) + 0.209721(\text{Dalc5}) + 0.779774(\text{health2}) \\
& - 0.904056(\text{health3}) + 0.246689(\text{health4}) - 0.438419(\text{health5}) - 0.540209(\text{romanticyes}) \\
& + 0.595347(\text{famsizeLE3}) + 2.62894(\text{goout2}) + 2.09955(\text{goout3}) + 2.32829(\text{goout4}) \\
& + 1.767(\text{goout5}) + 1.46286(\text{failures:schoolsupyes}) - 2.21384(\text{studytime2:health2}) \\
& - 1.49872(\text{studytime3:health2}) + 0.227675(\text{studytime4:health2}) - 0.570268(\text{studytime2:health3}) \\
& - 1.94258(\text{studytime3:health3}) - 1.15228(\text{studytime4:health3}) - 1.40046(\text{studytime2:health4}) \\
& - 2.73769(\text{studytime3:health4}) - 2.13216(\text{studytime4:health4}) - 0.62032(\text{studytime2:health5}) \\
& - 2.35092(\text{studytime3:health5}) - 0.214397(\text{studytime4:health5}) - 1.12391(\text{studytime2:goout2}) \\
& - 4.23562(\text{studytime3:goout2}) + 4.98029(\text{studytime4:goout2}) - 1.32132(\text{studytime2:goout3}) \\
& - 3.90003(\text{studytime3:goout3}) + 3.37694(\text{studytime4:goout3}) - 1.88884(\text{studytime2:goout4}) \\
& - 5.1949(\text{studytime3:goout4}) + 1.88472(\text{studytime4:goout4}) - 1.75839(\text{studytime2:goout5}) \\
& - 3.27269(\text{studytime3:goout5}) + 2.39711(\text{studytime4:goout5}) + 2.3211(\text{schoolsupyes:health2}) \\
& + 2.38419(\text{schoolsupyes:health3}) + 2.93603(\text{schoolsupyes:health4}) + 2.92593(\text{schoolsupyes:health5}) \\
& - 5.59801(\text{schoolsupyes:goout2}) - 4.50439(\text{schoolsupyes:goout3}) - 3.8287(\text{schoolsupyes:goout4}) \\
& - 5.54827(\text{schoolsupyes:goout5})
\end{aligned}$$

### 3.6 Assumption Verification

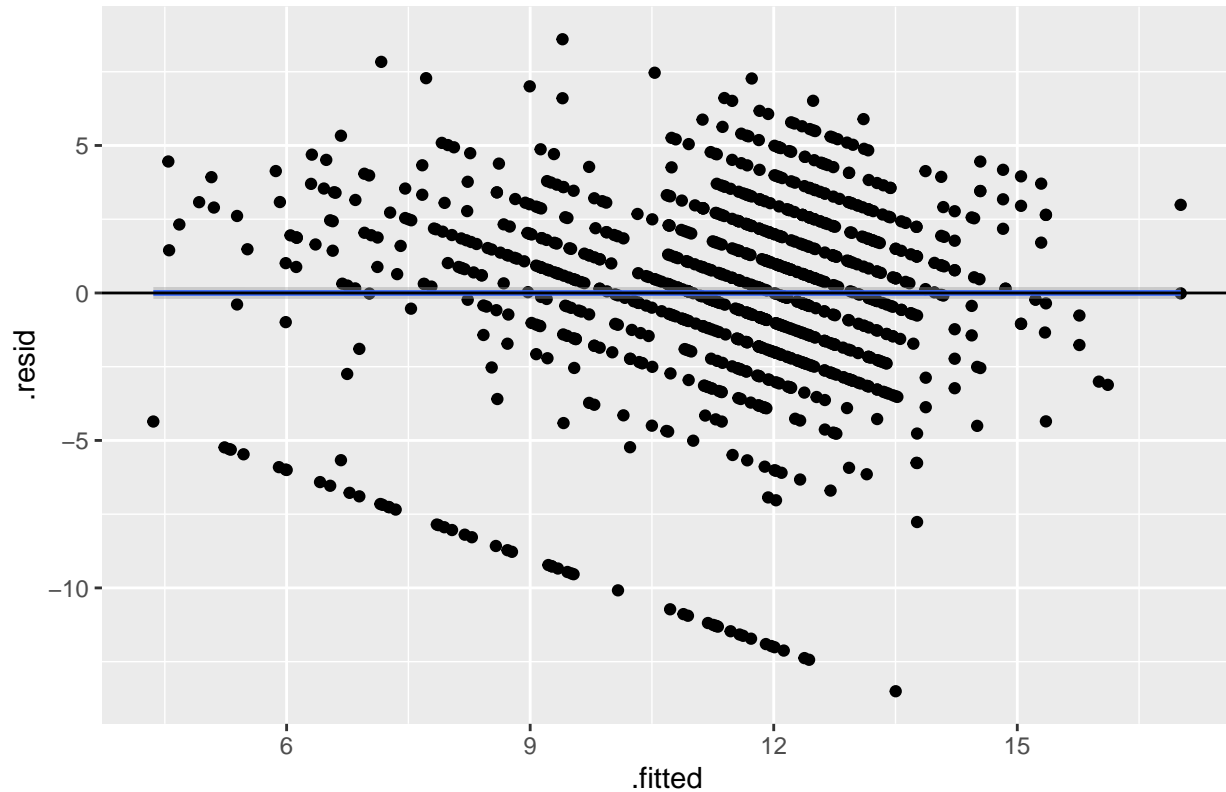
Our model building process is based on some assumption about the data. Hence it is very important to verify that these assumption are met or not. If our data do not fit the assumptions then the model we developed cannot be used for prediction.



### 3.6.1 Linearity Assumption

The linear regression model we build is based on the assumption that there is a linear relation between predictors and response variable. To confirm the linear relation we can use Residual plot.

Residual plot: Residual vs Fitted values

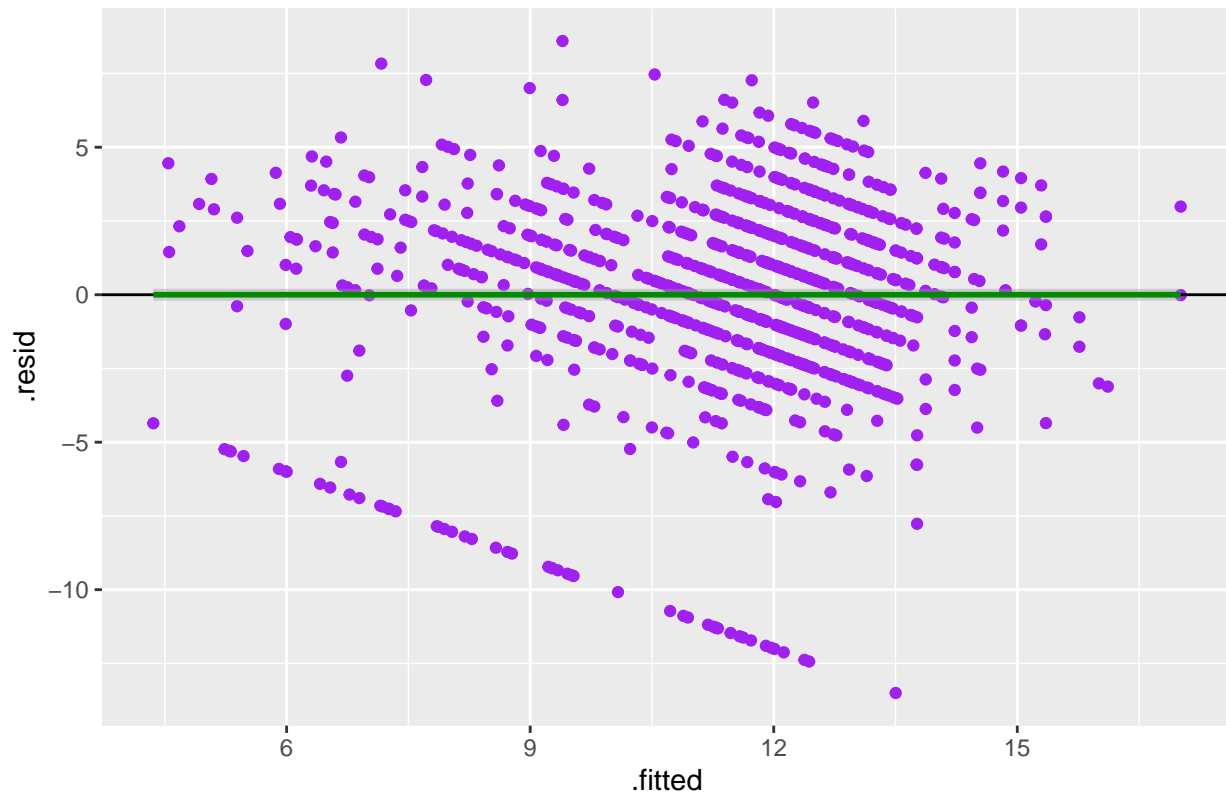


Comment: The residual vs fitted plot show a straight line which indicate that our linearity assumption is met.

### 3.6.2 Equal Variance Assumption

Another important assumption for our linear regression model is that the error term has a constant variance (**homoscedasticity**). To verify the homoscedasticity assumption we can again use the residual vs fitted plot and check if there is any pattern.

Residual plot: Residual vs Fitted values



To confirm homoscedasticity we can perform the Breusch-Pagan Test (bptest) using below hypothesis:

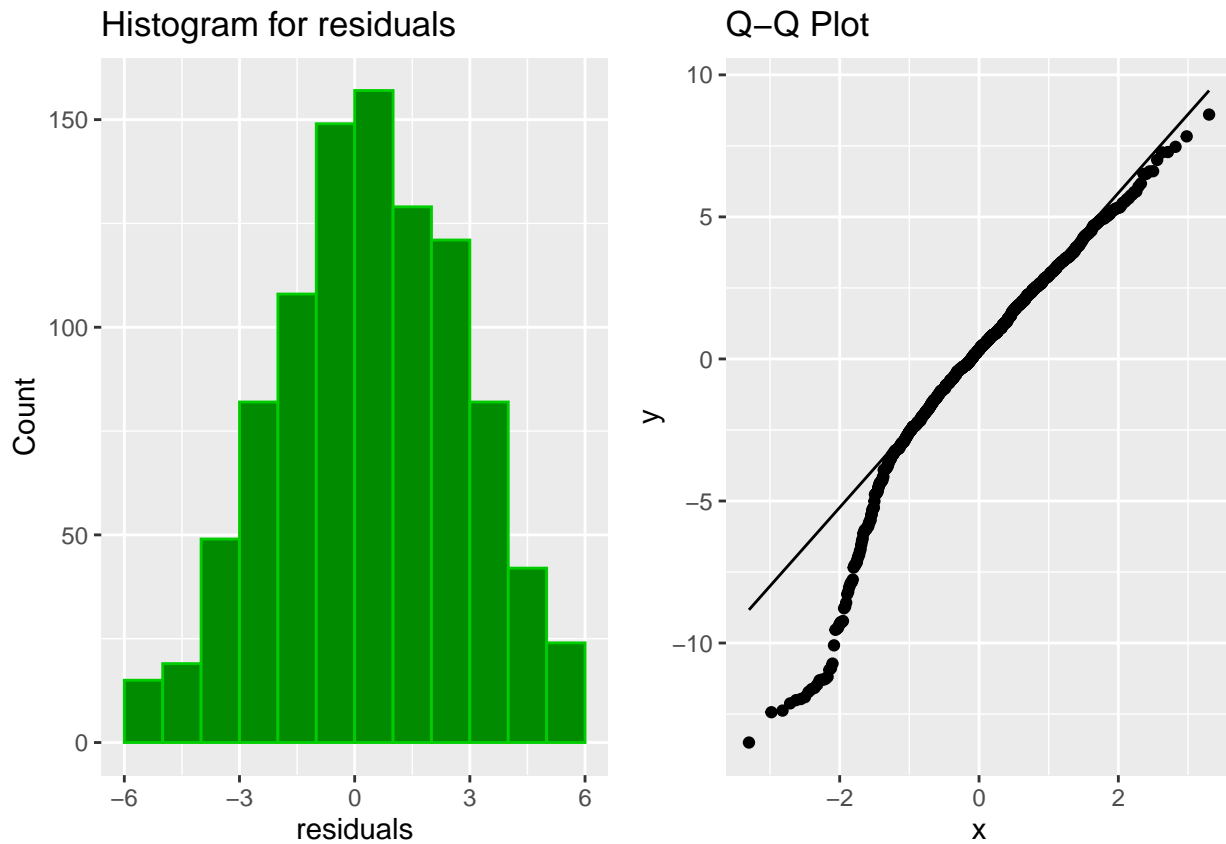
$H_0$  : Heteroscedasticity is not present (error term has common variance)  $H_a$  : Heteroscedasticity is present (error term does not have common variance)

```
bptest(student_performance_intMdl3) # It doesn't have heteroscedasticity.
```

```
##
## studentized Breusch-Pagan test
##
## data: student_performance_intMdl3
## BP = 67.517, df = 54, p-value = 0.1023
```

Comment: Since the p-value is 0.3235 higher than our assumed  $\alpha = 0.05$  we cannot reject the  $H_0$  and conclude that our model meets the assumption of common variance.

### 3.6.3 Normality Assumption



```
shapiro.test(residuals(student_performance_intMdl3)) # We don't have normality
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: residuals(student_performance_intMdl3)  
## W = 0.94218, p-value < 2.2e-16
```

### 3.6.4 Independence Assumption

Check if we need to add anything.

