Analisis Exploratorio: práctica con R

Patricia Sanchez Holgado

10 de Abril de 2018

#### ENUNCIADO: En esta práctica se pretende llegar sólo hasta el Exploratorio, pero pensando en una finalidad posterior (Modelo predictivo, Clasificación, ….). Aunque esta finalidad última no se incluya en la práctica, se debe plantear la “hipótesis” en la introducción. Con este dataset se podría tratar de predecir el valor de la nota final G3, o clasificar a los alumnos como aprobados o suspendidos en base a que esa nota sea mayor o menor que 5, pero se puede pensar otro objetivo si así se quiere).

La base de datos a utilizar se encuentra en: <https://archive.ics.uci.edu/ml/machine-learning-databases/00320/student.zip>

# INDICE

*1. Introduccion* 2. Carga de los Datos *3. Analisis Descriptivo* 4. Analisis Exploratorio \*5. Conclusiones

# 1. Introducción

El objetivo de este análisis es realizar un análisis descriptivo y exploratorio de los datos de los que disponemos sobre notas de alumnos de matemáticas y portugués, para obtener alguna conclusión. La Hipótesis con la que vamos a trabajar será predecir el valor de la nota final (Variable G3) basándonos en que dicha nota sea mayor o menor que 9. Para ello estudiaremos cómo afectan algunas de las variables a esta nota final y así predecir el comportamiento futuro con esta información.

Establecemos el directorio de trabajo

#Directorio de trabajo  
#ls()  
#getwd()  
#setwd("/Users/Patricia/Desktop/Exploratorio")  
#setwd("Exploratorio")  
#getwd()  
  
#if (!file.exists('./Exploratorio')){  
# dir.create('./Exploratorio')  
#}  
#list.files('./Exploratorio')  
  
#Cargamos las librerías que vamos a utilizar  
library(readr)  
library(knitr)  
library(caret)

## Loading required package: lattice

## Loading required package: ggplot2

## Warning in as.POSIXlt.POSIXct(Sys.time()): unknown timezone 'default/  
## Europe/Madrid'

library(ggplot2)  
library(gplots)

##   
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':  
##   
## lowess

library(corrplot)

## corrplot 0.84 loaded

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(cluster)  
library(Matrix)  
library(Hmisc)

## Loading required package: survival

##   
## Attaching package: 'survival'

## The following object is masked from 'package:caret':  
##   
## cluster

## Loading required package: Formula

##   
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:dplyr':  
##   
## src, summarize

## The following objects are masked from 'package:base':  
##   
## format.pval, units

library(lattice)  
library(gridExtra)

##   
## Attaching package: 'gridExtra'

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

# 2. Carga de los datos

Cargamos los dos datasets:

studentMat <- read.table ("student-mat.csv", row.names=NULL, sep=";", header=TRUE)  
studentPor <- read.table ("student-por.csv", row.names=NULL, sep=";", header=TRUE)

# 3. Análisis Descriptivo

Observamos que el dataset relativo a los estudiantes de matemáticas (student-mat), recoge datos de 395 estudiantes en base a 30 variables tales como género, edad, si viven en un entorno rural o urbano, información familiar, etc y 3 variables con la nota del primer (G1) y segundo ciclo (G2) y la nota final (G3). El dataset de los estudiantes de portugués (student-por) tiene datos de 649 estudiantes sobre las mismas variables. Por tanto en ambos la variable relativa a la nota final es ‘G3’, con un valor entre 0 y 20. Comprobamos los datos que presentan los estudiantes de matemáticas y portugués relacionados con esta nota final:

any(studentMat$G3 == 20)

## [1] TRUE

all(studentMat$G3 > 0)

## [1] FALSE

any(studentPor$G3 == 20)

## [1] FALSE

all(studentPor$G3 > 0)

## [1] FALSE

Los resultados nos indican que hay alumnos de matemáticas con una nota máxima de 20, pero no hay ninguno de Portugués, y además en todos los casos hay algún alumno con nota 0.

Vamos a cruzar ambas tablas para centrarnos en los que son alumnos comunes.

studentMatPor=merge(studentMat,studentPor,by=c("school","sex","age","address","famsize","Pstatus","Medu","Fedu","Mjob","Fjob","reason","nursery","internet"))  
print(nrow(studentMatPor)) # 382 students

## [1] 382

View(studentMatPor)

Obtenemos 382 registros.

Vamos a revisar el resumen de datos de la tabla conjunta StudentMatPor: summary, names, nrow, ncol, class

summary(studentMatPor)

## school sex age address famsize Pstatus  
## GP:342 F:198 Min. :15.00 R: 81 GT3:278 A: 38   
## MS: 40 M:184 1st Qu.:16.00 U:301 LE3:104 T:344   
## Median :17.00   
## Mean :16.59   
## 3rd Qu.:17.00   
## Max. :22.00   
## Medu Fedu Mjob Fjob   
## Min. :0.000 Min. :0.000 at\_home : 53 at\_home : 16   
## 1st Qu.:2.000 1st Qu.:2.000 health : 33 health : 17   
## Median :3.000 Median :3.000 other :138 other :211   
## Mean :2.806 Mean :2.565 services: 96 services:107   
## 3rd Qu.:4.000 3rd Qu.:4.000 teacher : 62 teacher : 31   
## Max. :4.000 Max. :4.000   
## reason nursery internet guardian.x traveltime.x   
## course :140 no : 72 no : 58 father: 91 Min. :1.000   
## home :110 yes:310 yes:324 mother:275 1st Qu.:1.000   
## other : 34 other : 16 Median :1.000   
## reputation: 98 Mean :1.442   
## 3rd Qu.:2.000   
## Max. :4.000   
## studytime.x failures.x schoolsup.x famsup.x paid.x   
## Min. :1.000 Min. :0.0000 no :331 no :144 no :205   
## 1st Qu.:1.000 1st Qu.:0.0000 yes: 51 yes:238 yes:177   
## Median :2.000 Median :0.0000   
## Mean :2.034 Mean :0.2906   
## 3rd Qu.:2.000 3rd Qu.:0.0000   
## Max. :4.000 Max. :3.0000   
## activities.x higher.x romantic.x famrel.x freetime.x   
## no :181 no : 18 no :261 Min. :1.00 Min. :1.000   
## yes:201 yes:364 yes:121 1st Qu.:4.00 1st Qu.:3.000   
## Median :4.00 Median :3.000   
## Mean :3.94 Mean :3.223   
## 3rd Qu.:5.00 3rd Qu.:4.000   
## Max. :5.00 Max. :5.000   
## goout.x Dalc.x Walc.x health.x   
## Min. :1.000 Min. :1.000 Min. :1.00 Min. :1.000   
## 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:3.000   
## Median :3.000 Median :1.000 Median :2.00 Median :4.000   
## Mean :3.113 Mean :1.474 Mean :2.28 Mean :3.579   
## 3rd Qu.:4.000 3rd Qu.:2.000 3rd Qu.:3.00 3rd Qu.:5.000   
## Max. :5.000 Max. :5.000 Max. :5.00 Max. :5.000   
## absences.x G1.x G2.x G3.x   
## Min. : 0.000 Min. : 3.00 Min. : 0.00 Min. : 0.00   
## 1st Qu.: 0.000 1st Qu.: 8.00 1st Qu.: 8.25 1st Qu.: 8.00   
## Median : 3.000 Median :10.50 Median :11.00 Median :11.00   
## Mean : 5.319 Mean :10.86 Mean :10.71 Mean :10.39   
## 3rd Qu.: 8.000 3rd Qu.:13.00 3rd Qu.:13.00 3rd Qu.:14.00   
## Max. :75.000 Max. :19.00 Max. :19.00 Max. :20.00   
## guardian.y traveltime.y studytime.y failures.y schoolsup.y  
## father: 91 Min. :1.000 Min. :1.000 Min. :0.0000 no :332   
## mother:275 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:0.0000 yes: 50   
## other : 16 Median :1.000 Median :2.000 Median :0.0000   
## Mean :1.445 Mean :2.039 Mean :0.1414   
## 3rd Qu.:2.000 3rd Qu.:2.000 3rd Qu.:0.0000   
## Max. :4.000 Max. :4.000 Max. :3.0000   
## famsup.y paid.y activities.y higher.y romantic.y famrel.y   
## no :143 no :356 no :182 no : 18 no :259 Min. :1.000   
## yes:239 yes: 26 yes:200 yes:364 yes:123 1st Qu.:4.000   
## Median :4.000   
## Mean :3.942   
## 3rd Qu.:5.000   
## Max. :5.000   
## freetime.y goout.y Dalc.y Walc.y   
## Min. :1.00 Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:3.00 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.000   
## Median :3.00 Median :3.000 Median :1.000 Median :2.000   
## Mean :3.23 Mean :3.118 Mean :1.476 Mean :2.291   
## 3rd Qu.:4.00 3rd Qu.:4.000 3rd Qu.:2.000 3rd Qu.:3.000   
## Max. :5.00 Max. :5.000 Max. :5.000 Max. :5.000   
## health.y absences.y G1.y G2.y   
## Min. :1.000 Min. : 0.000 Min. : 0.00 Min. : 5.00   
## 1st Qu.:3.000 1st Qu.: 0.000 1st Qu.:10.00 1st Qu.:11.00   
## Median :4.000 Median : 2.000 Median :12.00 Median :12.00   
## Mean :3.576 Mean : 3.673 Mean :12.11 Mean :12.24   
## 3rd Qu.:5.000 3rd Qu.: 6.000 3rd Qu.:14.00 3rd Qu.:14.00   
## Max. :5.000 Max. :32.000 Max. :19.00 Max. :19.00   
## G3.y   
## Min. : 0.00   
## 1st Qu.:11.00   
## Median :13.00   
## Mean :12.52   
## 3rd Qu.:14.00   
## Max. :19.00

names(studentMatPor)

## [1] "school" "sex" "age" "address"   
## [5] "famsize" "Pstatus" "Medu" "Fedu"   
## [9] "Mjob" "Fjob" "reason" "nursery"   
## [13] "internet" "guardian.x" "traveltime.x" "studytime.x"   
## [17] "failures.x" "schoolsup.x" "famsup.x" "paid.x"   
## [21] "activities.x" "higher.x" "romantic.x" "famrel.x"   
## [25] "freetime.x" "goout.x" "Dalc.x" "Walc.x"   
## [29] "health.x" "absences.x" "G1.x" "G2.x"   
## [33] "G3.x" "guardian.y" "traveltime.y" "studytime.y"   
## [37] "failures.y" "schoolsup.y" "famsup.y" "paid.y"   
## [41] "activities.y" "higher.y" "romantic.y" "famrel.y"   
## [45] "freetime.y" "goout.y" "Dalc.y" "Walc.y"   
## [49] "health.y" "absences.y" "G1.y" "G2.y"   
## [53] "G3.y"

nrow(studentMatPor)

## [1] 382

ncol(studentMatPor)

## [1] 53

class(studentMatPor)

## [1] "data.frame"

En la tabla cruzada hay 342 alumnos de la Escuela GP y 40 de la Escuela MS. En la variable G3 tenemos las columnas G3.x y G3.y que nos indican la nota final del período de grado, con medias de 10.39 y 12.52 respectivamente.

Podemos calcular la media, mediana o quantiles

mean(studentMatPor$G3.x)

## [1] 10.38743

median(studentMatPor$G3.x)

## [1] 11

quantile(studentMatPor$G3.x)

## 0% 25% 50% 75% 100%   
## 0 8 11 14 20

mean(studentMatPor$G3.y)

## [1] 12.51571

median(studentMatPor$G3.y)

## [1] 13

quantile(studentMatPor$G3.y)

## 0% 25% 50% 75% 100%   
## 0 11 13 14 19

Ahora observamos las frecuencias.

tabla<-table(studentMatPor$G3.x)  
frecrelativa<-(tabla)/margin.table(tabla)  
frecrelativa

##   
## 0 4 5 6 7 8   
## 0.102094241 0.002617801 0.018324607 0.039267016 0.018324607 0.081151832   
## 9 10 11 12 13 14   
## 0.070680628 0.146596859 0.112565445 0.078534031 0.065445026 0.070680628   
## 15 16 17 18 19 20   
## 0.083769634 0.044502618 0.015706806 0.034031414 0.013089005 0.002617801

frecabsol<-cumsum(tabla)  
frecabsol

## 0 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20   
## 39 40 47 62 69 100 127 183 226 256 281 308 340 357 363 376 381 382

Matriz de correlaciones. Queremos que todos los datos a manejar sean numéricos, por lo que convertimos los que no lo son. En este caso hemos ido convirtiendo cada una de las columnas en valor numérico y al finalizar usamos sapply.

schoolnum<-as.numeric(studentMatPor$school)  
sexnum<-as.numeric(studentMatPor$sex)  
adressnum<-as.numeric(studentMatPor$adress)  
famsizenum<-as.numeric(studentMatPor$famsize)  
statusnum<-as.numeric(studentMatPor$Pstatus)  
medunum<-as.numeric(studentMatPor$Medu)  
fedunum<-as.numeric(studentMatPor$Fedu)  
mjobnum<-as.numeric(studentMatPor$Mjob)  
fjobnum<-as.numeric(studentMatPor$Fjob)  
reasonnum<-as.numeric(studentMatPor$reason)  
nurserynum<-as.numeric(studentMatPor$nursery)  
internetnum<-as.numeric(studentMatPor$internet)  
guardianXnum<-as.numeric(studentMatPor$guardian.x)  
guardianYnum<-as.numeric(studentMatPor$guardian.y)  
schoolsupXnum<-as.numeric(studentMatPor$schoolsup.x)  
schoolsupYnum<-as.numeric(studentMatPor$schoolsup.y)  
famsupXnum<-as.numeric(studentMatPor$famsup.x)  
famsupYnum<-as.numeric(studentMatPor$famsup.y)  
paidXnum<-as.numeric(studentMatPor$paid.x)  
paidYnum<-as.numeric(studentMatPor$paid.y)  
activitiesXnum<-as.numeric(studentMatPor$activities.x)  
activitiesYnum<-as.numeric(studentMatPor$activities.y)  
higherXnum<-as.numeric(studentMatPor$higher.x)  
higherYnum<-as.numeric(studentMatPor$higher.y)  
romanticXnum<-as.numeric(studentMatPor$romantic.x)  
romanticYnum<-as.numeric(studentMatPor$romantic.y)  
  
studentMatPornum<-sapply(studentMatPor,as.numeric)  
is.numeric(studentMatPornum)

