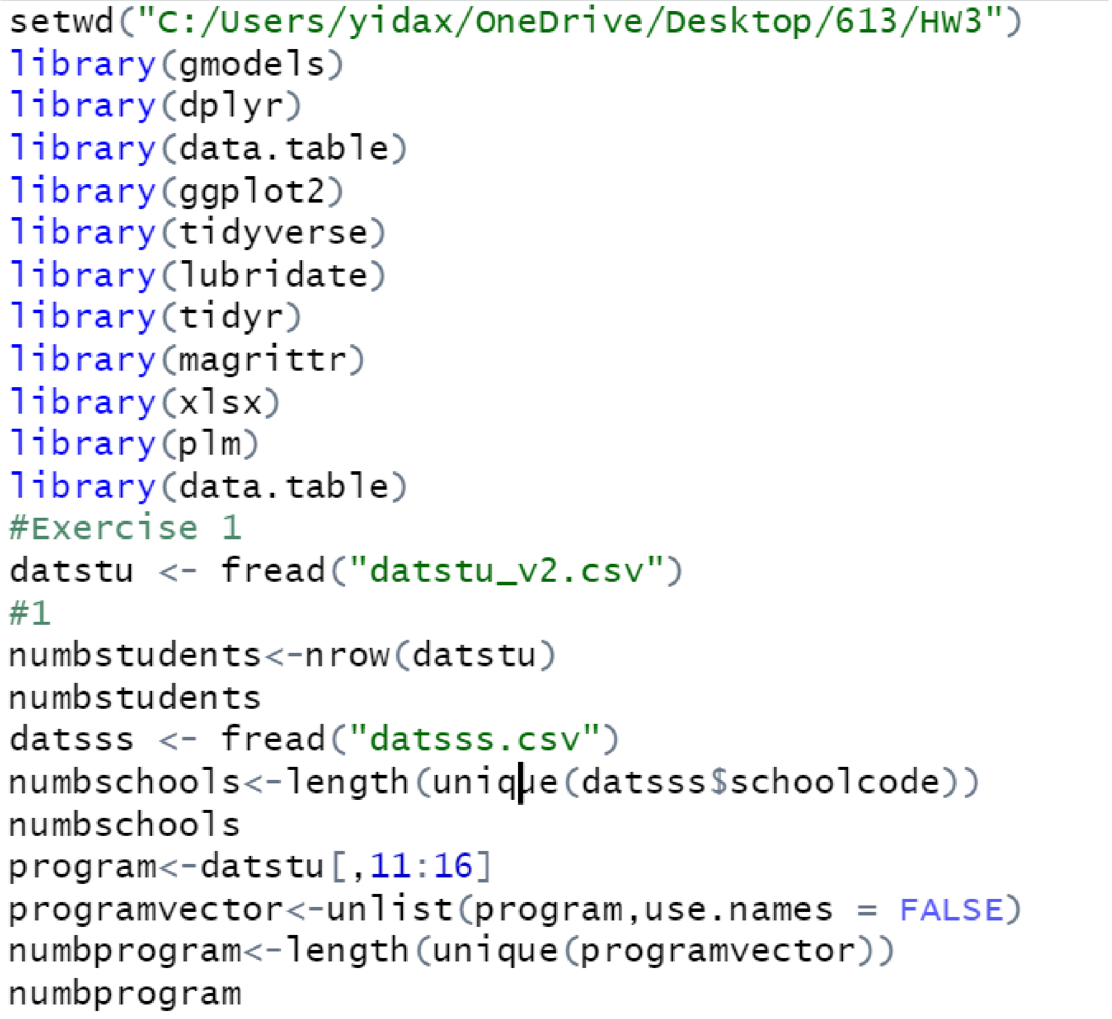
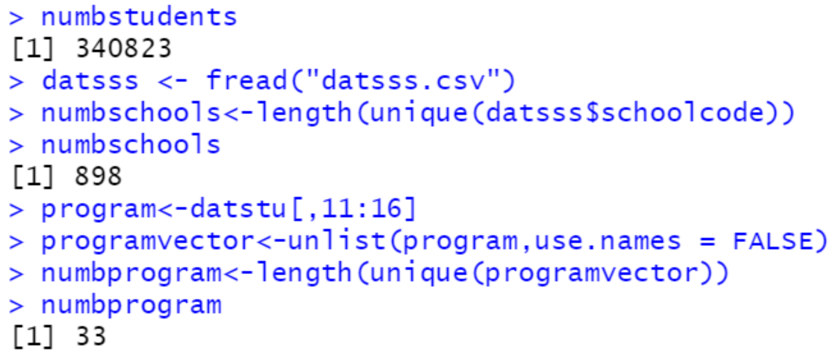
Assignment3

#Exercise1

1、

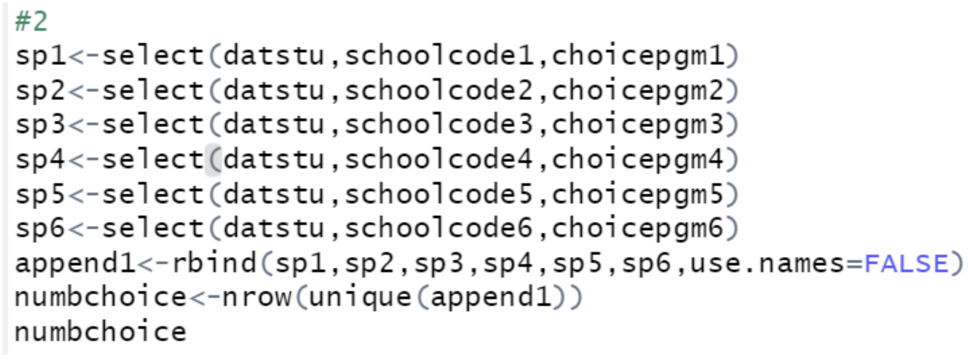


```
setwd("C:/Users/yidax/OneDrive/Desktop/613/HW3")
library(gmodels)
library(dplyr)
library(data.table)
library(ggplot2)
library(tidyverse)
library(lubridate)
library(tidyr)
library(magrittr)
library(xlsx)
library(plm)
library(data.table)
#Exercise 1
datstu <- fread("datstu_v2.csv")
#1
numbstudents<-nrow(datstu)
numbstudents
datsss <- fread("datsss.csv")
numbschools<-length(unique(datsss$schoolcode))
numbschools
program<-datstu[,11:16]
programvector<-unlist(program,use.names = FALSE)
numbprogram<-length(unique(programvector))
numbprogram
```

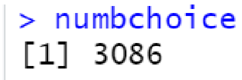


```
> numbstudents
[1] 340823
> datsss <- fread("datsss.csv")
> numbschools<-length(unique(datsss$schoolcode))
> numbschools
[1] 898
> program<-datstu[,11:16]
> programvector<-unlist(program,use.names = FALSE)
> numbprogram<-length(unique(programvector))
> numbprogram
[1] 33
```

2、

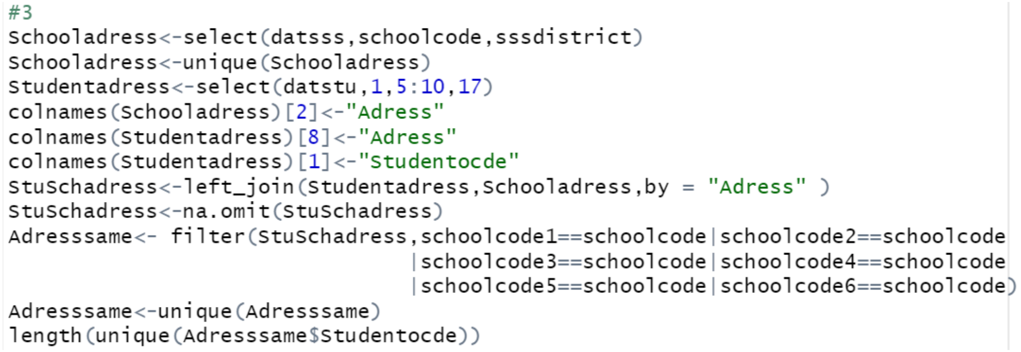


```
#2
sp1<-select(datstu,schoolcode1,choicepgm1)
sp2<-select(datstu,schoolcode2,choicepgm2)
sp3<-select(datstu,schoolcode3,choicepgm3)
sp4<-select(datstu,schoolcode4,choicepgm4)
sp5<-select(datstu,schoolcode5,choicepgm5)
sp6<-select(datstu,schoolcode6,choicepgm6)
append1<-rbind(sp1,sp2,sp3,sp4,sp5,sp6,use.names=FALSE)
numbchoice<-nrow(unique(append1))
numbchoice
```

