Final project

Haochen Shi

12/6/2023

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

## Including Plots

You can also embed plots, for example:

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.

library(readr)

## Warning: package 'readr' was built under R version 4.1.2

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.1.2

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.4.0 ✔ dplyr 1.0.10  
## ✔ tibble 3.1.8 ✔ stringr 1.4.0   
## ✔ tidyr 1.2.0 ✔ forcats 0.5.2   
## ✔ purrr 0.3.4

## Warning: package 'ggplot2' was built under R version 4.1.2

## Warning: package 'tibble' was built under R version 4.1.2

## Warning: package 'tidyr' was built under R version 4.1.2

## Warning: package 'dplyr' was built under R version 4.1.2

## Warning: package 'forcats' was built under R version 4.1.2

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

path<-"/Users/haochenshi/Downloads/Project\_1\_data.csv"  
Data<-read.csv(path) #把数据搞进来  
  
#把Ethnicgroup变成dummy variable  
Data$EthnicGroup <- as.factor(Data$EthnicGroup)  
Data$NumericEthnic <- as.numeric(Data$EthnicGroup)#把ethnic group用数字代替/dummy variable  
Data$NumericEthnic <- ifelse(Data$NumericEthnic == "1", NA, Data$NumericEthnic)#把空白的数据变成N/A. (RStudio不认空格,只认N/A)  
NumericEthnic <- Data$NumericEthnic  
mode\_value <- as.numeric(names(sort(table(NumericEthnic), decreasing = TRUE)[1]))  
Data$NumericEthnic[is.na(Data$NumericEthnic)] <- mode\_value#这三行我用mode把missing value给填上了,因为分组是catagoricaldata, 平均数不行(不能出现第2.64组,只能是整数,所以我用mode)  
Data$NumericEthnic=Data$NumericEthnic - 1 #从23456变成12345  
  
#把ParentEduc变成dummy variable 和以上同理mode  
Data$ParentEduc <- as.factor(Data$ParentEduc)  
Data$Numericparentedu <- as.numeric(Data$ParentEduc)  
Data$Numericparentedu <- ifelse(Data$Numericparentedu == "1", NA, Data$Numericparentedu)  
Numericparentedu <- Data$Numericparentedu  
mode\_of\_parentedu <- as.numeric(names(sort(table(Numericparentedu), decreasing = TRUE)[1]))  
Data$Numericparentedu[is.na(Data$Numericparentedu)] <- mode\_of\_parentedu  
Data$Numericparentedu=Data$Numericparentedu - 1  
  
#把Gender变成dummy variable,没有missing,真的太舒服了  
Data$Gender <- as.factor(Data$Gender)  
Data$NumericGender <- as.numeric(Data$Gender)  
NumericGender<- Data$NumericGender  
  
#把LunchType变成dummy variable, 也没有missing,也很舒服  
Data$LunchType <- as.factor(Data$LunchType)  
Data$NumericLunchType <- as.numeric(Data$LunchType)  
NumericLunchType<- Data$NumericLunchType  
  
#把TestPrep变成dummy variable 同理mode  
Data$TestPrep <- as.factor(Data$TestPrep)  
Data$NumericTestPrep <- as.numeric(Data$TestPrep)  
Data$NumericTestPrep <- ifelse(Data$NumericTestPrep == "1", NA, Data$NumericTestPrep)  
NumericTestPrep <- Data$NumericTestPrep  
mode\_of\_TestPrep <- as.numeric(names(sort(table(NumericTestPrep), decreasing = TRUE)[1]))  
Data$NumericTestPrep[is.na(Data$NumericTestPrep)] <- mode\_of\_TestPrep  
Data$NumericTestPrep=Data$NumericTestPrep - 1  
  
#把ParentMaritalStatus变成dummy variable 同理mode  
Data$ParentMaritalStatus <- as.factor(Data$ParentMaritalStatus)  
Data$NumericParentMaritalStatus <- as.numeric(Data$ParentMaritalStatus)  
Data$NumericParentMaritalStatus <- ifelse(Data$NumericParentMaritalStatus == "1", NA, Data$NumericParentMaritalStatus)  
NumericParentMaritalStatus <- Data$NumericParentMaritalStatus  
mode\_of\_ParentMaritalStatus <- as.numeric(names(sort(table(NumericParentMaritalStatus), decreasing = TRUE)[1]))  
Data$NumericParentMaritalStatus[is.na(Data$NumericParentMaritalStatus)] <- mode\_of\_ParentMaritalStatus  
Data$NumericParentMaritalStatus=Data$NumericParentMaritalStatus - 1  
  
#把PracticeSport变成dummy variable 同理mode  
Data$PracticeSport <- as.factor(Data$PracticeSport)  
Data$NumericPracticeSport <- as.numeric(Data$PracticeSport)  
Data$NumericPracticeSport <- ifelse(Data$NumericPracticeSport == "1", NA, Data$NumericPracticeSport)  
NumericPracticeSport <- Data$NumericPracticeSport  
mode\_of\_PracticeSport <- as.numeric(names(sort(table(NumericPracticeSport), decreasing = TRUE)[1]))  
Data$NumericPracticeSport[is.na(Data$NumericPracticeSport)] <- mode\_of\_PracticeSport  
Data$NumericPracticeSport=Data$NumericPracticeSport - 1  
  
#把IsFirstChild变成dummy variable 同理mode  
Data$IsFirstChild<- as.factor(Data$IsFirstChild)  
Data$NumericIsFirstChild <- as.numeric(Data$IsFirstChild)  
Data$NumericIsFirstChild <- ifelse(Data$NumericIsFirstChild == "1", NA, Data$NumericIsFirstChild)  
NumericIsFirstChild <- Data$NumericIsFirstChild  
mode\_of\_IsFirstChild <- as.numeric(names(sort(table(NumericIsFirstChild), decreasing = TRUE)[1]))  
Data$NumericIsFirstChild[is.na(Data$NumericIsFirstChild)] <- mode\_of\_IsFirstChild  
Data$NumericIsFirstChild=Data$NumericIsFirstChild - 1  
  
#填补他妈的number of siblings的缺失数据 还是用mode  
Data$NrSiblings <- ifelse(Data$NrSiblings == " ", NA, Data$NrSiblings)  
NrSiblings <- Data$NrSiblings  
mode\_of\_NrSiblings <- as.numeric(names(sort(table(NrSiblings), decreasing = TRUE)[1]))  
Data$NrSiblings[is.na(Data$NrSiblings)] <- mode\_of\_NrSiblings  
  
#把TransportMeans变成dummy variable 同理mode  
Data$TransportMeans<- as.factor(Data$TransportMeans)  
Data$NumericTransportMeans <- as.numeric(Data$TransportMeans)  
Data$NumericTransportMeans <- ifelse(Data$NumericTransportMeans == "1", NA, Data$NumericTransportMeans)  
NumericTransportMeans <- Data$NumericTransportMeans  
mode\_of\_TransportMeans <- as.numeric(names(sort(table(NumericTransportMeans), decreasing = TRUE)[1]))  
Data$NumericTransportMeans[is.na(Data$NumericTransportMeans)] <- mode\_of\_TransportMeans  
Data$NumericTransportMeans=Data$NumericTransportMeans - 1  
  