### 3.6.5 Multicollinearity

```
#finalMdl = lm(G3 ~ (failures+higher+studytime+schoolsup+Dalc+health+romantic+famsize+  
#                               +goout+failures:schoolsup+studytime:health+studytime:goout  
#                               +schoolsup:health+schoolsup:goout),  
#                               data = studentDataset)  
  
#finalMdl = lm(G3 ~ (failures+higher+schoolsup+Dalc+health+romantic+famsize+  
#                               +goout+failures:schoolsup+schoolsup:health+schoolsup:goout),  
#                               data = studentDataset)  
  
finalMdl = lm(G3 ~ (failures+higher+schoolsup+Dalc+health+romantic+famsize+  
#                               +goout+studytime),
```

```

                                data = studentDataset)
imcdiag(finalMdl, method="VIF")

```

```

##
## Call:
## imcdiag(mod = finalMdl, method = "VIF")
##
## VIF Multicollinearity Diagnostics
##
##           VIF detection
## failures      1.1247      0
## higheryes     1.1762      0
## schoolsupyes  1.0391      0
## Dalc2         1.0953      0
## Dalc3         1.1192      0
## Dalc4         1.0416      0
## Dalc5         1.0943      0
## health2       1.6986      0
## health3       2.0999      0
## health4       1.9511      0
## health5       2.4854      0
## romanticyes   1.0672      0
## famsizeLE3    1.0306      0
## goout2        3.5109      0
## goout3        4.0215      0
## goout4        3.4486      0
## goout5        2.9202      0
## studytime2    1.5008      0
## studytime3    1.4453      0
## studytime4    1.1850      0
##
## NOTE: VIF Method Failed to detect multicollinearity
##
##
## 0 --> COLLINEARITY is not detected by the test
##
## =====

```

```

summary(finalMdl)

```

```

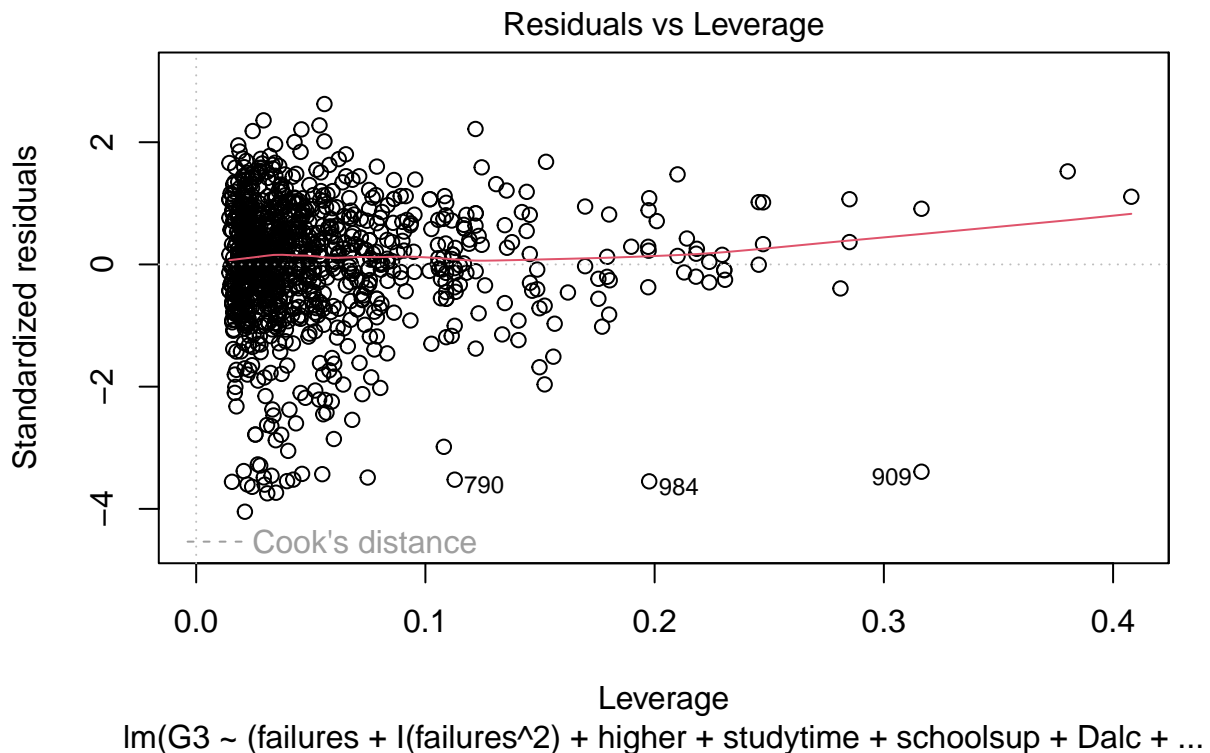
##
## Call:
## lm(formula = G3 ~ (failures + higher + schoolsup + Dalc + health +
##   romantic + famsize + goout + studytime), data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.1848  -1.5137   0.3029   2.1295   8.6106
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.68826    0.63642  16.794 < 2e-16 ***
## failures     -1.82887    0.17261 -10.595 < 2e-16 ***

```

```
## higheryes      1.54720    0.41456    3.732 0.000200 ***
## schoolsupyes -1.16127    0.34238   -3.392 0.000721 ***
## Dalc2         -0.73721    0.28608   -2.577 0.010107 *
## Dalc3         -0.14301    0.45454   -0.315 0.753108
## Dalc4        -1.65491    0.69909   -2.367 0.018107 *
## Dalc5         0.11089    0.71654    0.155 0.877041
## health2       -0.66440    0.43152   -1.540 0.123946
## health3       -1.21726    0.38251   -3.182 0.001505 **
## health4       -0.74820    0.40008   -1.870 0.061751 .
## health5       -0.99991    0.34699   -2.882 0.004038 **
## romanticyes  -0.59306    0.23040   -2.574 0.010190 *
## famsizeLE3    0.48561    0.23806    2.040 0.041625 *
## goout2        1.15597    0.46996    2.460 0.014070 *
## goout3        0.67978    0.45855    1.482 0.138527
## goout4        0.38442    0.48055    0.800 0.423919
## goout5        0.08698    0.50253    0.173 0.862617
## studytime2    0.26960    0.26170    1.030 0.303179
## studytime3    1.18603    0.35442    3.346 0.000849 ***
## studytime4    0.62526    0.49164    1.272 0.203738
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.449 on 1023 degrees of freedom
## Multiple R-squared:  0.2189, Adjusted R-squared:  0.2036
## F-statistic: 14.33 on 20 and 1023 DF, p-value: < 2.2e-16
```

### 3.6.6 Outlier

```
plot(student_performance_intMdl3,which=5)
```



```
studentDataset[cooks.distance(student_performance_intMd13)>0.5,]
```

```
## [1] school      sex      age      address   famsize   Pstatus
## [7] Medu        Fedu      Mjob      Fjob      reason    guardian
## [13] traveltime studytime failures schoolsup famsup    paid
## [19] activities nursery higher    internet  romantic  famrel
## [25] freetime    goout     Dalc      Walc      health    absences
## [31] G1          G2        G3
## <0 rows> (or 0-length row.names)
```

## 4. Conclusion

state final model give interpretation

## 5. Discussion

Future course of action

## Appendix

---

END

```
studentPerformance_p1 = lm(G3 ~ (sex+age+address+famsize+Pstatus+Medu+Fedu+traveltime+studytime+
                                failures+schoolsup+famsup+activities+higher+internet+romantic+
                                famrel+freetime+goout+Dalc+Walc+health+absences),
                           data = studentDataset)
```

```
summary(studentPerformance_p1)
```

```
##
## Call:
## lm(formula = G3 ~ (sex + age + address + famsize + Pstatus +
##   Medu + Fedu + traveltime + studytime + failures + schoolsup +
##   famsup + activities + higher + internet + romantic + famrel +
##   freetime + goout + Dalc + Walc + health + absences), data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.0843  -1.4289   0.4463   2.1270   7.5585
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.687870   2.591012   4.125 4.02e-05 ***
## sexM         -0.055044   0.252983  -0.218  0.82780
## age          0.042235   0.099746   0.423  0.67208
## addressU     0.550643   0.264811   2.079  0.03784 *
## famsizeLE3   0.433019   0.246931   1.754  0.07981 .
## PstatusT     -0.182639   0.359700  -0.508  0.61174
## Medu1       -1.677833   1.204934  -1.392  0.16409
## Medu2       -1.608633   1.202371  -1.338  0.18124
```

```
## Medu3      -1.159535    1.213198   -0.956    0.33942
## Medu4      -0.763292    1.224926   -0.623    0.53334
## Fedu1      -0.929902    1.203955   -0.772    0.44008
## Fedu2      -0.842970    1.210883   -0.696    0.48649
## Fedu3      -1.092206    1.222057   -0.894    0.37168
## Fedu4      -0.615769    1.233963   -0.499    0.61788
## traveltime2 -0.123818    0.251447   -0.492    0.62253
## traveltime3  0.246458    0.449547    0.548    0.58365
## traveltime4 -0.331842    0.768714   -0.432    0.66606
## studytime2   0.342980    0.272156    1.260    0.20788
## studytime3   1.155808    0.374877    3.083    0.00210 **
## studytime4   0.685403    0.505705    1.355    0.17562
## failures     -1.769215    0.181749   -9.734    < 2e-16 ***
## schoolsupyes -1.011981    0.356773   -2.836    0.00465 **
## famsupyes     -0.233627    0.231686   -1.008    0.31352
## activitiesyes 0.016151    0.225301    0.072    0.94287
## higheryes     1.252016    0.426002    2.939    0.00337 **
## internetyes   0.439557    0.284548    1.545    0.12272
## romanticyes  -0.653102    0.237145   -2.754    0.00599 **
## famrel2       0.583759    0.831103    0.702    0.48260
## famrel3       0.687533    0.710992    0.967    0.33378
## famrel4       1.110179    0.679117    1.635    0.10242
## famrel5       0.832581    0.694155    1.199    0.23065
## freetime2     1.066091    0.519722    2.051    0.04050 *
## freetime3    -0.044869    0.479245   -0.094    0.92543
## freetime4     0.372425    0.506624    0.735    0.46245
## freetime5     0.935673    0.581008    1.610    0.10762
## goout2        1.308530    0.475122    2.754    0.00599 **
## goout3         0.838256    0.472659    1.773    0.07645 .
## goout4         0.338196    0.498573    0.678    0.49772
## goout5        -0.088346    0.532722   -0.166    0.86832
## Dalc2         -0.724264    0.326990   -2.215    0.02699 *
## Dalc3         -0.327740    0.519089   -0.631    0.52794
## Dalc4         -1.982985    0.744620   -2.663    0.00787 **
## Dalc5         -0.726027    0.882460   -0.823    0.41086
## Walc2         -0.334080    0.308596   -1.083    0.27926
## Walc3          0.153191    0.342467    0.447    0.65474
## Walc4         -0.240007    0.435558   -0.551    0.58174
## Walc5          0.864545    0.649091    1.332    0.18319
## health2       -0.816591    0.436640   -1.870    0.06175 .
## health3       -1.165911    0.392086   -2.974    0.00301 **
## health4       -0.710288    0.405978   -1.750    0.08050 .
## health5       -1.106383    0.358159   -3.089    0.00206 **
## absences      0.007591    0.018036    0.421    0.67394
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 3.405 on 992 degrees of freedom
```

```
## Multiple R-squared:  0.2619, Adjusted R-squared:  0.2239
```

```
## F-statistic: 6.901 on 51 and 992 DF,  p-value: < 2.2e-16
```

```
subsets = ols_step_forward_p(studentPerformance_fm,p_val = 0.05, details = FALSE)
forwardMdl = subsets$model
summary(forwardMdl)
```

```

##
## Call:
## lm(formula = paste(response, "~", paste(preds, collapse = " + ")),
##     data = l)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.0125  -1.4391   0.3944   2.1255   8.2166
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.70133     1.38394   8.455 < 2e-16 ***
## failures     -1.76047     0.17299 -10.177 < 2e-16 ***
## higheryes     1.20109     0.41419   2.900 0.00381 **
## schoolsupyes -1.08027     0.34127  -3.165 0.00159 **
## schoolMS     -0.52091     0.27020  -1.928 0.05415 .
## romanticyes  -0.57404     0.22855  -2.512 0.01217 *
## studytime2    0.33184     0.26152   1.269 0.20476
## studytime3    1.19724     0.35335   3.388 0.00073 ***
## studytime4    0.53768     0.48696   1.104 0.26979
## addressU      0.45472     0.25698   1.769 0.07712 .
## Medu1        -1.68348     1.17642  -1.431 0.15273
## Medu2        -1.64019     1.16935  -1.403 0.16103
## Medu3        -1.22194     1.17050  -1.044 0.29676
## Medu4        -0.73276     1.16900  -0.627 0.53092
## Dalc2        -0.75785     0.28374  -2.671 0.00769 **
## Dalc3        -0.05715     0.45502  -0.126 0.90007
## Dalc4        -1.74603     0.68863  -2.536 0.01138 *
## Dalc5         0.04561     0.70762   0.064 0.94862
## famsizeLE3    0.51236     0.23591   2.172 0.03010 *
## freetime2     0.84199     0.50684   1.661 0.09697 .
## freetime3    -0.15376     0.46711  -0.329 0.74208
## freetime4     0.25650     0.48784   0.526 0.59915
## freetime5     0.87384     0.56062   1.559 0.11938
## goout2        1.20786     0.46546   2.595 0.00960 **
## goout3         0.82742     0.45858   1.804 0.07148 .
## goout4         0.40653     0.48356   0.841 0.40071
## goout5         0.01356     0.50561   0.027 0.97861
## health2       -0.71212     0.42745  -1.666 0.09603 .
## health3       -1.11917     0.37841  -2.958 0.00317 **
## health4       -0.62825     0.39712  -1.582 0.11396
## health5       -1.03268     0.34383  -3.003 0.00273 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.392 on 1013 degrees of freedom
## Multiple R-squared:  0.252, Adjusted R-squared:  0.2298
## F-statistic: 11.38 on 30 and 1013 DF, p-value: < 2.2e-16
bestSubsetMdl = ols_step_best_subset(forwardMdl, details = FALSE)
(bestSubsetMdl$metrics)

##      mindex  n
## 1         1  1
## 2         2  2

```



```

## 3      3 3
## 4      4 4
## 5      5 5
## 6      6 6
## 7      7 7
## 8      8 8
## 9      9 9
## 10     10 10
## 11     11 11
## 12     12 12
## 13     13 13
##
## 1
## 2
## 3
## 4
## 5
## 6
## 7
## 8
## 9
## 10
## 11
## 12
## 13 failures higher schoolsup school romantic studytime address Medu Dalc famsize freetime goout heal
##
## 1
## 2
## 3
## 4
## 5
## 6
## 7
## 8
## 9
## 10
## 11
## 12
## 13 failures higher schoolsup school romantic studytime address Medu Dalc famsize freetime goout heal
##
##      rsquare      adjr      predrsq      cp      aic      sbic      sbc      msep
## 1  0.1468003 0.1459815 0.1430529 115.45583 5624.782 2661.634 5639.634 13317.44
## 2  0.1688909 0.1648875 0.1598093  93.53936 5605.395 2636.206 5640.051 12985.11
## 3  0.1851195 0.1780267 0.1693287  79.56161 5592.808 2617.626 5647.267 12743.81
## 4  0.1980680 0.1879465 0.1758600  70.02599 5584.085 2602.952 5658.347 12553.39
## 5  0.2069598 0.1938198 0.1784832  65.98405 5580.445 2593.364 5674.510 12426.17
## 6  0.2152781 0.2014976 0.1855133  56.71894 5571.436 2584.445 5670.453 12307.70
## 7  0.2234061 0.2066724 0.1875748  53.71153 5568.566 2575.705 5687.386 12191.99
## 8  0.2299212 0.2110096 0.1892765  50.88843 5565.771 2569.046 5699.443 12101.40
## 9  0.2356582 0.2161175 0.1936972  45.11904 5559.964 2563.392 5698.587 12022.87
## 10 0.2410992 0.2209316 0.1980542  39.75043 5554.506 2558.109 5698.079 11948.85
## 11 0.2459738 0.2251731 0.2011428  35.14896 5549.778 2553.571 5698.303 11883.62
## 12 0.2496791 0.2282202 0.2042037  32.13104 5546.635 2550.603 5700.111 11836.70
## 13 0.2519911 0.2298388 0.2049181  31.00000 5545.414 2549.525 5703.840 11811.70
##
##      fpe      apc      hsp
## 1  12.78061 0.8564749 0.01225377
## 2  12.50962 0.8358993 0.01199402
## 3  12.32456 0.8211488 0.01181667
## 4  12.18745 0.8096503 0.01168531
## 5  12.11083 0.8022082 0.01161199
## 6  12.00692 0.7953160 0.01151250
## 7  11.94046 0.7885876 0.01144894
## 8  11.88646 0.7834715 0.01139736
## 9  11.82073 0.7791260 0.01133454
## 10 11.75928 0.7750633 0.01127584
## 11 11.70634 0.7715617 0.01122533
## 12 11.67135 0.7692427 0.01119204
## 13 11.65789 0.7683432 0.01117942