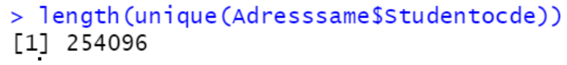


```
> numbchoice
[1] 3086
```

3、

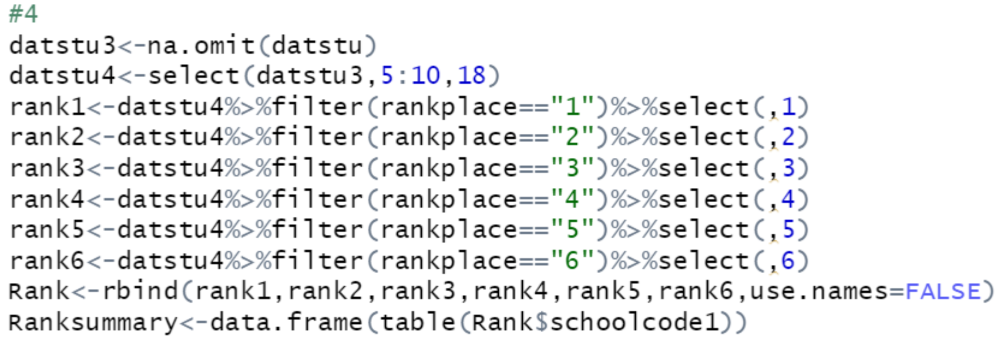


```
#3
Schooladress<-select(datsss,schoolcode,sssdistrict)
Schooladress<-unique(Schooladress)
Studentadress<-select(datstu,1,5:10,17)
colnames(Schooladress)[2]<-"Adress"
colnames(Studentadress)[8]<-"Adress"
colnames(Studentadress)[1]<-"Studentocde"
StuSchadress<-left_join(Studentadress,Schooladress,by = "Adress" )
StuSchadress<-na.omit(StuSchadress)
Adresssame<- filter(StuSchadress,schoolcode1==schoolcode|schoolcode2==schoolcode
                               |schoolcode3==schoolcode|schoolcode4==schoolcode
                               |schoolcode5==schoolcode|schoolcode6==schoolcode)
Adresssame<-unique(Adresssame)
length(unique(Adresssame$Studentocde))
```

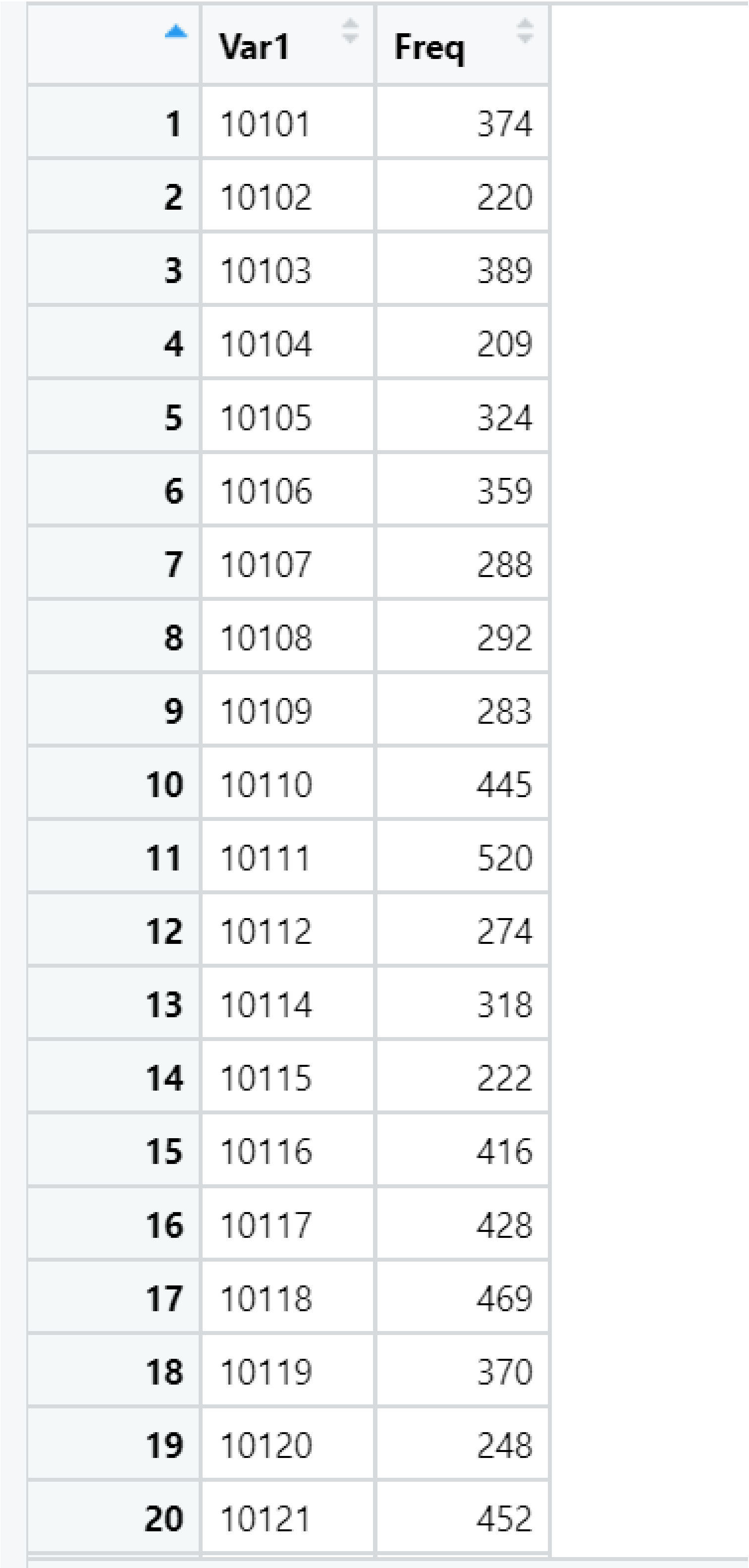


```
> length(unique(Adresssame$Studentocde))
[1] 254096
```

4、

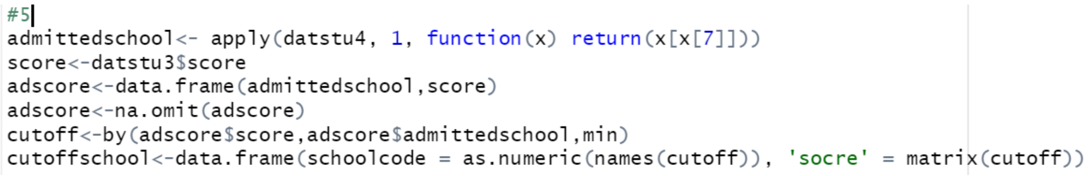


```
#4
datstu3<-na.omit(datstu)
datstu4<-select(datstu3,5:10,18)
rank1<-datstu4%>%filter(rankplace=="1")%>%select(,1)
rank2<-datstu4%>%filter(rankplace=="2")%>%select(,2)
rank3<-datstu4%>%filter(rankplace=="3")%>%select(,3)
rank4<-datstu4%>%filter(rankplace=="4")%>%select(,4)
rank5<-datstu4%>%filter(rankplace=="5")%>%select(,5)
rank6<-datstu4%>%filter(rankplace=="6")%>%select(,6)
Rank<-rbind(rank1,rank2,rank3,rank4,rank5,rank6,use.names=FALSE)
Ranksummary<-data.frame(table(Rank$schoolcode1))
```

