

FML_Assignment_1

Swetha Ramreddy

2023-09-10

#

Downloading and Activating the required Libraries

```
library(readr)
#install.packages("ggplot2")
library("ggplot2")
```

Reading the CSV file

```
data_shown <- read.csv("C:/Users/Public/Downloads/student-por.csv")
```

Summary of the CSV File

```
summary(data_shown)
```

```
##      school      sex      age      address
## Length:649    Length:649    Min.   :15.00    Length:649
## Class :character Class :character 1st Qu.:16.00    Class :character
## Mode  :character Mode  :character Median :17.00    Mode  :character
##                               Mean  :16.74
##                               3rd Qu.:18.00
##                               Max.   :22.00
##      famsize      Pstatus      Medu      Fedu
## Length:649    Length:649    Min.   :0.000    Min.   :0.000
## Class :character Class :character 1st Qu.:2.000    1st Qu.:1.000
## Mode  :character Mode  :character Median :2.000    Median :2.000
##                               Mean  :2.515    Mean  :2.307
##                               3rd Qu.:4.000    3rd Qu.:3.000
##                               Max.   :4.000    Max.   :4.000
##      Mjob      Fjob      reason      guardian
## Length:649    Length:649    Length:649    Length:649
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##      traveltime      studytime      failures      schoolsup
## Min.   :1.000    Min.   :1.000    Min.   :0.0000    Length:649
## 1st Qu.:1.000    1st Qu.:1.000    1st Qu.:0.0000    Class :character
## Median :1.000    Median :2.000    Median :0.0000    Mode  :character
## Mean   :1.569    Mean   :1.931    Mean   :0.2219
## 3rd Qu.:2.000    3rd Qu.:2.000    3rd Qu.:0.0000
## Max.   :4.000    Max.   :4.000    Max.   :3.0000
```

```
##      famsup          paid          activities          nursery
## Length:649      Length:649      Length:649      Length:649
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##      higher          internet          romantic          famrel
## Length:649      Length:649      Length:649      Min.    :1.000
## Class :character Class :character Class :character 1st Qu.:4.000
## Mode  :character Mode  :character Mode  :character Median :4.000
##                                     Mean  :3.931
##                                     3rd Qu.:5.000
##                                     Max.   :5.000
##
##      freetime          goout          Dalc          Walc          health
## Min.    :1.00      Min.    :1.000      Min.    :1.000      Min.    :1.00      Min.
## 1st Qu.:3.00      1st Qu.:2.000      1st Qu.:1.000      1st Qu.:1.00      1st
## Median :3.00      Median :3.000      Median :1.000      Median :2.00      Median
## Mean    :3.18      Mean    :3.185      Mean    :1.502      Mean    :2.28      Mean
## 3rd Qu.:4.00      3rd Qu.:4.000      3rd Qu.:2.000      3rd Qu.:3.00      3rd
## Max.    :5.00      Max.    :5.000      Max.    :5.000      Max.    :5.00      Max.
##
##      absences          G1          G2          G3
## Min.    : 0.000      Min.    : 0.0      Min.    : 0.00      Min.    : 0.00
## 1st Qu.: 0.000      1st Qu.:10.0      1st Qu.:10.00      1st Qu.:10.00
## Median : 2.000      Median :11.0      Median :11.00      Median :12.00
## Mean    : 3.659      Mean    :11.4      Mean    :11.57      Mean    :11.91
## 3rd Qu.: 6.000      3rd Qu.:13.0      3rd Qu.:13.00      3rd Qu.:14.00
## Max.    :32.000      Max.    :19.0      Max.    :19.00      Max.    :19.00
```

3.Print out descriptive statistics for a selection of Quantitative and Qualitative variables.