```

```
studentPerformance_p2 = lm(G3 ~ (failures+higher+schoolsup+school+romantic+studytime+address+Medu+Dalc+
                             data = studentDataset)
```

```
summary(studentPerformance_p2)
```

```
##
## Call:
## lm(formula = G3 ~ (failures + higher + schoolsup + school + romantic +
##   studytime + address + Medu + Dalc + famsize + freetime +
##   goout + health), data = studentDataset)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-13.0125	-1.4391	0.3944	2.1255	8.2166

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	11.70133	1.38394	8.455	< 2e-16 ***
failures	-1.76047	0.17299	-10.177	< 2e-16 ***
higheryes	1.20109	0.41419	2.900	0.00381 **
schoolsupyes	-1.08027	0.34127	-3.165	0.00159 **
schoolMS	-0.52091	0.27020	-1.928	0.05415 .
romanticyes	-0.57404	0.22855	-2.512	0.01217 *
studytime2	0.33184	0.26152	1.269	0.20476
studytime3	1.19724	0.35335	3.388	0.00073 ***
studytime4	0.53768	0.48696	1.104	0.26979
addressU	0.45472	0.25698	1.769	0.07712 .
Medu1	-1.68348	1.17642	-1.431	0.15273
Medu2	-1.64019	1.16935	-1.403	0.16103
Medu3	-1.22194	1.17050	-1.044	0.29676
Medu4	-0.73276	1.16900	-0.627	0.53092
Dalc2	-0.75785	0.28374	-2.671	0.00769 **
Dalc3	-0.05715	0.45502	-0.126	0.90007
Dalc4	-1.74603	0.68863	-2.536	0.01138 *
Dalc5	0.04561	0.70762	0.064	0.94862
famsizeLE3	0.51236	0.23591	2.172	0.03010 *
freetime2	0.84199	0.50684	1.661	0.09697 .
freetime3	-0.15376	0.46711	-0.329	0.74208
freetime4	0.25650	0.48784	0.526	0.59915
freetime5	0.87384	0.56062	1.559	0.11938
goout2	1.20786	0.46546	2.595	0.00960 **
goout3	0.82742	0.45858	1.804	0.07148 .
goout4	0.40653	0.48356	0.841	0.40071
goout5	0.01356	0.50561	0.027	0.97861
health2	-0.71212	0.42745	-1.666	0.09603 .
health3	-1.11917	0.37841	-2.958	0.00317 **
health4	-0.62825	0.39712	-1.582	0.11396
health5	-1.03268	0.34383	-3.003	0.00273 **

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.392 on 1013 degrees of freedom
## Multiple R-squared:  0.252, Adjusted R-squared:  0.2298
## F-statistic: 11.38 on 30 and 1013 DF, p-value: < 2.2e-16
```

```
studentPerformance_p3 = lm(G3 ~ (failures+higher+schoolsup+school+romantic+studytime+address+Dalc+famsi
                        data = studentDataset)
```

```
summary(studentPerformance_p3)
```

```
##
```

```
## Call:
```

```
## lm(formula = G3 ~ (failures + higher + schoolsup + school + romantic +
##      studytime + address + Dalc + famsize + freetime + goout +
##      health), data = studentDataset)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -13.2012  -1.5060   0.3486   2.1217   8.8832
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.206567   0.767864  13.292 < 2e-16 ***
## failures     -1.825667   0.170944 -10.680 < 2e-16 ***
## higheryes     1.391280   0.411564   3.380 0.000751 ***
## schoolsupyes -1.155056   0.342113  -3.376 0.000763 ***
## schoolMS     -0.643277   0.264528  -2.432 0.015196 *
## romanticyes  -0.591106   0.228859  -2.583 0.009938 **
## studytime2     0.322339   0.262345   1.229 0.219473
## studytime3     1.257265   0.354412   3.547 0.000407 ***
## studytime4     0.526773   0.488277   1.079 0.280914
## addressU       0.557077   0.256537   2.172 0.030122 *
## Dalc2         -0.685755   0.283446  -2.419 0.015723 *
## Dalc3          0.022110   0.453676   0.049 0.961140
## Dalc4        -1.638408   0.691086  -2.371 0.017936 *
## Dalc5          0.109440   0.710697   0.154 0.877648
## famsizeLE3     0.484782   0.236419   2.051 0.040570 *
## freetime2      0.927064   0.507925   1.825 0.068264 .
## freetime3     -0.134016   0.469134  -0.286 0.775191
## freetime4      0.309524   0.489214   0.633 0.527073
## freetime5      0.956691   0.562646   1.700 0.089373 .
## goout2         1.197306   0.466225   2.568 0.010368 *
## goout3         0.857997   0.460087   1.865 0.062489 .
## goout4         0.409206   0.485057   0.844 0.399078
## goout5         0.006222   0.507657   0.012 0.990224
## health2       -0.672061   0.427417  -1.572 0.116174
## health3       -1.216448   0.378853  -3.211 0.001365 **
## health4       -0.715844   0.397489  -1.801 0.072012 .
## health5       -1.043276   0.344777  -3.026 0.002541 **
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 3.408 on 1017 degrees of freedom
```

```
## Multiple R-squared:  0.2418, Adjusted R-squared:  0.2225
```

```
## F-statistic: 12.48 on 26 and 1017 DF,  p-value: < 2.2e-16
```

```
studentPerformance_p4 = lm(G3 ~ (failures+higher+schoolsup+school+romantic+studytime+address+Dalc+famsi
                        data = studentDataset)
```

```
summary(studentPerformance_p4)
```

```
##
## Call:
## lm(formula = G3 ~ (failures + higher + schoolsup + school + romantic +
##      studytime + address + Dalc + famsize + goout + health), data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.4761  -1.5432   0.3021   2.0677   8.7315
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.67659    0.67575  15.800 < 2e-16 ***
## failures     -1.80804    0.17159 -10.537 < 2e-16 ***
## higheryes      1.42592    0.41328   3.450 0.000583 ***
## schoolsupyes  -1.25532    0.34208  -3.670 0.000255 ***
## schoolMS      -0.60709    0.26501  -2.291 0.022178 *
## romanticyes   -0.54824    0.22933  -2.391 0.017001 *
## studytime2     0.18302    0.26096   0.701 0.483272
## studytime3     1.13226    0.35299   3.208 0.001380 **
## studytime4     0.49817    0.49038   1.016 0.309930
## addressU       0.56877    0.25748   2.209 0.027394 *
## Dalc2          -0.71648    0.28469  -2.517 0.011998 *
## Dalc3           0.04076    0.45434   0.090 0.928534
## Dalc4          -1.61287    0.69468  -2.322 0.020442 *
## Dalc5           0.10382    0.71195   0.146 0.884094
## famsizeLE3     0.46867    0.23721   1.976 0.048451 *
## goout2         1.09380    0.46756   2.339 0.019509 *
## goout3         0.62177    0.45596   1.364 0.172979
## goout4         0.29415    0.47806   0.615 0.538496
## goout5         0.03147    0.49952   0.063 0.949776
## health2       -0.67210    0.42879  -1.567 0.117318
## health3       -1.25563    0.38034  -3.301 0.000996 ***
## health4       -0.77580    0.39835  -1.948 0.051741 .
## health5       -1.05986    0.34533  -3.069 0.002204 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.427 on 1021 degrees of freedom
## Multiple R-squared:  0.2304, Adjusted R-squared:  0.2138
## F-statistic: 13.89 on 22 and 1021 DF,  p-value: < 2.2e-16

forwardMdl_int = lm(G3 ~ (failures+higher+schoolsup+school+romantic+studytime+address+Dalc+famsize+goou
summary(forwardMdl_int)
```

```
##
## Call:
## lm(formula = G3 ~ (failures + higher + schoolsup + school + romantic +
##      studytime + address + Dalc + famsize + goout + health)^2,
##      data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```

## -13.3916  -1.3662  0.2491  1.8013  7.9441
##
## Coefficients: (4 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.590e+00  3.385e+00   1.947  0.05191 .
## failures         -2.540e+00  1.528e+00  -1.663  0.09676 .
## higheryes         4.590e+00  2.784e+00   1.649  0.09955 .
## schoolsupyes      -2.996e+00  5.152e+00  -0.582  0.56104
## schoolMS          -3.367e+00  1.932e+00  -1.743  0.08178 .
## romanticyes       -2.552e+00  1.815e+00  -1.406  0.16025
## studytime2         4.348e+00  1.813e+00   2.398  0.01670 *
## studytime3         6.485e+00  3.734e+00   1.737  0.08277 .
## studytime4        -8.000e-01  6.536e+00  -0.122  0.90262
## addressU           8.358e-01  1.901e+00   0.440  0.66032
## Dalc2             -7.102e-01  2.472e+00  -0.287  0.77397
## Dalc3              5.633e+00  4.388e+00   1.284  0.19965
## Dalc4              4.875e+01  3.205e+01   1.521  0.12859
## Dalc5             -2.225e+00  1.632e+01  -0.136  0.89160
## famsizeLE3         8.180e-01  1.842e+00   0.444  0.65708
## goout2             3.457e+00  2.825e+00   1.224  0.22144
## goout3             3.277e+00  2.703e+00   1.212  0.22570
## goout4             8.148e-01  2.816e+00   0.289  0.77239
## goout5             1.175e+00  2.763e+00   0.425  0.67078
## health2           -1.102e+00  3.206e+00  -0.344  0.73112
## health3            1.602e+00  2.779e+00   0.576  0.56447
## health4            2.679e+00  3.101e+00   0.864  0.38784
## health5            9.115e-01  2.624e+00   0.347  0.72839
## failures:higheryes -5.093e-01  5.522e-01  -0.922  0.35663
## failures:schoolsupyes 1.645e+00  8.033e-01   2.048  0.04085 *
## failures:schoolMS    2.992e-01  5.473e-01   0.547  0.58472
## failures:romanticyes -5.744e-01  5.022e-01  -1.144  0.25305
## failures:studytime2 -5.070e-01  4.820e-01  -1.052  0.29315
## failures:studytime3 -9.167e-02  8.497e-01  -0.108  0.91411
## failures:studytime4  3.829e-01  3.146e+00   0.122  0.90317
## failures:addressU    9.893e-02  4.813e-01   0.206  0.83719
## failures:Dalc2      -5.519e-01  5.020e-01  -1.099  0.27191
## failures:Dalc3       2.944e-01  7.515e-01   0.392  0.69539
## failures:Dalc4       2.015e+01  1.145e+01   1.761  0.07870 .
## failures:Dalc5       1.292e+00  1.587e+00   0.814  0.41601
## failures:famsizeLE3  5.082e-01  4.926e-01   1.032  0.30253
## failures:goout2     -3.906e-01  1.088e+00  -0.359  0.71958
## failures:goout3      1.001e+00  1.072e+00   0.934  0.35038
## failures:goout4      6.547e-01  1.147e+00   0.571  0.56833
## failures:goout5      4.027e-01  1.103e+00   0.365  0.71527
## failures:health2     5.186e-01  1.117e+00   0.464  0.64252
## failures:health3     1.160e+00  9.213e-01   1.259  0.20830
## failures:health4     8.941e-02  1.004e+00   0.089  0.92904
## failures:health5     9.096e-01  8.911e-01   1.021  0.30767
## higheryes:schoolsupyes 3.486e+00  5.066e+00   0.688  0.49162
## higheryes:schoolMS  -7.438e-01  1.225e+00  -0.607  0.54386
## higheryes:romanticyes 1.408e-01  1.120e+00   0.126  0.89993
## higheryes:studytime2 -3.062e+00  1.189e+00  -2.577  0.01016 *
## higheryes:studytime3  2.444e+00  2.857e+00   0.856  0.39244
## higheryes:studytime4 -5.720e+00  6.182e+00  -0.925  0.35505

```

## higheryes:addressU	1.583e-01	1.204e+00	0.131	0.89544
## higheryes:Dalc2	1.906e+00	1.254e+00	1.520	0.12889
## higheryes:Dalc3	2.430e+00	2.344e+00	1.037	0.30007
## higheryes:Dalc4	-2.689e+01	1.930e+01	-1.393	0.16388
## higheryes:Dalc5	4.647e+00	7.762e+00	0.599	0.54956
## higheryes:famsizeLE3	-8.390e-01	1.240e+00	-0.677	0.49889
## higheryes:goout2	-3.631e-01	2.082e+00	-0.174	0.86163
## higheryes:goout3	-2.783e+00	1.929e+00	-1.442	0.14957
## higheryes:goout4	1.686e+00	2.046e+00	0.824	0.40999
## higheryes:goout5	1.206e+00	1.869e+00	0.645	0.51906
## higheryes:health2	-1.359e+00	2.248e+00	-0.604	0.54578
## higheryes:health3	-1.672e+00	2.128e+00	-0.786	0.43218
## higheryes:health4	-3.084e+00	2.213e+00	-1.393	0.16389
## higheryes:health5	-3.029e+00	1.988e+00	-1.524	0.12797
## schoolsupyes:schoolMS	2.068e+00	1.419e+00	1.457	0.14537
## schoolsupyes:romanticyes	4.895e-01	1.094e+00	0.447	0.65473
## schoolsupyes:studytime2	-1.496e-01	1.437e+00	-0.104	0.91712
## schoolsupyes:studytime3	-1.097e+00	1.773e+00	-0.619	0.53616
## schoolsupyes:studytime4	-1.797e+00	2.089e+00	-0.860	0.38998
## schoolsupyes:addressU	-1.224e+00	1.093e+00	-1.119	0.26332
## schoolsupyes:Dalc2	1.679e+00	1.269e+00	1.324	0.18604
## schoolsupyes:Dalc3	7.712e-01	2.853e+00	0.270	0.78698
## schoolsupyes:Dalc4	-5.983e+01	4.013e+01	-1.491	0.13641
## schoolsupyes:Dalc5	-4.871e+00	1.010e+01	-0.482	0.62975
## schoolsupyes:famsizeLE3	4.316e-01	1.168e+00	0.370	0.71177
## schoolsupyes:goout2	-4.142e+00	1.842e+00	-2.249	0.02477 *
## schoolsupyes:goout3	-4.519e+00	1.949e+00	-2.319	0.02067 *
## schoolsupyes:goout4	-3.644e+00	1.941e+00	-1.877	0.06088 .
## schoolsupyes:goout5	-5.252e+00	2.167e+00	-2.424	0.01557 *
## schoolsupyes:health2	2.858e+00	1.720e+00	1.662	0.09694 .
## schoolsupyes:health3	2.430e+00	1.533e+00	1.585	0.11327
## schoolsupyes:health4	4.464e+00	1.517e+00	2.943	0.00334 **
## schoolsupyes:health5	1.919e+00	1.431e+00	1.341	0.18036
## schoolMS:romanticyes	1.120e+00	6.585e-01	1.700	0.08944 .
## schoolMS:studytime2	9.869e-01	7.558e-01	1.306	0.19200
## schoolMS:studytime3	6.297e-01	1.051e+00	0.599	0.54916
## schoolMS:studytime4	5.127e+00	2.295e+00	2.234	0.02578 *
## schoolMS:addressU	-6.048e-01	6.437e-01	-0.940	0.34772
## schoolMS:Dalc2	1.465e-01	8.086e-01	0.181	0.85624
## schoolMS:Dalc3	-5.982e-01	1.467e+00	-0.408	0.68360
## schoolMS:Dalc4	-4.318e+01	2.379e+01	-1.815	0.06984 .
## schoolMS:Dalc5	-1.189e+00	5.933e+00	-0.200	0.84126
## schoolMS:famsizeLE3	6.475e-01	6.849e-01	0.945	0.34471
## schoolMS:goout2	1.426e+00	1.294e+00	1.102	0.27091
## schoolMS:goout3	2.300e+00	1.247e+00	1.844	0.06549 .
## schoolMS:goout4	2.906e+00	1.396e+00	2.082	0.03765 *
## schoolMS:goout5	2.657e+00	1.368e+00	1.943	0.05237 .
## schoolMS:health2	2.174e+00	1.424e+00	1.527	0.12710
## schoolMS:health3	9.158e-01	1.078e+00	0.850	0.39582
## schoolMS:health4	1.357e+00	1.178e+00	1.152	0.24967
## schoolMS:health5	-2.517e-01	9.622e-01	-0.262	0.79371
## romanticyes:studytime2	1.020e+00	6.763e-01	1.509	0.13179
## romanticyes:studytime3	6.736e-01	8.747e-01	0.770	0.44147
## romanticyes:studytime4	1.912e+00	2.190e+00	0.873	0.38280