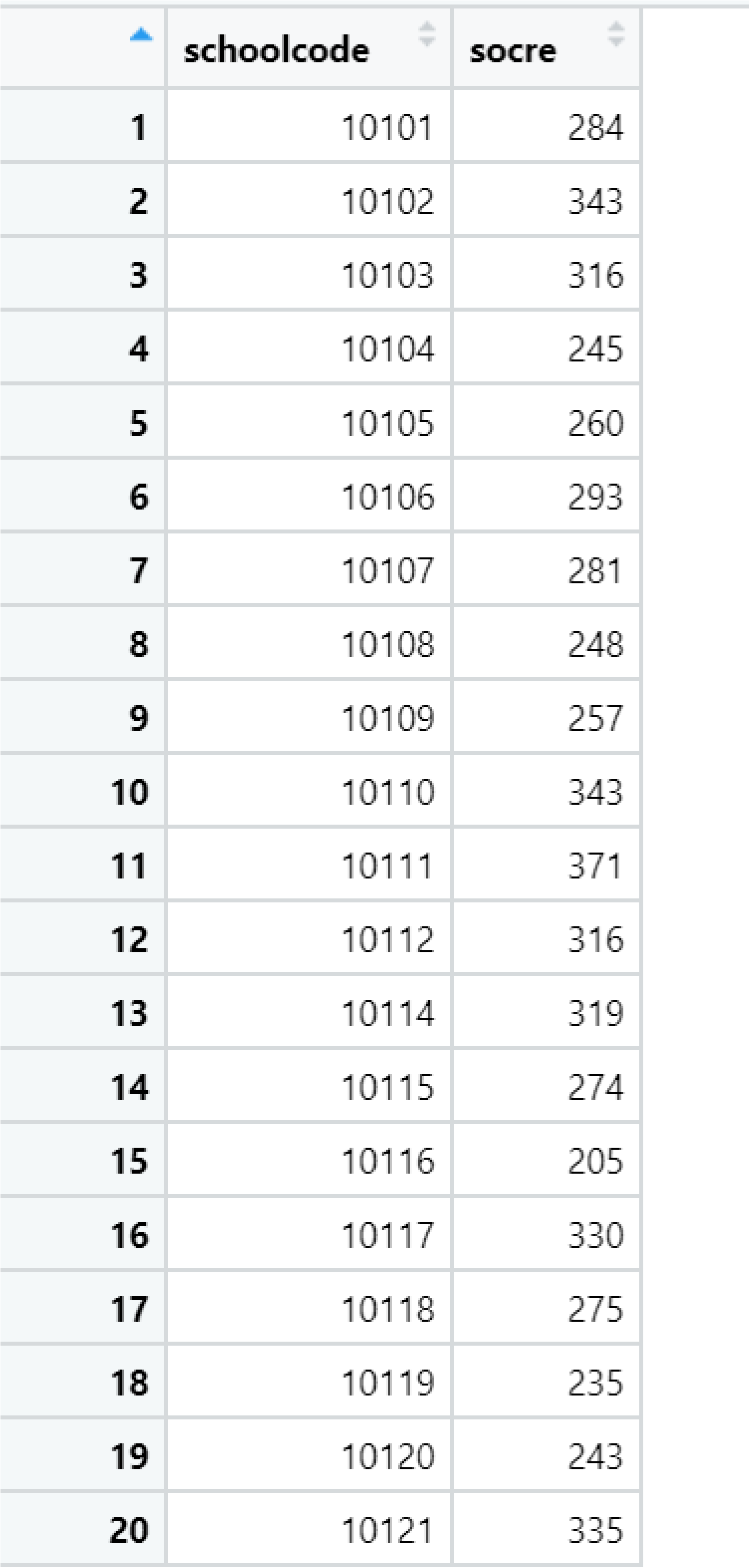


|  | Var1 | Freq |
|---|---|---|
| 1 | 10101 | 374 |
| 2 | 10102 | 220 |
| 3 | 10103 | 389 |
| 4 | 10104 | 209 |
| 5 | 10105 | 324 |
| 6 | 10106 | 359 |
| 7 | 10107 | 288 |
| 8 | 10108 | 292 |
| 9 | 10109 | 283 |
| 10 | 10110 | 445 |
| 11 | 10111 | 520 |
| 12 | 10112 | 274 |
| 13 | 10114 | 318 |
| 14 | 10115 | 222 |
| 15 | 10116 | 416 |
| 16 | 10117 | 428 |
| 17 | 10118 | 469 |
| 18 | 10119 | 370 |
| 19 | 10120 | 248 |
| 20 | 10121 | 452 |

5、



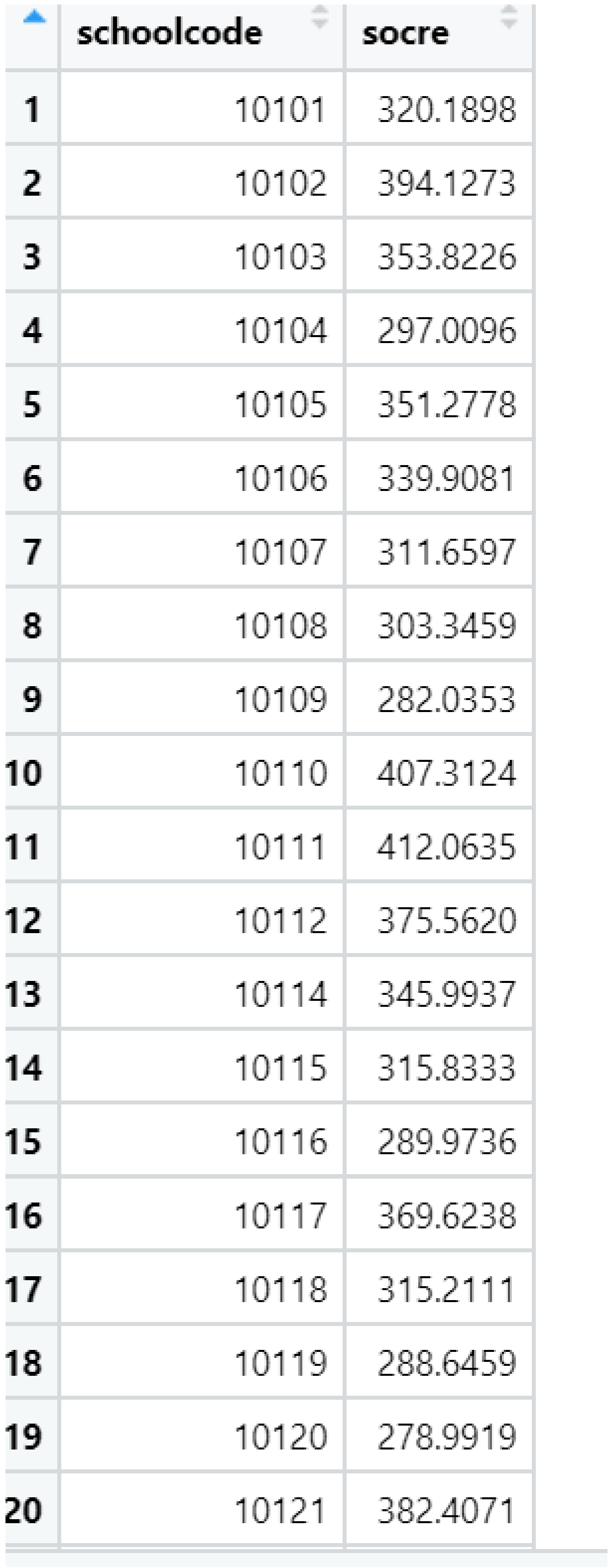
```
#5
admittedschool<- apply(datstu4, 1, function(x) return(x[x[7]]))
score<-datstu3$score
adscore<-data.frame(admittedschool,score)
adscore<-na.omit(adscore)
cutoff<-by(adscore$score,adscore$admittedschool,min)
cutoffschool<-data.frame(schoolcode = as.numeric(names(cutoff)), 'socre' = matrix(cutoff))
```



| | schoolcode | socre |
|---|---|---|
| 1 | 10101 | 284 |
| 2 | 10102 | 343 |
| 3 | 10103 | 316 |
| 4 | 10104 | 245 |
| 5 | 10105 | 260 |
| 6 | 10106 | 293 |
| 7 | 10107 | 281 |
| 8 | 10108 | 248 |
| 9 | 10109 | 257 |
| 10 | 10110 | 343 |
| 11 | 10111 | 371 |
| 12 | 10112 | 316 |
| 13 | 10114 | 319 |
| 14 | 10115 | 274 |
| 15 | 10116 | 205 |
| 16 | 10117 | 330 |
| 17 | 10118 | 275 |
| 18 | 10119 | 235 |
| 19 | 10120 | 243 |
| 20 | 10121 | 335 |