## [1] TRUE

correlacion<-cor(studentMatPornum)  
correlacion

## school sex age address  
## school 1.0000000000 -0.038791211 0.3541556502 -0.282759645  
## sex -0.0387912109 1.000000000 -0.0174138673 -0.012615968  
## age 0.3541556502 -0.017413867 1.0000000000 -0.155760079  
## address -0.2827596450 -0.012615968 -0.1557600794 1.000000000  
## famsize 0.0213198301 0.093052141 0.0101154559 0.058305045  
## Pstatus 0.0850997323 0.022820062 0.0543525484 -0.065425352  
## Medu -0.1280655955 0.104515999 -0.1206626661 0.137588840  
## Fedu -0.0751097355 0.047653975 -0.1278455189 0.074882999  
## Mjob -0.0452943978 0.209756583 -0.0593311060 0.116076906  
## Fjob 0.0245636224 0.081604451 -0.0009660828 -0.011752038  
## reason -0.0603107084 -0.080505945 -0.0276358132 -0.058541216  
## nursery -0.0756556403 -0.004278300 -0.0330130598 0.077504260  
## internet -0.0935538594 0.072079818 -0.0808544753 0.208831142  
## guardian.x -0.0199460848 -0.051959477 0.1498863240 -0.050860249  
## traveltime.x 0.2499650999 0.057311335 0.1026076175 -0.342732627  
## studytime.x -0.0846313055 -0.293137317 -0.0122254170 -0.024591483  
## failures.x 0.0044239635 0.054180296 0.1530331389 -0.048027368  
## schoolsup.x -0.1342418825 -0.147342472 -0.2358099882 0.034159255  
## famsup.x -0.1573935072 -0.158254450 -0.1224218501 0.006157785  
## paid.x -0.0091563747 -0.128773452 -0.0259391275 0.032511511  
## activities.x -0.1035455919 0.106849078 -0.0664954420 -0.056173679  
## higher.x -0.0046475352 -0.156504866 -0.2260668090 0.035762324  
## romantic.x 0.0611971397 -0.070754945 0.1250673497 -0.004720985  
## famrel.x -0.0890976583 0.068752732 0.0254506351 0.021730442  
## freetime.x 0.0008617613 0.217952450 0.0048821087 0.039087446  
## goout.x -0.0264911266 0.061583940 0.1596309469 0.074316630  
## Dalc.x 0.0777357158 0.247532348 0.1283708429 -0.105793067  
## Walc.x 0.0386775667 0.255419181 0.1521317124 -0.096533622  
## health.x -0.0619987218 0.140667879 -0.0632363278 -0.018952752  
## absences.x -0.0772150172 -0.063819717 0.1259715040 -0.037954493  
## G1.x -0.0318249561 0.124570520 -0.1174857027 0.062729830  
## G2.x -0.0591537991 0.116326160 -0.1753697694 0.121602321  
## G3.x -0.0666607975 0.121671254 -0.1945885369 0.092187335  
## guardian.y -0.0199460848 -0.051959477 0.1498863240 -0.050860249  
## traveltime.y 0.2595036851 0.053372402 0.1065453572 -0.338840020  
## studytime.y -0.0867744019 -0.286784720 -0.0126857245 -0.021381883  
## failures.y 0.0724834715 0.132766712 0.2585726791 -0.044351456  
## schoolsup.y -0.1327189396 -0.141102245 -0.2471595795 0.030420869  
## famsup.y -0.1594620119 -0.152863031 -0.1252984193 0.008972367  
## paid.y 0.0433675712 0.113922236 -0.0553978068 -0.088667976  
## activities.y -0.1017260349 0.111873912 -0.0683279508 -0.058878012  
## higher.y -0.0046475352 -0.156504866 -0.2260668090 0.035762324  
## romantic.y 0.0753978985 -0.081249770 0.1332598775 0.001112330  
## famrel.y -0.0724939010 0.066935946 0.0292860247 0.023531026  
## freetime.y 0.0242061586 0.210951259 0.0167970594 0.043352949  
## goout.y -0.0053768706 0.061656662 0.1669305260 0.076598533  
## Dalc.y 0.0863770783 0.244660760 0.1319275711 -0.104250072  
## Walc.y 0.0425654024 0.251686391 0.1567950971 -0.092314651  
## health.y -0.0611886875 0.142077859 -0.0669056310 -0.019868603  
## absences.y -0.0609205612 -0.009399842 0.0863029290 -0.016345975  
## G1.y -0.1657689550 -0.175880794 -0.1095486588 0.175881683  
## G2.y -0.1162890697 -0.190921913 -0.0420100300 0.185226128  
## G3.y -0.1820315528 -0.201063602 -0.0679774771 0.188918461  
## famsize Pstatus Medu Fedu  
## school 0.021319830 0.085099732 -0.128065596 -0.0751097355  
## sex 0.093052141 0.022820062 0.104515999 0.0476539748  
## age 0.010115456 0.054352548 -0.120662666 -0.1278455189  
## address 0.058305045 -0.065425352 0.137588840 0.0748829993  
## famsize 1.000000000 -0.150406734 -0.015467184 -0.0473060894  
## Pstatus -0.150406734 1.000000000 -0.131896360 -0.0839890780  
## Medu -0.015467184 -0.131896360 1.000000000 0.6497974864  
## Fedu -0.047306089 -0.083989078 0.649797486 1.0000000000  
## Mjob 0.067044329 -0.049629443 0.462829140 0.2661792267  
## Fjob -0.087934401 0.030316171 0.168693855 0.2120436753  
## reason -0.012324527 -0.029640799 0.105939081 0.0282514233  
## nursery 0.114316903 -0.070721069 0.179234430 0.1327446617  
## internet 0.045733205 0.029987505 0.193399774 0.1252385617  
## guardian.x -0.006954093 -0.114974290 0.007222486 -0.1147751613  
## traveltime.x 0.059191156 0.035409344 -0.205895329 -0.1878555999  
## studytime.x -0.059454409 0.034099562 0.050040187 -0.0038237458  
## failures.x -0.034065758 0.000502853 -0.230169469 -0.2453863636  
## schoolsup.x -0.032590749 -0.049545899 -0.064699791 0.0152013127  
## famsup.x -0.094605395 0.012189305 0.149882846 0.1698463262  
## paid.x 0.009571038 0.063272448 0.127198054 0.0763203133  
## activities.x -0.020288480 0.087492341 0.110838272 0.1165858280  
## higher.x -0.002760944 -0.032632397 0.142483563 0.1599896324  
## romantic.x 0.038653679 -0.055713760 0.038586461 0.0235464391  
## famrel.x -0.023885977 0.044777038 0.011913172 0.0130029499  
## freetime.x 0.028952451 0.057209543 0.035365607 -0.0074204872  
## goout.x 0.017132823 0.017621205 0.066870221 0.0331786948  
## Dalc.x 0.091183810 -0.019712592 0.043789396 0.0179777552  
## Walc.x 0.091206814 0.031703497 -0.013695586 0.0009282785  
## health.x -0.055372622 0.049933843 -0.041731304 0.0154503060  
## absences.x 0.057755381 -0.159489944 0.108559715 0.0336018123  
## G1.x 0.093944845 -0.013786609 0.229202048 0.2158709852  
## G2.x 0.109013214 -0.036430333 0.217917569 0.1800415572  
## G3.x 0.110190986 -0.054703437 0.205489379 0.1554440470  
## guardian.y -0.006954093 -0.114974290 0.007222486 -0.1147751613  
## traveltime.y 0.056562213 0.036453780 -0.210964862 -0.1921470477  
## studytime.y -0.063262509 0.036169359 0.048303417 -0.0041953922  
## failures.y -0.042470198 0.057533234 -0.204947316 -0.2124093063  
## schoolsup.y -0.028118205 -0.052543125 -0.052323426 0.0264274996  
## famsup.y -0.098045259 0.014004052 0.156042447 0.1771700557  
## paid.y -0.025185262 -0.014363967 0.105733533 0.1072679281  
## activities.y -0.017077331 0.085727448 0.105042500 0.1096874010  
## higher.y -0.002760944 -0.032632397 0.142483563 0.1599896324  
## romantic.y 0.031632224 -0.051747630 0.035255669 0.0227743826  
## famrel.y -0.025984651 0.046363016 0.007278681 0.0090613454  
## freetime.y 0.024162147 0.060044880 0.014830749 -0.0261498527  
## goout.y 0.014277559 0.019130391 0.046280313 0.0138432539  
## Dalc.y 0.089367319 -0.018727993 0.038862000 0.0137473847  
## Walc.y 0.086227305 0.034427681 -0.016006443 0.0004398100  
## health.y -0.054077601 0.049175188 -0.033346657 0.0231916341  
## absences.y -0.011965731 -0.059685556 0.023039780 0.0208293469  
## G1.y 0.090509954 -0.019603511 0.205381472 0.1579781054  
## G2.y 0.060180854 -0.060131537 0.215947891 0.1848249262  
## G3.y 0.042718250 -0.078506185 0.203553661 0.1720085869  
## Mjob Fjob reason nursery  
## school -0.045294398 2.456362e-02 -6.031071e-02 -0.0756556403  
## sex 0.209756583 8.160445e-02 -8.050594e-02 -0.0042782999  
## age -0.059331106 -9.660828e-04 -2.763581e-02 -0.0330130598  
## address 0.116076906 -1.175204e-02 -5.854122e-02 0.0775042600  
## famsize 0.067044329 -8.793440e-02 -1.232453e-02 0.1143169027  
## Pstatus -0.049629443 3.031617e-02 -2.964080e-02 -0.0707210695  
## Medu 0.462829140 1.686939e-01 1.059391e-01 0.1792344301  
## Fedu 0.266179227 2.120437e-01 2.825142e-02 0.1327446617  
## Mjob 1.000000000 2.261218e-01 -5.525017e-03 0.0889940665  
## Fjob 0.226121827 1.000000e+00 -2.398571e-02 -0.0188168451  
## reason -0.005525017 -2.398571e-02 1.000000e+00 0.0446337681  
## nursery 0.088994067 -1.881685e-02 4.463377e-02 1.0000000000  
## internet 0.246233588 5.354420e-02 1.627844e-02 -0.0173826666  
## guardian.x 0.003931720 -7.819078e-02 -2.377146e-02 -0.0290917377  
## traveltime.x -0.107325194 3.519100e-02 -5.939421e-02 -0.0399661862  
## studytime.x -0.022180544 -8.076594e-02 1.503527e-01 0.0828099676  
## failures.x -0.089683745 5.550276e-04 -6.665252e-02 -0.0466602141  
## schoolsup.x -0.042865903 -2.743843e-02 1.923345e-02 0.0514143514  
## famsup.x 0.042092732 -4.311966e-02 7.656273e-02 0.0532942879  
## paid.x 0.066372370 -5.329132e-02 1.375859e-01 0.0853877588  
## activities.x 0.104469420 1.768633e-02 9.941004e-02 0.0118613000  
## higher.x 0.038539624 -4.877490e-02 6.456107e-02 0.0191846529  
## romantic.x -0.067411568 -6.953438e-05 1.174290e-02 0.0259889693  
## famrel.x 0.046206617 2.425021e-02 -3.765670e-03 0.0121077625  
## freetime.x 0.114849592 -5.227808e-02 -4.447944e-02 -0.0134221203  
## goout.x -0.005901141 4.077695e-03 -3.515777e-02 -0.0171433383  
## Dalc.x 0.069188837 7.728229e-02 -1.893189e-02 -0.0898811344  
## Walc.x 0.002190278 5.880396e-02 -8.913106e-03 -0.1245115758  
## health.x 0.069058302 1.010788e-02 -1.693863e-01 -0.0351565896  
## absences.x 0.045834041 2.461388e-02 8.239899e-02 0.0237272548  
## G1.x 0.108899615 8.372380e-02 9.665789e-02 0.0940762262  
## G2.x 0.114777862 7.954471e-02 1.236565e-01 0.0739153645  
## G3.x 0.111351932 4.917070e-02 1.330521e-01 0.0470368723  
## guardian.y 0.003931720 -7.819078e-02 -2.377146e-02 -0.0290917377  
## traveltime.y -0.107364674 3.360150e-02 -6.293453e-02 -0.0379326778  
## studytime.y -0.023261015 -7.942780e-02 1.465742e-01 0.0858201832  
## failures.y -0.085363670 -2.388649e-02 -8.434338e-02 -0.0499090963  
## schoolsup.y -0.041882051 -2.479243e-02 2.092678e-02 0.0481071742  
## famsup.y 0.045634858 -3.880143e-02 7.107244e-02 0.0559751705  
## paid.y -0.012854437 2.247703e-02 -9.683418e-02 0.0239331396  
## activities.y 0.096782537 1.343940e-02 1.048071e-01 0.0093321960  
## higher.y 0.038539624 -4.877490e-02 6.456107e-02 0.0191846529  
## romantic.y -0.059902343 2.388522e-03 9.825145e-05 0.0312763826  
## famrel.y 0.046354261 2.352088e-02 -9.217352e-03 0.0136673196  
## freetime.y 0.109482190 -5.854392e-02 -5.065223e-02 -0.0096177605  
## goout.y -0.016142013 -3.099888e-03 -3.795081e-02 -0.0148878405  
## Dalc.y 0.068670527 7.617950e-02 -2.199029e-02 -0.0884483562  
## Walc.y 0.005785212 6.061027e-02 -1.566193e-02 -0.1205999129  
## health.y 0.072241961 1.297469e-02 -1.701121e-01 -0.0359588482  
## absences.y 0.019872064 -2.253028e-02 2.346626e-02 0.0006008188  
## G1.y 0.080351311 3.205441e-02 1.278127e-01 0.0474638113  
## G2.y 0.047481176 2.307573e-02 1.214338e-01 0.0655793309  
## G3.y 0.064901708 -1.563203e-02 9.433172e-02 0.0662866316  
## internet guardian.x traveltime.x studytime.x  
## school -0.09355386 -0.0199460848 0.2499650999 -0.084631305  
## sex 0.07207982 -0.0519594768 0.0573113352 -0.293137317  
## age -0.08085448 0.1498863240 0.1026076175 -0.012225417  
## address 0.20883114 -0.0508602494 -0.3427326273 -0.024591483  
## famsize 0.04573320 -0.0069540932 0.0591911556 -0.059454409  
## Pstatus 0.02998751 -0.1149742897 0.0354093445 0.034099562  
## Medu 0.19339977 0.0072224863 -0.2058953291 0.050040187  
## Fedu 0.12523856 -0.1147751613 -0.1878555999 -0.003823746  
## Mjob 0.24623359 0.0039317201 -0.1073251935 -0.022180544  
## Fjob 0.05354420 -0.0781907799 0.0351909992 -0.080765942  
## reason 0.01627844 -0.0237714576 -0.0593942130 0.150352716  
## nursery -0.01738267 -0.0290917377 -0.0399661862 0.082809968  
## internet 1.00000000 -0.0205928268 -0.1086163536 0.042954263  
## guardian.x -0.02059283 1.0000000000 0.0090549857 -0.021739600  
## traveltime.x -0.10861635 0.0090549857 1.0000000000 -0.101529805  
## studytime.x 0.04295426 -0.0217396001 -0.1015298050 1.000000000  
## failures.x -0.01148100 0.0642878196 0.1132762373 -0.198990122  