Prints the first 10 Rows of the Dataset.

```
head(data_shown, n=10)
```

```
##      school sex age address famsize Pstatus Medu Fedu      Mjob      Fjob
## 1      GP   F  18      U      GT3      A    4    4    at_home teacher
## 2      GP   F  17      U      GT3      T    1    1    at_home  other
## 3      GP   F  15      U      LE3      T    1    1    at_home  other
## 4      GP   F  15      U      GT3      T    4    2    health services
## 5      GP   F  16      U      GT3      T    3    3    other    other
## 6      GP   M  16      U      LE3      T    4    3    services other
## 7      GP   M  16      U      LE3      T    2    2    other    other
## 8      GP   F  17      U      GT3      A    4    4    other    teacher
```

```

## 9      GP  M  15      U      LE3      A  3  2 services  other
## 10     GP  M  15      U      GT3      T  3  4  other  other
##      reason guardian traveltime studytime failures schoolsup famsup paid
## 1      course  mother      2      2      0      yes  no  no
## 2      course  father      1      2      0      no   yes no
## 3      other  mother      1      2      0      yes  no  no
## 4      home  mother      1      3      0      no   yes no
## 5      home  father      1      2      0      no   yes no
## 6  reputation  mother      1      2      0      no   yes no
## 7      home  mother      1      2      0      no   no  no
## 8      home  mother      2      2      0      yes  yes no
## 9      home  mother      1      2      0      no   yes no
## 10     home  mother      1      2      0      no   yes no
##      activities nursery higher internet romantic famrel freetime goout Dalc
Walcl
## 1      no      yes  yes      no      no      4      3      4      1
1
## 2      no      no   yes      yes      no      5      3      3      1
1
## 3      no      yes  yes      yes      no      4      3      2      2
3
## 4      yes     yes  yes      yes      yes      3      2      2      1
1
## 5      no      yes  yes      no      no      4      3      2      1
2
## 6      yes     yes  yes      yes      no      5      4      2      1
2
## 7      no      yes  yes      yes      no      4      4      4      1
1
## 8      no      yes  yes      no      no      4      1      4      1
1
## 9      no      yes  yes      yes      no      4      2      2      1
1
## 10     yes     yes  yes      yes      no      5      5      1      1
1
##      health absences G1 G2 G3
## 1      3      4  0 11 11
## 2      3      2  9 11 11
## 3      3      6 12 13 12
## 4      5      0 14 14 14
## 5      5      0 11 13 13
## 6      5      6 12 12 13
## 7      3      0 13 12 13
## 8      1      2 10 13 13
## 9      1      0 15 16 17
## 10     5      0 12 12 13

```

Prints the last 10 entries of the dataset.

```
tail(data_shown, n=10)
```

##	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob
## 640	MS	M	19	R	GT3	T	1	1	other	services
## 641	MS	M	18	R	GT3	T	4	2	other	other
## 642	MS	F	18	R	GT3	T	2	2	at_home	other
## 643	MS	F	17	U	GT3	T	4	3	teacher	other
## 644	MS	F	18	R	GT3	T	4	4	teacher	at_home
## 645	MS	F	19	R	GT3	T	2	3	services	other
## 646	MS	F	18	U	LE3	T	3	1	teacher	services
## 647	MS	F	18	U	GT3	T	1	1	other	other
## 648	MS	M	17	U	LE3	T	3	1	services	services
## 649	MS	M	18	R	LE3	T	3	2	services	other
##	reason	guardian	traveltime	studytime	failures	schoolsup	famsup			
paid										
## 640	other	mother	2	1	1	no	no			
## 641	home	father	2	1	1	no	no			
## 642	other	mother	2	3	0	no	no			
## 643	other	mother	2	2	0	no	no			
## 644	reputation	mother	3	1	0	no	yes			
## 645	course	mother	1	3	1	no	no			
## 646	course	mother	1	2	0	no	yes			
## 647	course	mother	2	2	0	no	no			
## 648	course	mother	2	1	0	no	no			
## 649	course	mother	3	1	0	no	no			
##	activities	nursery	higher	internet	romantic	famrel	freetime	goout	Dalc	Walc
## 640	no	yes	yes	no	no	4	3	2	1	3
## 641	no	yes	yes	no	no	5	4	3	4	3
## 642	no	yes	yes	no	no	5	3	3	1	3
## 643	no	yes	yes	yes	no	5	5	4	1	1
## 644	yes	yes	yes	yes	yes	4	4	3	2	2
## 645	yes	no	yes	yes	no	5	4	2	1	2
## 646	no	yes	yes	yes	no	4	3	4	1	1
## 647	yes	yes	yes	no	no	1	1	1	1	1

```

1
## 648      no      no    yes      yes      no      2      4      5      3
4
## 649      no      no    yes      yes      no      4      4      1      3
4
##      health absences G1 G2 G3
## 640      5          0  5  8  0
## 641      3          0  7  7  0
## 642      4          0 14 17 15
## 643      1          0  6  9 11
## 644      5          4  7  9 10
## 645      5          4 10 11 10
## 646      1          4 15 15 16
## 647      5          6 11 12  9
## 648      2          6 10 10 10
## 649      5          4 10 11 11