#把WklyStudyHours变成dummy variable 同理mode  
Data$WklyStudyHours<- as.factor(Data$WklyStudyHours)  
Data$NumericWklyStudyHours <- as.numeric(Data$WklyStudyHours)  
Data$NumericWklyStudyHours <- ifelse(Data$NumericWklyStudyHours == "1", NA, Data$NumericWklyStudyHours)  
NumericWklyStudyHours <- Data$NumericWklyStudyHours  
mode\_of\_WklyStudyHours <- as.numeric(names(sort(table(NumericWklyStudyHours), decreasing = TRUE)[1]))  
Data$NumericWklyStudyHours[is.na(Data$NumericWklyStudyHours)] <- mode\_of\_WklyStudyHours  
Data$NumericWklyStudyHours=Data$NumericWklyStudyHours - 1  
  
  
#数据总算整理完了真离谱啊,所以现在除了NrSiblings是原先自己的,其他的我都给变成dummy variable了, 名字都叫做NumericXXXXX

#现在开始用所有variable来做三个成绩的regression  
regression\_of\_Math=lm(MathScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericParentMaritalStatus+NumericPracticeSport+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours +NrSiblings, data=Data)  
  
regression\_of\_Reading=lm(ReadingScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericParentMaritalStatus+NumericPracticeSport+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours +NrSiblings, data=Data)  
  
regression\_of\_Writing=lm(WritingScore~Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericParentMaritalStatus+NumericPracticeSport+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours +NrSiblings+NumericEthnic, data=Data)  
summary(regression\_of\_Math)

##   
## Call:  
## lm(formula = MathScore ~ NumericEthnic + Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericParentMaritalStatus +   
## NumericPracticeSport + NumericIsFirstChild + NumericTransportMeans +   
## NumericWklyStudyHours + NrSiblings, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -53.450 -8.769 0.313 9.788 35.081   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 33.1625 4.6565 7.122 2.12e-12 \*\*\*  
## NumericEthnic 2.6021 0.3946 6.595 7.11e-11 \*\*\*  
## Numericparentedu -0.6879 0.2446 -2.813 0.00501 \*\*   
## NumericGender 5.0682 0.8874 5.711 1.51e-08 \*\*\*  
## NumericLunchType 11.0270 0.9319 11.833 < 2e-16 \*\*\*  
## NumericTestPrep -5.8518 0.9369 -6.246 6.37e-10 \*\*\*  
## NumericParentMaritalStatus 0.4413 0.6511 0.678 0.49809   
## NumericPracticeSport 0.3835 0.6423 0.597 0.55063   
## NumericIsFirstChild 2.3145 0.9524 2.430 0.01527 \*   
## NumericTransportMeans 0.8384 0.9256 0.906 0.36527   
## NumericWklyStudyHours 1.3925 0.5130 2.714 0.00677 \*\*   
## NrSiblings 0.3452 0.3050 1.132 0.25803   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.62 on 936 degrees of freedom  
## Multiple R-squared: 0.2396, Adjusted R-squared: 0.2307   
## F-statistic: 26.82 on 11 and 936 DF, p-value: < 2.2e-16

summary(regression\_of\_Reading)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericEthnic + Numericparentedu +   
## NumericGender + NumericLunchType + NumericTestPrep + NumericParentMaritalStatus +   
## NumericPracticeSport + NumericIsFirstChild + NumericTransportMeans +   
## NumericWklyStudyHours + NrSiblings, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.347 -8.787 0.544 9.578 32.276   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 66.065922 4.536910 14.562 < 2e-16 \*\*\*  
## NumericEthnic 1.640552 0.384424 4.268 2.18e-05 \*\*\*  
## Numericparentedu -0.689804 0.238296 -2.895 0.00388 \*\*   
## NumericGender -7.295782 0.864622 -8.438 < 2e-16 \*\*\*  
## NumericLunchType 7.414695 0.907969 8.166 1.02e-15 \*\*\*  
## NumericTestPrep -7.129714 0.912800 -7.811 1.52e-14 \*\*\*  
## NumericParentMaritalStatus 0.396570 0.634408 0.625 0.53206   
## NumericPracticeSport 0.464966 0.625831 0.743 0.45769   
## NumericIsFirstChild 2.316538 0.927908 2.497 0.01271 \*   
## NumericTransportMeans 1.453229 0.901783 1.612 0.10741   
## NumericWklyStudyHours 0.979654 0.499870 1.960 0.05031 .   
## NrSiblings 0.006792 0.297149 0.023 0.98177   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.27 on 936 degrees of freedom  
## Multiple R-squared: 0.2052, Adjusted R-squared: 0.1958   
## F-statistic: 21.96 on 11 and 936 DF, p-value: < 2.2e-16

summary(regression\_of\_Writing)

##   
## Call:  
## lm(formula = WritingScore ~ Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericParentMaritalStatus +   
## NumericPracticeSport + NumericIsFirstChild + NumericTransportMeans +   
## NumericWklyStudyHours + NrSiblings + NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.581 -8.710 0.301 9.348 28.705   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 70.5496 4.4970 15.688 < 2e-16 \*\*\*  
## Numericparentedu -0.7645 0.2362 -3.237 0.00125 \*\*   
## NumericGender -9.2317 0.8570 -10.772 < 2e-16 \*\*\*  
## NumericLunchType 8.1419 0.9000 9.047 < 2e-16 \*\*\*  
## NumericTestPrep -9.6920 0.9048 -10.712 < 2e-16 \*\*\*  
## NumericParentMaritalStatus 0.3070 0.6288 0.488 0.62551   
## NumericPracticeSport 0.6789 0.6203 1.094 0.27406   
## NumericIsFirstChild 2.0614 0.9197 2.241 0.02524 \*   
## NumericTransportMeans 1.2121 0.8939 1.356 0.17541   
## NumericWklyStudyHours 0.9612 0.4955 1.940 0.05268 .   
## NrSiblings 0.1355 0.2945 0.460 0.64549   
## NumericEthnic 1.9791 0.3810 5.194 2.53e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.16 on 936 degrees of freedom  
## Multiple R-squared: 0.28, Adjusted R-squared: 0.2715   
## F-statistic: 33.08 on 11 and 936 DF, p-value: < 2.2e-16

#开始backward elimination  
regression\_of\_Math2=lm(MathScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericParentMaritalStatus+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours +NrSiblings, data=Data)  
  
regression\_of\_Reading2=lm(ReadingScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericPracticeSport+NumericParentMaritalStatus+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours, data=Data)  
  
regression\_of\_Writing2=lm(WritingScore~Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericPracticeSport+NumericParentMaritalStatus+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours+NumericEthnic, data=Data)  
summary(regression\_of\_Math2)

