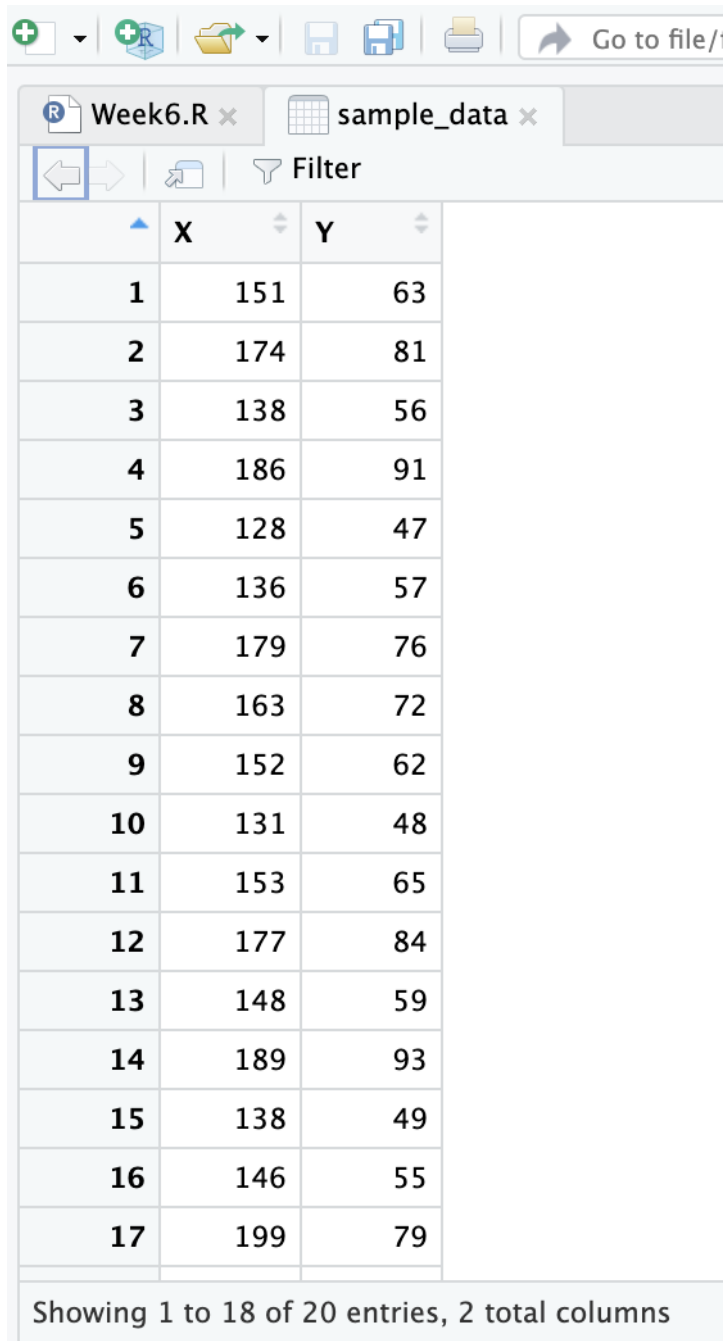


1) Uploading the data in R studio and viewing it

```
> data_file <- "/My System/simpliedata.csv"  
> sample_data <- read.csv(data_file)  
> View(sample_data)
```



The screenshot shows the R Studio interface. At the top, there is a toolbar with icons for file operations and a 'Go to file/' search bar. Below the toolbar, the 'Week6.R' script editor is open, and the 'sample_data' data frame is displayed in the viewer pane. The data frame has 20 rows and 2 columns, X and Y. The first 17 rows are visible, showing a sequence of values for X and Y. The bottom status bar indicates 'Showing 1 to 18 of 20 entries, 2 total columns'.

	X	Y
1	151	63
2	174	81
3	138	56
4	186	91
5	128	47
6	136	57
7	179	76
8	163	72
9	152	62
10	131	48
11	153	65
12	177	84
13	148	59
14	189	93
15	138	49
16	146	55
17	199	79

Showing 1 to 18 of 20 entries, 2 total columns

2) Summary of the dataset

```
> summary(sample_data)
```

Output:

```
> view(sample_data)
> summary(sample_data)
      X      Y
Min.   :128.0 Min.   :47.00
1st Qu.:138.0 1st Qu.:55.75
Median :152.5 Median :64.00
Mean    :156.9 Mean    :66.35
3rd Qu.:174.8 3rd Qu.:76.75
Max.    :199.0 Max.    :93.00
> |
```

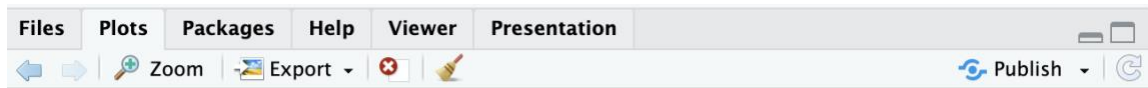
3) Structure of the dataset

```
> str(sample_data)
```