## romanticyes:addressU	-4.749e-02	6.723e-01	-0.071	0.94370
## romanticyes:Dalc2	-5.152e-01	7.150e-01	-0.721	0.47137
## romanticyes:Dalc3	2.469e+00	1.403e+00	1.759	0.07895 .
## romanticyes:Dalc4	-4.797e+00	6.026e+00	-0.796	0.42622
## romanticyes:Dalc5	4.022e+00	8.147e+00	0.494	0.62170
## romanticyes:famsizeLE3	2.428e-01	5.968e-01	0.407	0.68422
## romanticyes:goout2	1.710e+00	1.211e+00	1.412	0.15835
## romanticyes:goout3	7.695e-01	1.207e+00	0.638	0.52381
## romanticyes:goout4	1.403e+00	1.267e+00	1.108	0.26818
## romanticyes:goout5	1.283e+00	1.290e+00	0.995	0.32023
## romanticyes:health2	-2.631e-01	1.071e+00	-0.246	0.80605
## romanticyes:health3	3.073e-01	9.686e-01	0.317	0.75111
## romanticyes:health4	-8.619e-01	1.037e+00	-0.831	0.40605
## romanticyes:health5	-3.840e-01	8.692e-01	-0.442	0.65878
## studytime2:addressU	4.890e-01	7.496e-01	0.652	0.51442
## studytime3:addressU	-6.699e-01	9.309e-01	-0.720	0.47196
## studytime4:addressU	-1.427e+00	1.747e+00	-0.817	0.41423
## studytime2:Dalc2	-1.085e+00	8.029e-01	-1.352	0.17679
## studytime3:Dalc2	-8.729e-01	1.065e+00	-0.819	0.41274
## studytime4:Dalc2	-1.432e+00	2.546e+00	-0.562	0.57407
## studytime2:Dalc3	-8.172e-01	1.380e+00	-0.592	0.55390
## studytime3:Dalc3	-3.850e+00	3.566e+00	-1.079	0.28069
## studytime4:Dalc3	-3.944e+00	3.380e+00	-1.167	0.24353
## studytime2:Dalc4	-3.610e+00	4.189e+00	-0.862	0.38909
## studytime3:Dalc4	-5.217e+00	1.228e+01	-0.425	0.67111
## studytime4:Dalc4	NA	NA	NA	NA
## studytime2:Dalc5	-1.066e+01	1.139e+01	-0.936	0.34954
## studytime3:Dalc5	-2.548e+01	1.547e+01	-1.647	0.09998 .
## studytime4:Dalc5	-2.210e+01	1.276e+01	-1.732	0.08360 .
## studytime2:famsizeLE3	-1.088e-01	6.984e-01	-0.156	0.87629
## studytime3:famsizeLE3	4.497e-01	1.037e+00	0.434	0.66468
## studytime4:famsizeLE3	2.698e+00	1.636e+00	1.649	0.09946 .
## studytime2:goout2	-1.502e+00	1.261e+00	-1.191	0.23382
## studytime3:goout2	-5.390e+00	2.407e+00	-2.239	0.02540 *
## studytime4:goout2	7.232e+00	3.546e+00	2.039	0.04173 *
## studytime2:goout3	-8.344e-01	1.256e+00	-0.664	0.50673
## studytime3:goout3	-4.804e+00	2.444e+00	-1.965	0.04973 *
## studytime4:goout3	7.450e+00	3.855e+00	1.933	0.05364 .
## studytime2:goout4	-1.882e+00	1.319e+00	-1.427	0.15390
## studytime3:goout4	-7.911e+00	2.599e+00	-3.044	0.00241 **
## studytime4:goout4	3.517e+00	4.210e+00	0.835	0.40371
## studytime2:goout5	-2.311e+00	1.396e+00	-1.656	0.09816 .
## studytime3:goout5	-5.438e+00	2.717e+00	-2.001	0.04571 *
## studytime4:goout5	2.670e+00	3.600e+00	0.742	0.45849
## studytime2:health2	-1.170e+00	1.233e+00	-0.949	0.34296
## studytime3:health2	-1.007e+00	1.909e+00	-0.528	0.59797
## studytime4:health2	4.485e+00	3.186e+00	1.408	0.15962
## studytime2:health3	-6.531e-01	1.096e+00	-0.596	0.55127
## studytime3:health3	-2.312e+00	1.681e+00	-1.375	0.16940
## studytime4:health3	1.703e-01	1.847e+00	0.092	0.92658
## studytime2:health4	-2.390e+00	1.171e+00	-2.041	0.04159 *
## studytime3:health4	-3.038e+00	1.709e+00	-1.777	0.07588 .
## studytime4:health4	-1.194e+00	2.437e+00	-0.490	0.62416
## studytime2:health5	2.602e-01	9.322e-01	0.279	0.78023

## studytime3:health5	-1.725e+00	1.572e+00	-1.097	0.27292
## studytime4:health5	4.150e+00	2.191e+00	1.894	0.05852
## addressU:Dalc2	-7.901e-01	7.464e-01	-1.059	0.29013
## addressU:Dalc3	-2.424e+00	1.408e+00	-1.722	0.08544
## addressU:Dalc4	-4.823e+00	1.021e+01	-0.473	0.63665
## addressU:Dalc5	1.803e+00	8.092e+00	0.223	0.82370
## addressU:famsizeLE3	1.596e-01	7.176e-01	0.222	0.82401
## addressU:goout2	-8.941e-01	1.366e+00	-0.655	0.51292
## addressU:goout3	-6.488e-01	1.340e+00	-0.484	0.62853
## addressU:goout4	-1.350e+00	1.440e+00	-0.937	0.34881
## addressU:goout5	-4.220e-01	1.523e+00	-0.277	0.78185
## addressU:health2	1.330e+00	1.456e+00	0.913	0.36132
## addressU:health3	1.669e+00	1.139e+00	1.466	0.14314
## addressU:health4	1.046e+00	1.158e+00	0.903	0.36654
## addressU:health5	1.438e+00	1.009e+00	1.426	0.15432
## Dalc2:famsizeLE3	-8.212e-01	7.084e-01	-1.159	0.24669
## Dalc3:famsizeLE3	2.061e-03	1.252e+00	0.002	0.99869
## Dalc4:famsizeLE3	-3.082e+01	2.248e+01	-1.371	0.17089
## Dalc5:famsizeLE3	-5.829e-02	4.943e+00	-0.012	0.99059
## Dalc2:goout2	4.688e-01	1.950e+00	0.240	0.81004
## Dalc3:goout2	-6.719e+00	3.155e+00	-2.130	0.03347 *
## Dalc4:goout2	5.584e+01	3.976e+01	1.404	0.16061
## Dalc5:goout2	9.979e+00	9.898e+00	1.008	0.31369
## Dalc2:goout3	6.826e-01	1.939e+00	0.352	0.72489
## Dalc3:goout3	-7.388e+00	3.132e+00	-2.359	0.01856 *
## Dalc4:goout3	-2.648e+01	2.118e+01	-1.250	0.21158
## Dalc5:goout3	1.304e+01	5.633e+00	2.315	0.02086 *
## Dalc2:goout4	-1.139e-01	1.952e+00	-0.058	0.95348
## Dalc3:goout4	-6.916e+00	3.106e+00	-2.226	0.02626 *
## Dalc4:goout4	1.300e+01	8.623e+00	1.507	0.13208
## Dalc5:goout4	-5.658e+00	5.782e+00	-0.979	0.32805
## Dalc2:goout5	-1.164e+00	2.053e+00	-0.567	0.57099
## Dalc3:goout5	-5.000e+00	3.181e+00	-1.572	0.11634
## Dalc4:goout5	NA	NA	NA	NA
## Dalc5:goout5	NA	NA	NA	NA
## Dalc2:health2	-1.378e+00	1.286e+00	-1.072	0.28413
## Dalc3:health2	-1.165e+00	2.557e+00	-0.456	0.64882
## Dalc4:health2	-2.776e+01	2.240e+01	-1.239	0.21566
## Dalc5:health2	NA	NA	NA	NA
## Dalc2:health3	-9.810e-01	1.349e+00	-0.727	0.46729
## Dalc3:health3	-7.091e-01	2.500e+00	-0.284	0.77678
## Dalc4:health3	2.313e+01	1.728e+01	1.339	0.18109
## Dalc5:health3	1.134e+00	7.022e+00	0.162	0.87169
## Dalc2:health4	9.383e-01	1.231e+00	0.763	0.44597
## Dalc3:health4	1.367e+00	4.434e+00	0.308	0.75788
## Dalc4:health4	4.486e+00	5.601e+00	0.801	0.42343
## Dalc5:health4	-6.022e+00	1.209e+01	-0.498	0.61843
## Dalc2:health5	-3.466e-01	1.054e+00	-0.329	0.74234
## Dalc3:health5	-3.504e-01	1.993e+00	-0.176	0.86045
## Dalc4:health5	-3.072e+01	1.880e+01	-1.634	0.10259
## Dalc5:health5	1.806e+00	9.745e+00	0.185	0.85301
## famsizeLE3:goout2	1.657e+00	1.246e+00	1.329	0.18407
## famsizeLE3:goout3	8.139e-01	1.213e+00	0.671	0.50228
## famsizeLE3:goout4	3.187e-01	1.266e+00	0.252	0.80132



```

## famsizeLE3:goout5          6.095e-01  1.365e+00  0.446  0.65547
## famsizeLE3:health2        -1.685e+00  1.082e+00 -1.557  0.11993
## famsizeLE3:health3        -9.085e-01  9.872e-01 -0.920  0.35769
## famsizeLE3:health4        -7.637e-04  1.015e+00 -0.001  0.99940
## famsizeLE3:health5        -1.248e+00  8.771e-01 -1.422  0.15528
## goout2:health2            5.819e-01  2.182e+00  0.267  0.78984
## goout3:health2            2.263e+00  2.082e+00  1.087  0.27725
## goout4:health2            2.329e+00  2.118e+00  1.099  0.27191
## goout5:health2            1.052e+00  2.258e+00  0.466  0.64155
## goout2:health3           -2.469e+00  1.889e+00 -1.307  0.19142
## goout3:health3           -1.966e+00  1.861e+00 -1.056  0.29130
## goout4:health3           -2.487e+00  1.928e+00 -1.290  0.19742
## goout5:health3           -4.238e+00  2.134e+00 -1.986  0.04741 *
## goout2:health4           -1.220e+00  2.394e+00 -0.509  0.61055
## goout3:health4            6.162e-01  2.246e+00  0.274  0.78384
## goout4:health4            3.440e-01  2.288e+00  0.150  0.88054
## goout5:health4            2.667e-01  2.484e+00  0.107  0.91453
## goout2:health5           -1.196e+00  1.712e+00 -0.699  0.48503
## goout3:health5            1.719e+00  1.668e+00  1.031  0.30284
## goout4:health5            3.672e-01  1.683e+00  0.218  0.82732
## goout5:health5           -3.529e-01  1.835e+00 -0.192  0.84754
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.327 on 815 degrees of freedom
## Multiple R-squared:  0.4211, Adjusted R-squared:  0.2591
## F-statistic: 2.6 on 228 and 815 DF, p-value: < 2.2e-16

forwardMdl_int1 = lm(G3 ~ (failures+higher+schoolsup+school+romantic+studytime+address+Dalc+famsize+goout+health+failures:schoolsup +
summary(forwardMdl_int1)

##
## Call:
## lm(formula = G3 ~ (failures + higher + schoolsup + school + romantic +
##   studytime + address + Dalc + famsize + goout + health + failures:schoolsup +
##   higher:studytime + schoolsup:goout + schoolsup:health + school:studytime +
##   school:goout + studytime:goout + studytime:health + Dalc:goout),
##   data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.4595  -1.4824   0.2603   2.0490   7.7728
##
## Coefficients: (2 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      8.79533     1.19827   7.340 4.53e-13 ***
## failures         -2.07407     0.18757 -11.057 < 2e-16 ***
## higheryes         1.96842     0.53774   3.661 0.000265 ***
## schoolsupyes     -0.24336     1.55566  -0.156 0.875722
## schoolMS         -3.10170     1.01766  -3.048 0.002367 **
## romanticyes      -0.52367     0.23737  -2.206 0.027611 *
## studytime2        3.50748     1.44097   2.434 0.015109 *
## studytime3        7.65334     2.74575   2.787 0.005419 **
## studytime4       -3.23124     3.42016  -0.945 0.345017
## addressU          0.68023     0.27412   2.481 0.013253 *

