ML_Assignment_1_Sjain15

Sargam Jain

2/2/2022

#Read the Data

```
student_marks <- read.csv(file = "~/Downloads/student_marks.csv")
```

```
library(fBasics)
```

```
## Loading required package: timeDate
```

```
## Loading required package: timeSeries
```

Descriptive analytics

Calculating summary stats for all the columns of data set

```
summary(student_marks)
```

```
##      Name               Gender              DOB                Maths
##  Length:5           Length:5           Length:5           Min.   :25.0
##  Class :character   Class :character   Class :character   1st Qu.:55.0
##  Mode  :character   Mode  :character   Mode  :character   Median :58.0
##                                                           Mean   :58.2
##                                                           3rd Qu.:75.0
##                                                           Max.   :78.0
##
##     Physics      Chemistry        English        Biology       Economics
##  Min.   :45   Min.   :56.00   Min.   :46.0   Min.   :21.0   Min.   :52.0
##  1st Qu.:54   1st Qu.:72.50   1st Qu.:63.0   1st Qu.:54.0   1st Qu.:61.0
##  Median :55   Median :82.00   Median :64.0   Median :90.0   Median :77.0
##  Mean   :61   Mean   :77.25   Mean   :67.2   Mean   :71.2   Mean   :73.2
##  3rd Qu.:55   3rd Qu.:86.75   3rd Qu.:76.0   3rd Qu.:95.0   3rd Qu.:87.0
##  Max.   :96   Max.   :89.00   Max.   :87.0   Max.   :96.0   Max.   :89.0
##               NA's   :1
##     History         Civics
##  Min.   :56.0   Min.   : 2.0
##  1st Qu.:58.0   1st Qu.:45.0
##  Median :75.0   Median :53.0
##  Mean   :72.2   Mean   :47.8
##  3rd Qu.:83.0   3rd Qu.:65.0
##  Max.   :89.0   Max.   :74.0
##
```

Calculating specific summary statistics

##mean

```
mean(student_marks$Maths)
```

```
## [1] 58.2
```

```
mean(student_marks$Chemistry, na.rm = TRUE)
```

```
## [1] 77.25
```

##median

```
median(student_marks$Physics)
```

```
## [1] 55
```

```
median(student_marks$Chemistry, na.rm = TRUE)
```

```
## [1] 82
```

##standard deviation

```
sd(student_marks$Maths)
```

```
## [1] 21.13528
```

##max value in specific column of dataset

```
max(student_marks$Economics)
```

```
## [1] 89
```

##min value in specific column of dataset

```
min(student_marks$Biology)
```

```
## [1] 21
```

##Range

```
range(student_marks$History)
```

```
## [1] 56 89
```

##mode

```
mode(student_marks$Maths)
```

```
## [1] "numeric"
```

##Transforming a variable

#Example 1

```
transform(student_marks, Maths = Maths + 5)
```

```
##       Name Gender       DOB Maths Physics Chemistry English Biology Economics
## 1     John      M    5/4/88    60      45        56      87      21        52
## 2   Suresh      M    4/5/87    80      55        NA      64      90        61
## 3   Ramesh      M 25/5/1989    30      54        89      76      95        87
## 4  Jessica      F   12/8/90    83      55        86      63      54        89
## 5 Jennifer      F    2/9/89    63      96        78      46      96        77
##   History Civics
## 1      89     65
## 2      58      2
## 3      56     74
## 4      75     45
## 5      83     53
```