## schoolsup.x -0.02694981 -0.0467749314 0.0048450918 0.029744491  
## famsup.x 0.07731126 -0.0029921972 -0.0256150697 0.159236013  
## paid.x 0.14445135 0.0614317791 -0.0627900940 0.161442772  
## activities.x 0.06601110 -0.0163918279 0.0005731197 0.100299354  
## higher.x -0.02523391 -0.0385600218 -0.0540201356 0.184467146  
## romantic.x 0.08423216 0.0430096157 -0.0043059375 0.032524084  
## famrel.x 0.07535520 0.0143762854 -0.0197589487 0.039673641  
## freetime.x 0.10278222 0.0306996026 -0.0252259949 -0.131548836  
## goout.x 0.10020791 0.0774716412 0.0299322575 -0.053358979  
## Dalc.x 0.06990662 0.0032295369 0.1572597217 -0.186142153  
## Walc.x 0.04125762 -0.0041246812 0.1402321501 -0.245865498  
## health.x -0.07013014 -0.0480252256 0.0060045484 -0.078714218  
## absences.x 0.09533546 0.0887941193 -0.0103819098 -0.073315012  
## G1.x 0.07186861 -0.0993764890 -0.1043048450 0.153626339  
## G2.x 0.10920347 -0.1191145074 -0.1608586144 0.129342867  
## G3.x 0.08800197 -0.0841340955 -0.1260043156 0.091338684  
## guardian.y -0.02059283 0.9674884394 0.0013853526 -0.002822682  
## traveltime.y -0.11685756 0.0028747355 0.9865306778 -0.096666433  
## studytime.y 0.04558900 -0.0003467384 -0.0965791737 0.977999632  
## failures.y -0.12525214 0.1101680262 0.0743482653 -0.198540526  
## schoolsup.y -0.03046083 -0.0502696811 -0.0125220188 0.030307888  
## famsup.y 0.07970726 0.0211761483 -0.0213075711 0.140027937  
## paid.y 0.02744798 0.0656598091 -0.0224897612 -0.010887902  
## activities.y 0.06377623 -0.0078168889 -0.0036353963 0.088074056  
## higher.y -0.02523391 -0.0134236254 -0.0362307461 0.169841482  
## romantic.y 0.07299490 0.0245000696 -0.0033579893 0.051829984  
## famrel.y 0.06155807 0.0392016601 -0.0094145530 0.026456237  
## freetime.y 0.08424236 0.0502281992 -0.0150654411 -0.144889940  
## goout.y 0.08267811 0.0839013057 0.0502444004 -0.067147345  
## Dalc.y 0.06291073 0.0044107082 0.1638798873 -0.186245838  
## Walc.y 0.03903064 -0.0091764918 0.1379988473 -0.241411308  
## health.y -0.07072574 -0.0448399603 0.0044891837 -0.078421079  
## absences.y 0.06256393 0.0722469204 0.0202349616 -0.151647551  
## G1.y 0.07865363 -0.0616614491 -0.1860606149 0.247058545  
## G2.y 0.04088684 -0.0391807201 -0.1838936733 0.265146693  
## G3.y 0.08905539 -0.0023937231 -0.1539727185 0.272129631  
## failures.x schoolsup.x famsup.x paid.x  
## school 0.0044239635 -0.134241883 -0.157393507 -0.009156375  
## sex 0.0541802963 -0.147342472 -0.158254450 -0.128773452  
## age 0.1530331389 -0.235809988 -0.122421850 -0.025939128  
## address -0.0480273677 0.034159255 0.006157785 0.032511511  
## famsize -0.0340657584 -0.032590749 -0.094605395 0.009571038  
## Pstatus 0.0005028530 -0.049545899 0.012189305 0.063272448  
## Medu -0.2301694688 -0.064699791 0.149882846 0.127198054  
## Fedu -0.2453863636 0.015201313 0.169846326 0.076320313  
## Mjob -0.0896837448 -0.042865903 0.042092732 0.066372370  
## Fjob 0.0005550276 -0.027438429 -0.043119657 -0.053291324  
## reason -0.0666525167 0.019233451 0.076562731 0.137585853  
## nursery -0.0466602141 0.051414351 0.053294288 0.085387759  
## internet -0.0114809996 -0.026949806 0.077311264 0.144451354  
## guardian.x 0.0642878196 -0.046774931 -0.002992197 0.061431779  
## traveltime.x 0.1132762373 0.004845092 -0.025615070 -0.062790094  
## studytime.x -0.1989901221 0.029744491 0.159236013 0.161442772  
## failures.x 1.0000000000 0.023037641 -0.023408289 -0.197673462  
## schoolsup.x 0.0230376412 1.000000000 0.082983496 -0.025172445  
## famsup.x -0.0234082893 0.082983496 1.000000000 0.267807472  
## paid.x -0.1976734624 -0.025172445 0.267807472 1.000000000  
## activities.x -0.0604916014 0.048785403 -0.024128624 -0.022431418  
## higher.x -0.3691642430 0.014643156 0.081948630 0.181855944  
## romantic.x 0.0528346170 -0.085277349 0.018723975 0.021830714  
## famrel.x -0.0676045327 0.008953196 0.001935774 -0.013363289  
## freetime.x 0.0921171258 -0.033909156 0.011175612 -0.076515701  
## goout.x 0.1255733484 -0.052703518 -0.013334442 0.009640443  
## Dalc.x 0.1275042141 -0.018826354 -0.016903465 0.066037649  
## Walc.x 0.1427806262 -0.115855368 -0.082910671 0.083678000  
## health.x 0.0816606288 -0.035800861 0.051404600 -0.091593597  
## absences.x 0.0172461203 0.025986784 0.002120548 0.024452734  
## G1.x -0.3949111131 -0.211525438 -0.083944270 0.038543805  
## G2.x -0.3849417367 -0.121283092 -0.066987359 0.100081223  
## G3.x -0.3807414784 -0.075236887 -0.049860955 0.100286569  
## guardian.y 0.0715989083 -0.046774931 0.018988944 0.040069067  
## traveltime.y 0.1214268703 -0.007673462 -0.022554036 -0.073432257  
## studytime.y -0.2057716431 0.036433086 0.132118393 0.161941876  
## failures.y 0.4998884951 -0.003144596 -0.038401082 -0.164082326  
## schoolsup.y 0.0263275256 0.965835516 0.093661132 -0.018172608  
## famsup.y -0.0330231504 0.080975484 0.972111559 0.284848343  
## paid.y 0.0348840818 0.046719920 0.102971712 0.082395162  
## activities.y -0.0583852433 0.035419689 -0.028199339 -0.028066349  
## higher.y -0.3691642430 0.050965790 0.081948630 0.181855944  
## romantic.y 0.0558263873 -0.089305576 0.027358751 0.011323866  
## famrel.y -0.0657434447 0.007946596 -0.013646823 -0.004663203  
## freetime.y 0.0965314278 -0.037150571 -0.015523203 -0.078840748  
## goout.y 0.1362260676 -0.020446768 -0.028800054 -0.013218070  
## Dalc.y 0.1263140079 -0.019985689 -0.026806859 0.069214996  
## Walc.y 0.1479681000 -0.107072803 -0.076570568 0.067903794  
## health.y 0.0847412846 -0.034969029 0.053662046 -0.082118788  
## absences.y 0.1872513824 -0.020910924 0.030736983 -0.137235472  
## G1.y -0.3060969358 -0.180090293 0.063913236 0.105019679  
## G2.y -0.3694327292 -0.144088466 0.048873169 0.114647813  
## G3.y -0.3838642136 -0.144695012 0.084951667 0.101225784  
## activities.x higher.x romantic.x famrel.x  
## school -0.1035455919 -0.0046475352 6.119714e-02 -0.089097658  
## sex 0.1068490779 -0.1565048662 -7.075494e-02 0.068752732  
## age -0.0664954420 -0.2260668090 1.250673e-01 0.025450635  
## address -0.0561736789 0.0357623240 -4.720985e-03 0.021730442  
## famsize -0.0202884799 -0.0027609439 3.865368e-02 -0.023885977  
## Pstatus 0.0874923405 -0.0326323972 -5.571376e-02 0.044777038  
## Medu 0.1108382720 0.1424835628 3.858646e-02 0.011913172  
## Fedu 0.1165858280 0.1599896324 2.354644e-02 0.013002950  
## Mjob 0.1044694204 0.0385396244 -6.741157e-02 0.046206617  
## Fjob 0.0176863345 -0.0487749040 -6.953438e-05 0.024250208  
## reason 0.0994100357 0.0645610738 1.174290e-02 -0.003765670  
## nursery 0.0118613000 0.0191846529 2.598897e-02 0.012107763  
## internet 0.0660111005 -0.0252339121 8.423216e-02 0.075355197  
## guardian.x -0.0163918279 -0.0385600218 4.300962e-02 0.014376285  
## traveltime.x 0.0005731197 -0.0540201356 -4.305938e-03 -0.019758949  
## studytime.x 0.1002993545 0.1844671459 3.252408e-02 0.039673641  
## failures.x -0.0604916014 -0.3691642430 5.283462e-02 -0.067604533  
## schoolsup.x 0.0487854033 0.0146431561 -8.527735e-02 0.008953196  
## famsup.x -0.0241286239 0.0819486301 1.872397e-02 0.001935774  
## paid.x -0.0224314180 0.1818559442 2.183071e-02 -0.013363289  
## activities.x 1.0000000000 0.1106276293 2.628618e-02 0.063238976  
## higher.x 0.1106276293 1.0000000000 -1.407054e-01 0.039142821  
## romantic.x 0.0262861783 -0.1407053546 1.000000e+00 -0.071619944  
## famrel.x 0.0632389762 0.0391428209 -7.161994e-02 1.000000000  
## freetime.x 0.0864549729 -0.0500050674 1.753765e-02 0.161720215  
## goout.x 0.0295630347 -0.0434282361 6.867284e-03 0.079476832  
## Dalc.x -0.0724943377 -0.0903277743 5.510691e-02 -0.090305814  
## Walc.x -0.0380611991 -0.0863808825 1.803992e-02 -0.123334683  
## health.x 0.0251719130 -0.0493484639 4.022563e-02 0.096206015  
## absences.x 0.0177665533 -0.0312311781 1.155358e-01 -0.056266628  
## G1.x 0.0734939405 0.2308573156 -4.746287e-02 0.029598967  
## G2.x 0.0751715845 0.2188922003 -1.266655e-01 -0.010866049  
## G3.x 0.0415803979 0.2321765504 -1.417025e-01 0.068602816  
## guardian.y -0.0057245736 -0.0134236254 2.011112e-02 0.037523771  
## traveltime.y -0.0033798695 -0.0351913116 -6.833463e-03 -0.023473616  
## studytime.y 0.1000012898 0.1712658810 4.163854e-02 0.033353869  
## failures.y -0.0553716116 -0.2519968209 -3.408362e-02 -0.031897758  
## schoolsup.y 0.0262870080 0.0496693837 -8.071142e-02 0.008520933  
## famsup.y -0.0190290994 0.0832653149 2.669437e-02 -0.003584618  
## paid.y 0.0899155810 0.0110433245 -5.264024e-03 0.074142928  
## activities.y 0.9737706601 0.1094331923 4.111489e-02 0.051490311  
## higher.y 0.1106276293 0.8833943834 -1.141493e-01 -0.001124390  
## romantic.y 0.0480261754 -0.1111619506 9.639424e-01 -0.064488851  
## famrel.y 0.0610859367 -0.0004988136 -5.598908e-02 0.970333717  
## freetime.y 0.0836472376 -0.0483891687 2.359775e-02 0.128065710  
## goout.y 0.0431656603 -0.1077999805 1.364791e-02 0.036950187  
## Dalc.y -0.0637587049 -0.0757053381 4.673130e-02 -0.096531290  
## Walc.y -0.0303122475 -0.1038723245 1.247784e-02 -0.122827732  
## health.y 0.0308077148 -0.0672454112 4.941018e-02 0.087705332  
## absences.y -0.0344855671 -0.1560553369 3.973156e-02 -0.052550007  
## G1.y 0.0767457550 0.2517404828 -1.018325e-02 -0.011597495  
## G2.y 0.0831957433 0.2921163639 -4.297115e-02 0.021320515  
## G3.y 0.1004202778 0.3035755910 -2.372173e-02 0.045309273  
## freetime.x goout.x Dalc.x Walc.x  
## school 0.0008617613 -0.026491127 0.077735716 0.0386775667  
## sex 0.2179524497 0.061583940 0.247532348 0.2554191812  
## age 0.0048821087 0.159630947 0.128370843 0.1521317124  
## address 0.0390874455 0.074316630 -0.105793067 -0.0965336221  
## famsize 0.0289524508 0.017132823 0.091183810 0.0912068137  
## Pstatus 0.0572095432 0.017621205 -0.019712592 0.0317034972  
## Medu 0.0353656074 0.066870221 0.043789396 -0.0136955864  
## Fedu -0.0074204872 0.033178695 0.017977755 0.0009282785  
## Mjob 0.1148495922 -0.005901141 0.069188837 0.0021902778  
## Fjob -0.0522780789 0.004077695 0.077282290 0.0588039594  
## reason -0.0444794390 -0.035157767 -0.018931895 -0.0089131064  
## nursery -0.0134221203 -0.017143338 -0.089881134 -0.1245115758  
## internet 0.1027822196 0.100207910 0.069906619 0.0412576189  
## guardian.x 0.0306996026 0.077471641 0.003229537 -0.0041246812  
## traveltime.x -0.0252259949 0.029932258 0.157259722 0.1402321501  
## studytime.x -0.1315488361 -0.053358979 -0.186142153 -0.2458654976  
## failures.x 0.0921171258 0.125573348 0.127504214 0.1427806262  
## schoolsup.x -0.0339091556 -0.052703518 -0.018826354 -0.1158553679  
## famsup.x 0.0111756118 -0.013334442 -0.016903465 -0.0829106711  
## paid.x -0.0765157009 0.009640443 0.066037649 0.0836779997  
## activities.x 0.0864549729 0.029563035 -0.072494338 -0.0380611991  
## higher.x -0.0500050674 -0.043428236 -0.090327774 -0.0863808825  
## romantic.x 0.0175376483 0.006867284 0.055106906 0.0180399189  
## famrel.x 0.1617202147 0.079476832 -0.090305814 -0.1233346834  
## freetime.x 1.0000000000 0.294310179 0.199967314 0.1349654701  
## goout.x 0.2943101792 1.000000000 0.260664387 0.4174491737  
## Dalc.x 0.1999673136 0.260664387 1.000000000 0.6471015587  
## Walc.x 0.1349654701 0.417449174 0.647101559 1.0000000000  
## health.x 0.0698415844 -0.027945456 0.057705697 0.0717311531  
## absences.x -0.0488139590 0.080056917 0.148442541 0.2089684631  
## G1.x 0.0101451339 -0.129491236 -0.080370378 -0.0996680309  
## G2.x -0.0142226119 -0.155863156 -0.050135813 -0.0652281523  
## G3.x 0.0176040438 -0.109654363 -0.044940500 -0.0272613329  
## guardian.y 0.0468900051 0.082183338 0.003229537 -0.0124393419  
## traveltime.y -0.0335229940 0.032704466 0.158594022 0.1327645964  
## studytime.y -0.1424050014 -0.062217444 -0.189511363 -0.2472871283  
## failures.y 0.0620149163 0.090000716 0.175495144 0.1669205922  
## schoolsup.y -0.0481735045 -0.031777197 -0.014829901 -0.1030188145  
## famsup.y -0.0009899849 -0.013891589 -0.025934378 -0.0715423519  
## paid.y 0.0127924208 0.009868691 0.101944227 0.0382708131  
## activities.y 0.0716817574 0.034713844 -0.063747360 -0.0409964086  
## higher.y -0.0500050674 -0.108999723 -0.076369356 -0.1056663863  
## romantic.y 0.0206120151 0.010677873 0.042538603 0.0155117727  
## famrel.y 0.1399586574 0.049688683 -0.089856664 -0.1279444921  
## freetime.y 0.9744238010 0.287391024 0.196330603 0.1294962410  
## goout.y 0.2810912936 0.936608021 0.265615810 0.3887119281  
## Dalc.y 0.1932906417 0.262964452 0.991650524 0.6510179975  
## Walc.y 0.1393651285 0.400458236 0.649798402 0.9856767876  
## health.y 0.0700690655 -0.044194709 0.060654281 0.0733976241  
## absences.y 0.0285917889 0.112994785 0.146830956 0.1538902363  
## G1.y -0.0909725257 -0.095996758 -0.233282420 -0.1913028405  
## G2.y -0.1132471079 -0.128926738 -0.256907363 -0.2200573917  
## G3.y -0.1089579060 -0.118223941 -0.264789491 -0.2230967330  
## health.x absences.x G1.x G2.x  
## school -0.061998722 -0.077215017 -0.031824956 -0.05915380  
## sex 0.140667879 -0.063819717 0.124570520 0.11632616  
## age -0.063236328 0.125971504 -0.117485703 -0.17536977  
## address -0.018952752 -0.037954493 0.062729830 0.12160232  
## famsize -0.055372622 0.057755381 0.093944845 0.10901321  
## Pstatus 0.049933843 -0.159489944 -0.013786609 -0.03643033  
## Medu -0.041731304 0.108559715 0.229202048 0.21791757  
## Fedu 0.015450306 0.033601812 0.215870985 0.18004156  
## Mjob 0.069058302 0.045834041 0.108899615 0.11477786  
## Fjob 0.010107881 0.024613884 0.083723801 0.07954471  
## reason -0.169386321 0.082398994 0.096657887 0.12365650  
## nursery -0.035156590 0.023727255 0.094076226 0.07391536  
## internet -0.070130136 0.095335456 0.071868610 0.10920347  
## guardian.x -0.048025226 0.088794119 -0.099376489 -0.11911451  
## traveltime.x 0.006004548 -0.010381910 -0.104304845 -0.16085861  
## studytime.x -0.078714218 -0.073315012 0.153626339 0.12934287  
## failures.x 0.081660629 0.017246120 -0.394911113 -0.38494174  
## schoolsup.x -0.035800861 0.025986784 -0.211525438 -0.12128309  
## famsup.x 0.051404600 0.002120548 -0.083944270 -0.06698736  
## paid.x -0.091593597 0.024452734 0.038543805 0.10008122  
## activities.x 0.025171913 0.017766553 0.073493940 0.07517158  
## higher.x -0.049348464 -0.031231178 0.230857316 0.21889220  
## romantic.x 0.040225633 0.115535838 -0.047462866 -0.12666546  
## famrel.x 0.096206015 -0.056266628 0.029598967 -0.01086605  
## freetime.x 0.069841584 -0.048813959 0.010145134 -0.01422261  
## goout.x -0.027945456 0.080056917 -0.129491236 -0.15586316  
## Dalc.x 0.057705697 0.148442541 -0.080370378 -0.05013581  
## Walc.x 0.071731153 0.208968463 -0.099668031 -0.06522815  
## health.x 1.000000000 -0.017103021 -0.067903833 -0.09798512  
## absences.x -0.017103021 1.000000000 -0.055608370 -0.04184040  
## G1.x -0.067903833 -0.055608370 1.000000000 0.86039517  
## G2.x -0.097985124 -0.041840401 0.860395165 1.00000000  
## G3.x -0.059429782 0.028987245 0.805128705 0.90302671  
## guardian.y -0.044216707 0.074805603 -0.078674964 -0.09824087  
## traveltime.y 0.012459997 -0.024753298 -0.111400474 -0.17239234  
## studytime.y -0.074650369 -0.070745248 0.143729991 0.11931529  
## failures.y 0.090414779 0.017271595 -0.151937047 -0.15404553  
## schoolsup.y -0.027342092 0.025512110 -0.213630198 -0.13505606  
## famsup.y 0.053106135 0.013971914 -0.088688098 -0.07373960  
## paid.y 0.074010784 0.028248593 -0.044726208 -0.01497144  
## activities.y 0.031081405 0.013852060 0.057586882 0.05421667  
## higher.y -0.067015861 -0.031231178 0.212389504 0.21566450  
## romantic.y 0.055446360 0.113828608 -0.036741660 -0.11648983  
## famrel.y 0.096361062 -0.053010153 0.008577284 -0.03039207  
## freetime.y 0.080080262 -0.049304194 -0.015744110 -0.04217040  
## goout.y -0.034774090 0.066074532 -0.140155773 -0.15406198  
## Dalc.y 0.062821629 0.155685085 -0.078472598 -0.04836377  
## Walc.y 0.074210955 0.194449021 -0.103017524 -0.06622970  
## health.y 0.984656260 -0.020409051 -0.065560949 -0.09346507  
## absences.y 0.004323473 0.496594774 -0.142383138 -0.16764967  
## G1.y -0.146536898 -0.064725243 0.568008116 0.52594436  
## G2.y -0.185767329 -0.107105660 0.588829103 0.56965643  
## G3.y -0.177518998 -0.094297337 0.550843715 0.50749963  
## G3.x guardian.y traveltime.y studytime.y  
## school -0.066660798 -0.0199460848 0.259503685 -0.0867744019  
## sex 0.121671254 -0.0519594768 0.053372402 -0.2867847201  
## age -0.194588537 0.1498863240 0.106545357 -0.0126857245  
## address 0.092187335 -0.0508602494 -0.338840020 -0.0213818830  
## famsize 0.110190986 -0.0069540932 0.056562213 -0.0632625086  
## Pstatus -0.054703437 -0.1149742897 0.036453780 0.0361693590  
## Medu 0.205489379 0.0072224863 -0.210964862 0.0483034173  
## Fedu 0.155444047 -0.1147751613 -0.192147048 -0.0041953922  
## Mjob 0.111351932 0.0039317201 -0.107364674 -0.0232610153  
## Fjob 0.049170701 -0.0781907799 0.033601505 -0.0794278024  
## reason 0.133052099 -0.0237714576 -0.062934533 0.1465742181  
## nursery 0.047036872 -0.0290917377 -0.037932678 0.0858201832  
## internet 0.088001968 -0.0205928268 -0.116857562 0.0455889960  
## guardian.x -0.084134095 0.9674884394 0.002874736 -0.0003467384  
## traveltime.x -0.126004316 0.0013853526 0.986530678 -0.0965791737  
## studytime.x 0.091338684 -0.0028226816 -0.096666433 0.9779996319  
## failures.x -0.380741478 0.0715989083 0.121426870 -0.2057716431  
## schoolsup.x -0.075236887 -0.0467749314 -0.007673462 0.0364330856  
## famsup.x -0.049860955 0.0189889439 -0.022554036 0.1321183926  
## paid.x 0.100286569 0.0400690667 -0.073432257 0.1619418756  
## activities.x 0.041580398 -0.0057245736 -0.003379870 0.1000012898  
## higher.x 0.232176550 -0.0134236254 -0.035191312 0.1712658810  
## romantic.x -0.141702491 0.0201111165 -0.006833463 0.0416385390  
## famrel.x 0.068602816 0.0375237713 -0.023473616 0.0333538689  
## freetime.x 0.017604044 0.0468900051 -0.033522994 -0.1424050014  
## goout.x -0.109654363 0.0821833379 0.032704466 -0.0622174435  
## Dalc.x -0.044940500 0.0032295369 0.158594022 -0.1895113629  
## Walc.x -0.027261333 -0.0124393419 0.132764596 -0.2472871283  
## health.x -0.059429782 -0.0442167066 0.012459997 -0.0746503694  
## absences.x 0.028987245 0.0748056033 -0.024753298 -0.0707452480  
## G1.x 0.805128705 -0.0786749644 -0.111400474 0.1437299912  
## G2.x 0.903026711 -0.0982408721 -0.172392340 0.1193152879  
## G3.x 1.000000000 -0.0693422695 -0.148818057 0.0855522893  
## guardian.y -0.069342269 1.0000000000 0.010500770 -0.0192687463  
## traveltime.y -0.148818057 0.0105007701 1.000000000 -0.1006428642  
## studytime.y 0.085552289 -0.0192687463 -0.100642864 1.0000000000  
## failures.y -0.148290672 0.1101680262 0.109452807 -0.2003037737  
## schoolsup.y -0.091809129 -0.0502696811 -0.002792711 0.0279100071  
## famsup.y -0.044600202 -0.0008355215 -0.033776670 0.1512668336  
## paid.y -0.004603647 0.0656598091 -0.023374826 -0.0248749164  
## activities.y 0.014010889 -0.0184813587 0.007465205 0.1064257233  
## higher.y 0.216341621 -0.0385600218 -0.052879576 0.1858954800  
## romantic.y -0.128849100 0.0472991600 0.002099937 0.0276674632  
## famrel.y 0.045297638 0.0157298197 -0.021510924 0.0439327454  
## freetime.y -0.012559091 0.0339862364 -0.012047669 -0.1306257482  
## goout.y -0.122706501 0.0791970199 0.042948088 -0.0541207995  
## Dalc.y -0.035704634 0.0044107082 0.160930846 -0.1896331366  
## Walc.y -0.039732027 -0.0008599559 0.139290256 -0.2477230250  
## health.y -0.059908270 -0.0486379348 0.005597072 -0.0765673423  
## absences.y -0.158033197 0.0852921752 0.007365494 -0.1651937684  
## G1.y 0.495304739 -0.0825229082 -0.193976204 0.2699206119  
## G2.y 0.535551523 -0.0543054848 -0.186251424 0.2797093042  
## G3.y 0.480349362 -0.0168793027 -0.156302894 0.2837619952  
## failures.y schoolsup.y famsup.y paid.y  
## school 0.072483471 -0.132718940 -0.1594620119 0.0433675712  
## sex 0.132766712 -0.141102245 -0.1528630309 0.1139222360  
## age 0.258572679 -0.247159579 -0.1252984193 -0.0553978068  
## address -0.044351456 0.030420869 0.0089723672 -0.0886679759  
## famsize -0.042470198 -0.028118205 -0.0980452589 -0.0251852621  
## Pstatus 0.057533234 -0.052543125 0.0140040522 -0.0143639669  
## Medu -0.204947316 -0.052323426 0.1560424466 0.1057335328  
## Fedu -0.212409306 0.026427500 0.1771700557 0.1072679281  
## Mjob -0.085363670 -0.041882051 0.0456348583 -0.0128544371  
## Fjob -0.023886490 -0.024792425 -0.0388014325 0.0224770277  
## reason -0.084343380 0.020926785 0.0710724371 -0.0968341836  
## nursery -0.049909096 0.048107174 0.0559751705 0.0239331396  
## internet -0.125252139 -0.030460830 0.0797072633 0.0274479804  
## guardian.x 0.110168026 -0.050269681 0.0211761483 0.0656598091  
## traveltime.x 0.074348265 -0.012522019 -0.0213075711 -0.0224897612  
## studytime.x -0.198540526 0.030307888 0.1400279366 -0.0108879020  
## failures.x 0.499888495 0.026327526 -0.0330231504 0.0348840818  
## schoolsup.x -0.003144596 0.965835516 0.0809754838 0.0467199199  
## famsup.x -0.038401082 0.093661132 0.9721115588 0.1029717121  
## paid.x -0.164082326 -0.018172608 0.2848483430 0.0823951617  
## activities.x -0.055371612 0.026287008 -0.0190290994 0.0899155810  
## higher.x -0.251996821 0.049669384 0.0832653149 0.0110433245  
## romantic.x -0.034083619 -0.080711419 0.0266943706 -0.0052640244  
## famrel.x -0.031897758 0.008520933 -0.0035846175 0.0741429280  
## freetime.x 0.062014916 -0.048173504 -0.0009899849 0.0127924208  
## goout.x 0.090000716 -0.031777197 -0.0138915895 0.0098686909  
## Dalc.x 0.175495144 -0.014829901 -0.0259343778 0.1019442268  
## Walc.x 0.166920592 -0.103018814 -0.0715423519 0.0382708131  
## health.x 0.090414779 -0.027342092 0.0531061345 0.0740107838  
## absences.x 0.017271595 0.025512110 0.0139719141 0.0282485931  
## G1.x -0.151937047 -0.213630198 -0.0886880984 -0.0447262079  
## G2.x -0.154045530 -0.135056059 -0.0737396003 -0.0149714396  
## G3.x -0.148290672 -0.091809129 -0.0446002023 -0.0046036467  
## guardian.y 0.110168026 -0.050269681 -0.0008355215 0.0656598091  
## traveltime.y 0.109452807 -0.002792711 -0.0337766702 -0.0233748258  
## studytime.y -0.200303774 0.027910007 0.1512668336 -0.0248749164  
## failures.y 1.000000000 0.044395395 -0.0399461814 0.1282505401  
## schoolsup.y 0.044395395 1.000000000 0.0916920370 0.0492112184  
## famsup.y -0.039946181 0.091692037 1.0000000000 0.1016529026  
## paid.y 0.128250540 0.049211218 0.1016529026 1.0000000000  
## activities.y -0.023235016 0.043854486 -0.0230784557 0.0704972071  
## higher.y -0.324302368 0.013040610 0.0832653149 0.0110433245  
## romantic.y -0.015165173 -0.084709569 0.0120921490 0.0139764582  
## famrel.y -0.027513637 -0.001029677 0.0045552590 0.0744024596  
## freetime.y 0.106731383 -0.043535532 -0.0058148925 0.0001106307  
## goout.y 0.061519869 -0.060942093 -0.0341803412 -0.0189373094  
## Dalc.y 0.186204541 -0.015976399 -0.0236445559 0.1011365004  
## Walc.y 0.160716612 -0.112273955 -0.0863499671 0.0441849298  
## health.y 0.068829279 -0.037610813 0.0630864559 0.0817219588  
## absences.y 0.082002800 -0.050119866 0.0245168040 -0.0307443819  
## G1.y -0.304200516 -0.187345830 0.0764751825 -0.1381172899  
## G2.y -0.308409148 -0.144552846 0.0550012347 -0.1230953687  
## G3.y -0.343498836 -0.139276734 0.0896379591 -0.1286497474  
## activities.y higher.y romantic.y famrel.y  
## school -0.101726035 -0.004647535 7.539790e-02 -0.0724939010  
## sex 