```

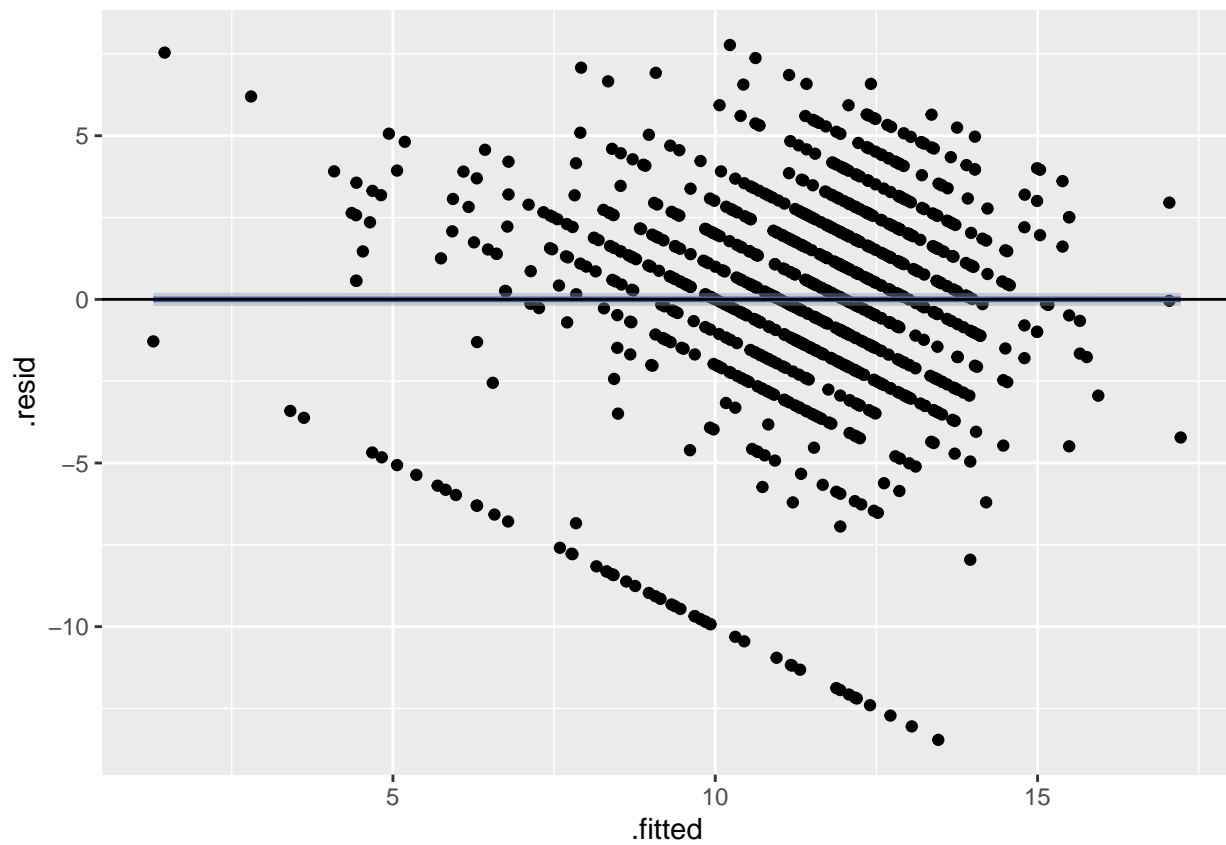
## Dalc2	-1.65770	1.64376	-1.008	0.313475	
## Dalc3	4.51811	2.24473	2.013	0.044417	*
## Dalc4	-1.57331	1.29849	-1.212	0.225945	
## Dalc5	-0.05161	1.03734	-0.050	0.960331	
## famsizeLE3	0.50905	0.24361	2.090	0.036914	*
## goout2	2.06314	1.01507	2.033	0.042376	*
## goout3	1.30185	1.00859	1.291	0.197095	
## goout4	1.45447	1.05340	1.381	0.167679	
## goout5	1.75621	1.12264	1.564	0.118060	
## health2	0.60464	0.86639	0.698	0.485415	
## health3	-0.66097	0.76239	-0.867	0.386174	
## health4	0.49567	0.77876	0.636	0.524608	
## health5	-0.67830	0.66380	-1.022	0.307110	
## failures:schoolsupyes	1.59277	0.53384	2.984	0.002920	**
## higheryes:studytime2	-1.51932	0.88081	-1.725	0.084863	.
## higheryes:studytime3	-0.18739	1.94542	-0.096	0.923284	
## higheryes:studytime4	0.96047	3.02806	0.317	0.751169	
## schoolsupyes:goout2	-4.88833	1.42645	-3.427	0.000636	***
## schoolsupyes:goout3	-3.56036	1.44200	-2.469	0.013720	*
## schoolsupyes:goout4	-2.87103	1.44514	-1.987	0.047240	*
## schoolsupyes:goout5	-5.25140	1.62367	-3.234	0.001261	**
## schoolsupyes:health2	2.11051	1.42348	1.483	0.138498	
## schoolsupyes:health3	2.24864	1.27336	1.766	0.077728	.
## schoolsupyes:health4	2.58365	1.28290	2.014	0.044296	*
## schoolsupyes:health5	2.62579	1.19243	2.202	0.027899	*
## schoolMS:studytime2	0.59786	0.57304	1.043	0.297063	
## schoolMS:studytime3	0.84418	0.83506	1.011	0.312305	
## schoolMS:studytime4	3.30040	1.75216	1.884	0.059917	.
## schoolMS:goout2	1.34253	1.11335	1.206	0.228172	
## schoolMS:goout3	2.54857	1.05669	2.412	0.016058	*
## schoolMS:goout4	3.16334	1.11144	2.846	0.004519	**
## schoolMS:goout5	1.78553	1.13484	1.573	0.115961	
## studytime2:goout2	-0.96641	1.14262	-0.846	0.397881	
## studytime3:goout2	-4.64492	1.88962	-2.458	0.014141	*
## studytime4:goout2	5.12257	2.44913	2.092	0.036736	*
## studytime2:goout3	-1.27448	1.11501	-1.143	0.253311	
## studytime3:goout3	-4.55559	1.87034	-2.436	0.015043	*
## studytime4:goout3	3.23950	2.53300	1.279	0.201234	
## studytime2:goout4	-1.85335	1.14688	-1.616	0.106423	
## studytime3:goout4	-5.71197	2.00793	-2.845	0.004539	**
## studytime4:goout4	2.15436	2.99562	0.719	0.472210	
## studytime2:goout5	-2.09542	1.19911	-1.747	0.080873	.
## studytime3:goout5	-4.23842	2.09215	-2.026	0.043053	*
## studytime4:goout5	1.52248	2.32038	0.656	0.511895	
## studytime2:health2	-1.88244	1.04041	-1.809	0.070711	.
## studytime3:health2	-1.26928	1.65687	-0.766	0.443819	
## studytime4:health2	0.08011	2.21686	0.036	0.971179	
## studytime2:health3	-0.85884	0.91957	-0.934	0.350561	
## studytime3:health3	-2.27637	1.44852	-1.572	0.116393	
## studytime4:health3	-1.49903	1.57469	-0.952	0.341359	
## studytime2:health4	-1.65074	0.95678	-1.725	0.084790	.
## studytime3:health4	-2.76377	1.45821	-1.895	0.058350	.
## studytime4:health4	-2.14841	2.06784	-1.039	0.299081	
## studytime2:health5	-0.40386	0.81421	-0.496	0.620000	

```
## studytime3:health5      -2.05327      1.37166     -1.497  0.134741
## studytime4:health5     -0.19113      1.61964     -0.118  0.906084
## Dalc2:goout2            1.00381      1.75203      0.573  0.566820
## Dalc3:goout2           -4.85899      2.55174     -1.904  0.057184 .
## Dalc4:goout2           -0.13359      2.07082     -0.065  0.948578
## Dalc5:goout2            1.19243      2.68011      0.445  0.656479
## Dalc2:goout3            1.32103      1.71856      0.769  0.442270
## Dalc3:goout3           -4.78638      2.42498     -1.974  0.048691 *
## Dalc4:goout3           -0.89921      2.19804     -0.409  0.682558
## Dalc5:goout3            1.40799      1.78638      0.788  0.430785
## Dalc2:goout4            0.85312      1.73013      0.493  0.622059
## Dalc3:goout4           -5.34689      2.38842     -2.239  0.025405 *
## Dalc4:goout4            0.14390      1.76188      0.082  0.934924
## Dalc5:goout4           -2.83252      2.73779     -1.035  0.301114
## Dalc2:goout5           -0.21352      1.80651     -0.118  0.905936
## Dalc3:goout5           -3.83476      2.42813     -1.579  0.114594
## Dalc4:goout5            NA            NA            NA            NA
## Dalc5:goout5            NA            NA            NA            NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.371 on 964 degrees of freedom
## Multiple R-squared:  0.2967, Adjusted R-squared:  0.2391
## F-statistic: 5.148 on 79 and 964 DF,  p-value: < 2.2e-16
anova(studentPerformance_p4,forwardMdl_int1)

## Analysis of Variance Table
##
## Model 1: G3 ~ (failures + higher + schoolsup + school + romantic + studytime +
##      address + Dalc + famsize + goout + health)
## Model 2: G3 ~ (failures + higher + schoolsup + school + romantic + studytime +
##      address + Dalc + famsize + goout + health + failures:schoolsup +
##      higher:studytime + schoolsup:goout + schoolsup:health + school:studytime +
##      school:goout + studytime:goout + studytime:health + Dalc:goout)
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1    1021 11989
## 2     964 10956 57      1033 1.5945 0.004019 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### Assumption code

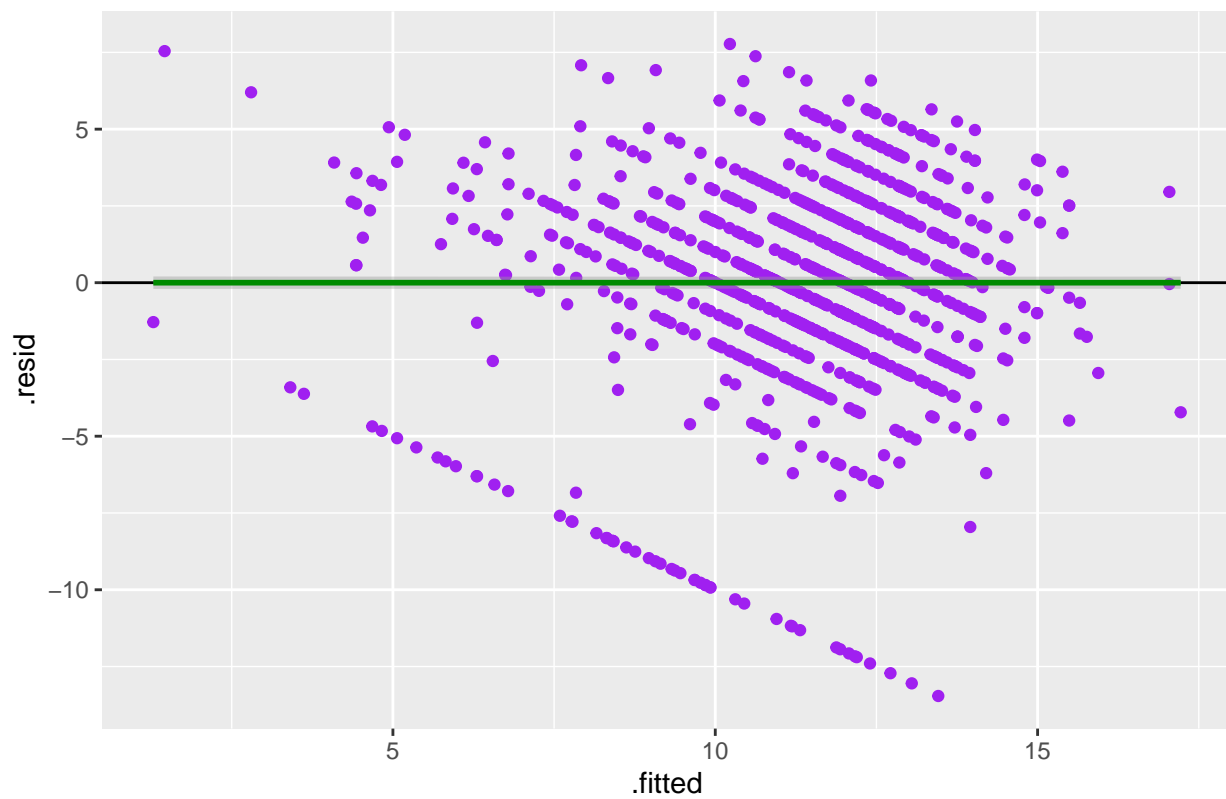
```
#Linearity Assumption
library(ggplot2)
ggplot(forwardMdl_int1, aes(x=.fitted, y=.resid)) +
  geom_point() +geom_smooth()+
  geom_hline(yintercept = 0)
```



*# Equal Variance Assumption*

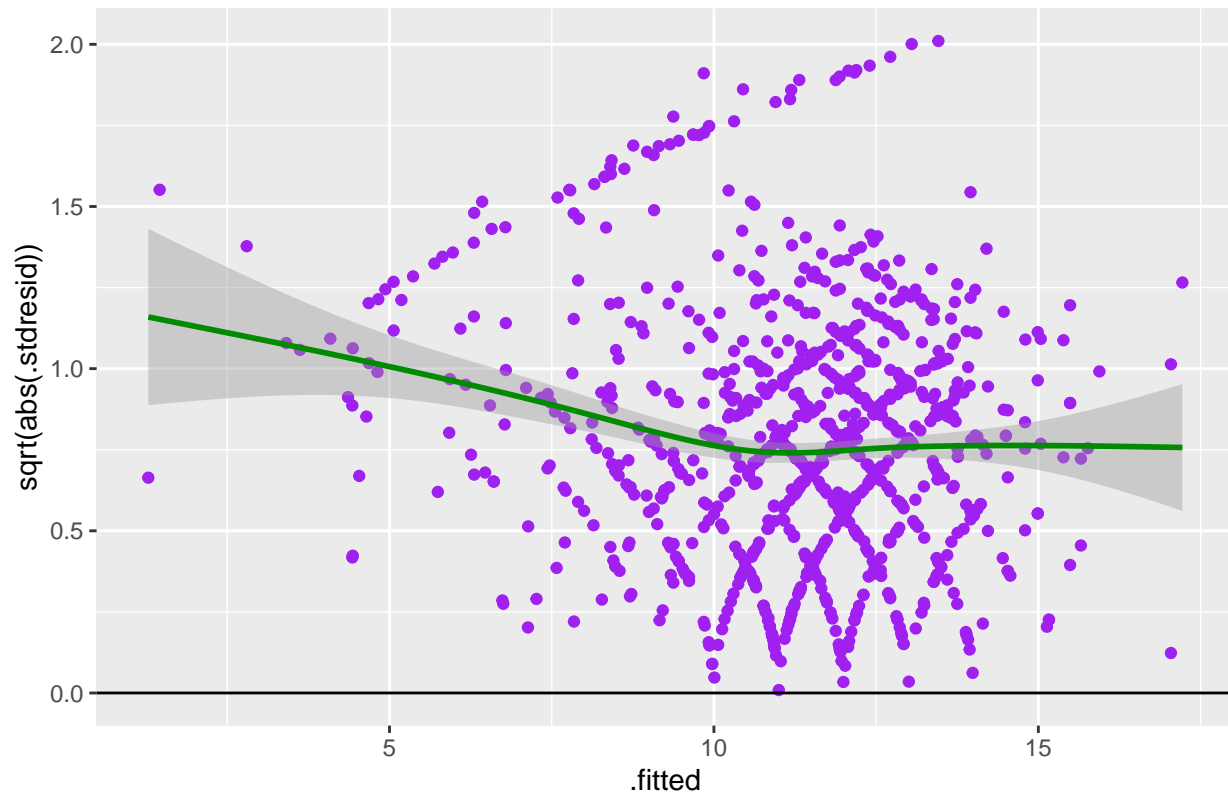
```
ggplot(forwardMdl_int1, aes(x=.fitted, y=.resid)) +  
  geom_point(colour = "purple") +  
  geom_hline(yintercept = 0) +  
  geom_smooth(colour = "green4")+  
  ggtitle("Residual plot: Residual vs Fitted values")
```

Residual plot: Residual vs Fitted values