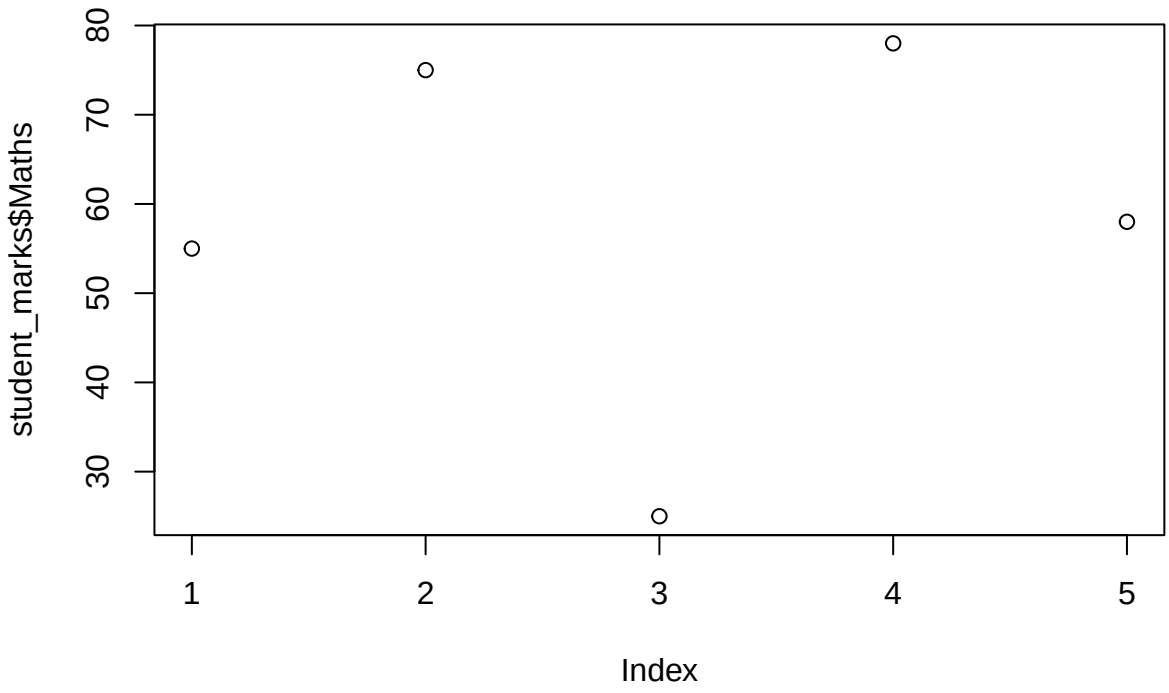
#Example 2

```
transform(student_marks, Accountancy = c(89, 88, 45, 37, 76))
```

```
##       Name Gender       DOB Maths Physics Chemistry English Biology Economics
## 1     John      M    5/4/88    55      45        56      87      21        52
## 2   Suresh      M    4/5/87    75      55        NA      64      90        61
## 3   Ramesh      M 25/5/1989    25      54        89      76      95        87
## 4  Jessica      F   12/8/90    78      55        86      63      54        89
## 5 Jennifer      F    2/9/89    58      96        78      46      96        77
##   History Civics Accountancy
## 1      89     65          89
## 2      58      2          88
## 3      56     74          45
## 4      75     45          37
## 5      83     53          76
```

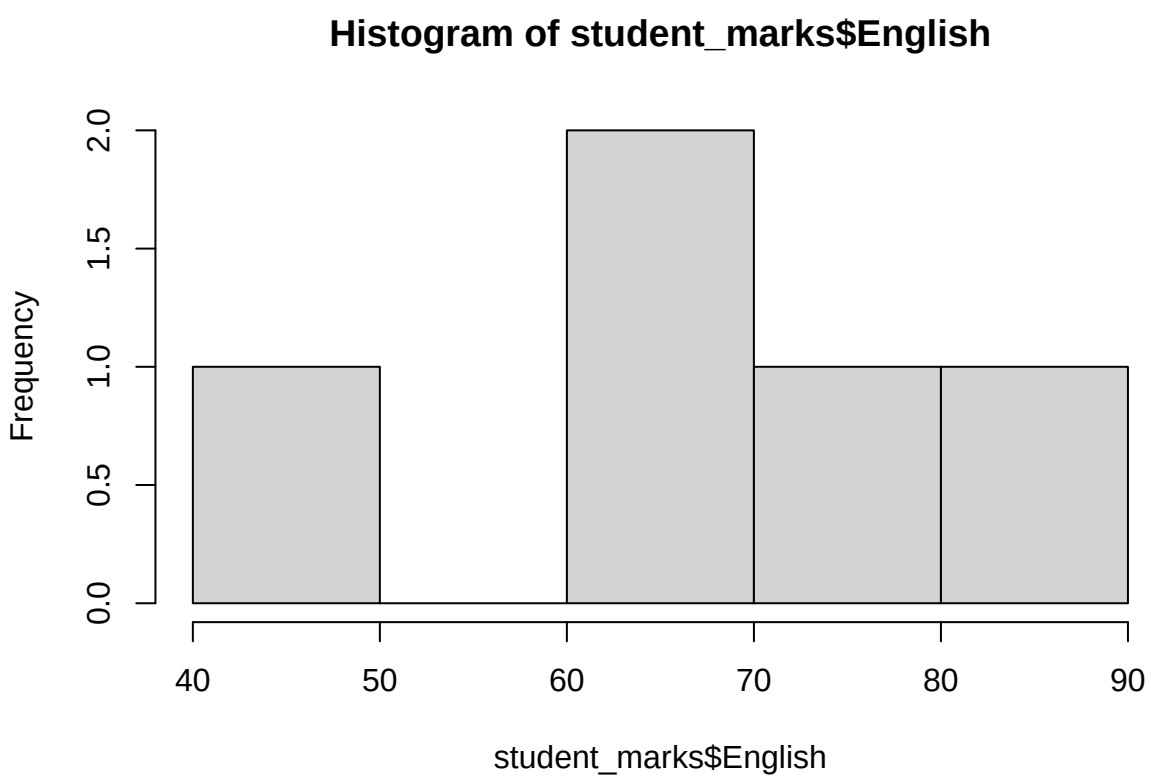
#Plot

```
plot(student_marks$Maths)
```