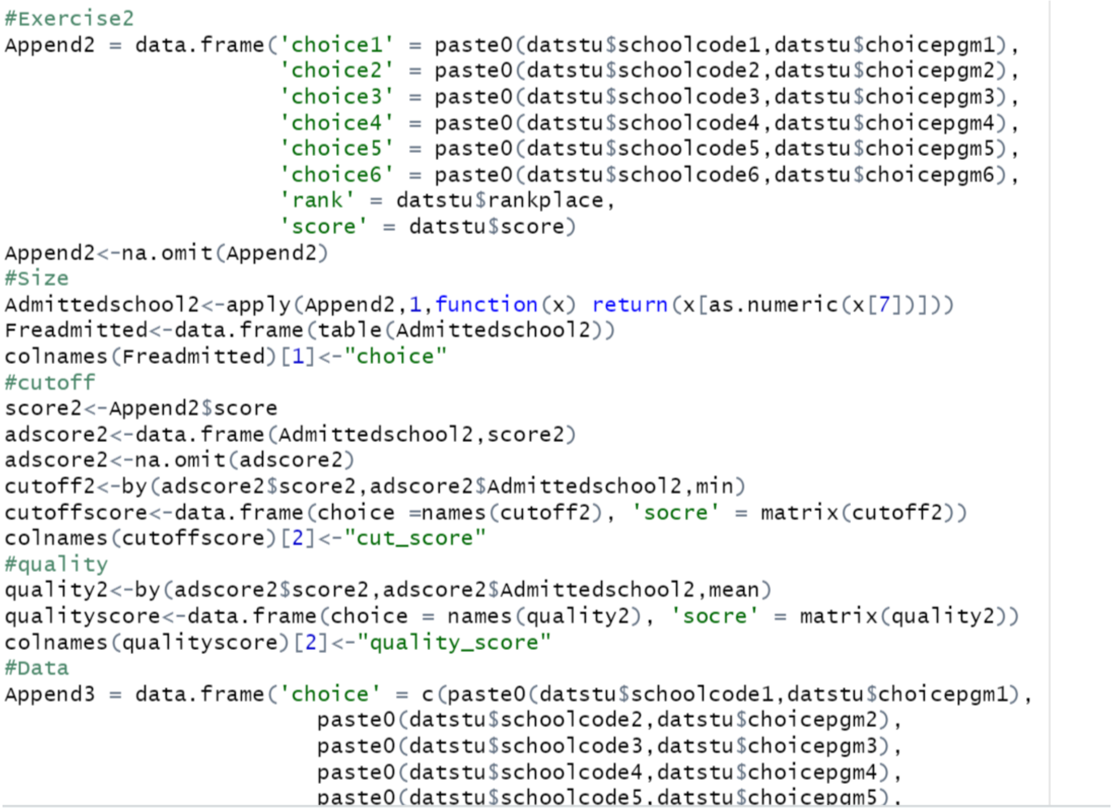
6、

```
#6
quality<-by(adscore$score,adscore$admittedschool,mean)
qualityschool<-data.frame(schoolcode = as.numeric(names(quality)), 'socre' = matrix(quality))
```

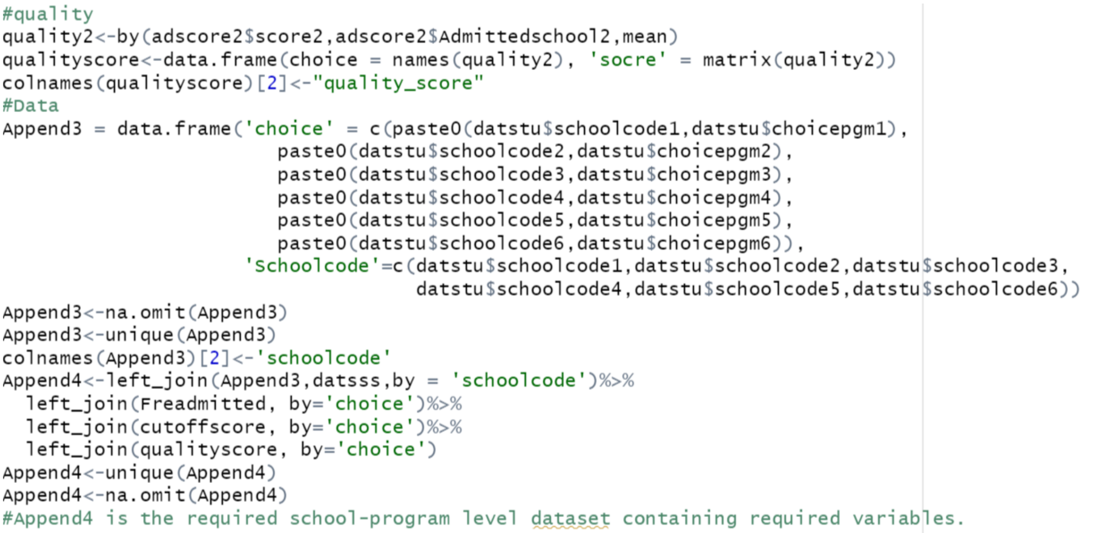


| | schoolcode | socre |
|---|---|---|
| 1 | 10101 | 320.1898 |
| 2 | 10102 | 394.1273 |
| 3 | 10103 | 353.8226 |
| 4 | 10104 | 297.0096 |
| 5 | 10105 | 351.2778 |
| 6 | 10106 | 339.9081 |
| 7 | 10107 | 311.6597 |
| 8 | 10108 | 303.3459 |
| 9 | 10109 | 282.0353 |
| 10 | 10110 | 407.3124 |
| 11 | 10111 | 412.0635 |
| 12 | 10112 | 375.5620 |
| 13 | 10114 | 345.9937 |
| 14 | 10115 | 315.8333 |
| 15 | 10116 | 289.9736 |
| 16 | 10117 | 369.6238 |
| 17 | 10118 | 315.2111 |
| 18 | 10119 | 288.6459 |
| 19 | 10120 | 278.9919 |
| 20 | 10121 | 382.4071 |