0.111873912 -0.156504866 -8.124977e-02 0.0669359464  
## age -0.068327951 -0.226066809 1.332599e-01 0.0292860247  
## address -0.058878012 0.035762324 1.112330e-03 0.0235310258  
## famsize -0.017077331 -0.002760944 3.163222e-02 -0.0259846509  
## Pstatus 0.085727448 -0.032632397 -5.174763e-02 0.0463630156  
## Medu 0.105042500 0.142483563 3.525567e-02 0.0072786810  
## Fedu 0.109687401 0.159989632 2.277438e-02 0.0090613454  
## Mjob 0.096782537 0.038539624 -5.990234e-02 0.0463542615  
## Fjob 0.013439401 -0.048774904 2.388522e-03 0.0235208834  
## reason 0.104807127 0.064561074 9.825145e-05 -0.0092173515  
## nursery 0.009332196 0.019184653 3.127638e-02 0.0136673196  
## internet 0.063776231 -0.025233912 7.299490e-02 0.0615580719  
## guardian.x -0.007816889 -0.013423625 2.450007e-02 0.0392016601  
## traveltime.x -0.003635396 -0.036230746 -3.357989e-03 -0.0094145530  
## studytime.x 0.088074056 0.169841482 5.182998e-02 0.0264562373  
## failures.x -0.058385243 -0.369164243 5.582639e-02 -0.0657434447  
## schoolsup.x 0.035419689 0.050965790 -8.930558e-02 0.0079465959  
## famsup.x -0.028199339 0.081948630 2.735875e-02 -0.0136468234  
## paid.x -0.028066349 0.181855944 1.132387e-02 -0.0046632032  
## activities.x 0.973770660 0.110627629 4.802618e-02 0.0610859367  
## higher.x 0.109433192 0.883394383 -1.111620e-01 -0.0004988136  
## romantic.x 0.041114891 -0.114149304 9.639424e-01 -0.0559890818  
## famrel.x 0.051490311 -0.001124390 -6.448885e-02 0.9703337168  
## freetime.x 0.071681757 -0.050005067 2.061202e-02 0.1399586574  
## goout.x 0.034713844 -0.108999723 1.067787e-02 0.0496886832  
## Dalc.x -0.063747360 -0.076369356 4.253860e-02 -0.0898566638  
## Walc.x -0.040996409 -0.105666386 1.551177e-02 -0.1279444921  
## health.x 0.031081405 -0.067015861 5.544636e-02 0.0963610619  
## absences.x 0.013852060 -0.031231178 1.138286e-01 -0.0530101531  
## G1.x 0.057586882 0.212389504 -3.674166e-02 0.0085772842  
## G2.x 0.054216671 0.215664499 -1.164898e-01 -0.0303920733  
## G3.x 0.014010889 0.216341621 -1.288491e-01 0.0452976379  
## guardian.y -0.018481359 -0.038560022 4.729916e-02 0.0157298197  
## traveltime.y 0.007465205 -0.052879576 2.099937e-03 -0.0215109242  
## studytime.y 0.106425723 0.185895480 2.766746e-02 0.0439327454  
## failures.y -0.023235016 -0.324302368 -1.516517e-02 -0.0275136372  
## schoolsup.y 0.043854486 0.013040610 -8.470957e-02 -0.0010296769  
## famsup.y -0.023078456 0.083265315 1.209215e-02 0.0045552590  
## paid.y 0.070497207 0.011043325 1.397646e-02 0.0744024596  
## activities.y 1.000000000 0.109433192 2.918994e-02 0.0607374325  
## higher.y 0.109433192 1.000000000 -1.376027e-01 0.0403326420  
## romantic.y 0.029189944 -0.137602713 1.000000e+00 -0.0735524264  
## famrel.y 0.060737432 0.040332642 -7.355243e-02 1.0000000000  
## freetime.y 0.095508133 -0.048389169 1.517654e-02 0.1379793744  
## goout.y 0.043699731 -0.042331630 -2.422383e-03 0.0549983388  
## Dalc.y -0.066842357 -0.089662596 5.315641e-02 -0.0961780467  
## Walc.y -0.037299292 -0.084582471 1.862978e-02 -0.1252031054  
## health.y 0.018002350 -0.049626929 5.258416e-02 0.1042017182  
## absences.y -0.066919991 -0.125797420 4.373799e-02 -0.0519164918  
## G1.y 0.093379993 0.285611573 -4.354906e-02 -0.0062392784  
## G2.y 0.100690501 0.307151186 -7.114107e-02 0.0213405671  
## G3.y 0.119131368 0.311975263 -4.843913e-02 0.0434776962  
## freetime.y goout.y Dalc.y Walc.y  
## school 0.0242061586 -0.005376871 0.086377078 0.0425654024  
## sex 0.2109512588 0.061656662 0.244660760 0.2516863907  
## age 0.0167970594 0.166930526 0.131927571 0.1567950971  
## address 0.0433529490 0.076598533 -0.104250072 -0.0923146515  
## famsize 0.0241621471 0.014277559 0.089367319 0.0862273050  
## Pstatus 0.0600448801 0.019130391 -0.018727993 0.0344276811  
## Medu 0.0148307491 0.046280313 0.038862000 -0.0160064432  
## Fedu -0.0261498527 0.013843254 0.013747385 0.0004398100  
## Mjob 0.1094821905 -0.016142013 0.068670527 0.0057852116  
## Fjob -0.0585439172 -0.003099888 0.076179497 0.0606102750  
## reason -0.0506522275 -0.037950806 -0.021990291 -0.0156619333  
## nursery -0.0096177605 -0.014887841 -0.088448356 -0.1205999129  
## internet 0.0842423596 0.082678109 0.062910729 0.0390306378  
## guardian.x 0.0502281992 0.083901306 0.004410708 -0.0091764918  
## traveltime.x -0.0150654411 0.050244400 0.163879887 0.1379988473  
## studytime.x -0.1448899404 -0.067147345 -0.186245838 -0.2414113078  
## failures.x 0.0965314278 0.136226068 0.126314008 0.1479681000  
## schoolsup.x -0.0371505710 -0.020446768 -0.019985689 -0.1070728033  
## famsup.x -0.0155232026 -0.028800054 -0.026806859 -0.0765705678  
## paid.x -0.0788407481 -0.013218070 0.069214996 0.0679037939  
## activities.x 0.0836472376 0.043165660 -0.063758705 -0.0303122475  
## higher.x -0.0483891687 -0.107799981 -0.075705338 -0.1038723245  
## romantic.x 0.0235977486 0.013647914 0.046731300 0.0124778370  
## famrel.x 0.1280657095 0.036950187 -0.096531290 -0.1228277317  
## freetime.x 0.9744238010 0.281091294 0.193290642 0.1393651285  
## goout.x 0.2873910243 0.936608021 0.262964452 0.4004582357  
## Dalc.x 0.1963306031 0.265615810 0.991650524 0.6497984023  
## Walc.x 0.1294962410 0.388711928 0.651017998 0.9856767876  
## health.x 0.0800802624 -0.034774090 0.062821629 0.0742109547  
## absences.x -0.0493041937 0.066074532 0.155685085 0.1944490206  
## G1.x -0.0157441097 -0.140155773 -0.078472598 -0.1030175241  
## G2.x -0.0421703966 -0.154061979 -0.048363770 -0.0662296967  
## G3.x -0.0125590913 -0.122706501 -0.035704634 -0.0397320272  
## guardian.y 0.0339862364 0.079197020 0.004410708 -0.0008599559  
## traveltime.y -0.0120476692 0.042948088 0.160930846 0.1392902559  
## studytime.y -0.1306257482 -0.054120800 -0.189633137 -0.2477230250  
## failures.y 0.1067313833 0.061519869 0.186204541 0.1607166120  
## schoolsup.y -0.0435355322 -0.060942093 -0.015976399 -0.1122739551  
## famsup.y -0.0058148925 -0.034180341 -0.023644556 -0.0863499671  
## paid.y 0.0001106307 -0.018937309 0.101136500 0.0441849298  
## activities.y 0.0955081329 0.043699731 -0.066842357 -0.0372992915  
## higher.y -0.0483891687 -0.042331630 -0.089662596 -0.0845824714  
## romantic.y 0.0151765353 -0.002422383 0.053156407 0.0186297756  
## famrel.y 0.1379793744 0.054998339 -0.096178047 -0.1252031054  
## freetime.y 1.0000000000 0.290556490 0.198627942 0.1276112226  
## goout.y 0.2905564902 1.000000000 0.260061817 0.4041945072  
## Dalc.y 0.1986279420 0.260061817 1.000000000 0.6421467171  
## Walc.y 0.1276112226 0.404194507 0.642146717 1.0000000000  
## health.y 0.0708087716 -0.029537641 0.061543577 0.0802581300  
## absences.y 0.0080354611 0.145686877 0.148827157 0.1582250521  
## G1.y -0.0936985544 -0.090616221 -0.246135333 -0.1901054774  
## G2.y -0.1143791770 -0.132922513 -0.267969490 -0.2217260759  
## G3.y -0.1116089682 -0.132210695 -0.275340080 -0.2266645528  
## health.y absences.y G1.y G2.y  
## school -0.061188688 -0.0609205612 -0.165768955 -0.11628907  
## sex 0.142077859 -0.0093998418 -0.175880794 -0.19092191  
## age -0.066905631 0.0863029290 -0.109548659 -0.04201003  
## address -0.019868603 -0.0163459754 0.175881683 0.18522613  
## famsize -0.054077601 -0.0119657308 0.090509954 0.06018085  
## Pstatus 0.049175188 -0.0596855565 -0.019603511 -0.06013154  
## Medu -0.033346657 0.0230397799 0.205381472 0.21594789  
## Fedu 0.023191634 0.0208293469 0.157978105 0.18482493  
## Mjob 0.072241961 0.0198720644 0.080351311 0.04748118  
## Fjob 0.012974694 -0.0225302765 0.032054407 0.02307573  
## reason -0.170112148 0.0234662568 0.127812730 0.12143380  
## nursery -0.035958848 0.0006008188 0.047463811 0.06557933  
## internet -0.070725743 0.0625639278 0.078653634 0.04088684  
## guardian.x -0.044839960 0.0722469204 -0.061661449 -0.03918072  
## traveltime.x 0.004489184 0.0202349616 -0.186060615 -0.18389367  
## studytime.x -0.078421079 -0.1516475505 0.247058545 0.26514669  
## failures.x 0.084741285 0.1872513824 -0.306096936 -0.36943273  
## schoolsup.x -0.034969029 -0.0209109243 -0.180090293 -0.14408847  
## famsup.x 0.053662046 0.0307369832 0.063913236 0.04887317  
## paid.x -0.082118788 -0.1372354722 0.105019679 0.11464781  
## activities.x 0.030807715 -0.0344855671 0.076745755 0.08319574  
## higher.x -0.067245411 -0.1560553369 0.251740483 0.29211636  
## romantic.x 0.049410182 0.0397315571 -0.010183252 -0.04297115  
## famrel.x 0.087705332 -0.0525500071 -0.011597495 0.02132051  
## freetime.x 0.070069066 0.0285917889 -0.090972526 -0.11324711  
## goout.x -0.044194709 0.1129947853 -0.095996758 -0.12892674  
## Dalc.x 0.060654281 0.1468309563 -0.233282420 -0.25690736  
## Walc.x 0.073397624 0.1538902363 -0.191302841 -0.22005739  
## health.x 0.984656260 0.0043234731 -0.146536898 -0.18576733  
## absences.x -0.020409051 0.4965947743 -0.064725243 -0.10710566  
## G1.x -0.065560949 -0.1423831381 0.568008116 0.58882910  
## G2.x -0.093465074 -0.1676496705 0.525944359 0.56965643  
## G3.x -0.059908270 -0.1580331972 0.495304739 0.53555152  
## guardian.y -0.048637935 0.0852921752 -0.082522908 -0.05430548  
## traveltime.y 0.005597072 0.0073654941 -0.193976204 -0.18625142  
## studytime.y -0.076567342 -0.1651937684 0.269920612 0.27970930  
## failures.y 0.068829279 0.0820028002 -0.304200516 -0.30840915  
## schoolsup.y -0.037610813 -0.0501198662 -0.187345830 -0.14455285  
## famsup.y 0.063086456 0.0245168040 0.076475182 0.05500123  
## paid.y 0.081721959 -0.0307443819 -0.138117290 -0.12309537  
## activities.y 0.018002350 -0.0669199912 0.093379993 0.10069050  
## higher.y -0.049626929 -0.1257974203 0.285611573 0.30715119  
## romantic.y 0.052584156 0.0437379939 -0.043549059 -0.07114107  
## famrel.y 0.104201718 -0.0519164918 -0.006239278 0.02134057  
## freetime.y 0.070808772 0.0080354611 -0.093698554 -0.11437918  
## goout.y -0.029537641 0.1456868772 -0.090616221 -0.13292251  
## Dalc.y 0.061543577 0.1488271568 -0.246135333 -0.26796949  
## Walc.y 0.080258130 0.1582250521 -0.190105477 -0.22172608  
## health.y 1.000000000 0.0285698146 -0.147511098 -0.18885875  
## absences.y 0.028569815 1.0000000000 -0.157981318 -0.20031888  
## G1.y -0.147511098 -0.1579813178 1.000000000 0.88790615  
## G2.y -0.188858747 -0.2003188826 0.887906146 1.00000000  
## G3.y -0.183680550 -0.1448609397 0.819048197 0.89244365  
## G3.y  
## school -0.182031553  
## sex -0.201063602  
## age -0.067977477  
## address 0.188918461  
## famsize 0.042718250  
## Pstatus -0.078506185  
## Medu 0.203553661  
## Fedu 0.172008587  
## Mjob 0.064901708  
## Fjob -0.015632027  
## reason 0.094331716  
## nursery 0.066286632  
## internet 0.089055393  
## guardian.x -0.002393723  
## traveltime.x -0.153972718  
## studytime.x 0.272129631  
## failures.x -0.383864214  
## schoolsup.x -0.144695012  
## famsup.x 0.084951667  
## paid.x 0.101225784  
## activities.x 0.100420278  
## higher.x 0.303575591  
## romantic.x -0.023721731  
## famrel.x 0.045309273  
## freetime.x -0.108957906  
## goout.x -0.118223941  
## Dalc.x -0.264789491  
## Walc.x -0.223096733  
## health.x -0.177518998  
## absences.x -0.094297337  
## G1.x 0.550843715  
## G2.x 0.507499632  
## G3.x 0.480349362  
## guardian.y -0.016879303  
## traveltime.y -0.156302894  
## studytime.y 0.283761995  
## failures.y -0.343498836  
## schoolsup.y -0.139276734  
## famsup.y 0.089637959  
## paid.y -0.128649747  
## activities.y 0.119131368  
## higher.y 0.311975263  
## romantic.y -0.048439126  
## famrel.y 0.043477696  
## freetime.y -0.111608968  
## goout.y -0.132210695  
## Dalc.y -0.275340080  
## Walc.y -0.226664553  
## health.y -0.183680550  
## absences.y -0.144860940  
## G1.y 0.819048197  
## G2.y 0.892443645  
## G3.y 1.000000000