```
ggplot(forwardMdl_int1, aes(x=.fitted, y=sqrt(abs(.stdresid)))) +  
  geom_point(colour = "purple") +  
  geom_hline(yintercept = 0) +  
  geom_smooth( colour = "green4")+  
  ggtitle("Scale-Location plot : Standardized Residual vs Fitted values")
```

Scale–Location plot : Standardized Residual vs Fitted values

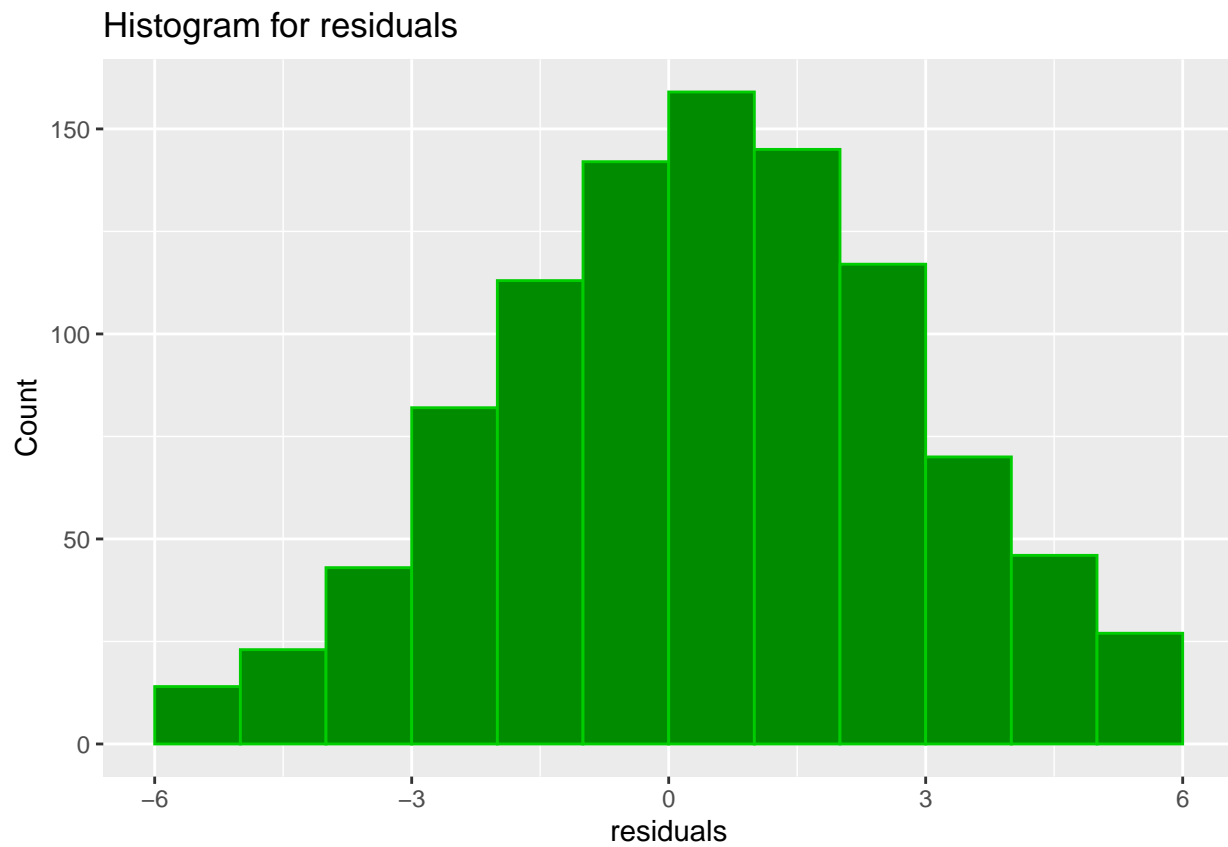


```
library(lmtest)

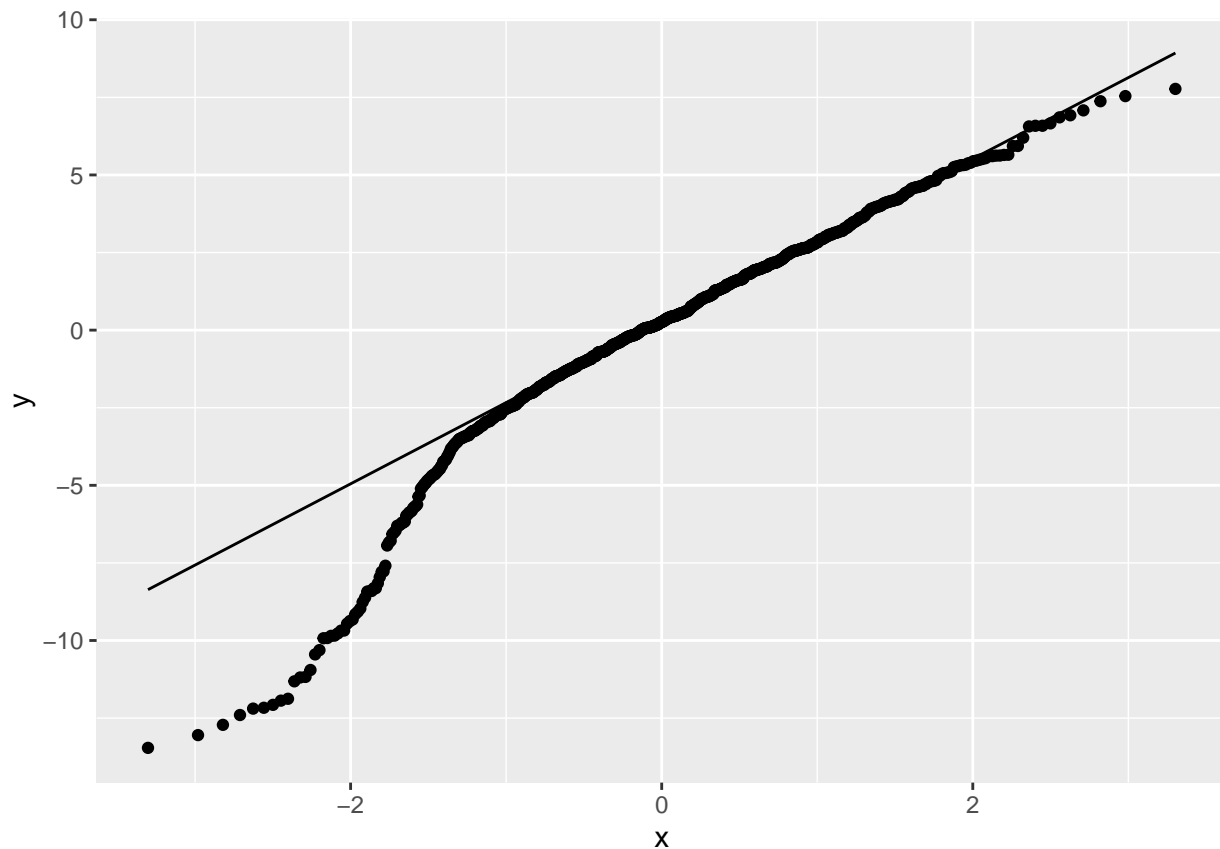
bptest(forwardMdl_int1) # It doesn't have heteroscedasticity.

##
## studentized Breusch-Pagan test
##
## data: forwardMdl_int1
## BP = 84.201, df = 79, p-value = 0.3235
# Normality Assumption

ggplot(data=studentDataset, aes(residuals(forwardMdl_int1))) +
  geom_histogram(breaks = seq(-6,6,by=1), col="green3", fill="green4") +
  labs(title="Histogram for residuals") +
  labs(x="residuals", y="Count")
```



```
ggplot(studentDataset, aes(sample=forwardMdl_int1$residuals)) +  
  stat_qq() +  
  stat_qq_line()
```



```
#Testing for Normality
shapiro.test(residuals(forwardMdl_int1)) # We don't have normality
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(forwardMdl_int1)
## W = 0.94129, p-value < 2.2e-16
```

```
#Multicollinearity
library(mctest)
#pairs(~G3+studytime+goout+health,data=studentDataset)
```

```
imcdiag(forwardMdl_int1, method="VIF")
```

```
##
## Call:
## imcdiag(mod = forwardMdl_int1, method = "VIF")
##
##
##  VIF Multicollinearity Diagnostics
##
##              VIF detection
## failures          1.3901      0
## higheryes         2.0713      0
## schoolsupyes      22.4504      1
## schoolMS         18.3272      1
## romanticyes       1.1856      0
```



## studytime2	47.6192	1
## studytime3	90.7846	1
## studytime4	60.0211	1
## addressU	1.3699	0
## Dalc2	37.8472	1
## Dalc3	28.5685	1
## Dalc4	Inf	1
## Dalc5	Inf	1
## famsizeLE3	1.1295	0
## goout2	17.1422	1
## goout3	20.3623	1
## goout4	17.3436	1
## goout5	15.2527	1
## health2	7.1663	0
## health3	8.7307	0
## health4	7.7371	0
## health5	9.5196	0
## failures:schoolsupyes	1.5300	0
## higheryes:studytime2	17.6776	1
## higheryes:studytime3	44.6505	1
## higheryes:studytime4	45.6230	1
## schoolsupyes:goout2	5.7212	0
## schoolsupyes:goout3	5.6750	0
## schoolsupyes:goout4	5.0070	0
## schoolsupyes:goout5	3.2038	0
## schoolsupyes:health2	2.6358	0
## schoolsupyes:health3	3.6169	0
## schoolsupyes:health4	3.3955	0
## schoolsupyes:health5	4.9279	0
## schoolMS:studytime2	3.0907	0
## schoolMS:studytime3	2.0181	0
## schoolMS:studytime4	2.1444	0
## schoolMS:goout2	5.0932	0
## schoolMS:goout3	8.3228	0
## schoolMS:goout4	5.8569	0
## schoolMS:goout5	5.3943	0
## studytime2:goout2	11.2137	1
## studytime3:goout2	15.5228	1
## studytime4:goout2	9.8448	0
## studytime2:goout3	14.3606	1
## studytime3:goout3	20.3667	1
## studytime4:goout3	11.6164	1
## studytime2:goout4	13.7744	1
## studytime3:goout4	7.2996	0
## studytime4:goout4	4.7101	0
## studytime2:goout5	8.3714	0
## studytime3:goout5	4.9442	0
## studytime4:goout5	5.6194	0
## studytime2:health2	6.3843	0
## studytime3:health2	3.5710	0
## studytime4:health2	2.5795	0
## studytime2:health3	6.7875	0
## studytime3:health3	7.1015	0
## studytime4:health3	4.4894	0

```

## studytime2:health4      5.6756      0
## studytime3:health4      7.3694      0
## studytime4:health4      2.9867      0
## studytime2:health5      8.4187      0
## studytime3:health5      8.7729      0
## studytime4:health5      3.8598      0
## Dalc2:goout2            9.6388      0
## Dalc3:goout2            4.5481      0
## Dalc4:goout2             Inf      1
## Dalc5:goout2             Inf      1
## Dalc2:goout3           14.9252      1
## Dalc3:goout3            8.1514      0
## Dalc4:goout3             Inf      1
## Dalc5:goout3             Inf      1
## Dalc2:goout4           14.6606      1
## Dalc3:goout4           11.2897      1
## Dalc4:goout4             Inf      1
## Dalc5:goout4             Inf      1
## Dalc2:goout5            9.1760      0
## Dalc3:goout5            9.6767      0
## Dalc4:goout5             Inf      1
## Dalc5:goout5             Inf      1
##
## Multicollinearity may be due to schoolsupyes schoolMS studytime2 studytime3 studytime4 Dalc2 Dalc3 D
##
## 1 --> COLLINEARITY is detected by the test
## 0 --> COLLINEARITY is not detected by the test
##
## =====
forwardMdl_int2 = lm(G3 ~ (failures+higher+schoolsup+school+romantic+studytime+address+Dalc+famsize+goout
summary(forwardMdl_int2)

##
## Call:
## lm(formula = G3 ~ (failures + higher + schoolsup + school + romantic +
##      studytime + address + Dalc + famsize + goout + health + failures:schoolsup +
##      schoolsup:goout + schoolsup:health + school:studytime + school:goout +
##      studytime:health), data = studentDataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.5199  -1.5709   0.3203   2.0301   8.0638
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    10.32087    0.91709  11.254 < 2e-16 ***
## failures        -2.06027    0.18339 -11.235 < 2e-16 ***
## higheryes        1.52800    0.41286   3.701 0.000227 ***
## schoolsupyes     -0.97777    1.46673  -0.667 0.505162
## schoolMS        -2.89238    0.95357  -3.033 0.002483 **
## romanticyes     -0.53566    0.23155  -2.313 0.020905 *
## studytime2       0.72607    0.71905   1.010 0.312859
## studytime3       2.93350    1.26460   2.320 0.020559 *
## studytime4       1.05524    1.28695   0.820 0.412441

```

```

## addressU          0.59024    0.26131    2.259 0.024116 *
## Dalc2             -0.78128    0.28736   -2.719 0.006665 **
## Dalc3              0.20127    0.45662    0.441 0.659472
## Dalc4             -1.54100    0.69631   -2.213 0.027118 *
## Dalc5              0.36431    0.72658    0.501 0.616199
## famsizeLE3        0.54342    0.23867    2.277 0.023008 *
## goout2             1.30773    0.61556    2.124 0.033880 *
## goout3             0.35476    0.61467    0.577 0.563969
## goout4            -0.10565    0.63639   -0.166 0.868186
## goout5             0.07413    0.66255    0.112 0.910934
## health2            0.56064    0.85172    0.658 0.510530
## health3           -1.00463    0.74558   -1.347 0.178144
## health4            0.38872    0.76348    0.509 0.610765
## health5           -0.72147    0.64509   -1.118 0.263660
## failures:schoolsupyes 1.51599    0.52839    2.869 0.004204 **
## schoolsupyes:goout2 -3.88775    1.33111   -2.921 0.003572 **
## schoolsupyes:goout3 -2.56820    1.34428   -1.910 0.056361 .
## schoolsupyes:goout4 -2.05501    1.35840   -1.513 0.130645
## schoolsupyes:goout5 -3.99757    1.53982   -2.596 0.009567 **
## schoolsupyes:health2 1.71672    1.40788    1.219 0.222994
## schoolsupyes:health3 2.26558    1.25923    1.799 0.072295 .
## schoolsupyes:health4 2.34196    1.27080    1.843 0.065641 .
## schoolsupyes:health5 2.42643    1.17344    2.068 0.038918 *
## schoolMS:studytime2 0.60873    0.55564    1.096 0.273535
## schoolMS:studytime3 1.04362    0.80769    1.292 0.196621
## schoolMS:studytime4 1.33055    1.47473    0.902 0.367152
## schoolMS:goout2     1.01842    1.06078    0.960 0.337257
## schoolMS:goout3     2.27018    0.99703    2.277 0.023001 *
## schoolMS:goout4     2.70094    1.05574    2.558 0.010665 *
## schoolMS:goout5     1.77044    1.07634    1.645 0.100313
## studytime2:health2 -1.81858    1.01863   -1.785 0.074513 .
## studytime3:health2 -1.35672    1.63997   -0.827 0.408274
## studytime4:health2 -1.17298    1.99290   -0.589 0.556278
## studytime2:health3 -0.48780    0.90252   -0.540 0.588983
## studytime3:health3 -1.87334    1.41588   -1.323 0.186108
## studytime4:health3 -0.41088    1.53611   -0.267 0.789156
## studytime2:health4 -1.49437    0.93728   -1.594 0.111170
## studytime3:health4 -2.82163    1.42904   -1.974 0.048602 *
## studytime4:health4 -2.45577    1.83317   -1.340 0.180671
## studytime2:health5 -0.35775    0.79546   -0.450 0.652996
## studytime3:health5 -2.15527    1.34447   -1.603 0.109239
## studytime4:health5 -0.32237    1.53311   -0.210 0.833500
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.378 on 993 degrees of freedom
## Multiple R-squared:  0.2725, Adjusted R-squared:  0.2359
## F-statistic: 7.439 on 50 and 993 DF,  p-value: < 2.2e-16
imcdiag(forwardMdl_int2, method="VIF")