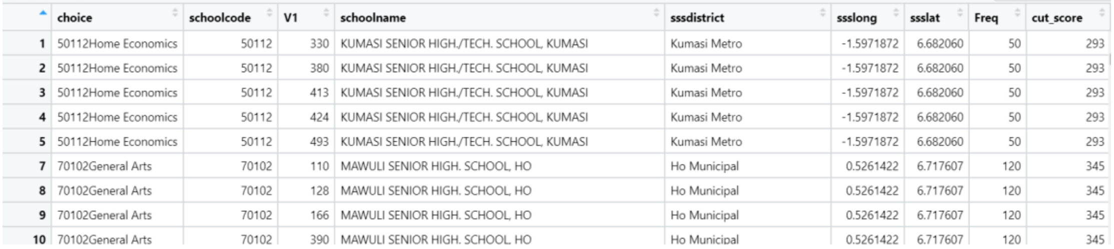
#Exercise2



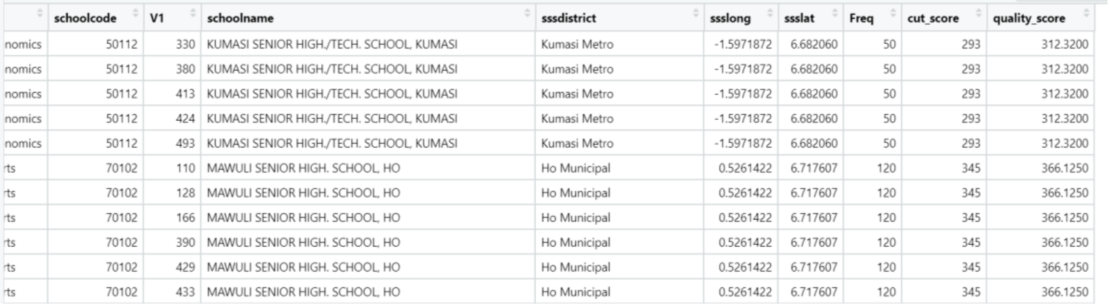
```
#Exercise2
Append2 = data.frame('choice1' = paste0(datstu$schoolcode1,datstu$choicepgm1),
                     'choice2' = paste0(datstu$schoolcode2,datstu$choicepgm2),
                     'choice3' = paste0(datstu$schoolcode3,datstu$choicepgm3),
                     'choice4' = paste0(datstu$schoolcode4,datstu$choicepgm4),
                     'choice5' = paste0(datstu$schoolcode5,datstu$choicepgm5),
                     'choice6' = paste0(datstu$schoolcode6,datstu$choicepgm6),
                     'rank' = datstu$rankplace,
                     'score' = datstu$score)
Append2<-na.omit(Append2)
#Size
Admittedschool2<-apply(Append2,1,function(x) return(x[as.numeric(x[7])]))
Freadmitted<-data.frame(table(Admittedschool2))
colnames(Freadmitted)[1]<-"choice"
#cutoff
score2<-Append2$score
adscore2<-data.frame(Admittedschool2,score2)
adscore2<-na.omit(adscore2)
cutoff2<-by(adscore2$score2,adscore2$Admittedschool2,min)
cutoffscore<-data.frame(choice =names(cutoff2), 'socre' = matrix(cutoff2))
colnames(cutoffscore)[2]<-"cut_score"
#quality
quality2<-by(adscore2$score2,adscore2$Admittedschool2,mean)
qualityscore<-data.frame(choice = names(quality2), 'socre' = matrix(quality2))
colnames(qualityscore)[2]<-"quality_score"
#Data
Append3 = data.frame('choice' = c(paste0(datstu$schoolcode1,datstu$choicepgm1),
                        paste0(datstu$schoolcode2,datstu$choicepgm2),
                        paste0(datstu$schoolcode3,datstu$choicepgm3),
                        paste0(datstu$schoolcode4,datstu$choicepgm4),
                        paste0(datstu$schoolcode5,datstu$choicepgm5),
```



```
#quality
quality2<-by(adscore2$score2,adscore2$Admittedschool2,mean)
qualityscore<-data.frame(choice = names(quality2), 'socre' = matrix(quality2))
colnames(qualityscore)[2]<-"quality_score"
#Data
Append3 = data.frame('choice' = c(paste0(datstu$schoolcode1,datstu$choicepgm1),
                        paste0(datstu$schoolcode2,datstu$choicepgm2),
                        paste0(datstu$schoolcode3,datstu$choicepgm3),
                        paste0(datstu$schoolcode4,datstu$choicepgm4),
                        paste0(datstu$schoolcode5,datstu$choicepgm5),
                        paste0(datstu$schoolcode6,datstu$choicepgm6)),
                     'Schoolcode'=c(datstu$schoolcode1,datstu$schoolcode2,datstu$schoolcode3,
                                    datstu$schoolcode4,datstu$schoolcode5,datstu$schoolcode6))
Append3<-na.omit(Append3)
Append3<-unique(Append3)
colnames(Append3)[2]<-'schoolcode'
Append4<-left_join(Append3,datsss,by = 'schoolcode')%>%
  left_join(Freadmitted, by='choice')%>%
  left_join(cutoffscore, by='choice')%>%
  left_join(qualityscore, by='choice')
Append4<-unique(Append4)
Append4<-na.omit(Append4)
#Append4 is the required school-program level dataset containing required variables.
```

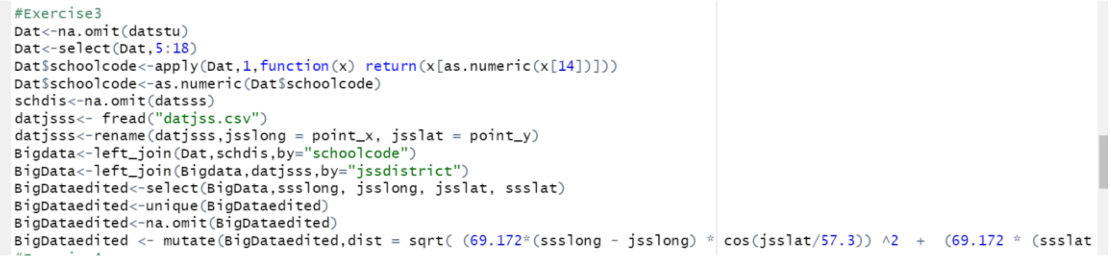


| | choice | schoolcode | V1 | schoolname | sssdistrict | ssslong | ssslat | Freq | cut_score |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 50112Home Economics | 50112 | 330 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 |
| 2 | 50112Home Economics | 50112 | 380 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 |
| 3 | 50112Home Economics | 50112 | 413 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 |
| 4 | 50112Home Economics | 50112 | 424 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 |
| 5 | 50112Home Economics | 50112 | 493 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 |
| 7 | 70102General Arts | 70102 | 110 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 |
| 8 | 70102General Arts | 70102 | 128 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 |
| 9 | 70102General Arts | 70102 | 166 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 |
| 10 | 70102General Arts | 70102 | 390 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 |

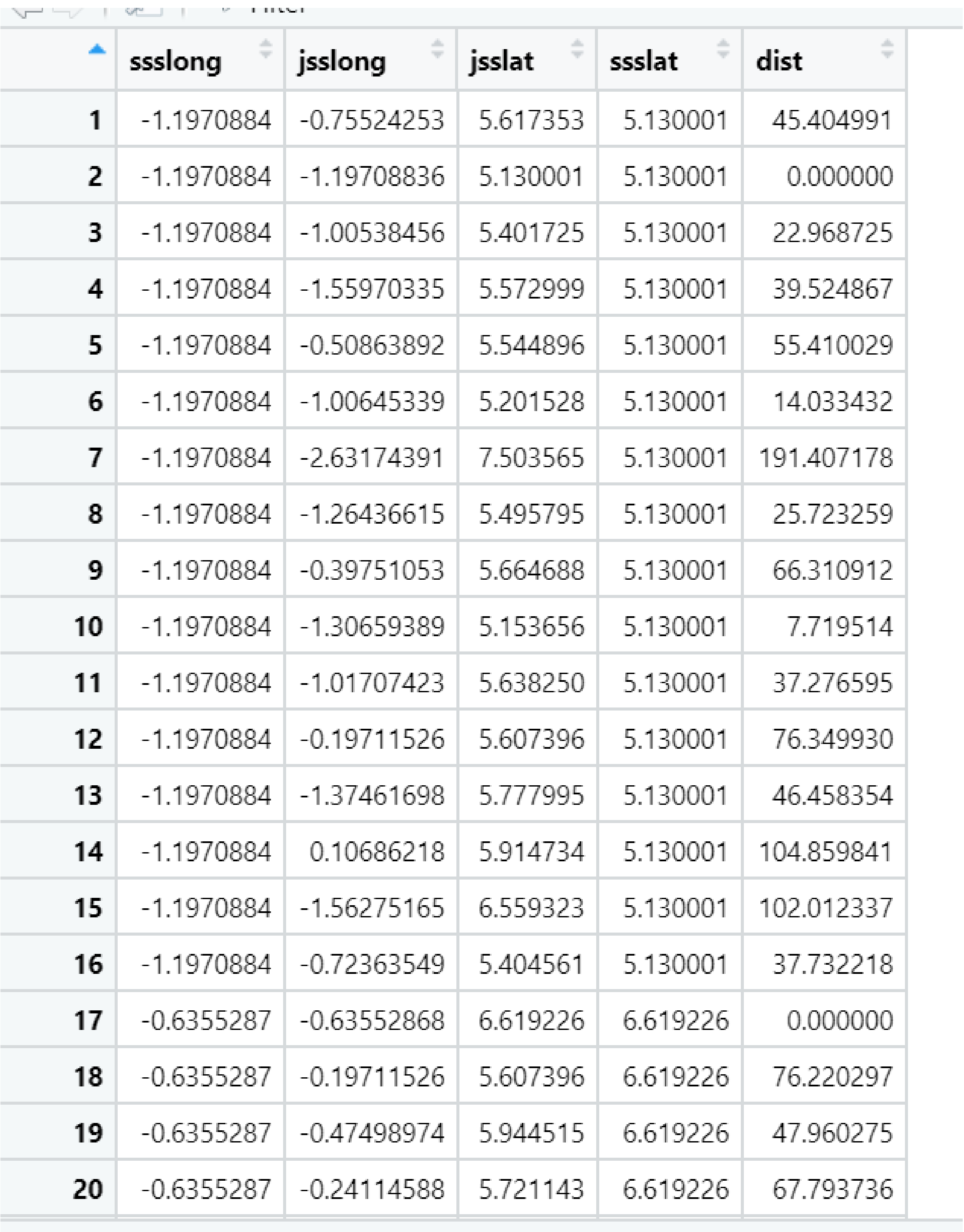


| | schoolcode | V1 | schoolname | sssdistrict | ssslong | ssslat | Freq | cut_score | quality_score |
|---|---|---|---|---|---|---|---|---|---|
| nomics | 50112 | 330 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 | 312.3200 |
| nomics | 50112 | 380 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 | 312.3200 |
| nomics | 50112 | 413 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 | 312.3200 |
| nomics | 50112 | 424 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 | 312.3200 |
| nomics | 50112 | 493 | KUMASI SENIOR HIGH./TECH. SCHOOL, KUMASI | Kumasi Metro | -1.5971872 | 6.682060 | 50 | 293 | 312.3200 |
| rts | 70102 | 110 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 | 366.1250 |
| rts | 70102 | 128 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 | 366.1250 |
| rts | 70102 | 166 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 | 366.1250 |
| rts | 70102 | 390 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 | 366.1250 |
| rts | 70102 | 429 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 | 366.1250 |
| rts | 70102 | 433 | MAWULI SENIOR HIGH. SCHOOL, HO | Ho Municipal | 0.5261422 | 6.717607 | 120 | 345 | 366.1250 |