La variable que es objeto de este estudio es G3, por lo que nos centraremos en ella.

# Se clasifican los alumnos como aprobados o suspendidos. Se considera G3 como la ultima evaluacion. Aprobados mayor a 9 (G3 tiene un valor entre 0 y 20).

sMatAprobados <- studentMat[!studentMat$G3<9,]  
 dim(sMatAprobados)

## [1] 293 33

sPorAprobados <- studentPor[!studentPor$G3<9,]  
 dim(sPorAprobados)

## [1] 584 33

Convertimos el dataframe a una matriz para trabajar. A la hora de revisar las relaciones, las que sean positivas nos indican que a medida que la variable aumenta, también aumenta G3. En las negativas, sucede que a medida que disminuya la variable, también disminuye G3. Las correlaciones perfectas son =0 (negativa) o =1 (positiva)

studentMatPormatriz<-as.matrix(correlacion)  
is.matrix(correlacion)

## [1] TRUE

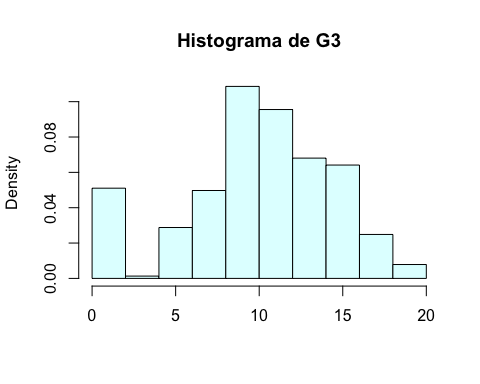
correlacionG3<-correlacion[, 33]  
correlacionG3

## school sex age address famsize   
## -0.066660798 0.121671254 -0.194588537 0.092187335 0.110190986   
## Pstatus Medu Fedu Mjob Fjob   
## -0.054703437 0.205489379 0.155444047 0.111351932 0.049170701   
## reason nursery internet guardian.x traveltime.x   
## 0.133052099 0.047036872 0.088001968 -0.084134095 -0.126004316   
## studytime.x failures.x schoolsup.x famsup.x paid.x   
## 0.091338684 -0.380741478 -0.075236887 -0.049860955 0.100286569   
## activities.x higher.x romantic.x famrel.x freetime.x   
## 0.041580398 0.232176550 -0.141702491 0.068602816 0.017604044   
## goout.x Dalc.x Walc.x health.x absences.x   
## -0.109654363 -0.044940500 -0.027261333 -0.059429782 0.028987245   
## G1.x G2.x G3.x guardian.y traveltime.y   
## 0.805128705 0.903026711 1.000000000 -0.069342269 -0.148818057   
## studytime.y failures.y schoolsup.y famsup.y paid.y   
## 0.085552289 -0.148290672 -0.091809129 -0.044600202 -0.004603647   
## activities.y higher.y romantic.y famrel.y freetime.y   
## 0.014010889 0.216341621 -0.128849100 0.045297638 -0.012559091   
## goout.y Dalc.y Walc.y health.y absences.y   
## -0.122706501 -0.035704634 -0.039732027 -0.059908270 -0.158033197   
## G1.y G2.y G3.y   
## 0.495304739 0.535551523 0.480349362

Podemos revisar cuáles pueden ser potenciales variables a estudiar según su correlación con G3 (positiva o negativa).

Representacion Gráfica de G3

histograma<-hist(studentMatPor$G3.x, freq=FALSE,col="lightcyan",main="Histograma de G3",xlab="")

 # 4. Análisis Exploratorio Vamos a realizar una agrupación de alumnos aprobados o suspensos, a partir de un cluster jerarquico.

Lo primero que debemos hacer es crear una variable binaria que sea capaz de explicar si el alumno aprueba o no.

binaria <-cut2(studentMatPor$G3.x, c(10,20))  
levels(binaria) <- c(0,1)  
binaria2<-cbind(studentMatPor,binaria)

Como queremos aplicar cluster jerarquico podemos utilizar diversas distancias.

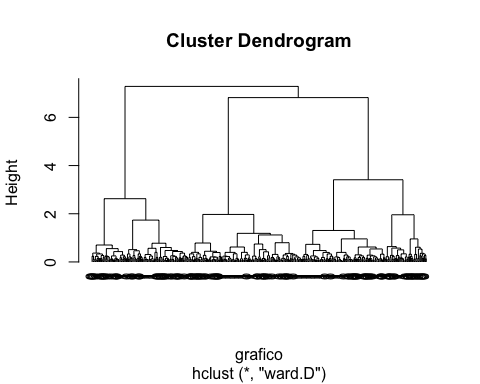
jerarquico <- binaria2 %>% select(-c(studentMatPor$G3.x,binaria))   
idsamp <- sample(1:dim(jerarquico)[1], 360)   
jerarquicosamp <- jerarquico[idsamp, ]  
head(jerarquicosamp)

## school sex age activities.x higher.x romantic.x famrel.x freetime.x  
## 58 GP F 16 yes yes no 4 4  
## 29 GP F 15 no yes no 3 3  
## 196 GP M 15 no yes no 5 4  
## 352 MS F 18 no yes no 5 3  
## 229 GP M 16 yes yes yes 3 3  
## 274 GP M 17 yes yes no 5 2  
## goout.x Dalc.x Walc.x health.x absences.x G1.x G2.x G3.x guardian.y  
## 58 3 1 3 4 6 8 10 10 mother  
## 29 3 1 2 2 0 10 8 9 mother  
## 196 3 1 2 3 2 10 10 11 mother  
## 352 3 1 3 4 2 10 9 10 mother  
## 229 3 1 2 3 2 11 12 11 mother  
## 274 4 1 4 5 20 9 7 8 father  
## traveltime.y studytime.y failures.y schoolsup.y famsup.y paid.y  
## 58 1 2 0 yes yes no  
## 29 1 2 0 no yes no  
## 196 2 2 0 no yes no  
## 352 2 3 0 no no no  
## 229 1 1 0 no yes no  
## 274 3 2 1 no yes no  
## activities.y higher.y romantic.y famrel.y freetime.y goout.y Dalc.y  
## 58 yes yes no 4 4 3 1  
## 29 no yes no 3 3 3 1  
## 196 no yes no 5 4 3 1  
## 352 no yes no 5 3 3 1  
## 229 yes yes yes 3 3 3 1  
## 274 yes yes no 5 2 4 1  
## Walc.y health.y absences.y G1.y G2.y G3.y  
## 58 3 4 4 11 12 12  
## 29 2 2 2 14 14 14  
## 196 2 3 0 12 12 13  
## 352 3 4 0 14 17 15  
## 229 2 3 0 10 9 11  
## 274 4 5 14 12 11 11

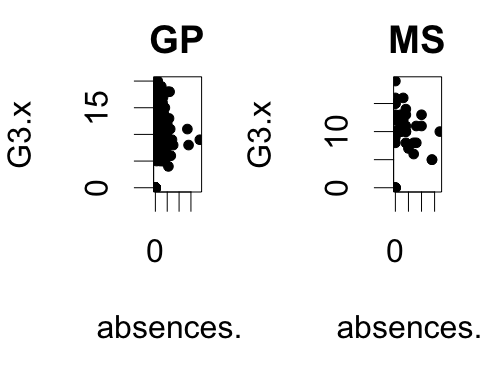
grafico<-daisy(jerarquicosamp,"gower")