##   
## Call:  
## lm(formula = MathScore ~ NumericEthnic + Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericParentMaritalStatus +   
## NumericIsFirstChild + NumericTransportMeans + NumericWklyStudyHours +   
## NrSiblings, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -53.603 -8.849 0.327 9.938 34.953   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 34.1635 4.3427 7.867 9.99e-15 \*\*\*  
## NumericEthnic 2.5958 0.3943 6.584 7.63e-11 \*\*\*  
## Numericparentedu -0.6879 0.2445 -2.814 0.00500 \*\*   
## NumericGender 5.0661 0.8871 5.711 1.51e-08 \*\*\*  
## NumericLunchType 11.0118 0.9312 11.825 < 2e-16 \*\*\*  
## NumericTestPrep -5.8581 0.9365 -6.255 6.02e-10 \*\*\*  
## NumericParentMaritalStatus 0.4341 0.6508 0.667 0.50495   
## NumericIsFirstChild 2.2909 0.9512 2.408 0.01622 \*   
## NumericTransportMeans 0.8444 0.9252 0.913 0.36167   
## NumericWklyStudyHours 1.3995 0.5127 2.730 0.00646 \*\*   
## NrSiblings 0.3486 0.3048 1.143 0.25313   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.62 on 937 degrees of freedom  
## Multiple R-squared: 0.2393, Adjusted R-squared: 0.2312   
## F-statistic: 29.48 on 10 and 937 DF, p-value: < 2.2e-16

summary(regression\_of\_Reading2)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericEthnic + Numericparentedu +   
## NumericGender + NumericLunchType + NumericTestPrep + NumericPracticeSport +   
## NumericParentMaritalStatus + NumericIsFirstChild + NumericTransportMeans +   
## NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.341 -8.783 0.535 9.571 32.279   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 66.0847 4.4593 14.819 < 2e-16 \*\*\*  
## NumericEthnic 1.6405 0.3842 4.270 2.16e-05 \*\*\*  
## Numericparentedu -0.6898 0.2382 -2.896 0.00386 \*\*   
## NumericGender -7.2962 0.8640 -8.445 < 2e-16 \*\*\*  
## NumericLunchType 7.4138 0.9067 8.177 9.43e-16 \*\*\*  
## NumericTestPrep -7.1298 0.9123 -7.815 1.47e-14 \*\*\*  
## NumericPracticeSport 0.4652 0.6254 0.744 0.45712   
## NumericParentMaritalStatus 0.3972 0.6335 0.627 0.53086   
## NumericIsFirstChild 2.3138 0.9195 2.516 0.01202 \*   
## NumericTransportMeans 1.4530 0.9013 1.612 0.10725   
## NumericWklyStudyHours 0.9800 0.4994 1.963 0.05000 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.26 on 937 degrees of freedom  
## Multiple R-squared: 0.2051, Adjusted R-squared: 0.1967   
## F-statistic: 24.18 on 10 and 937 DF, p-value: < 2.2e-16

summary(regression\_of\_Writing2)

##   
## Call:  
## lm(formula = WritingScore ~ Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericPracticeSport +   
## NumericParentMaritalStatus + NumericIsFirstChild + NumericTransportMeans +   
## NumericWklyStudyHours + NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.460 -8.702 0.401 9.475 28.588   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 70.9248 4.4206 16.044 < 2e-16 \*\*\*  
## Numericparentedu -0.7647 0.2361 -3.239 0.00124 \*\*   
## NumericGender -9.2401 0.8565 -10.789 < 2e-16 \*\*\*  
## NumericLunchType 8.1247 0.8988 9.039 < 2e-16 \*\*\*  
## NumericTestPrep -9.6935 0.9044 -10.718 < 2e-16 \*\*\*  
## NumericPracticeSport 0.6842 0.6200 1.104 0.27003   
## NumericParentMaritalStatus 0.3190 0.6280 0.508 0.61161   
## NumericIsFirstChild 2.0062 0.9115 2.201 0.02798 \*   
## NumericTransportMeans 1.2080 0.8934 1.352 0.17667   
## NumericWklyStudyHours 0.9683 0.4950 1.956 0.05074 .   
## NumericEthnic 1.9785 0.3809 5.195 2.52e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.15 on 937 degrees of freedom  
## Multiple R-squared: 0.2798, Adjusted R-squared: 0.2721   
## F-statistic: 36.4 on 10 and 937 DF, p-value: < 2.2e-16

#继续  
regression\_of\_Math3=lm(MathScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours +NrSiblings, data=Data)  
  
regression\_of\_Reading3=lm(ReadingScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericPracticeSport+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours, data=Data)  
  
regression\_of\_Writing3=lm(WritingScore~Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericPracticeSport+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours+NumericEthnic, data=Data)  
summary(regression\_of\_Math3)

##   
## Call:  
## lm(formula = MathScore ~ NumericEthnic + Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericTransportMeans + NumericWklyStudyHours + NrSiblings,   
## data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -53.225 -8.580 0.177 9.935 34.905   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 35.1756 4.0677 8.647 < 2e-16 \*\*\*  
## NumericEthnic 2.5910 0.3941 6.574 8.10e-11 \*\*\*  
## Numericparentedu -0.6855 0.2444 -2.805 0.00513 \*\*   
## NumericGender 5.0469 0.8864 5.694 1.66e-08 \*\*\*  
## NumericLunchType 11.0249 0.9307 11.845 < 2e-16 \*\*\*  
## NumericTestPrep -5.8765 0.9358 -6.280 5.18e-10 \*\*\*  
## NumericIsFirstChild 2.2644 0.9501 2.383 0.01736 \*   
## NumericTransportMeans 0.8467 0.9249 0.915 0.36019   
## NumericWklyStudyHours 1.3890 0.5123 2.711 0.00683 \*\*   
## NrSiblings 0.3569 0.3045 1.172 0.24139   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.61 on 938 degrees of freedom  
## Multiple R-squared: 0.239, Adjusted R-squared: 0.2317   
## F-statistic: 32.73 on 9 and 938 DF, p-value: < 2.2e-16

summary(regression\_of\_Reading3)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericEthnic + Numericparentedu +   
## NumericGender + NumericLunchType + NumericTestPrep + NumericPracticeSport +   
## NumericIsFirstChild + NumericTransportMeans + NumericWklyStudyHours,   
## data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -46.991 -8.829 0.517 9.715 32.236   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 67.0526 4.1822 16.033 < 2e-16 \*\*\*  
## NumericEthnic 1.6359 0.3840 4.260 2.25e-05 \*\*\*  
## Numericparentedu -0.6876 0.2381 -2.888 0.00396 \*\*   
## NumericGender -7.3143 0.8632 -8.474 < 2e-16 \*\*\*  
## NumericLunchType 7.4246 0.9062 8.193 8.32e-16 \*\*\*  
## NumericTestPrep -7.1468 0.9116 -7.840 1.22e-14 \*\*\*  
## NumericPracticeSport 0.4582 0.6251 0.733 0.46372   
## NumericIsFirstChild 2.2859 0.9181 2.490 0.01295 \*   
## NumericTransportMeans 1.4551 0.9010 1.615 0.10665   
## NumericWklyStudyHours 0.9709 0.4990 1.946 0.05199 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.26 on 938 degrees of freedom  
## Multiple R-squared: 0.2048, Adjusted R-squared: 0.1972   
## F-statistic: 26.84 on 9 and 938 DF, p-value: < 2.2e-16

summary(regression\_of\_Writing3)