#Exercise3

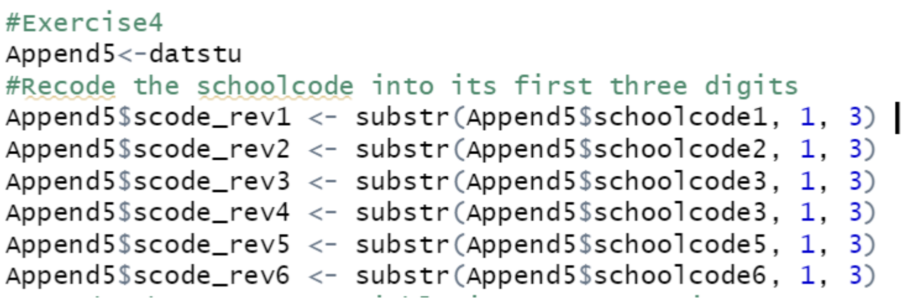


```
#Exercise3
Dat<-na.omit(datstu)
Dat<-select(Dat,5:18)
Dat$schoolcode<-apply(Dat,1,function(x) return(x[as.numeric(x[14])]))
Dat$schoolcode<-as.numeric(Dat$schoolcode)
schdis<-na.omit(datsss)
datjsss<- fread("datjss.csv")
datjsss<-rename(datjsss,jsslong = point_x, jsslat = point_y)
Bigdata<-left_join(Dat,schdis,by="schoolcode")
BigData<-left_join(Bigdata,datjsss,by="jssdistrict")
BigDataedited<-select(BigData,ssslong, jsslong, jsslat, ssslat)
BigDataedited<-unique(BigDataedited)
BigDataedited<-na.omit(BigDataedited)
BigDataedited <- mutate(BigDataedited,dist = sqrt( (69.172*(ssslong - jsslong) * cos(jsslat/57.3)) ^2  +  (69.172 * (ssslat
```

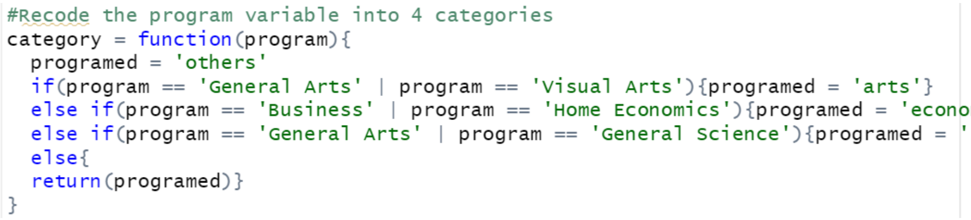


| | ssslong | jsslong | jsslat | ssslat | dist |
|---|---|---|---|---|---|
| 1 | -1.1970884 | -0.75524253 | 5.617353 | 5.130001 | 45.404991 |
| 2 | -1.1970884 | -1.19708836 | 5.130001 | 5.130001 | 0.000000 |
| 3 | -1.1970884 | -1.00538456 | 5.401725 | 5.130001 | 22.968725 |
| 4 | -1.1970884 | -1.55970335 | 5.572999 | 5.130001 | 39.524867 |
| 5 | -1.1970884 | -0.50863892 | 5.544896 | 5.130001 | 55.410029 |
| 6 | -1.1970884 | -1.00645339 | 5.201528 | 5.130001 | 14.033432 |
| 7 | -1.1970884 | -2.63174391 | 7.503565 | 5.130001 | 191.407178 |
| 8 | -1.1970884 | -1.26436615 | 5.495795 | 5.130001 | 25.723259 |
| 9 | -1.1970884 | -0.39751053 | 5.664688 | 5.130001 | 66.310912 |
| 10 | -1.1970884 | -1.30659389 | 5.153656 | 5.130001 | 7.719514 |
| 11 | -1.1970884 | -1.01707423 | 5.638250 | 5.130001 | 37.276595 |
| 12 | -1.1970884 | -0.19711526 | 5.607396 | 5.130001 | 76.349930 |
| 13 | -1.1970884 | -1.37461698 | 5.777995 | 5.130001 | 46.458354 |
| 14 | -1.1970884 | 0.10686218 | 5.914734 | 5.130001 | 104.859841 |
| 15 | -1.1970884 | -1.56275165 | 6.559323 | 5.130001 | 102.012337 |
| 16 | -1.1970884 | -0.72363549 | 5.404561 | 5.130001 | 37.732218 |
| 17 | -0.6355287 | -0.63552868 | 6.619226 | 6.619226 | 0.000000 |
| 18 | -0.6355287 | -0.19711526 | 5.607396 | 6.619226 | 76.220297 |
| 19 | -0.6355287 | -0.47498974 | 5.944515 | 6.619226 | 47.960275 |
| 20 | -0.6355287 | -0.24114588 | 5.721143 | 6.619226 | 67.793736 |

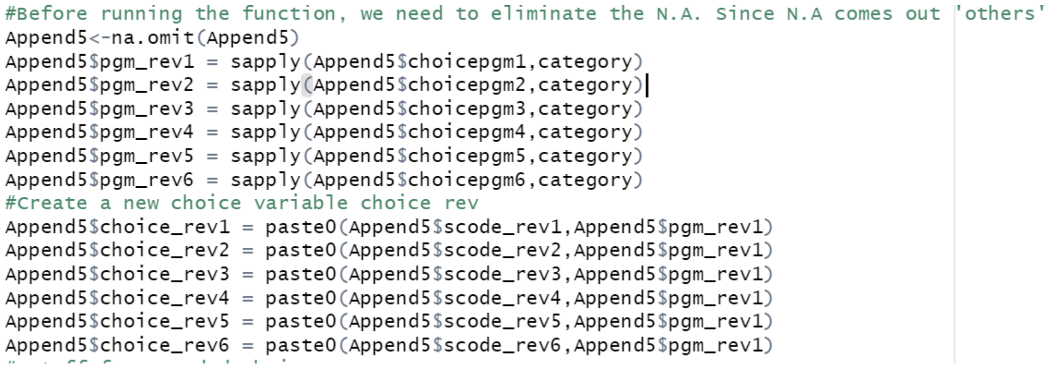
#Exercise4



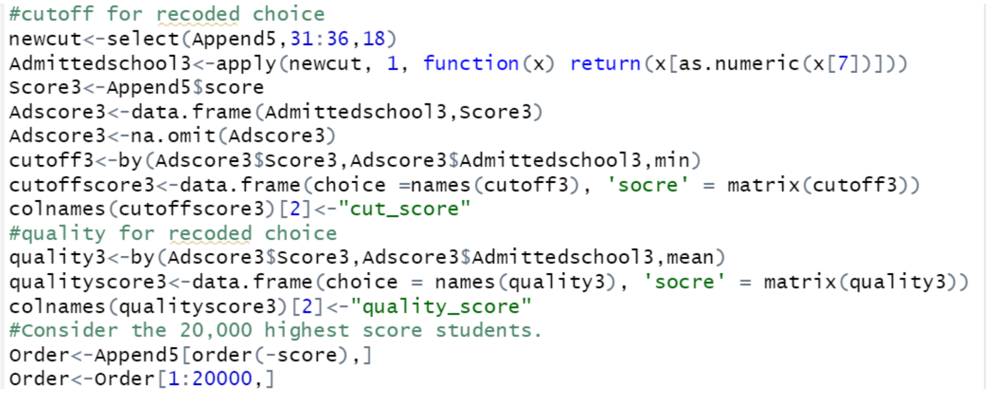
```
#Exercise4
Append5<-datstu
#Recode the schoolcode into its first three digits
Append5$scode_rev1 <- substr(Append5$schoolcode1, 1, 3)
Append5$scode_rev2 <- substr(Append5$schoolcode2, 1, 3)
Append5$scode_rev3 <- substr(Append5$schoolcode3, 1, 3)
Append5$scode_rev4 <- substr(Append5$schoolcode3, 1, 3)
Append5$scode_rev5 <- substr(Append5$schoolcode5, 1, 3)
Append5$scode_rev6 <- substr(Append5$schoolcode6, 1, 3)
```