##
## Call:
## imcdiag(mod = forwardMdl_int2, method = "VIF")
##

```

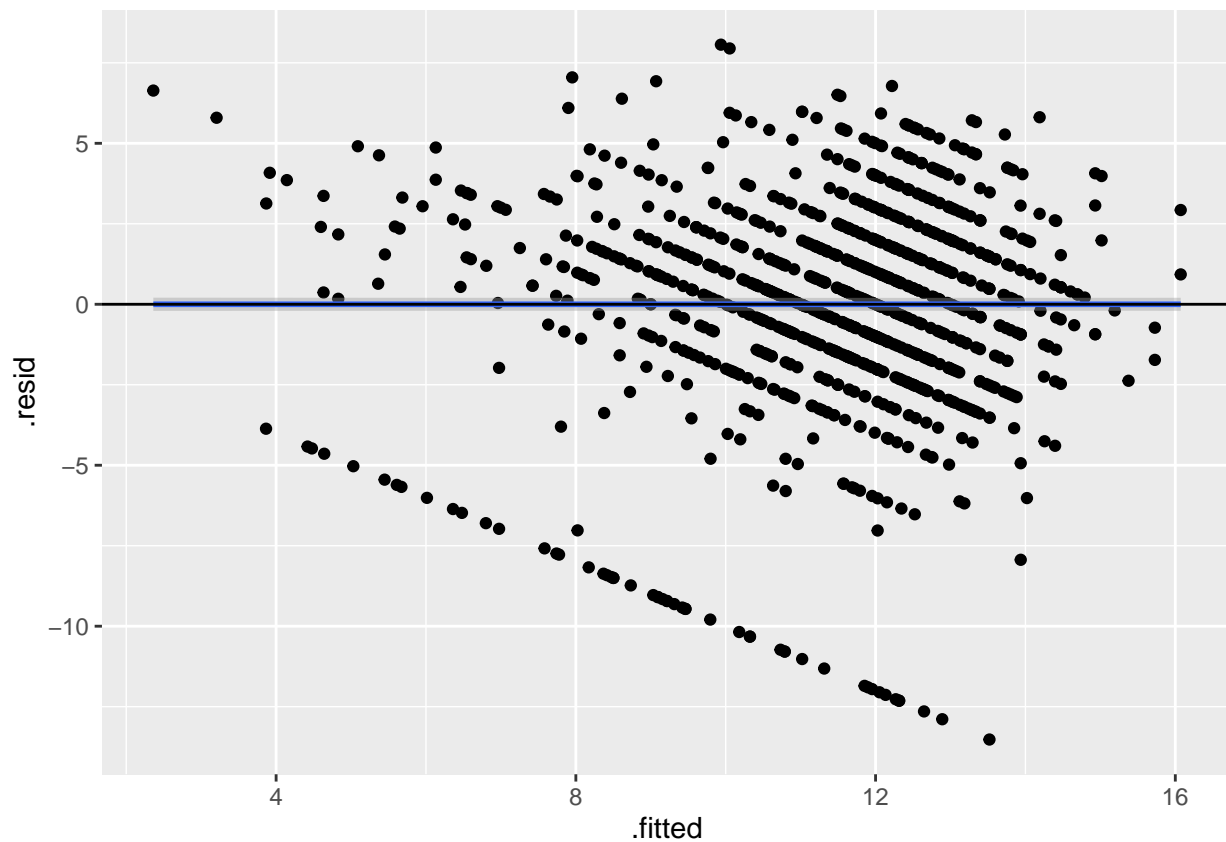
```

##
## VIF Multicollinearity Diagnostics
##
## VIF detection
## failures 1.3231 0
## higheryes 1.2158 0
## schoolsupyes 19.8730 1
## schoolMS 16.0238 1
## romanticyes 1.1234 0
## studytime2 11.8076 1
## studytime3 19.1762 1
## studytime4 8.4626 0
## addressU 1.2396 0
## Dalc2 1.1518 0
## Dalc3 1.1772 0
## Dalc4 1.0770 0
## Dalc5 1.1726 0
## famsizeLE3 1.0796 0
## goout2 6.2775 0
## goout3 7.5310 0
## goout4 6.3034 0
## goout5 5.2902 0
## health2 6.8966 0
## health3 8.3150 0
## health4 7.4052 0
## health5 8.9526 0
## failures:schoolsupyes 1.4926 0
## schoolsupyes:goout2 4.9610 0
## schoolsupyes:goout3 4.9112 0
## schoolsupyes:goout4 4.4054 0
## schoolsupyes:goout5 2.8693 0
## schoolsupyes:health2 2.5675 0
## schoolsupyes:health3 3.5222 0
## schoolsupyes:health4 3.3177 0
## schoolsupyes:health5 4.7521 0
## schoolMS:studytime2 2.8936 0
## schoolMS:studytime3 1.8800 0
## schoolMS:studytime4 1.5127 0
## schoolMS:goout2 4.6041 0
## schoolMS:goout3 7.3782 0
## schoolMS:goout4 5.2624 0
## schoolMS:goout5 4.8320 0
## studytime2:health2 6.0940 0
## studytime3:health2 3.4838 0
## studytime4:health2 2.0758 0
## studytime2:health3 6.5106 0
## studytime3:health3 6.7565 0
## studytime4:health3 4.2542 0
## studytime2:health4 5.4237 0
## studytime3:health4 7.0477 0
## studytime4:health4 2.3374 0
## studytime2:health5 8.0016 0
## studytime3:health5 8.3931 0
## studytime4:health5 3.4438 0

```

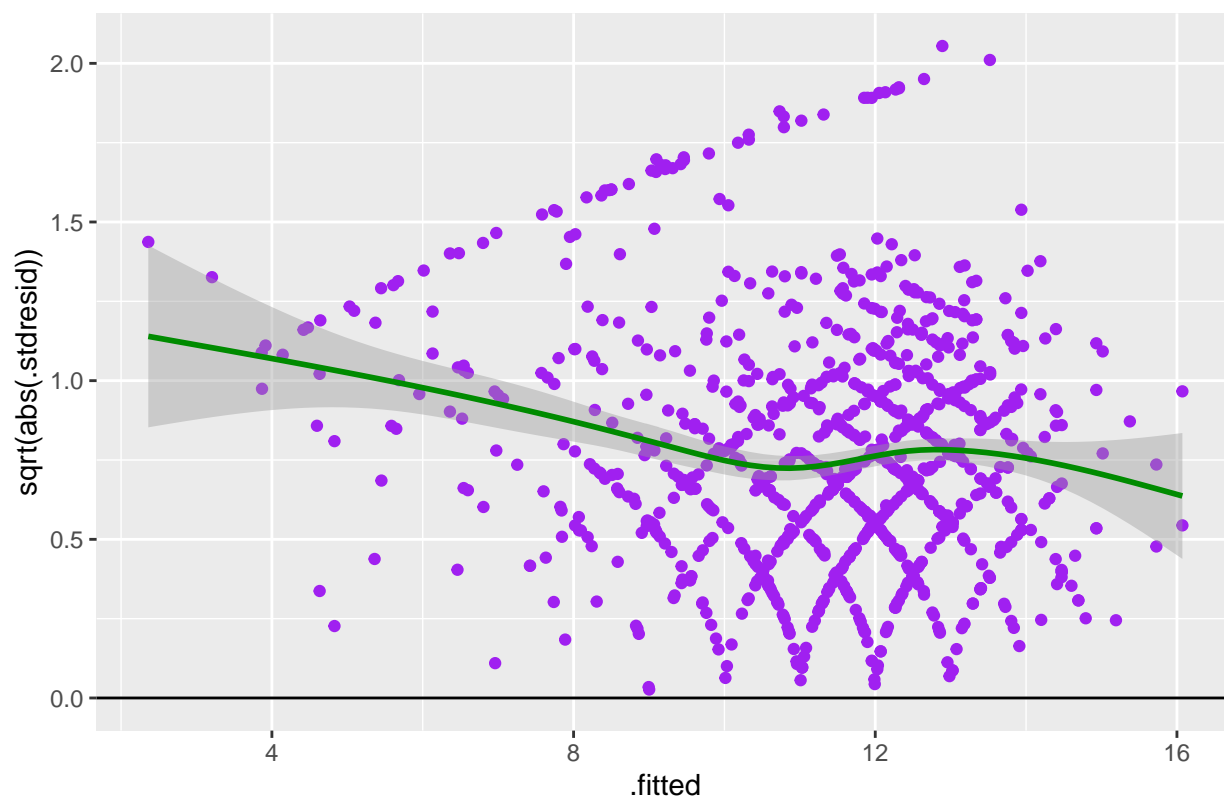
```
##
## Multicollinearity may be due to schoolsupyes schoolMS studytime2 studytime3 regressors
##
## 1 --> COLLINEARITY is detected by the test
## 0 --> COLLINEARITY is not detected by the test
##
## =====
```

```
#Linearity Assumption
library(ggplot2)
ggplot(forwardMdl_int2, aes(x=.fitted, y=.resid)) +
  geom_point() +geom_smooth()+
  geom_hline(yintercept = 0)
```



```
ggplot(forwardMdl_int2, aes(x=.fitted, y=sqrt(abs(.stdresid)))) +
  geom_point(colour = "purple") +
  geom_hline(yintercept = 0) +
  geom_smooth( colour = "green4")+
  ggtitle("Scale-Location plot : Standardized Residual vs Fitted values")
```

Scale–Location plot : Standardized Residual vs Fitted values



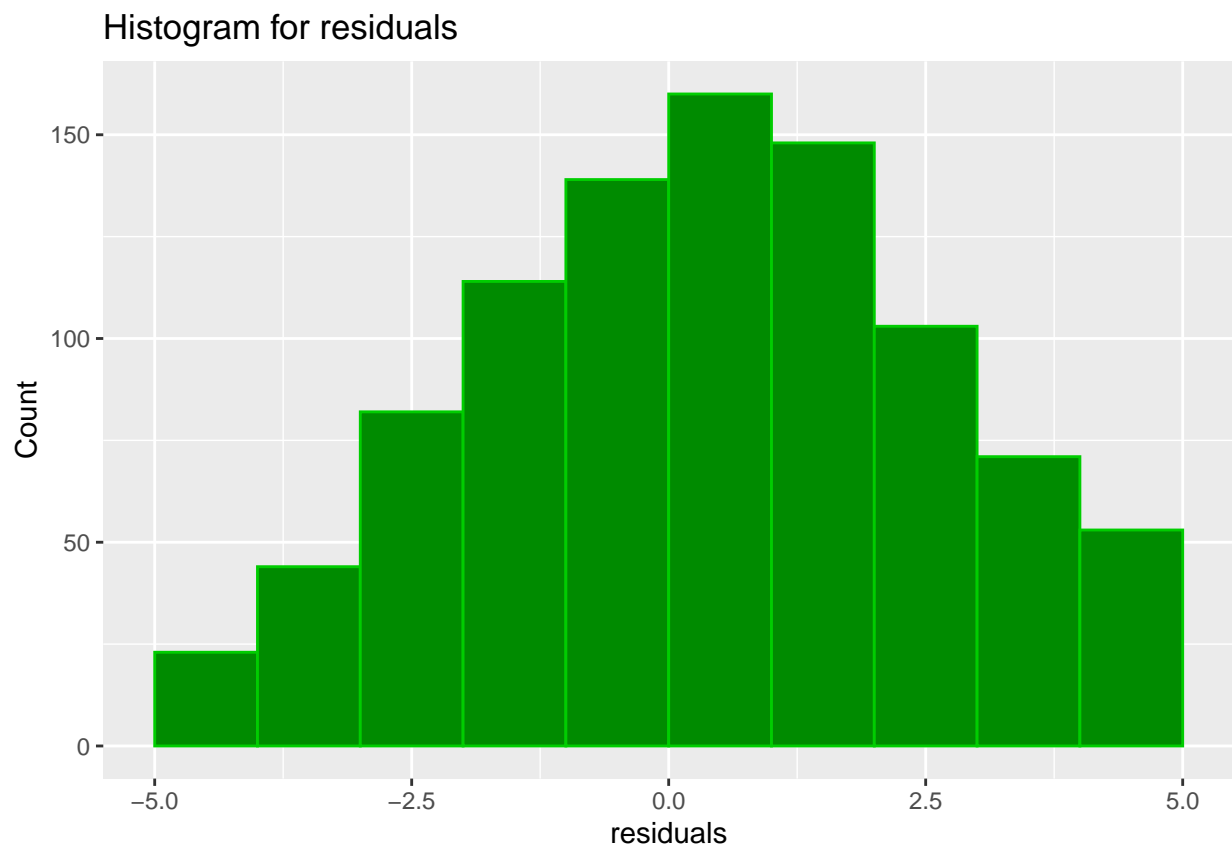
```
library(lmtest)
```

```
bptest(forwardMdl_int2) # It doesn't have heteroscedasticity.
```

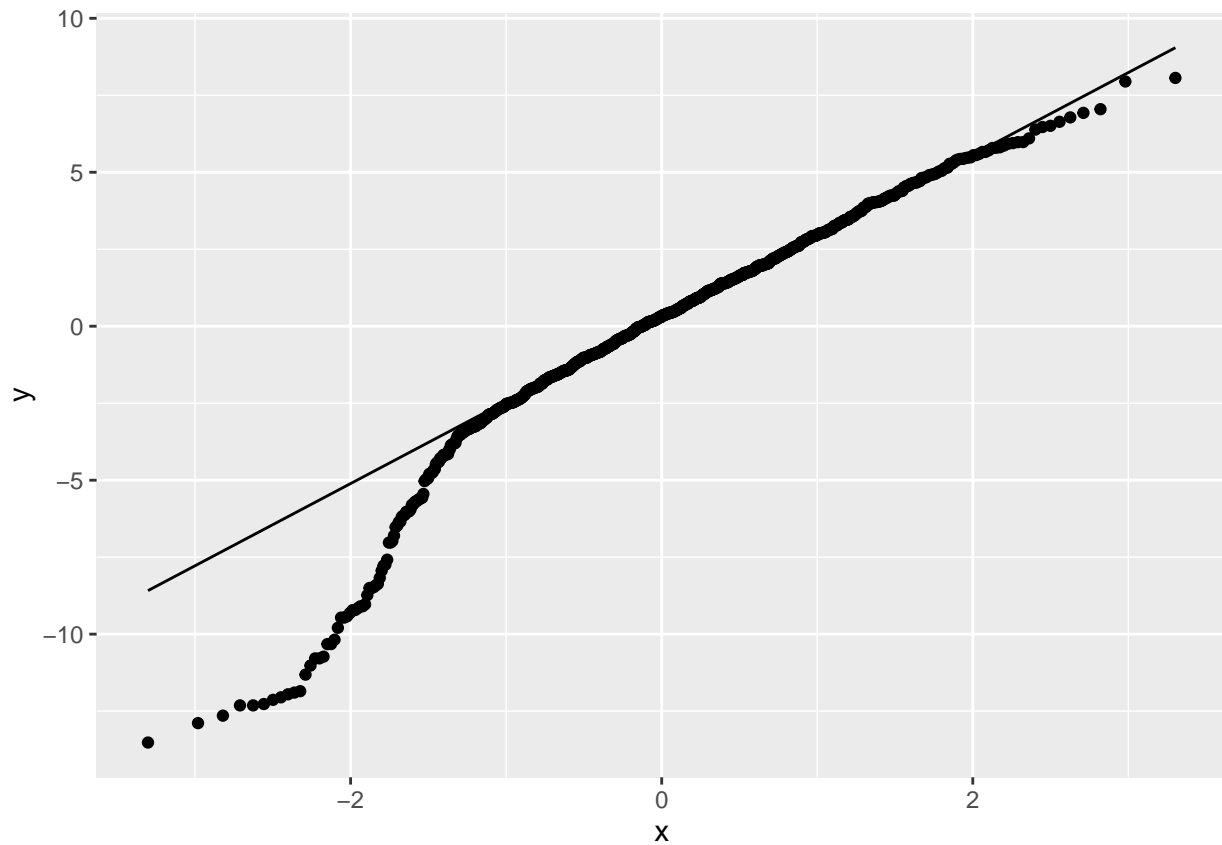
```
##
## studentized Breusch-Pagan test
##
## data: forwardMdl_int2
## BP = 66.114, df = 50, p-value = 0.063
```

```
# Normality Assumption
```

```
ggplot(data=studentDataset, aes(residuals(forwardMdl_int2))) +
  geom_histogram(breaks = seq(-5,5,by=1), col="green3", fill="green4") +
  labs(title="Histogram for residuals") +
  labs(x="residuals", y="Count")
```