### Clustering jerarquico

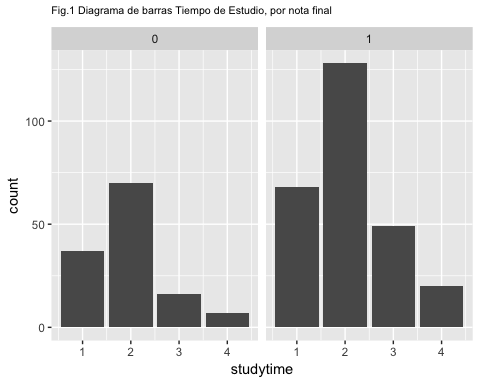
clusterjer <- hclust(grafico, method="ward.D")  
plot(clusterjer, hang = -1, cex=0.8, labels=binaria2$binaria[idsamp])

 Es interesante ver el Grafico de Ausencias de alumnos por Escuela (GP o MS)

par(mfrow = c(1, 2), mar = c(5, 4, 2, 1), cex=2)  
with(subset(studentMatPor, school == "GP"),  
plot(absences.x, G3.x, main = "GP", pch=20))  
with(subset(studentMatPor, school == "MS"),  
plot(absences.x, G3.x, main = "MS", pch=20))

 Analizamos la relación de algunas variables, por ejemplo elegimos ‘Tiempo de Estudio’ para comparar con la variable objetivo: considerando que G3 = evaluación final: Suponemos que “fail” es una nota final de 0 a 9 y que “pass” es superior a 9:

studentMat$pass <- ifelse(studentMat$G3>9, 1, 0)  
ggplot(studentMat, aes(studytime)) + geom\_bar() +  
 facet\_wrap(~ pass) +  
 ggtitle('Fig.1 Diagrama de barras Tiempo de Estudio, por nota final') +  
 theme(plot.title = element\_text(vjust = +1.4, size = 8))

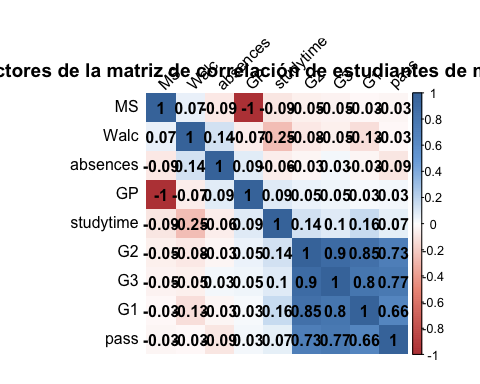
 Vemos que según el diagrama de barras los estudiantes de matemáticas que le dedican entre 2 a 5 horas aprueban con una gran diferencia a los estudiantes que le dedican menos horas de estudio (o incluso más de 5 horas de estudio). A pesar de esto la distribución de esta variable es muy similar entre los que suspenden y los que aprueban.

Podemos calcular una matriz de correlación para averiguar la relacion de unas variables con otras y con el target. Tomaremos sólo algunas variables. Primero creamos variables dummies y posteriormente la matriz:

studentMat$GP <- ifelse(studentMat$school == "GP", 1, 0)  
studentMat$MS <- ifelse(studentMat$school == "MS", 1, 0)  
matCor <- cor(studentMat[, c('GP','MS','absences','studytime','Walc','G1','G2','G3','pass')])  
matCor[is.na(matCor)] <- 0  
col <- colorRampPalette(c('#BB4444', '#EE9988', '#FFFFFF', '#77AADD', '#4477AA'))  
corrplot(matCor, method = 'shade', shade.col = NA, tl.col = 'black',  
tl.srt = 45, col = col(200), addCoef.col="black",  
order='AOE',  
mar = c(1,0,2,0), line=-2,  
main = 'Fig. 3 Predictores de la matriz de correlación de estudiantes de matemáticas')

## Warning in text.default(pos.xlabel[, 1], pos.xlabel[, 2], newcolnames, srt  
## = tl.srt, : "line" is not a graphical parameter

## Warning in text.default(pos.ylabel[, 1], pos.ylabel[, 2], newrownames, col  
## = tl.col, : "line" is not a graphical parameter

 Llevamos a cabo el mismo análisis para los estudiantes de portugués:

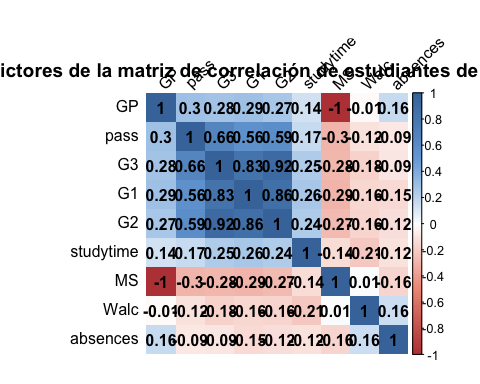
studentPor$pass <- ifelse(studentPor$G3>9, 1, 0)

Dibujamos también su matriz de correlación:

studentPor$GP <- ifelse(studentPor$school == "GP", 1, 0)  
studentPor$MS <- ifelse(studentPor$school == "MS", 1, 0)  
matCor <- cor(studentPor[, c('GP','MS','absences','studytime','Walc','G1','G2','G3','pass')])  
matCor[is.na(matCor)] <- 0  
col <- colorRampPalette(c('#BB4444', '#EE9988', '#FFFFFF', '#77AADD', '#4477AA'))  
corrplot(matCor, method = 'shade', shade.col = NA, tl.col = 'black',  
tl.srt = 45, col = col(200), addCoef.col="black",  
order='AOE',  
mar = c(1,0,2,0), line=-2,  
main = 'Fig. 4 Predictores de la matriz de correlación de estudiantes de portugués')

## Warning in text.default(pos.xlabel[, 1], pos.xlabel[, 2], newcolnames, srt  
## = tl.srt, : "line" is not a graphical parameter

## Warning in text.default(pos.ylabel[, 1], pos.ylabel[, 2], newrownames, col  
## = tl.col, : "line" is not a graphical parameter

 Los valores más cercanos a 1, G1, G2, G3, pertenecen a las variables más correladas. Esto significaría que las notas de otros periodos anteriores va a influir en la nota final.

## Análisis exploratorio apoyado en algún método NO supervisado

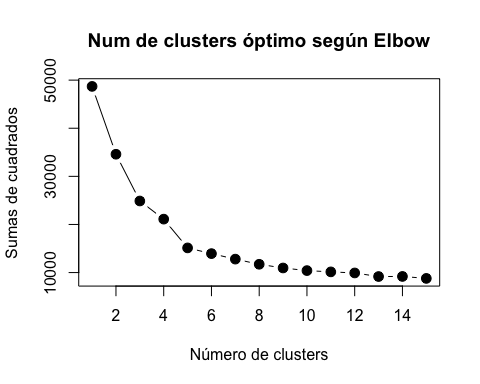
El método no supervisado que utilizaremos será el de agrupaciones o clustering particionado. Vamos a quitar las variables que no queremos que se incluyan en el clustering: Para los estudiantes de matemáticas:

studentMat.mod <- studentMat %>% select (-school, -sex, -address, -famsize, -Pstatus, -Mjob, -Fjob, -reason, -guardian, -schoolsup, -famsup, -paid, -activities, -nursery, -higher, -internet, -romantic, -pass)  
head(studentMat.mod)

## age Medu Fedu traveltime studytime failures famrel freetime goout Dalc  
## 1 18 4 4 2 2 0 4 3 4 1  
## 2 17 1 1 1 2 0 5 3 3 1  
## 3 15 1 1 1 2 3 4 3 2 2  
## 4 15 4 2 1 3 0 3 2 2 1  
## 5 16 3 3 1 2 0 4 3 2 1  
## 6 16 4 3 1 2 0 5 4 2 1  
## Walc health absences G1 G2 G3 GP MS  
## 1 1 3 6 5 6 6 1 0  
## 2 1 3 4 5 5 6 1 0  
## 3 3 3 10 7 8 10 1 0  
## 4 1 5 2 15 14 15 1 0  
## 5 2 5 4 6 10 10 1 0  
## 6 2 5 10 15 15 15 1 0

Aplicamos el clustering con k-means ya que todas las variables son numéricas:

mydata <- studentMat.mod  
wss <- (nrow(mydata)-1)\*sum(apply(mydata,2,var))  
for (i in 2:15) wss[i] <- sum(kmeans(mydata,  
 centers=i)$withinss)  
plot(1:15, wss, type='b', xlab='Número de clusters',  
 ylab='Sumas de cuadrados',  
 main='Num de clusters óptimo según Elbow',  
 pch=20, cex=2)

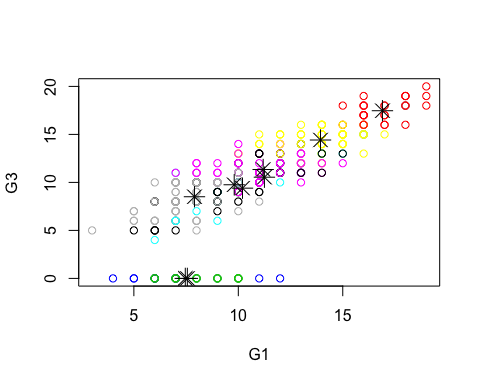
 Podemos ver que el cluster final K para los estudiantes de matemáticas podríamos tomarlo a 9 ya que a partir de ese número la variación es muy pequeña. Aplicamos pues el kmeans() con k=9

set.seed(1234)  
studentMatkmeans.clust <- kmeans(studentMat.mod, 9)  
studentMatkmeans.clust

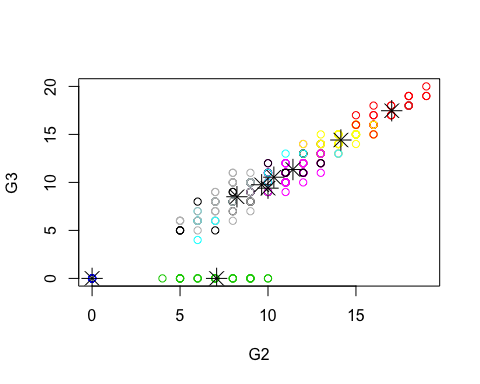
## K-means clustering with 9 clusters of sizes 5, 30, 25, 13, 24, 79, 77, 81, 61  
##   
## Cluster means:  
## age Medu Fedu traveltime studytime failures famrel  
## 1 17.80000 3.000000 3.000000 1.400000 2.000000 0.40000000 4.200000  
## 2 16.30000 3.333333 2.900000 1.300000 2.300000 0.00000000 4.233333  
## 3 17.36000 2.160000 2.160000 1.440000 2.080000 0.96000000 3.560000  
## 4 16.53846 2.615385 2.538462 1.923077 1.769231 0.84615385 4.384615  
## 5 17.04167 2.916667 2.541667 1.500000 1.708333 0.45833333 3.791667  
## 6 16.54430 2.645570 2.696203 1.468354 1.987342 0.10126582 4.088608  
## 7 16.35065 3.051948 2.740260 1.298701 2.207792 0.05194805 3.870130  
## 8 16.70370 2.308642 2.123457 1.580247 2.012346 0.50617284 4.012346  
## 9 17.04918 2.983607 2.459016 1.393443 1.950820 0.50819672 3.721311  
## freetime goout Dalc Walc health absences G1  
## 1 2.200000 2.600000 1.400000 2.000000 3.600000 52.600000 10.200000  
## 2 3.333333 2.833333 1.166667 1.700000 3.400000 3.533333 16.900000  
## 3 2.960000 2.920000 1.400000 2.000000 3.240000 0.000000 7.560000  
## 4 3.461538 3.769231 1.230769 1.769231 4.307692 0.000000 7.461538  
## 5 3.291667 3.125000 1.416667 2.750000 3.375000 20.708333 11.250000  
## 6 3.202532 2.898734 1.443038 2.215190 3.898734 1.746835 11.189873  
## 7 3.298701 3.000000 1.428571 2.155844 3.194805 3.415584 13.935065  
## 8 3.185185 3.296296 1.432099 2.456790 3.691358 3.530864 7.901235  
## 9 3.344262 3.377049 1.934426 2.704918 3.491803 11.508197 9.803279  
## G2 G3 GP MS  
## 1 10.000000 9.400000 1.0000000 0.00000000  
## 2 17.033333 17.466667 0.9333333 0.06666667  
## 3 7.080000 0.000000 0.8400000 0.16000000  
## 4 0.000000 0.000000 1.0000000 0.00000000  
## 5 10.333333 10.541667 0.9583333 0.04166667  
## 6 11.417722 11.341772 0.8481013 0.15189873  
## 7 14.142857 14.415584 0.9090909 0.09090909  
## 8 8.234568 8.506173 0.8271605 0.17283951  
## 9 9.639344 9.754098 0.9016393 0.09836066  
##   
## Clustering vector:  
## [1] 8 8 9 7 8 7 6 8 2 7 8 6 7 6 7 7 7 8 9 8 7 7 7 6 8 9 6 7 6 9 6 2 2 6 7  
## [36] 8 2 7 6 7 5 9 2 8 9 8 9 2 7 8 6 6 6 6 7 9 7 7 8 2 6 8 8 8 6 7 7 8 8 2  
## [71] 7 6 8 6 1 8 9 6 8 9 6 6 9 7 8 8 8 7 9 5 8 2 8 6 7 8 7 8 7 8 9 2 6 5 2  
## [106] 9 8 2 7 7 2 8 6 2 9 7 7 6 5 7 7 7 6 5 8 6 6 8 3 2 4 4 9 5 4 4 4 4 6 2  
## [141] 3 9 6 7 4 6 3 6 3 8 3 7 9 4 6 8 7 8 7 6 3 8 4 6 8 5 8 7 3 7 3 7 6 3 6  
## [176] 8 6 8 9 6 9 6 2 1 9 9 6 7 8 8 9 8 9 9 7 7 2 9 5 6 2 8 8 5 8 5 8 9 8 8  
## [211] 9 9 6 9 9 7 5 9 8 8 8 3 2 6 7 9 7 6 9 9 9 6 9 7 5 9 7 5 6 3 9 6 4 6 4  
## [246] 2 6 9 8 7 8 8 8 8 6 8 7 9 7 3 5 8 6 8 3 2 8 9 9 4 9 7 6 7 6 6 1 5 9 9  
## [281] 5 5 6 8 8 6 2 7 7 7 9 7 9 2 7 6 3 9 7 7 9 6 7 2 5 7 2 1 6 5 3 5 6 5 9  
## [316] 1 3 9 6 6 5 9 6 7 7 6 7 9 8 7 8 7 4 3 3 5 9 3 2 8 6 3 7 3 6 7 2 6 7 9  
## [351] 8 7 8 8 6 6 7 6 8 2 6 6 6 7 6 8 6 3 6 9 8 6 9 9 2 8 7 8 7 5 7 8 6 3 9  
## [386] 8 8 3 8 3 9 7 8 6 8  
##   
## Within cluster sum of squares by cluster:  
## [1] 965.6000 920.4667 453.6800 211.6923 1461.8333 1493.0127 1976.8571  
## [8] 2032.5926 2208.7541  
## (between\_SS / total\_SS = 75.9 %)  
##   
## Available components:  
##   
## [1] "cluster" "centers" "totss" "withinss"   
## [5] "tot.withinss" "betweenss" "size" "iter"   
## [9] "ifault"