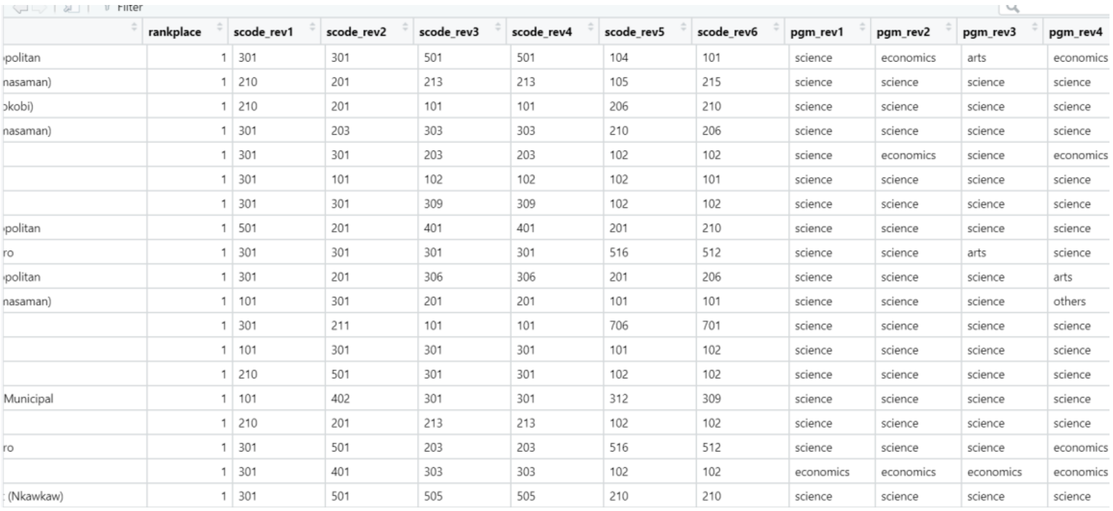
```
#Recode the program variable into 4 categories
category = function(program){
  programed = 'others'
  if(program == 'General Arts' | program == 'Visual Arts'){programed = 'arts'}
  else if(program == 'Business' | program == 'Home Economics'){programed = 'economics'}
  else if(program == 'General Arts' | program == 'General Science'){programed = 'science'}
  else{
  return(programed)}
}
```



```
#Before running the function, we need to eliminate the N.A. Since N.A comes out 'others'
Append5<-na.omit(Append5)
Append5$pgm_rev1 = sapply(Append5$choicepgm1,category)
Append5$pgm_rev2 = sapply(Append5$choicepgm2,category)
Append5$pgm_rev3 = sapply(Append5$choicepgm3,category)
Append5$pgm_rev4 = sapply(Append5$choicepgm4,category)
Append5$pgm_rev5 = sapply(Append5$choicepgm5,category)
Append5$pgm_rev6 = sapply(Append5$choicepgm6,category)
#Create a new choice variable choice rev
Append5$choice_rev1 = paste0(Append5$scode_rev1,Append5$pgm_rev1)
Append5$choice_rev2 = paste0(Append5$scode_rev2,Append5$pgm_rev1)
Append5$choice_rev3 = paste0(Append5$scode_rev3,Append5$pgm_rev1)
Append5$choice_rev4 = paste0(Append5$scode_rev4,Append5$pgm_rev1)
Append5$choice_rev5 = paste0(Append5$scode_rev5,Append5$pgm_rev1)
Append5$choice_rev6 = paste0(Append5$scode_rev6,Append5$pgm_rev1)
```



```
#cutoff for recoded choice
newcut<-select(Append5,31:36,18)
Admittedschool3<-apply(newcut, 1, function(x) return(x[as.numeric(x[7])]))
Score3<-Append5$score
Adscore3<-data.frame(Admittedschool3,Score3)
Adscore3<-na.omit(Adscore3)
cutoff3<-by(Adscore3$Score3,Adscore3$Admittedschool3,min)
cutoffscore3<-data.frame(choice =names(cutoff3), 'socre' = matrix(cutoff3))
colnames(cutoffscore3)[2]<-"cut_score"
#quality for recoded choice
quality3<-by(Adscore3$Score3,Adscore3$Admittedschool3,mean)
qualityscore3<-data.frame(choice = names(quality3), 'socre' = matrix(quality3))
colnames(qualityscore3)[2]<-"quality_score"
#Consider the 20,000 highest score students.
Order<-Append5[order(-score),]
Order<-Order[1:20000,]
```



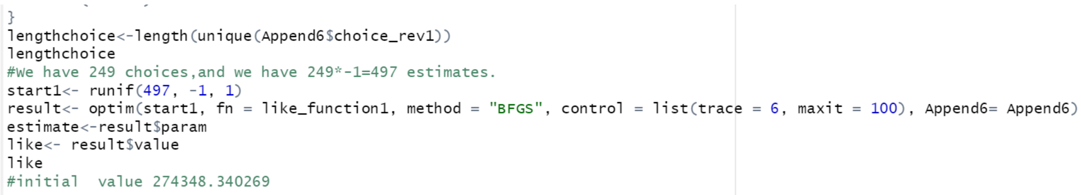
| | rankplace | scode_rev1 | scode_rev2 | scode_rev3 | scode_rev4 | scode_rev5 | scode_rev6 | pgm_rev1 | pgm_rev2 | pgm_rev3 | pgm_rev4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| politan | 1 | 301 | 301 | 501 | 501 | 104 | 101 | science | economics | arts | economics |
| nasaman) | 1 | 210 | 201 | 213 | 213 | 105 | 215 | science | science | science | science |
| okobi) | 1 | 210 | 201 | 101 | 101 | 206 | 210 | science | science | science | science |
| nasaman) | 1 | 301 | 203 | 303 | 303 | 210 | 206 | science | science | science | science |
| | 1 | 301 | 301 | 203 | 203 | 102 | 102 | science | economics | science | economics |
| | 1 | 301 | 101 | 102 | 102 | 102 | 101 | science | science | science | science |
| | 1 | 301 | 301 | 309 | 309 | 102 | 102 | science | science | science | science |
| politan | 1 | 501 | 201 | 401 | 401 | 201 | 210 | science | science | science | science |
| ro | 1 | 301 | 301 | 301 | 301 | 516 | 512 | science | science | arts | science |
| politan | 1 | 301 | 201 | 306 | 306 | 201 | 206 | science | science | science | arts |
| nasaman) | 1 | 101 | 301 | 201 | 201 | 101 | 101 | science | science | science | others |
| | 1 | 301 | 211 | 101 | 101 | 706 | 701 | science | science | science | science |
| | 1 | 101 | 301 | 301 | 301 | 101 | 102 | science | science | science | science |
| | 1 | 210 | 501 | 301 | 301 | 102 | 102 | science | science | science | science |
| Municipal | 1 | 101 | 402 | 301 | 301 | 312 | 309 | science | science | science | science |
| | 1 | 210 | 201 | 213 | 213 | 102 | 102 | science | science | science | science |
| ro | 1 | 301 | 501 | 203 | 203 | 516 | 512 | science | science | science | economics |
| | 1 | 301 | 401 | 303 | 303 | 102 | 102 | economics | economics | economics | economics |
| (Nkawkaw) | 1 | 301 | 501 | 505 | 505 | 210 | 210 | science | science | science | science |