```
ggplot(studentDataset, aes(sample=forwardMdl_int2$residuals)) +  
  stat_qq() +  
  stat_qq_line()
```



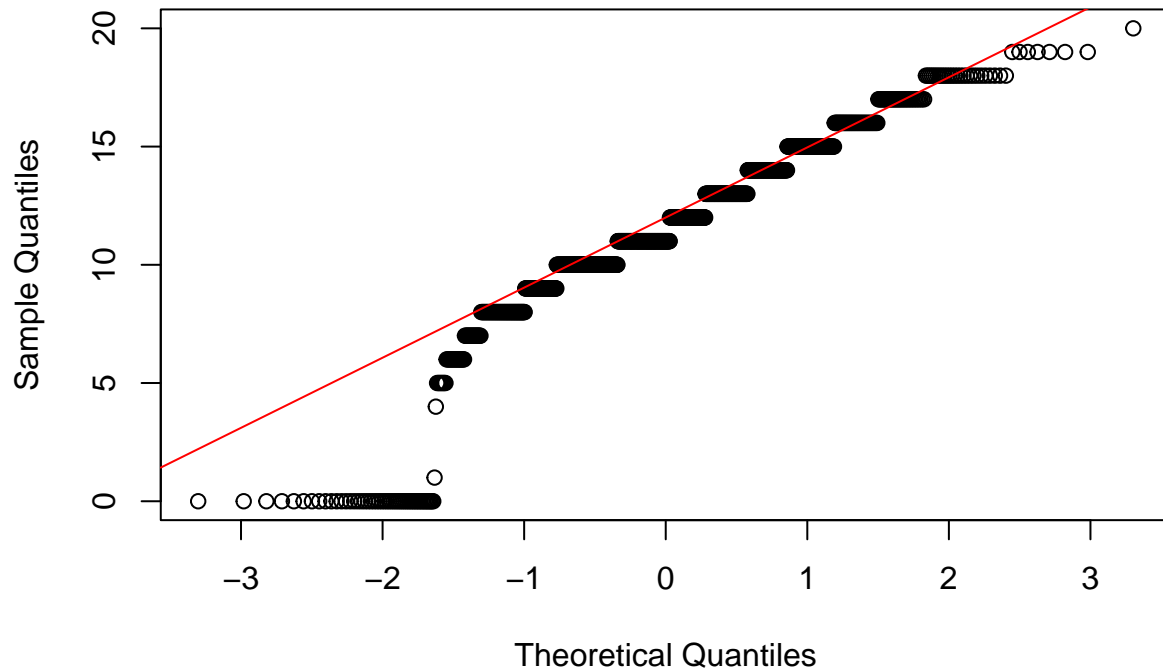
```
#Testing for Normality  
shapiro.test(residuals(forwardMdl_int2)) # We don't have normality
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  residuals(forwardMdl_int2)  
## W = 0.94152, p-value < 2.2e-16
```

```
qqnorm(studentDataset$G3)  
qqline(studentDataset$G3, col="red")
```



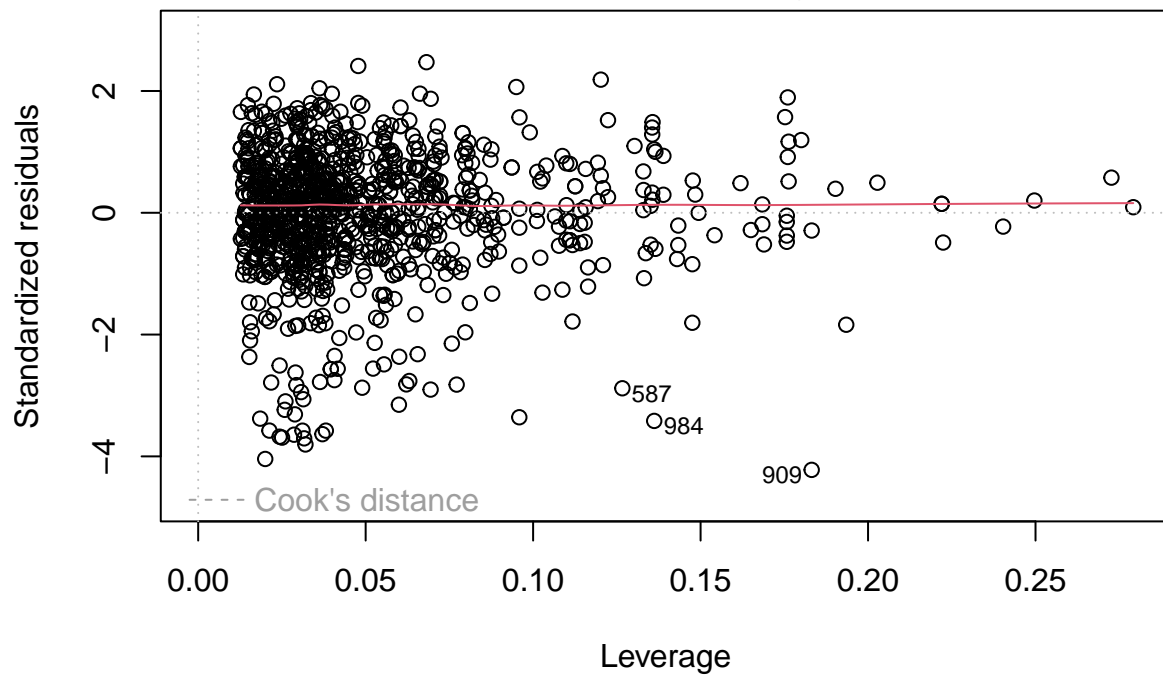
# Normal Q-Q Plot



# Outlier

```
plot(forwardMdl_int2, which=5)
```

# Residuals vs Leverage

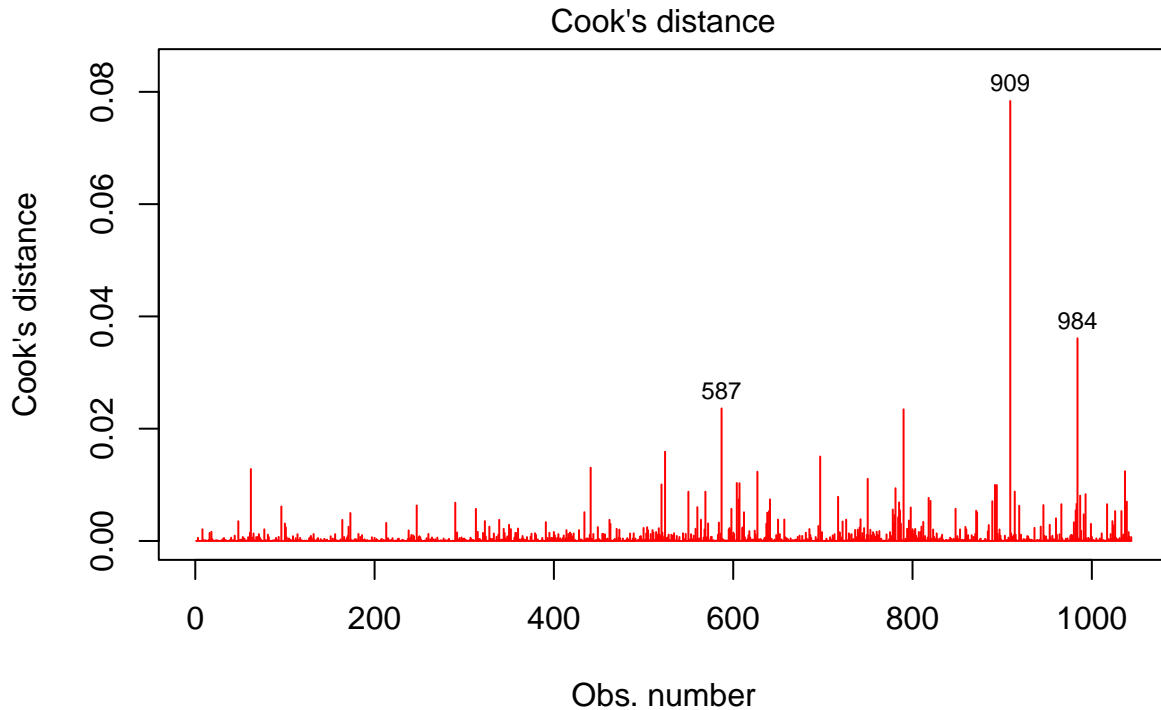


lm(G3 ~ (failures + higher + schoolsup + school + romantic + studytime + ad ...

```
studentDataset[cooks.distance(forwardMdl_int2)>0.5,]
```

```
## [1] school      sex      age      address   famsize   Pstatus
## [7] Medu        Fedu      Mjob      Fjob      reason    guardian
## [13] traveltime studytime failures schoolsup famsup    paid
## [19] activities nursery higher internet romantic famrel
## [25] freetime goout Dalc Walc health absences
## [31] G1          G2          G3
## <0 rows> (or 0-length row.names)
```

```
plot(forwardMdl_int2,pch=18,col="red",which=c(4))
```



lm(G3 ~ (failures + higher + schoolsup + school + romantic + studytime + ad ...

```
fanyi <- function(model,dep) {
  coe=coefficients(model)
  for(a in names(coe)){
    if (coe[a]>0){
      print(paste('For one unit increases in ',a,' ',dep,' increases by ', signif(coe[a])))
    }else{
      print(paste('For one unit increases in ',a,' ',dep,' decreases by ', signif(-1*coe[a])))
    }
  }
}
```

```
}
```

```
fanyi(studentPerformance_fm,'the grade of student')
```

```
## [1] "For one unit increases in (Intercept) the grade of student increases by 11.4817"
## [1] "For one unit increases in schoolMS the grade of student decreases by 0.309885"
## [1] "For one unit increases in sexM the grade of student decreases by 0.104478"
## [1] "For one unit increases in age the grade of student increases by 0.0500721"
## [1] "For one unit increases in addressU the grade of student increases by 0.444262"
## [1] "For one unit increases in famsizeLE3 the grade of student increases by 0.483175"
## [1] "For one unit increases in PstatusT the grade of student decreases by 0.110171"
## [1] "For one unit increases in Medu1 the grade of student decreases by 1.59966"
## [1] "For one unit increases in Medu2 the grade of student decreases by 1.63335"
## [1] "For one unit increases in Medu3 the grade of student decreases by 1.27121"
## [1] "For one unit increases in Medu4 the grade of student decreases by 0.693972"
## [1] "For one unit increases in Fedu1 the grade of student decreases by 1.17035"
## [1] "For one unit increases in Fedu2 the grade of student decreases by 1.08347"
## [1] "For one unit increases in Fedu3 the grade of student decreases by 1.32433"
## [1] "For one unit increases in Fedu4 the grade of student decreases by 1.15362"
## [1] "For one unit increases in Mjobhealth the grade of student increases by 0.64477"
## [1] "For one unit increases in Mjbother the grade of student decreases by 0.0972596"
## [1] "For one unit increases in Mjobservices the grade of student increases by 0.573808"
## [1] "For one unit increases in Mjobteacher the grade of student decreases by 0.537369"
## [1] "For one unit increases in Fjobhealth the grade of student decreases by 0.0240838"
## [1] "For one unit increases in Fjbother the grade of student increases by 0.146518"
## [1] "For one unit increases in Fjobservices the grade of student decreases by 0.169258"
## [1] "For one unit increases in Fjobteacher the grade of student increases by 1.19945"
## [1] "For one unit increases in reasonhome the grade of student decreases by 0.118871"
## [1] "For one unit increases in reasonother the grade of student decreases by 0.209278"
## [1] "For one unit increases in reasonreputation the grade of student increases by 0.19445"
## [1] "For one unit increases in guardianmother the grade of student decreases by 0.329956"
## [1] "For one unit increases in guardianother the grade of student increases by 0.118881"
## [1] "For one unit increases in traveltime2 the grade of student decreases by 0.112876"
## [1] "For one unit increases in traveltime3 the grade of student increases by 0.332622"
## [1] "For one unit increases in traveltime4 the grade of student decreases by 0.530444"
## [1] "For one unit increases in studytime2 the grade of student increases by 0.345119"
## [1] "For one unit increases in studytime3 the grade of student increases by 1.18931"
## [1] "For one unit increases in studytime4 the grade of student increases by 0.563988"
## [1] "For one unit increases in failures the grade of student decreases by 1.83833"
## [1] "For one unit increases in schoolsupyes the grade of student decreases by 1.11426"
## [1] "For one unit increases in famsupyes the grade of student decreases by 0.29653"
## [1] "For one unit increases in activitiesyes the grade of student decreases by 0.00929773"
## [1] "For one unit increases in nurseryyes the grade of student decreases by 0.0614225"
## [1] "For one unit increases in higheryes the grade of student increases by 1.19099"
## [1] "For one unit increases in internetyes the grade of student increases by 0.363775"
## [1] "For one unit increases in romanticyes the grade of student decreases by 0.648752"
## [1] "For one unit increases in famrel2 the grade of student increases by 0.297864"
## [1] "For one unit increases in famrel3 the grade of student increases by 0.544597"
## [1] "For one unit increases in famrel4 the grade of student increases by 0.95647"
## [1] "For one unit increases in famrel5 the grade of student increases by 0.729534"
## [1] "For one unit increases in freetime2 the grade of student increases by 0.974211"
## [1] "For one unit increases in freetime3 the grade of student decreases by 0.0213376"
## [1] "For one unit increases in freetime4 the grade of student increases by 0.332933"
## [1] "For one unit increases in freetime5 the grade of student increases by 0.968995"
```

```
## [1] "For one unit increases in goout2 the grade of student increases by 1.33574"
## [1] "For one unit increases in goout3 the grade of student increases by 0.87664"
## [1] "For one unit increases in goout4 the grade of student increases by 0.278556"
## [1] "For one unit increases in goout5 the grade of student decreases by 0.0791345"
## [1] "For one unit increases in Dalc2 the grade of student decreases by 0.709801"
## [1] "For one unit increases in Dalc3 the grade of student decreases by 0.0923023"
## [1] "For one unit increases in Dalc4 the grade of student decreases by 2.00016"
## [1] "For one unit increases in Dalc5 the grade of student decreases by 0.617704"
## [1] "For one unit increases in Walc2 the grade of student decreases by 0.304902"
## [1] "For one unit increases in Walc3 the grade of student increases by 0.126708"
## [1] "For one unit increases in Walc4 the grade of student decreases by 0.126274"
## [1] "For one unit increases in Walc5 the grade of student increases by 0.944524"
## [1] "For one unit increases in health2 the grade of student decreases by 1.03765"
## [1] "For one unit increases in health3 the grade of student decreases by 1.30066"
## [1] "For one unit increases in health4 the grade of student decreases by 0.853415"
## [1] "For one unit increases in health5 the grade of student decreases by 1.24919"
## [1] "For one unit increases in absences the grade of student increases by 0.00501298"
```

```
formula <- function(model,dep){

  coe=coefficients(model)

  b=paste(dep, '=')

  for(a in names(coe)){

    if (coe[a]>0){

      b=paste(b, ' +', signif(coe[a]), '(', a, ')', sep="")

    }else{

      b=paste(b, ' ', signif(coe[a]), '(', a, ')', sep="")

    }

  }

  print(b)

}

formula(student_performance_intMdl3, 'G3')
```

```
## [1] "G3 = +8.67068((Intercept)) -3.84533(failures) +0.757609(I(failures^2)) +1.49398(higheryes) +2.5"
```