**1**.Out of the demographic variable only **Age** was appropriate to average out given it was a numerical data type. When querying the questionnaire\_data dataset using **summary(age)** we are able to extrapolate the given data:

Min = 18

1st Quartile = 18

Median = 19

Mean = 20.44

3rd Quartile = 21

Max = 65

The youngest and oldest members of the dataset are 18 and 65 respectfully while the median age is 19 and the mean 20.44. We can safely conclude that the majority of the data as it refers to age is skewed left towards the younger end.

Sex, employment status, student status, are all categorical therefore we would make better use of frequency as opposed to mean.

* By using **table(questionnaire\_data$sex)** we create a frequency table that shows us our dataset consists of 428 female, 250 male, and 1 “.”
* By using **table(questionnaire\_data$employed)** we create a frequency table that shows us our dataset consists of 124 Full Time employees, 301 Part Time, 248 Unemployed, 5 “.” and 1 N/A
* By using **table(questionnaire\_data$student)** we create a frequency table that shows us our dataset consists of 638 full time students, 39 part time students, and 2 “.”

**2.**

**Studying Sleep**

Men= 10.07 Men= 6.46

Woman= 8.80 Woman= 6.5

Graphical user interface

Description automatically generated with low confidence

Graphical user interface

Description automatically generated with medium confidence

I chose studying and sleep. First I turned the dataset into a dataframe called **df** by using

**df <- questionnaire\_data**

After the dataframe was created I executed the queries in the above screenshot to find the mean for

**Studying :**

**Mean(df$studying[df$sex==’male])**

**Mean(df$studying[df$sex==’female])**

**Sleep:**

**Mean(df$sleep[df$sex==’male])**

**Mean(df$sleep[df$sex==’female])**