##   
## Call:  
## lm(formula = WritingScore ~ Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericPracticeSport +   
## NumericIsFirstChild + NumericTransportMeans + NumericWklyStudyHours +   
## NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.179 -8.719 0.375 9.443 28.757   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 71.7022 4.1456 17.296 < 2e-16 \*\*\*  
## Numericparentedu -0.7630 0.2360 -3.233 0.00127 \*\*   
## NumericGender -9.2547 0.8556 -10.816 < 2e-16 \*\*\*  
## NumericLunchType 8.1333 0.8983 9.054 < 2e-16 \*\*\*  
## NumericTestPrep -9.7072 0.9036 -10.742 < 2e-16 \*\*\*  
## NumericPracticeSport 0.6786 0.6196 1.095 0.27372   
## NumericIsFirstChild 1.9838 0.9101 2.180 0.02952 \*   
## NumericTransportMeans 1.2096 0.8931 1.354 0.17591   
## NumericWklyStudyHours 0.9610 0.4946 1.943 0.05233 .   
## NumericEthnic 1.9748 0.3807 5.188 2.61e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.14 on 938 degrees of freedom  
## Multiple R-squared: 0.2796, Adjusted R-squared: 0.2727   
## F-statistic: 40.45 on 9 and 938 DF, p-value: < 2.2e-16

#继续  
regression\_of\_Math4=lm(MathScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericWklyStudyHours +NrSiblings, data=Data)  
  
regression\_of\_Reading4=lm(ReadingScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours, data=Data)  
  
regression\_of\_Writing4=lm(WritingScore~Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericTransportMeans+NumericWklyStudyHours+NumericEthnic, data=Data)  
summary(regression\_of\_Math4)

##   
## Call:  
## lm(formula = MathScore ~ NumericEthnic + Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericWklyStudyHours + NrSiblings, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.910 -8.783 0.288 10.107 34.321   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 36.6043 3.7561 9.745 < 2e-16 \*\*\*  
## NumericEthnic 2.5794 0.3939 6.549 9.53e-11 \*\*\*  
## Numericparentedu -0.6929 0.2442 -2.837 0.00465 \*\*   
## NumericGender 5.0482 0.8863 5.696 1.64e-08 \*\*\*  
## NumericLunchType 11.0272 0.9307 11.849 < 2e-16 \*\*\*  
## NumericTestPrep -5.8596 0.9355 -6.263 5.73e-10 \*\*\*  
## NumericIsFirstChild 2.2652 0.9500 2.384 0.01731 \*   
## NumericWklyStudyHours 1.3879 0.5123 2.709 0.00687 \*\*   
## NrSiblings 0.3542 0.3044 1.164 0.24489   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.61 on 939 degrees of freedom  
## Multiple R-squared: 0.2383, Adjusted R-squared: 0.2318   
## F-statistic: 36.72 on 8 and 939 DF, p-value: < 2.2e-16

summary(regression\_of\_Reading4)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericEthnic + Numericparentedu +   
## NumericGender + NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericTransportMeans + NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.178 -8.962 0.374 9.573 32.029   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 68.2404 3.8546 17.704 < 2e-16 \*\*\*  
## NumericEthnic 1.6286 0.3838 4.243 2.42e-05 \*\*\*  
## Numericparentedu -0.6877 0.2380 -2.889 0.00395 \*\*   
## NumericGender -7.3167 0.8630 -8.478 < 2e-16 \*\*\*  
## NumericLunchType 7.4056 0.9057 8.177 9.38e-16 \*\*\*  
## NumericTestPrep -7.1540 0.9113 -7.850 1.13e-14 \*\*\*  
## NumericIsFirstChild 2.2566 0.9170 2.461 0.01404 \*   
## NumericTransportMeans 1.4620 0.9007 1.623 0.10487   
## NumericWklyStudyHours 0.9797 0.4987 1.964 0.04978 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.26 on 939 degrees of freedom  
## Multiple R-squared: 0.2044, Adjusted R-squared: 0.1976   
## F-statistic: 30.15 on 8 and 939 DF, p-value: < 2.2e-16

summary(regression\_of\_Writing4)

##   
## Call:  
## lm(formula = WritingScore ~ Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericTransportMeans + NumericWklyStudyHours + NumericEthnic,   
## data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.455 -8.657 0.448 9.569 28.483   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 73.4613 3.8222 19.220 < 2e-16 \*\*\*  
## Numericparentedu -0.7630 0.2360 -3.233 0.00127 \*\*   
## NumericGender -9.2582 0.8557 -10.819 < 2e-16 \*\*\*  
## NumericLunchType 8.1052 0.8980 9.025 < 2e-16 \*\*\*  
## NumericTestPrep -9.7178 0.9037 -10.754 < 2e-16 \*\*\*  
## NumericIsFirstChild 1.9404 0.9093 2.134 0.03311 \*   
## NumericTransportMeans 1.2200 0.8931 1.366 0.17227   
## NumericWklyStudyHours 0.9741 0.4945 1.970 0.04917 \*   
## NumericEthnic 1.9639 0.3806 5.160 3.01e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.15 on 939 degrees of freedom  
## Multiple R-squared: 0.2787, Adjusted R-squared: 0.2725   
## F-statistic: 45.35 on 8 and 939 DF, p-value: < 2.2e-16

#继续  
regression\_of\_Math5=lm(MathScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericWklyStudyHours, data=Data)  
  
regression\_of\_Reading5=lm(ReadingScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericWklyStudyHours, data=Data)  
  
regression\_of\_Writing5=lm(WritingScore~Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericWklyStudyHours+NumericEthnic, data=Data)  
summary(regression\_of\_Math5)