#Exercise5

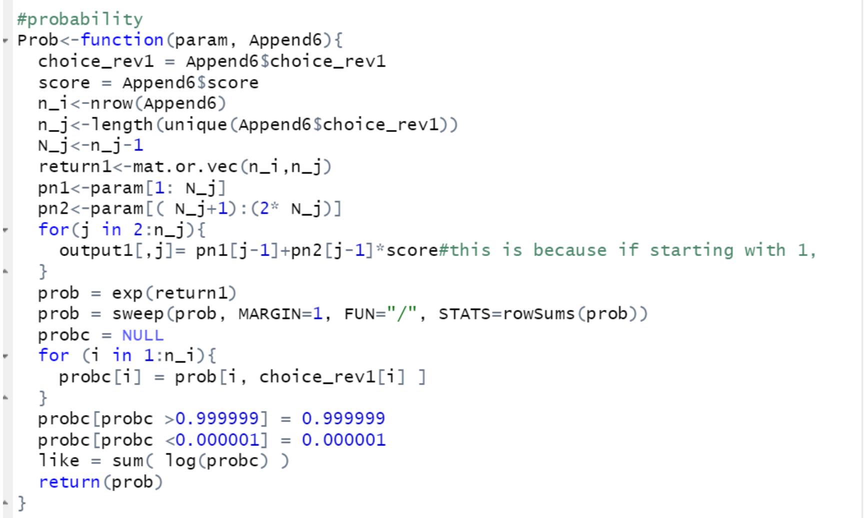
```
#Propose a model specification.
#We use the recoded choices and the 20,000 highest score students.
Append6<-Order
#Write the likelihood function
Append6$choice_rev1 <- as.numeric( as.factor(Append6$choice_rev1) )

like_function1<-function(param, Append6){
  choice_rev1 = Append6$choice_rev1
  score = Append6$score
  n_i<-nrow(Append6)
  n_j<-length(unique(Append6$choice_rev1))
  N_j<-n_j-1
  return1<-mat.or.vec(n_i,n_j)
  pn1<-param[1: N_j]
  pn2<-param[( N_j+1):(2* N_j)]
  for(j in 2:n_j){
    return1[,j]= pn1[j-1]+pn2[j-1]*score
  }
  prob = exp(return1)
  prob = sweep(prob, MARGIN=1, FUN="/", STATS=rowSums(prob))

  probc = NULL
  for (i in 1:n_i){
    probc[i] = prob[i, choice_rev1[i] ]
  }
  probc[probc >0.999999] = 0.999999
  probc[probc <0.000001] = 0.000001
  like = sum( log(probc) )
  return(- like)
```

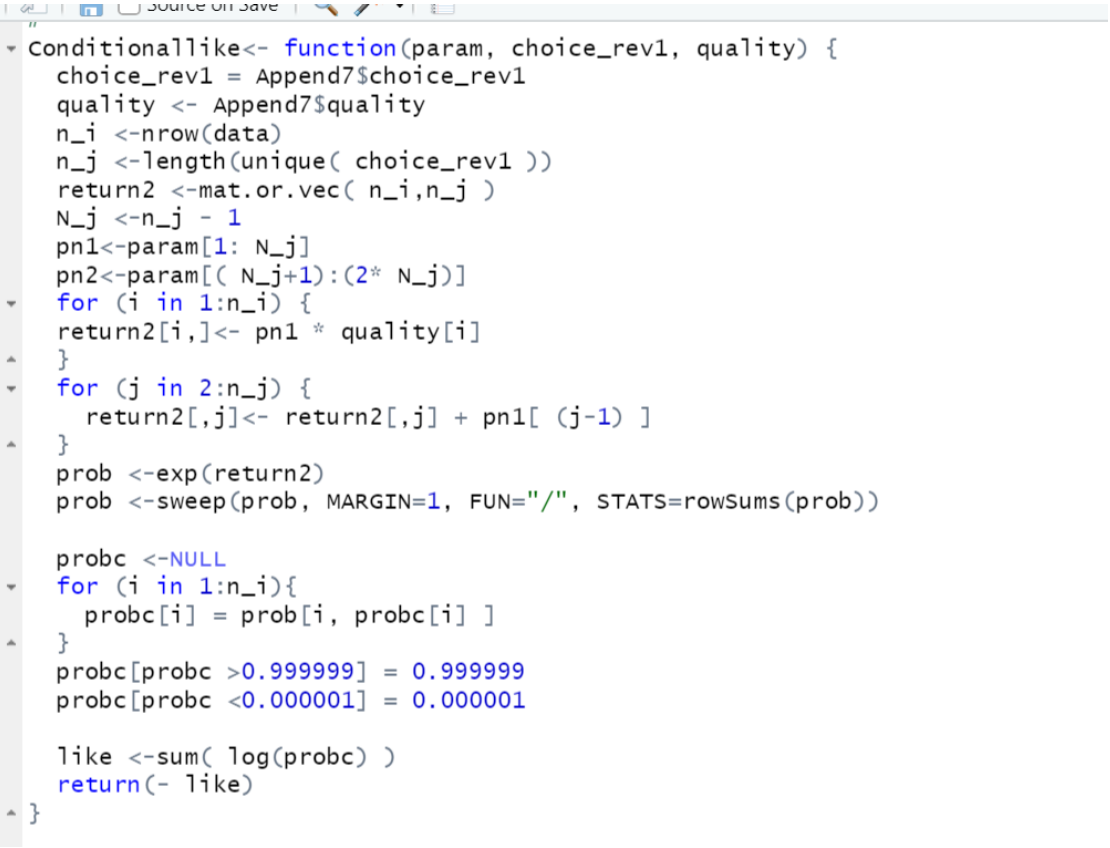


```
}
lengthchoice<-length(unique(Append6$choice_rev1))
lengthchoice
#We have 249 choices,and we have 249*-1=497 estimates.
start1<- runif(497, -1, 1)
result<- optim(start1, fn = like_function1, method = "BFGS", control = list(trace = 6, maxit = 100), Append6= Append6)
estimate<-result$param
like<- result$value
like
#initial  value 274348.340269
```



```
#probability
Prob<-function(param, Append6){
  choice_rev1 = Append6$choice_rev1
  score = Append6$score
  n_i<-nrow(Append6)
  n_j<-length(unique(Append6$choice_rev1))
  N_j<-n_j-1
  return1<-mat.or.vec(n_i,n_j)
  pn1<-param[1: N_j]
  pn2<-param[( N_j+1):(2* N_j)]
  for(j in 2:n_j){
    output1[,j]= pn1[j-1]+pn2[j-1]*score#this is because if starting with 1,
  }
  prob = exp(return1)
  prob = sweep(prob, MARGIN=1, FUN="/", STATS=rowSums(prob))
  probc = NULL
  for (i in 1:n_i){
    probc[i] = prob[i, choice_rev1[i] ]
  }
  probc[probc >0.999999] = 0.999999
  probc[probc <0.000001] = 0.000001
  like = sum( log(probc) )
  return(prob)
}
```

#Exercise6

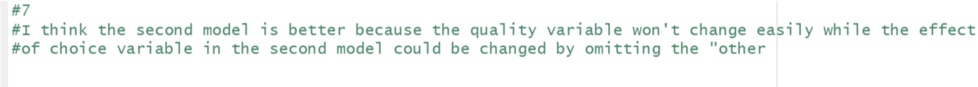


```
Conditionallike<- function(param, choice_rev1, quality) {
  choice_rev1 = Append7$choice_rev1
  quality <- Append7$quality
  n_i <-nrow(data)
  n_j <-length(unique( choice_rev1 ))
  return2 <-mat.or.vec( n_i,n_j )
  N_j <-n_j - 1
  pn1<-param[1: N_j]
  pn2<-param[( N_j+1):(2* N_j)]
  for (i in 1:n_i) {
  return2[i,]<- pn1 * quality[i]
  }
  for (j in 2:n_j) {
    return2[,j]<- return2[,j] + pn1[ (j-1) ]
  }
  prob <-exp(return2)
  prob <-sweep(prob, MARGIN=1, FUN="/", STATS=rowSums(prob))

  probc <-NULL
  for (i in 1:n_i){
    probc[i] = prob[i, probc[i] ]
  }
  probc[probc >0.999999] = 0.999999
  probc[probc <0.000001] = 0.000001

  like <-sum( log(probc) )
  return(- like)
}
```

#Exercise7



```
#7
#I think the second model is better because the quality variable won't change easily while the effect
#of choice variable in the second model could be changed by omitting the "other
```