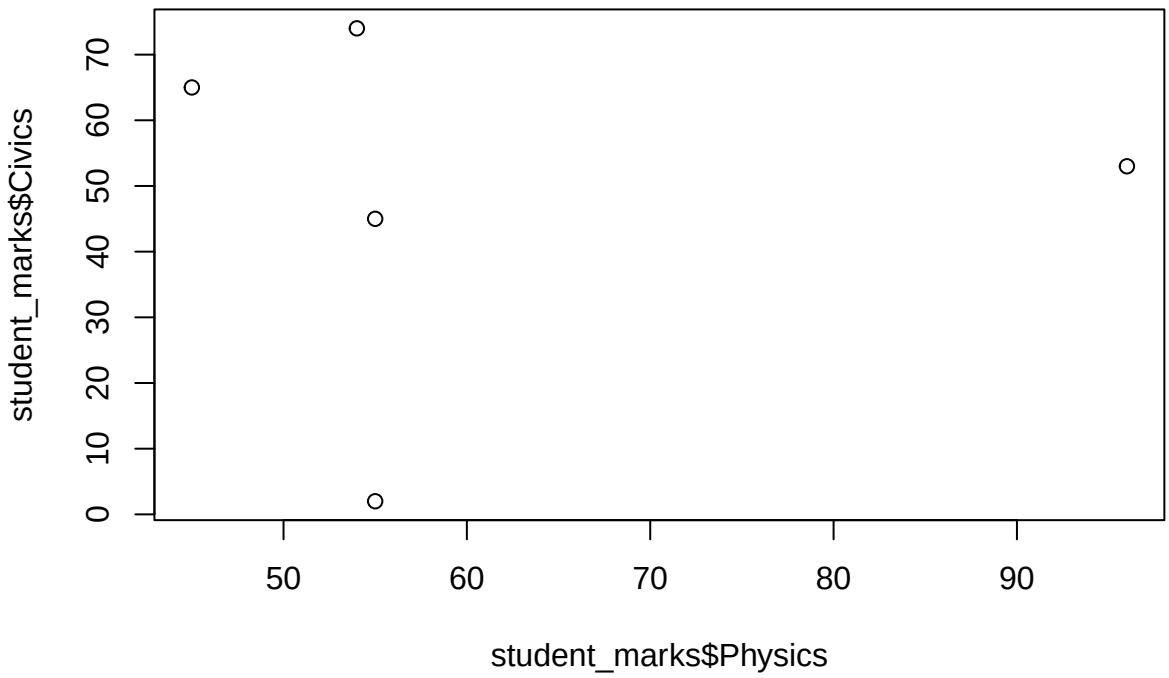
Plot Histogram

```
hist(student_marks$English)
```



##Scatter Plot

```
plot(student_marks$Physics, student_marks$Civics)
```


student_marks$Civics
0
10
20
30
40
50
60
70
50
60
70
80
90
student_marks$Physics