##   
## Call:  
## lm(formula = MathScore ~ NumericEthnic + Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.574 -8.722 0.145 9.976 33.653   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.6787 3.6415 10.347 < 2e-16 \*\*\*  
## NumericEthnic 2.5773 0.3939 6.543 9.92e-11 \*\*\*  
## Numericparentedu -0.6934 0.2443 -2.838 0.00463 \*\*   
## NumericGender 5.0245 0.8862 5.670 1.90e-08 \*\*\*  
## NumericLunchType 10.9825 0.9300 11.809 < 2e-16 \*\*\*  
## NumericTestPrep -5.8652 0.9357 -6.268 5.55e-10 \*\*\*  
## NumericIsFirstChild 2.1178 0.9417 2.249 0.02475 \*   
## NumericWklyStudyHours 1.4060 0.5122 2.745 0.00616 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.61 on 940 degrees of freedom  
## Multiple R-squared: 0.2372, Adjusted R-squared: 0.2315   
## F-statistic: 41.76 on 7 and 940 DF, p-value: < 2.2e-16

summary(regression\_of\_Reading5)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericEthnic + Numericparentedu +   
## NumericGender + NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -46.638 -8.841 0.666 9.682 31.115   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 70.6934 3.5491 19.919 < 2e-16 \*\*\*  
## NumericEthnic 1.6086 0.3839 4.190 3.05e-05 \*\*\*  
## Numericparentedu -0.7004 0.2381 -2.942 0.00334 \*\*   
## NumericGender -7.3141 0.8637 -8.468 < 2e-16 \*\*\*  
## NumericLunchType 7.4101 0.9064 8.175 9.52e-16 \*\*\*  
## NumericTestPrep -7.1249 0.9119 -7.813 1.49e-14 \*\*\*  
## NumericIsFirstChild 2.2599 0.9178 2.462 0.01399 \*   
## NumericWklyStudyHours 0.9775 0.4992 1.958 0.05048 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.27 on 940 degrees of freedom  
## Multiple R-squared: 0.2021, Adjusted R-squared: 0.1962   
## F-statistic: 34.02 on 7 and 940 DF, p-value: < 2.2e-16

summary(regression\_of\_Writing5)

##   
## Call:  
## lm(formula = WritingScore ~ Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericWklyStudyHours + NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.004 -8.764 0.332 9.504 28.915   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 75.5082 3.5178 21.465 < 2e-16 \*\*\*  
## Numericparentedu -0.7737 0.2360 -3.279 0.00108 \*\*   
## NumericGender -9.2560 0.8561 -10.812 < 2e-16 \*\*\*  
## NumericLunchType 8.1090 0.8984 9.026 < 2e-16 \*\*\*  
## NumericTestPrep -9.6935 0.9039 -10.724 < 2e-16 \*\*\*  
## NumericIsFirstChild 1.9431 0.9097 2.136 0.03294 \*   
## NumericWklyStudyHours 0.9722 0.4948 1.965 0.04969 \*   
## NumericEthnic 1.9472 0.3805 5.117 3.77e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.15 on 940 degrees of freedom  
## Multiple R-squared: 0.2772, Adjusted R-squared: 0.2719   
## F-statistic: 51.51 on 7 and 940 DF, p-value: < 2.2e-16

#marginal distribution of 11variables and 3 score 这一格个都是 math score  
Mathregression\_of\_eth=lm(MathScore~NumericEthnic,data=Data)  
summary(Mathregression\_of\_eth)

##   
## Call:  
## lm(formula = MathScore ~ NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -65.523 -9.393 0.390 10.477 39.130   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 57.0438 1.4743 38.692 < 2e-16 \*\*\*  
## NumericEthnic 2.8264 0.4392 6.435 1.96e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.21 on 946 degrees of freedom  
## Multiple R-squared: 0.04193, Adjusted R-squared: 0.04092   
## F-statistic: 41.4 on 1 and 946 DF, p-value: 1.963e-10

Mathregression\_of\_gender=lm(MathScore~NumericGender,data=Data)  
summary(Mathregression\_of\_gender)

##   
## Call:  
## lm(formula = MathScore ~ NumericGender, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -63.375 -9.748 0.252 10.625 36.625   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 58.0022 1.5585 37.216 < 2e-16 \*\*\*  
## NumericGender 5.3728 0.9946 5.402 8.33e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.3 on 946 degrees of freedom  
## Multiple R-squared: 0.02993, Adjusted R-squared: 0.0289   
## F-statistic: 29.18 on 1 and 946 DF, p-value: 8.329e-08

Mathregression\_of\_parentedu=lm(MathScore~Numericparentedu,data=Data)  
summary(Mathregression\_of\_parentedu)

##   
## Call:  
## lm(formula = MathScore ~ Numericparentedu, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -64.300 -9.828 0.005 11.005 35.005   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 68.4668 1.1139 61.47 <2e-16 \*\*\*  
## Numericparentedu -0.6944 0.2778 -2.50 0.0126 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.49 on 946 degrees of freedom  
## Multiple R-squared: 0.006565, Adjusted R-squared: 0.005514   
## F-statistic: 6.251 on 1 and 946 DF, p-value: 0.01258

Mathregression\_of\_lunchtype=lm(MathScore~NumericLunchType,data=Data)  
summary(Mathregression\_of\_lunchtype)

##   
## Call:  
## lm(formula = MathScore ~ NumericLunchType, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -58.471 -10.011 -0.011 10.989 41.529   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 46.9313 1.7010 27.59 <2e-16 \*\*\*  
## NumericLunchType 11.5400 0.9899 11.66 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.53 on 946 degrees of freedom  
## Multiple R-squared: 0.1256, Adjusted R-squared: 0.1247   
## F-statistic: 135.9 on 1 and 946 DF, p-value: < 2.2e-16

Mathregression\_of\_testprep=lm(MathScore~NumericTestPrep,data=Data)  
summary(Mathregression\_of\_testprep)

##   
## Call:  
## lm(formula = MathScore ~ NumericTestPrep, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -63.899 -9.899 -0.031 10.101 36.101   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 76.163 1.807 42.140 < 2e-16 \*\*\*  
## NumericTestPrep -6.132 1.047 -5.858 6.48e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.26 on 946 degrees of freedom  
## Multiple R-squared: 0.035, Adjusted R-squared: 0.03398   
## F-statistic: 34.31 on 1 and 946 DF, p-value: 6.477e-09

Mathregression\_of\_parentmarriage=lm(MathScore~NumericParentMaritalStatus,data=Data)  
summary(Mathregression\_of\_parentmarriage)

##   
## Call:  
## lm(formula = MathScore ~ NumericParentMaritalStatus, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -66.307 -9.937 0.063 10.063 34.432   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 65.1978 1.6485 39.55 <2e-16 \*\*\*  
## NumericParentMaritalStatus 0.3697 0.7398 0.50 0.617   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.54 on 946 degrees of freedom  
## Multiple R-squared: 0.0002639, Adjusted R-squared: -0.0007929   
## F-statistic: 0.2498 on 1 and 946 DF, p-value: 0.6174

Mathregression\_of\_sport=lm(MathScore~NumericPracticeSport,data=Data)  
summary(Mathregression\_of\_sport)