```

Total number of rows and columns in the dataset

```
nrow(data_shown)
```

```
## [1] 649
```

```
ncol(data_shown)
```

```
## [1] 33
```

Viewing the dataset

```
View(data_shown)
```

Changing column name of a specific column from the dataframe

```
colnames(data_shown)[colnames(data_shown) == "famsize"] <- "Family_Size"
colnames(data_shown)
```

```

## [1] "school"      "sex"         "age"         "address"     "Family_Size"
## [6] "Pstatus"    "Medu"        "Fedu"        "Mjob"        "Fjob"
## [11] "reason"     "guardian"    "traveltime"  "studytime"   "failures"
## [16] "schoolsup"  "famsup"      "paid"        "activities"  "nursery"
## [21] "higher"     "internet"    "romantic"    "famrel"      "freetime"
## [26] "goout"      "Dalc"        "Walc"        "health"      "absences"
## [31] "G1"         "G2"          "G3"

```

4.Transform at least one variable. It doesn't matter what the transformation is.

Transforming a column whose data type is character to a factor data type. This is Qualitative Transformation.

```
data_shown$Family_Size <- factor(data_shown$Family_Size)
levels(data_shown$Family_Size)
```

```
## [1] "GT3" "LE3"
```

```
is.factor(data_shown$Family_Size)
```

```
## [1] TRUE
```

Quantitative Transformation for a column

```
sqrt(data_shown$Medu)
```

```
## [1] 2.000000 1.000000 1.000000 2.000000 1.732051 2.000000 1.414214
2.000000
## [9] 1.732051 1.732051 2.000000 1.414214 2.000000 2.000000 1.414214
2.000000
## [17] 2.000000 1.732051 1.732051 2.000000 2.000000 2.000000 2.000000
1.414214
## [25] 1.414214 1.414214 1.414214 2.000000 1.732051 2.000000 2.000000
2.000000
## [33] 2.000000 1.732051 1.732051 1.414214 2.000000 2.000000 1.732051
1.414214
## [41] 1.414214 2.000000 2.000000 1.414214 1.414214 2.000000 1.732051
2.000000
## [49] 2.000000 2.000000 1.414214 2.000000 2.000000 2.000000 1.732051
1.414214
## [57] 2.000000 2.000000 1.000000 2.000000 2.000000 1.000000 1.000000
2.000000
## [65] 2.000000 2.000000 2.000000 1.732051 1.414214 1.732051 1.732051
2.000000
## [73] 1.000000 1.732051 1.732051 2.000000 2.000000 1.414214 1.414214
1.732051
## [81] 1.414214 1.414214 1.732051 1.414214 1.000000 2.000000 1.414214
2.000000
## [89] 1.414214 2.000000 1.732051 2.000000 1.732051 2.000000 1.414214
1.000000
## [97] 2.000000 1.414214 2.000000 2.000000 2.000000 2.000000 2.000000
1.732051
## [105] 1.732051 1.732051 1.414214 1.732051 2.000000 2.000000 2.000000
1.732051
## [113] 1.414214 2.000000 1.414214 2.000000 2.000000 1.732051 1.000000
1.732051
## [121] 1.000000 1.414214 1.414214 2.000000 1.414214 1.732051 1.732051
0.000000
## [129] 2.000000 1.414214 1.732051 1.414214 1.414214 1.000000 1.414214
1.732051
## [137] 1.414214 1.414214 1.732051 2.000000 1.732051 1.000000 1.732051
1.732051
## [145] 1.000000 1.732051 1.000000 2.000000 1.000000 2.000000 1.732051
2.000000
## [153] 1.414214 2.000000 1.000000 1.414214 1.000000 1.414214 1.732051
1.000000
## [161] 1.000000 2.000000 1.414214 1.000000 1.414214 1.732051 1.732051
2.000000
## [169] 1.414214 1.000000 1.000000 1.414214 1.732051 1.000000 1.414214
```