Comparamos el resultado de la clusterización con las variables G1 y G2 ya que en la matriz de correlación habíamos comprobado ya que tanto las variables G1 como G2 influían en G3

plot(studentMat.mod %>% select(G1, G3), col = studentMatkmeans.clust$cluster)  
points(as.data.frame(studentMatkmeans.clust$centers) %>% select(G1, G3), pch = 8, cex = 2)



#Hacemos lo mismo con G2 y G3:  
plot(studentMat.mod %>% select(G2, G3), col = studentMatkmeans.clust$cluster)  
points(as.data.frame(studentMatkmeans.clust$centers) %>% select(G2, G3), pch = 8, cex = 2)

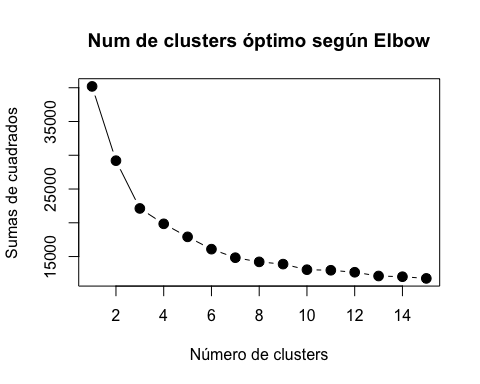
 ``` Aplicamos también este análisis para los estudiantes de portugués:

studentPor.mod <- studentPor %>% select (-school, -sex, -address, -famsize, -Pstatus, -Mjob, -Fjob, -reason, -guardian, -schoolsup, -famsup, -paid, -activities, -nursery, -higher, -internet, -romantic, -pass)  
head(studentPor.mod)

## age Medu Fedu traveltime studytime failures famrel freetime goout Dalc  
## 1 18 4 4 2 2 0 4 3 4 1  
## 2 17 1 1 1 2 0 5 3 3 1  
## 3 15 1 1 1 2 0 4 3 2 2  
## 4 15 4 2 1 3 0 3 2 2 1  
## 5 16 3 3 1 2 0 4 3 2 1  
## 6 16 4 3 1 2 0 5 4 2 1  
## Walc health absences G1 G2 G3 GP MS  
## 1 1 3 4 0 11 11 1 0  
## 2 1 3 2 9 11 11 1 0  
## 3 3 3 6 12 13 12 1 0  
## 4 1 5 0 14 14 14 1 0  
## 5 2 5 0 11 13 13 1 0  
## 6 2 5 6 12 12 13 1 0

Realizamos también la gráfica Elbow:

mydataPor <- studentPor.mod  
wss <- (nrow(mydataPor)-1)\*sum(apply(mydataPor,2,var))  
for (i in 2:15) wss[i] <- sum(kmeans(mydataPor,  
 centers=i)$withinss)  
plot(1:15, wss, type='b', xlab='Número de clusters',  
 ylab='Sumas de cuadrados',  
 main='Num de clusters óptimo según Elbow',  
 pch=20, cex=2)

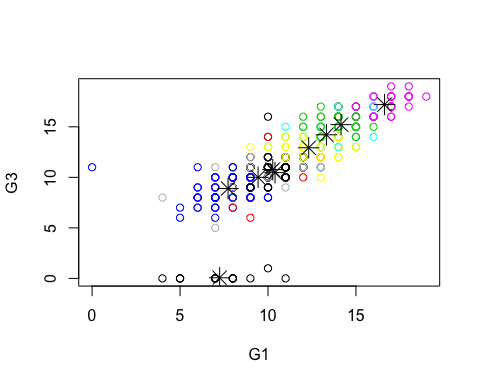
 Tambien consideramos k = 9

set.seed(1234)  
studentPorkmeans.clust <- kmeans(studentPor.mod, 9)  
studentPorkmeans.clust

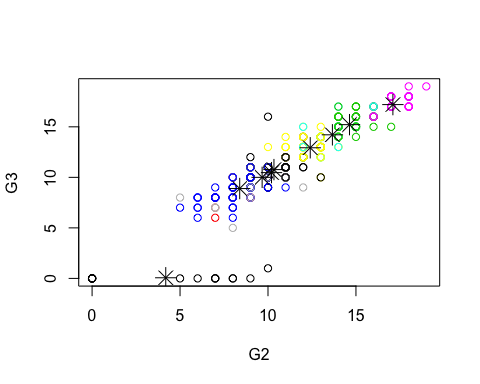
## K-means clustering with 9 clusters of sizes 133, 30, 79, 74, 62, 40, 137, 78, 16  
##   
## Cluster means:  
## age Medu Fedu traveltime studytime failures famrel  
## 1 16.46617 2.270677 2.187970 1.699248 1.819549 0.15037594 4.157895  
## 2 17.16667 2.500000 2.633333 1.600000 1.633333 0.50000000 3.833333  
## 3 16.62025 3.126582 2.683544 1.354430 2.227848 0.02531646 4.126582  
## 4 17.12162 2.000000 1.837838 1.837838 1.689189 0.78378378 3.662162  
## 5 16.61290 2.725806 2.612903 1.451613 1.983871 0.01612903 3.854839  
## 6 17.00000 3.075000 2.600000 1.375000 2.375000 0.00000000 4.075000  
## 7 16.37956 2.583942 2.401460 1.467153 2.094891 0.01459854 3.970803  
## 8 17.24359 2.269231 2.000000 1.628205 1.692308 0.42307692 3.564103  
## 9 17.68750 2.312500 1.750000 1.750000 1.500000 0.81250000 3.875000  
## freetime goout Dalc Walc health absences G1  
## 1 3.360902 3.398496 1.646617 2.503759 4.045113 1.586466 10.270677  
## 2 3.166667 3.200000 2.066667 2.733333 3.566667 18.000000 10.400000  
## 3 3.240506 3.177215 1.139241 1.911392 3.443038 1.632911 14.151899  
## 4 3.337838 3.094595 1.391892 2.175676 3.351351 2.581081 7.729730  
## 5 3.000000 3.274194 1.193548 2.241935 3.564516 7.645161 13.322581  
## 6 3.000000 2.925000 1.375000 1.850000 3.075000 0.600000 16.625000  
## 7 2.992701 2.854015 1.306569 1.934307 3.291971 1.051095 12.313869  
## 8 3.141026 3.551282 2.064103 2.961538 3.500000 8.487179 9.435897  
## 9 3.625000 3.187500 2.000000 2.750000 3.875000 0.000000 7.250000  
## G2 G3 GP MS  
## 1 10.338346 10.766917 0.6090226 0.3909774  
## 2 10.200000 10.466667 1.0000000 0.0000000  
## 3 14.632911 15.215190 0.7721519 0.2278481  
## 4 8.391892 8.891892 0.2432432 0.7567568  
## 5 13.661290 14.209677 0.8709677 0.1290323  
## 6 17.100000 17.200000 0.6500000 0.3500000  
## 7 12.401460 12.927007 0.7664234 0.2335766  
## 8 9.666667 10.000000 0.5897436 0.4102564  
## 9 4.187500 0.062500 0.1250000 0.8750000  
##   
## Clustering vector:  
## [1] 4 1 5 3 7 5 7 7 6 7 3 7 7 7 3 6 5 3 4 5 7 7 7 1 1 8 5 1 7 1 1 3 3 7 7  
## [36] 1 3 7 7 5 2 8 3 1 2 1 5 6 7 7 7 3 1 7 7 7 3 5 7 3 6 1 7 7 7 3 7 1 1 5  
## [71] 7 1 1 7 1 1 7 7 4 2 7 1 1 7 7 1 5 3 8 8 1 5 7 3 7 7 8 7 7 7 8 6 7 2 6  
## [106] 8 1 3 8 3 3 1 8 5 1 5 3 5 2 3 7 5 7 7 1 1 1 1 5 7 1 8 8 5 7 5 4 8 1 3  
## [141] 1 7 1 8 8 4 1 5 4 8 2 3 1 7 1 2 7 1 1 5 8 2 1 9 8 7 8 7 4 2 8 1 9 1 8  
## [176] 4 1 4 4 8 1 6 7 6 4 6 1 7 3 1 7 1 8 7 1 7 6 2 5 5 1 3 5 8 5 7 2 1 7 5  
## [211] 3 2 2 1 3 1 3 2 5 2 1 3 7 7 7 2 7 8 5 3 2 8 7 7 1 5 7 8 3 4 6 1 1 3 7  
## [246] 7 7 7 4 7 6 1 7 2 2 2 2 8 3 8 1 7 2 2 8 3 3 5 7 5 3 1 3 8 8 8 5 1 7 8  
## [281] 8 1 8 1 4 7 7 1 7 3 1 1 7 7 7 7 6 1 7 8 3 1 1 7 7 8 5 7 8 1 8 2 5 7 6  
## [316] 5 5 5 2 5 7 5 4 1 8 2 2 6 7 5 5 7 6 7 3 6 3 6 6 3 3 7 3 6 6 3 7 4 6 6  
## [351] 1 3 1 1 7 7 3 3 5 3 5 8 7 7 3 5 7 7 1 8 4 1 7 1 6 3 3 7 6 1 7 3 1 1 1  
## [386] 1 3 5 8 1 5 5 7 3 3 7 5 2 7 3 3 3 7 5 3 2 5 1 8 5 3 3 6 2 5 8 5 3 5 8  
## [421] 5 4 7 8 7 8 1 6 4 1 1 1 4 7 4 1 4 7 1 3 9 4 7 8 1 8 1 8 5 7 3 1 1 4 4  
## [456] 1 3 1 7 1 1 7 7 7 8 8 1 7 7 7 7 1 3 4 1 7 4 8 4 4 4 1 1 1 4 1 1 4 4 4  
## [491] 4 8 4 8 4 7 3 1 7 6 8 5 4 7 7 1 8 1 1 6 6 1 8 4 4 8 3 3 8 9 4 4 4 4 4  
## [526] 3 3 1 4 1 8 1 4 3 7 8 5 5 1 8 8 1 1 4 1 8 3 1 1 6 7 7 4 1 1 5 1 4 1 7  
## [561] 1 8 7 9 1 1 4 9 4 8 4 4 4 8 8 8 8 8 4 1 4 4 4 9 1 4 9 4 4 4 4 7 7 7 6  
## [596] 6 6 9 8 5 7 4 7 9 4 9 6 1 4 7 9 5 1 1 7 3 7 6 3 5 3 7 4 6 4 1 9 4 8 4  
## [631] 6 1 4 3 6 8 6 9 3 9 9 3 4 4 1 3 8 8 1  
##   
## Within cluster sum of squares by cluster:  
## [1] 2461.293 1428.000 1264.405 1733.689 1531.242 625.100 2225.504 2074.321  
## [9] 529.875  
## (between\_SS / total\_SS = 65.5 %)  
##   
## Available components:  
##   
## [1] "cluster" "centers" "totss" "withinss"   
## [5] "tot.withinss" "betweenss" "size" "iter"   
## [9] "ifault"

Comparamos el resultado de la clusterización con las variables G1 y G2

plot(studentPor.mod %>% select(G1, G3), col = studentPorkmeans.clust$cluster)  
points(as.data.frame(studentPorkmeans.clust$centers) %>% select(G1, G3), pch = 8, cex = 2)



plot(studentPor.mod %>% select(G2, G3), col = studentPorkmeans.clust$cluster)  
points(as.data.frame(studentPorkmeans.clust$centers) %>% select(G2, G3), pch = 8, cex = 2)

 # 5. Conclusiones

Podemos ver que donde más alumnos de matemáticas hay concentrados es en el cluster 6, 7 y 8. Es en el cluster 7 donde 81 estudiantes aprueban esta asignatura, con una nota de 14 y cuya nota en el los distintos exámenes anteriores ha sido similar a la conseguida finalmente. Con referencia a los alumnos del cluster 6 (79 estudiantes) han aprobado con una nota cercana al 11.50 y vemos que es prácticamente la misma que la nota sacada en los examenes realizados con anterioridad. El cluster 4 es que suspende con una nota final de 0 habiendo suspendido también los anteriores exámenes. Por último vemos que hay un grupo de estudiantes de 30 personas, incluido en el cluster 2, que sacan muy buenas notas finales, cercanas al 17.50 y en los que vemos que los exámenes anteriores tienen notas similares.

En relación a los estudiantes portugueses se muestra que las notas son superiores a los de matemáticas en la mayoría de los cluster establecidos, en concreto en el cluster 6, hay 40 estudiantes que tienen una nota de 17, similar en los exámenes anteriores. El cluster más numeroso es el 7 con 137 estudiantes con una nota final cercana a los 13 puntos, aceptable para tan alto número de estudiantes.

# FIN