##   
## Call:  
## lm(formula = MathScore ~ NumericPracticeSport, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -65.956 -9.956 -0.020 10.107 34.044   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 65.82881 1.82681 36.035 <2e-16 \*\*\*  
## NumericPracticeSport 0.06381 0.73097 0.087 0.93   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.54 on 946 degrees of freedom  
## Multiple R-squared: 8.054e-06, Adjusted R-squared: -0.001049   
## F-statistic: 0.007619 on 1 and 946 DF, p-value: 0.9305

Mathregression\_of\_firstchild=lm(MathScore~NumericIsFirstChild,data=Data)  
summary(Mathregression\_of\_firstchild)

##   
## Call:  
## lm(formula = MathScore ~ NumericIsFirstChild, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -66.669 -9.669 0.331 11.331 35.404   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 62.522 1.856 33.695 <2e-16 \*\*\*  
## NumericIsFirstChild 2.073 1.070 1.937 0.053 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.51 on 946 degrees of freedom  
## Multiple R-squared: 0.003952, Adjusted R-squared: 0.002899   
## F-statistic: 3.753 on 1 and 946 DF, p-value: 0.053

Mathregression\_of\_NrSibling=lm(MathScore~NrSiblings,data=Data)  
summary(Mathregression\_of\_NrSibling)

##   
## Call:  
## lm(formula = MathScore ~ NrSiblings, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -66.067 -9.878 -0.020 10.216 34.122   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 65.78386 0.88090 74.678 <2e-16 \*\*\*  
## NrSiblings 0.09442 0.34396 0.275 0.784   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.54 on 946 degrees of freedom  
## Multiple R-squared: 7.965e-05, Adjusted R-squared: -0.0009773   
## F-statistic: 0.07536 on 1 and 946 DF, p-value: 0.7837

Mathregression\_of\_transport=lm(MathScore~NumericTransportMeans,data=Data)  
summary(Mathregression\_of\_transport)

##   
## Call:  
## lm(formula = MathScore ~ NumericTransportMeans, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -66.206 -9.576 -0.206 10.424 34.424   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 64.9451 1.8055 35.971 <2e-16 \*\*\*  
## NumericTransportMeans 0.6306 1.0541 0.598 0.55   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.54 on 946 degrees of freedom  
## Multiple R-squared: 0.0003781, Adjusted R-squared: -0.0006786   
## F-statistic: 0.3578 on 1 and 946 DF, p-value: 0.5499

Mathregression\_of\_studytiem=lm(MathScore~NumericWklyStudyHours,data=Data)  
summary(Mathregression\_of\_studytiem)

##   
## Call:  
## lm(formula = MathScore ~ NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -65.461 -9.461 0.231 10.616 35.231   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 62.0762 1.4328 43.327 < 2e-16 \*\*\*  
## NumericWklyStudyHours 1.6923 0.5814 2.911 0.00369 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.47 on 946 degrees of freedom  
## Multiple R-squared: 0.008878, Adjusted R-squared: 0.00783   
## F-statistic: 8.474 on 1 and 946 DF, p-value: 0.003688

#marginal distribution of 11variables and 3 score 这一格个都是 writing score  
writingregression\_of\_eth=lm(WritingScore~NumericEthnic,data=Data)  
summary(writingregression\_of\_eth)

##   
## Call:  
## lm(formula = WritingScore ~ NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -57.560 -10.560 0.895 10.508 33.985   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 60.7421 1.4741 41.207 < 2e-16 \*\*\*  
## NumericEthnic 2.2727 0.4392 5.175 2.79e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.21 on 946 degrees of freedom  
## Multiple R-squared: 0.02753, Adjusted R-squared: 0.0265   
## F-statistic: 26.78 on 1 and 946 DF, p-value: 2.785e-07

writingregression\_of\_gender=lm(WritingScore~NumericGender,data=Data)  
summary(writingregression\_of\_gender)

##   
## Call:  
## lm(formula = WritingScore ~ NumericGender, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -62.258 -10.258 0.742 10.663 36.663   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 81.1794 1.5032 54.01 <2e-16 \*\*\*  
## NumericGender -8.9212 0.9592 -9.30 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.76 on 946 degrees of freedom  
## Multiple R-squared: 0.08378, Adjusted R-squared: 0.08281   
## F-statistic: 86.5 on 1 and 946 DF, p-value: < 2.2e-16

writingregression\_of\_parentedu=lm(WritingScore~Numericparentedu,data=Data)  
summary(writingregression\_of\_parentedu)

##   
## Call:  
## lm(formula = WritingScore ~ Numericparentedu, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -56.198 -10.949 0.227 10.657 33.802   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 70.4878 1.1051 63.782 < 2e-16 \*\*\*  
## Numericparentedu -0.7151 0.2756 -2.595 0.00961 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.37 on 946 degrees of freedom  
## Multiple R-squared: 0.007067, Adjusted R-squared: 0.006017   
## F-statistic: 6.733 on 1 and 946 DF, p-value: 0.009612

writingregression\_of\_lunchtype=lm(WritingScore~NumericLunchType,data=Data)  
summary(writingregression\_of\_lunchtype)

##   
## Call:  
## lm(formula = WritingScore ~ NumericLunchType, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.731 -9.718 0.282 11.269 37.269   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 54.744 1.749 31.295 < 2e-16 \*\*\*  
## NumericLunchType 7.987 1.018 7.845 1.16e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.94 on 946 degrees of freedom  
## Multiple R-squared: 0.06109, Adjusted R-squared: 0.0601   
## F-statistic: 61.55 on 1 and 946 DF, p-value: 1.163e-14

writingregression\_of\_testprep=lm(WritingScore~NumericTestPrep,data=Data)  
summary(writingregression\_of\_testprep)

##   
## Call:  
## lm(formula = WritingScore ~ NumericTestPrep, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -54.65 -10.30 1.35 10.35 35.35   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 83.959 1.744 48.15 <2e-16 \*\*\*  
## NumericTestPrep -9.654 1.010 -9.56 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.73 on 946 degrees of freedom  
## Multiple R-squared: 0.0881, Adjusted R-squared: 0.08713   
## F-statistic: 91.39 on 1 and 946 DF, p-value: < 2.2e-16

writingregression\_of\_parentmarriage=lm(WritingScore~NumericParentMaritalStatus,data=Data)  
summary(writingregression\_of\_parentmarriage)

##   
## Call:  
## lm(formula = WritingScore ~ NumericParentMaritalStatus, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -58.458 -10.856 0.343 10.542 32.745   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 66.6541 1.6356 40.752 <2e-16 \*\*\*  
## NumericParentMaritalStatus 0.6011 0.7340 0.819 0.413   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.42 on 946 degrees of freedom  
## Multiple R-squared: 0.0007085, Adjusted R-squared: -0.0003479   
## F-statistic: 0.6707 on 1 and 946 DF, p-value: 0.413

writingregression\_of\_sport=lm(WritingScore~NumericPracticeSport,data=Data)  
summary(writingregression\_of\_sport)