1.732051
[177] 1.000000 1.000000 1.000000 1.732051 1.414214 2.000000 1.414214
2.000000
[185] 1.732051 1.000000 2.000000 1.000000 1.732051 2.000000 1.414214
1.732051
[193] 2.000000 2.000000 2.000000 1.732051 1.414214 1.732051 1.732051
1.732051
[201] 1.000000 1.414214 1.732051 1.000000 1.414214 1.000000 1.000000
1.732051
[209] 1.414214 1.414214 2.000000 1.732051 2.000000 2.000000 2.000000
1.414214
[217] 1.000000 1.414214 1.414214 1.732051 1.732051 2.000000 1.000000
2.000000
[225] 1.732051 2.000000 1.414214 1.414214 2.000000 1.732051 2.000000
1.732051
[233] 1.414214 1.414214 1.414214 1.000000 1.414214 1.414214 2.000000
1.732051
[241] 1.732051 1.414214 1.414214 1.414214 2.000000 1.414214 2.000000
2.000000
[249] 1.000000 1.732051 1.414214 1.414214 2.000000 1.414214 1.414214
1.732051
[257] 1.414214 2.000000 1.732051 1.000000 1.000000 1.414214 1.000000
1.414214
[265] 1.414214 1.000000 2.000000 2.000000 2.000000 2.000000 2.000000
2.000000
[273] 1.000000 1.732051 1.000000 1.414214 1.414214 1.000000 1.414214
1.732051
[281] 1.732051 0.000000 1.732051 1.732051 1.414214 1.414214 1.414214
1.000000
[289] 1.414214 2.000000 1.414214 1.414214 2.000000 1.000000 1.414214
2.000000
[297] 1.414214 1.414214 1.414214 1.414214 2.000000 2.000000 1.732051
1.732051
[305] 1.414214 1.414214 1.732051 1.732051 2.000000 2.000000 1.414214
1.732051
[313] 1.414214 1.000000 1.000000 1.414214 1.414214 1.414214 1.732051
2.000000
[321] 2.000000 1.414214 1.732051 2.000000 1.000000 2.000000 1.732051
1.000000
[329] 1.000000 1.414214 1.414214 1.000000 1.414214 1.000000 1.414214
2.000000
[337] 2.000000 2.000000 1.732051 1.732051 1.732051 2.000000 2.000000
2.000000
[345] 2.000000 2.000000 2.000000 2.000000 2.000000 1.732051 1.414214
1.732051
[353] 1.732051 1.000000 1.414214 1.414214 2.000000 2.000000 2.000000
1.732051
[361] 1.732051 2.000000 2.000000 1.732051 2.000000 2.000000 1.732051
1.414214
[369] 2.000000 1.000000 1.000000 1.414214 1.414214 1.414214 1.732051

0.000000
[377] 1.000000 2.000000 1.732051 1.414214 2.000000 2.000000 1.732051
1.414214
[385] 1.732051 1.414214 1.414214 1.414214 1.732051 1.414214 1.414214
1.732051
[393] 1.732051 1.732051 1.732051 2.000000 1.732051 1.414214 1.414214
1.732051
[401] 2.000000 2.000000 2.000000 1.414214 1.414214 1.414214 1.732051
2.000000
[409] 1.414214 1.732051 1.414214 2.000000 2.000000 1.000000 1.000000
2.000000
[417] 1.732051 1.732051 1.414214 1.414214 1.414214 1.000000 1.732051
1.000000
[425] 1.414214 1.000000 1.732051 1.414214 1.414214 2.000000 1.000000
1.732051
[433] 1.000000 2.000000 1.414214 1.000000 1.000000 0.000000 1.414214
1.732051
[441] 1.000000 1.000000 2.000000 2.000000 1.732051 1.000000 1.732051
1.414214
[449] 2.000000 1.000000 1.414214 1.000000 1.414214 1.000000 1.000000
1.414214
[457] 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000
1.414214
[465] 1.000000 1.000000 1.732051 1.414214 1.414214 2.000000 1.732051
1.414214
[473] 1.414214 1.414214 1.000000 1.414214 1.414214 1.732051 1.000000
1.414214
[481] 1.414214 1.414214 1.000000 1.414214 1.414214 1.414214 1.414214
1.732051
[489] 1.000000 1.000000 1.000000 1.000000 1.414214 0.000000 1.000000
1.732051
[497] 1.000000 1.732051 1.414214 1.414214 1.000000 2.000000 1.414214
1.000000
[505] 1.414214 1.000000 1.732051 1.000000 1.414214 1.000000 1.414214
1.000000
[513] 1.000000 1.732051 1.732051 1.000000 2.000000 1.000000 1.414214
1.414214
[521] 1.000000 1.414214 2.000000 2.000000 1.732051 2.000000 2.000000
1.414214
[529] 1.414214 2.000000 1.732051 1.000000 1.414214 1.414214 2.000000
1.000000
[537] 1.414214 1.414214 1.414214 2.000000 1.732051 1.414214 2.000000
2.000000
[545] 1.000000 1.732051 2.000000 2.000000 1.000000 2.000000 1.414214
1.732051
[553] 1.000000 1.000000 1.000000 1.000000 1.000000 1.732051 1.414214
1.000000
[561] 1.000000 1.000000 1.414214 1.414214 1.732051 1.000000 1.000000
1.732051
[569] 1.732051 1.732051 1.414214 1.414214 1.732051 1.000000 1.000000


```

2.000000
## [577] 1.732051 1.000000 1.732051 1.000000 1.000000 1.000000 1.000000
1.414214
## [585] 0.000000 1.732051 2.000000 1.414214 2.000000 1.000000 1.000000
1.414214
## [593] 1.732051 1.000000 2.000000 2.000000 2.000000 1.414214 1.000000
2.000000
## [601] 1.000000 1.414214 1.414214 2.000000 1.000000 1.000000 2.000000
1.414214
## [609] 1.000000 1.000000 1.000000 1.414214 2.000000 1.732051 2.000000
1.732051
## [617] 1.000000 1.000000 1.000000 1.732051 2.000000 1.000000 1.000000
2.000000
## [625] 1.000000 1.414214 2.000000 1.000000 1.414214 1.000000 2.000000
1.000000
## [633] 1.000000 2.000000 1.732051 1.732051 2.000000 1.414214 1.414214
1.000000
## [641] 2.000000 1.414214 2.000000 2.000000 1.414214 1.732051 1.000000
1.732051
## [649] 1.732051