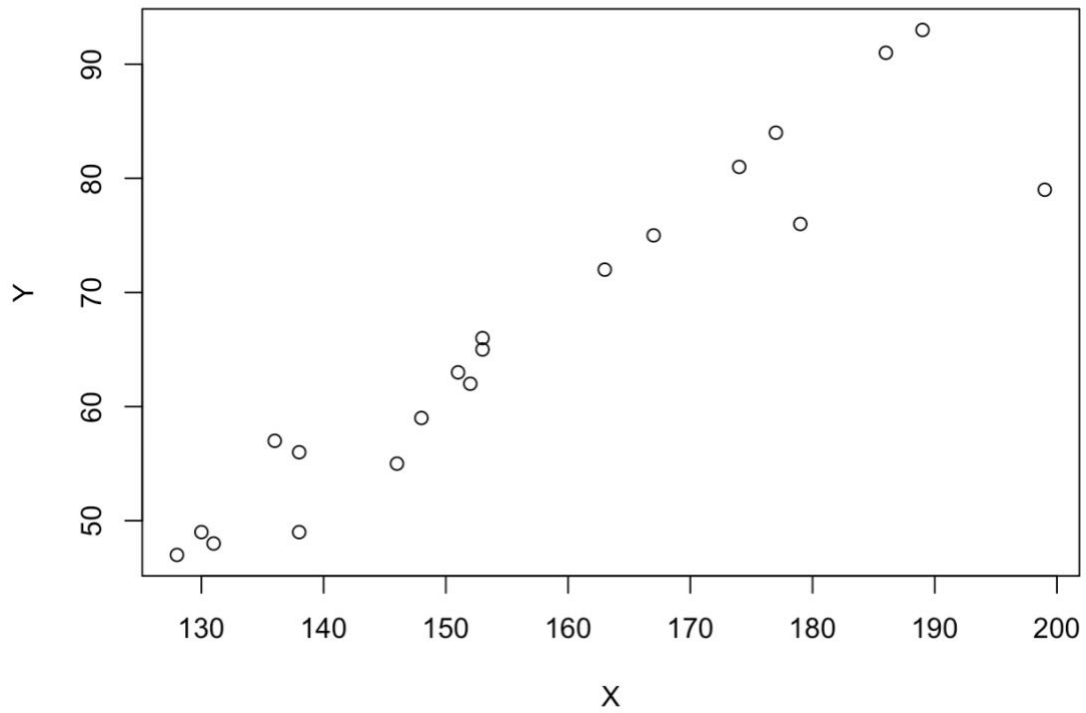
```
> str(sample_data)
'data.frame':  20 obs. of  2 variables:
 $ X: int  151 174 138 186 128 136 179 163 152 131 ...
 $ Y: int   63 81 56 91 47 57 76 72 62 48 ...
. |
```

4) Scatterplot

```
> plot(sample_data$X, sample_data$Y, main = "Scatterplot b/w X and Y", xlab = "X",
       ylab = "Y")
```



Scatterplot b/w X and Y



5) Simple Linear regression and it's summary

```
> simple_reg <- lm(sample_data$Y ~ sample_data$X, data = sample_data)
> summary(simple_reg)
```

```
> simple_reg <- lm(sample_data$Y ~ sample_data$X, data = sample_data)
> summary(simple_reg)
```

Call:

```
lm(formula = sample_data$Y ~ sample_data$X, data = sample_data)
```

Residuals:

Min	1Q	Median	3Q	Max
-14.1573	-1.7267	0.7701	2.6045	6.2102

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-33.55669	8.25032	-4.067	0.000723 ***
sample_data\$X	0.63675	0.05213	12.215	3.79e-10 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.846 on 18 degrees of freedom

Multiple R-squared: 0.8924, Adjusted R-squared: 0.8864

F-statistic: 149.2 on 1 and 18 DF, p-value: 3.788e-10

6) Intercept

```
> intercept = simple_reg$coefficients[1]
> intercept
```

```
> intercept = simple_reg$coefficients[1]
```

```
> intercept
```

```
(Intercept)
```

```
-33.55669
```

```
└─
```

7) Slope

```
> slope <- simple_reg$coefficients[2]
> slope
```

```

> slope <- simple_reg$coefficients[2]
> slope
sample_data$X
0.6367539
> |

```

8) Linear equation, by putting value of X we can get the value of Y

```

> X <- 151
> Y <- intercept + slope*X
> Y

```

```

> X <- 151
> Y <- intercept + slope*X
> Y
(Intercept)
62.59315
> |

```

9) Correlation

```

> cor(sample_data$X, sample_data$Y)
> cor(sample_data$X, sample_data$Y)
[1] 0.944644
> |

```