##   
## Call:  
## lm(formula = WritingScore ~ NumericPracticeSport, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -57.736 -10.736 0.747 10.756 32.264   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 66.7710 1.8125 36.838 <2e-16 \*\*\*  
## NumericPracticeSport 0.4823 0.7253 0.665 0.506   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.42 on 946 degrees of freedom  
## Multiple R-squared: 0.0004672, Adjusted R-squared: -0.0005894   
## F-statistic: 0.4422 on 1 and 946 DF, p-value: 0.5062

writingregression\_of\_firstchild=lm(WritingScore~NumericIsFirstChild,data=Data)  
summary(writingregression\_of\_firstchild)

##   
## Call:  
## lm(formula = WritingScore ~ NumericIsFirstChild, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -58.809 -10.809 0.847 10.191 33.847   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 63.497 1.839 34.528 <2e-16 \*\*\*  
## NumericIsFirstChild 2.656 1.061 2.504 0.0124 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.37 on 946 degrees of freedom  
## Multiple R-squared: 0.006587, Adjusted R-squared: 0.005536   
## F-statistic: 6.272 on 1 and 946 DF, p-value: 0.01243

writingregression\_of\_NrSibling=lm(WritingScore~NrSiblings,data=Data)  
summary(writingregression\_of\_NrSibling)

##   
## Call:  
## lm(formula = WritingScore ~ NrSiblings, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -57.948 -10.906 0.115 10.338 32.115   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 67.8855 0.8743 77.649 <2e-16 \*\*\*  
## NrSiblings 0.0209 0.3414 0.061 0.951   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.42 on 946 degrees of freedom  
## Multiple R-squared: 3.961e-06, Adjusted R-squared: -0.001053   
## F-statistic: 0.003747 on 1 and 946 DF, p-value: 0.9512

writingregression\_of\_transport=lm(WritingScore~NumericTransportMeans,data=Data)  
summary(writingregression\_of\_transport)

##   
## Call:  
## lm(formula = WritingScore ~ NumericTransportMeans, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -58.268 -10.315 0.685 10.697 32.685   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 66.3607 1.7913 37.045 <2e-16 \*\*\*  
## NumericTransportMeans 0.9539 1.0459 0.912 0.362   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.41 on 946 degrees of freedom  
## Multiple R-squared: 0.0008785, Adjusted R-squared: -0.0001777   
## F-statistic: 0.8318 on 1 and 946 DF, p-value: 0.362

writingregression\_of\_studytiem=lm(WritingScore~NumericWklyStudyHours,data=Data)  
summary(writingregression\_of\_studytiem)

##   
## Call:  
## lm(formula = WritingScore ~ NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -57.548 -10.787 0.692 10.213 33.692   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 65.0685 1.4248 45.669 <2e-16 \*\*\*  
## NumericWklyStudyHours 1.2395 0.5781 2.144 0.0323 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.38 on 946 degrees of freedom  
## Multiple R-squared: 0.004836, Adjusted R-squared: 0.003784   
## F-statistic: 4.597 on 1 and 946 DF, p-value: 0.03229

#marginal distribution of 11variables and 3 score 这一格个都是 reading score  
Readingregression\_of\_eth=lm(ReadingScore~NumericEthnic,data=Data)  
summary(Readingregression\_of\_eth)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -51.536 -9.906 0.583 10.464 35.224   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 62.8956 1.4206 44.274 < 2e-16 \*\*\*  
## NumericEthnic 1.8802 0.4232 4.442 9.95e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.66 on 946 degrees of freedom  
## Multiple R-squared: 0.02044, Adjusted R-squared: 0.0194   
## F-statistic: 19.74 on 1 and 946 DF, p-value: 9.945e-06

Readingregression\_of\_gender=lm(ReadingScore~NumericGender,data=Data)  
summary(Readingregression\_of\_gender)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericGender, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -55.26 -10.21 0.74 10.74 34.78   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 79.3053 1.4646 54.148 < 2e-16 \*\*\*  
## NumericGender -7.0450 0.9346 -7.538 1.12e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.38 on 946 degrees of freedom  
## Multiple R-squared: 0.05666, Adjusted R-squared: 0.05566   
## F-statistic: 56.82 on 1 and 946 DF, p-value: 1.116e-13

Readingregression\_of\_parentedu=lm(ReadingScore~Numericparentedu,data=Data)  
summary(Readingregression\_of\_parentedu)

##   
## Call:  
## lm(formula = ReadingScore ~ Numericparentedu, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -50.244 -9.888 0.766 10.457 32.756   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 71.2024 1.0615 67.079 <2e-16 \*\*\*  
## Numericparentedu -0.6597 0.2647 -2.493 0.0129 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.76 on 946 degrees of freedom  
## Multiple R-squared: 0.006525, Adjusted R-squared: 0.005474   
## F-statistic: 6.213 on 1 and 946 DF, p-value: 0.01285

Readingregression\_of\_lunchtype=lm(ReadingScore~NumericLunchType,data=Data)  
summary(Readingregression\_of\_lunchtype)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericLunchType, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.073 -9.154 0.600 10.600 35.927   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 56.7447 1.6845 33.687 < 2e-16 \*\*\*  
## NumericLunchType 7.3278 0.9803 7.475 1.76e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.39 on 946 degrees of freedom  
## Multiple R-squared: 0.05577, Adjusted R-squared: 0.05477   
## F-statistic: 55.87 on 1 and 946 DF, p-value: 1.758e-13

Readingregression\_of\_testprep=lm(ReadingScore~NumericTestPrep,data=Data)  
summary(Readingregression\_of\_testprep)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericTestPrep, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -49.42 -9.42 0.58 10.45 33.58   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 80.6793 1.7070 47.264 < 2e-16 \*\*\*  
## NumericTestPrep -7.1296 0.9887 -7.211 1.13e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.42 on 946 degrees of freedom  
## Multiple R-squared: 0.05211, Adjusted R-squared: 0.05111   
## F-statistic: 52 on 1 and 946 DF, p-value: 1.134e-12

Readingregression\_of\_parentmarriage=lm(ReadingScore~NumericParentMaritalStatus,data=Data)  
summary(Readingregression\_of\_parentmarriage)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericParentMaritalStatus, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.358 -9.770 0.817 10.817 31.817   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 67.5951 1.5705 43.040 <2e-16 \*\*\*  
## NumericParentMaritalStatus 0.5877 0.7048 0.834 0.405   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.8 on 946 degrees of freedom  
## Multiple R-squared: 0.0007344, Adjusted R-squared: -0.0003219   
## F-statistic: 0.6952 on 1 and 946 DF, p-value: 0.4046