```

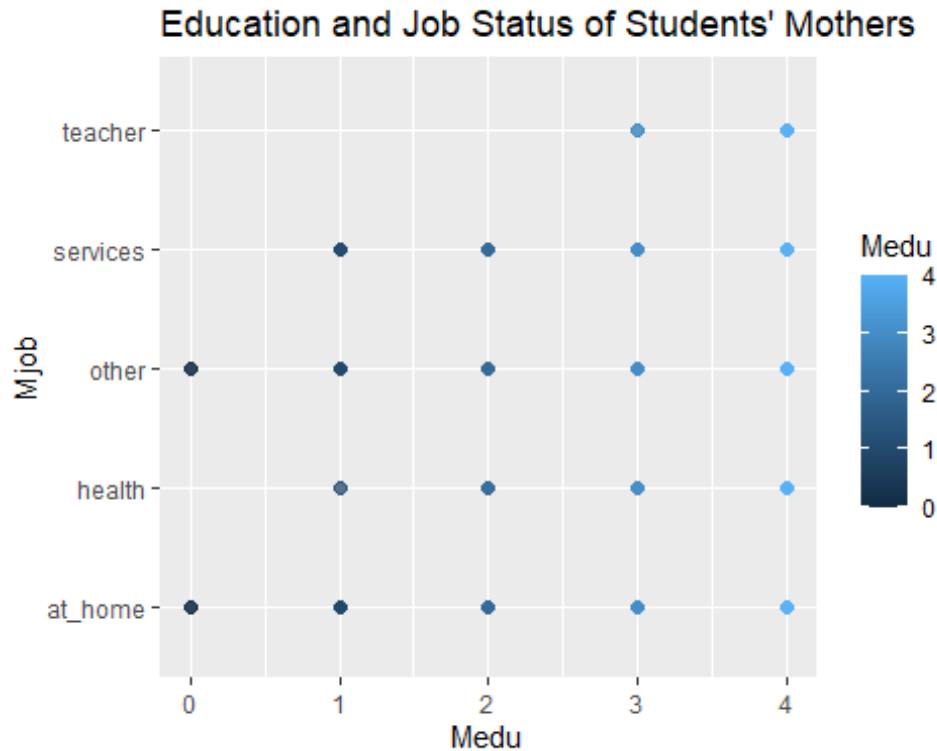
5a. Plot scatter plot - Quantitative Variable

```

qplot(data=data_shown, x = Medu, y = Mjob,
      size = I(2), color=Medu, alpha=I(0.5),
      main = "Education and Job Status of Students' Mothers")

## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

```



5b. Create a bar plot for Study Time Frequency for one Qualitative variable

```
ggplot(data_shown, aes(x = studytime, fill = studytime)) +
  geom_bar() +
  labs(
    x = "Studytime",
    y = "Frequency",
    title = "StudyTime and its Frequency among the students"
  ) +
  theme_minimal()
```

Warning: The following aesthetics were dropped during statistical transformation: fill

i This can happen when ggplot fails to infer the correct grouping structure in

the data.

i Did you forget to specify a `group` aesthetic or to convert a numerical

variable into a factor?