Readingregression\_of\_sport=lm(ReadingScore~NumericPracticeSport,data=Data)  
summary(Readingregression\_of\_sport)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericPracticeSport, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -51.735 -10.001 0.765 10.999 31.265   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 68.2024 1.7407 39.180 <2e-16 \*\*\*  
## NumericPracticeSport 0.2662 0.6965 0.382 0.702   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.81 on 946 degrees of freedom  
## Multiple R-squared: 0.0001544, Adjusted R-squared: -0.0009026   
## F-statistic: 0.1461 on 1 and 946 DF, p-value: 0.7024

Readingregression\_of\_firstchild=lm(ReadingScore~NumericIsFirstChild,data=Data)  
summary(Readingregression\_of\_firstchild)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericIsFirstChild, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.767 -9.975 1.025 10.233 33.025   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 64.182 1.765 36.370 <2e-16 \*\*\*  
## NumericIsFirstChild 2.792 1.018 2.743 0.0062 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.75 on 946 degrees of freedom  
## Multiple R-squared: 0.007892, Adjusted R-squared: 0.006844   
## F-statistic: 7.525 on 1 and 946 DF, p-value: 0.006198

Readingregression\_of\_NrSibling=lm(ReadingScore~NrSiblings,data=Data)  
summary(Readingregression\_of\_NrSibling)

##   
## Call:  
## lm(formula = ReadingScore ~ NrSiblings, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -51.733 -9.975 0.828 10.904 31.752   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 69.0960 0.8394 82.313 <2e-16 \*\*\*  
## NrSiblings -0.1211 0.3278 -0.369 0.712   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.81 on 946 degrees of freedom  
## Multiple R-squared: 0.0001443, Adjusted R-squared: -0.0009126   
## F-statistic: 0.1365 on 1 and 946 DF, p-value: 0.7119

Readingregression\_of\_transport=lm(ReadingScore~NumericTransportMeans,data=Data)  
summary(Readingregression\_of\_transport)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericTransportMeans, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.293 -10.024 0.707 10.707 31.976   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 66.755 1.719 38.824 <2e-16 \*\*\*  
## NumericTransportMeans 1.269 1.004 1.264 0.206   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.79 on 946 degrees of freedom  
## Multiple R-squared: 0.001687, Adjusted R-squared: 0.0006316   
## F-statistic: 1.598 on 1 and 946 DF, p-value: 0.2064

Readingregression\_of\_studytiem=lm(ReadingScore~NumericWklyStudyHours,data=Data)  
summary(Readingregression\_of\_studytiem)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -51.47 -10.26 0.74 10.32 31.74   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 66.0509 1.3680 48.282 <2e-16 \*\*\*  
## NumericWklyStudyHours 1.2092 0.5551 2.178 0.0296 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.77 on 946 degrees of freedom  
## Multiple R-squared: 0.004992, Adjusted R-squared: 0.00394   
## F-statistic: 4.746 on 1 and 946 DF, p-value: 0.02962

NEWregression\_of\_Math=lm(MathScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericWklyStudyHours, data=Data)  
  
NEWregression\_of\_Reading=lm(ReadingScore~NumericEthnic+Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericWklyStudyHours, data=Data)  
  
NEWregression\_of\_Writing=lm(WritingScore~Numericparentedu+NumericGender+NumericLunchType+NumericTestPrep+NumericIsFirstChild+NumericWklyStudyHours +NumericEthnic, data=Data)  
summary(NEWregression\_of\_Math)

##   
## Call:  
## lm(formula = MathScore ~ NumericEthnic + Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.574 -8.722 0.145 9.976 33.653   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.6787 3.6415 10.347 < 2e-16 \*\*\*  
## NumericEthnic 2.5773 0.3939 6.543 9.92e-11 \*\*\*  
## Numericparentedu -0.6934 0.2443 -2.838 0.00463 \*\*   
## NumericGender 5.0245 0.8862 5.670 1.90e-08 \*\*\*  
## NumericLunchType 10.9825 0.9300 11.809 < 2e-16 \*\*\*  
## NumericTestPrep -5.8652 0.9357 -6.268 5.55e-10 \*\*\*  
## NumericIsFirstChild 2.1178 0.9417 2.249 0.02475 \*   
## NumericWklyStudyHours 1.4060 0.5122 2.745 0.00616 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.61 on 940 degrees of freedom  
## Multiple R-squared: 0.2372, Adjusted R-squared: 0.2315   
## F-statistic: 41.76 on 7 and 940 DF, p-value: < 2.2e-16

summary(NEWregression\_of\_Reading)

##   
## Call:  
## lm(formula = ReadingScore ~ NumericEthnic + Numericparentedu +   
## NumericGender + NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericWklyStudyHours, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -46.638 -8.841 0.666 9.682 31.115   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 70.6934 3.5491 19.919 < 2e-16 \*\*\*  
## NumericEthnic 1.6086 0.3839 4.190 3.05e-05 \*\*\*  
## Numericparentedu -0.7004 0.2381 -2.942 0.00334 \*\*   
## NumericGender -7.3141 0.8637 -8.468 < 2e-16 \*\*\*  
## NumericLunchType 7.4101 0.9064 8.175 9.52e-16 \*\*\*  
## NumericTestPrep -7.1249 0.9119 -7.813 1.49e-14 \*\*\*  
## NumericIsFirstChild 2.2599 0.9178 2.462 0.01399 \*   
## NumericWklyStudyHours 0.9775 0.4992 1.958 0.05048 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.27 on 940 degrees of freedom  
## Multiple R-squared: 0.2021, Adjusted R-squared: 0.1962   
## F-statistic: 34.02 on 7 and 940 DF, p-value: < 2.2e-16

summary(NEWregression\_of\_Writing)

##   
## Call:  
## lm(formula = WritingScore ~ Numericparentedu + NumericGender +   
## NumericLunchType + NumericTestPrep + NumericIsFirstChild +   
## NumericWklyStudyHours + NumericEthnic, data = Data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.004 -8.764 0.332 9.504 28.915   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 75.5082 3.5178 21.465 < 2e-16 \*\*\*  
## Numericparentedu -0.7737 0.2360 -3.279 0.00108 \*\*   
## NumericGender -9.2560 0.8561 -10.812 < 2e-16 \*\*\*  
## NumericLunchType 8.1090 0.8984 9.026 < 2e-16 \*\*\*  
## NumericTestPrep -9.6935 0.9039 -10.724 < 2e-16 \*\*\*  
## NumericIsFirstChild 1.9431 0.9097 2.136 0.03294 \*   
## NumericWklyStudyHours 0.9722 0.4948 1.965 0.04969 \*   
## NumericEthnic 1.9472 0.3805 5.117 3.77e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.15 on 940 degrees of freedom  
## Multiple R-squared: 0.2772, Adjusted R-squared: 0.2719   
## F-statistic: 51.51 on 7 and 940 DF, p-value: < 2.2e-16

#直接去掉P>0.05的和backward elimination的结果完全一致, 并且三个score的预测variable一样

```