

## Practical 3

Exploring data: View() or print() (R).

The screenshot shows the RStudio interface with the following content:

```
1 install.packages(c("readr", "psych"))
2 library(readr)
3 library(psych)
4
5 my_data <- read.csv("C:/Users/mvluc/Downloads/simranpython/tips.csv")
6
7 head(my_data)
8
9 tail(my_data)
10
11
12 cat("Dimensions (Rows, columns): ", paste(dim(my_data), collapse = ", "), "\n")
13
14
15 str(my_data)
16
17 summary(my_data)
18
19
20 cat("Column Names: ", paste(names(my_data), collapse = ", "), "\n")
21
22 describe(my_data)
23
24
```

The console output shows the first six rows of the dataset:

```
> my_data <- read.csv("C:/Users/mvluc/Downloads/simranpython/tips.csv")
> my_data <- read.csv("C:/Users/mvluc/Downloads/simranpython/tips.csv")
> head(my_data)
  total_bill tip    sex smoker day time size
1    16.99 1.01 Female   No  Sun  Dinner    2
2    10.34 1.66   Male   No  Sun  Dinner    3
3    21.01 3.50   Male   No  Sun  Dinner    3
4    23.68 3.31   Male   No  Sun  Dinner    2
5    24.59 3.61 Female   No  Sun  Dinner    4
6    25.29 4.71   Male   No  Sun  Dinner    4
```

The Environment pane on the right shows the loaded objects: my\_data (244 obs. of 7 variables), student\_m (101 obs. of 11 variables), and tips (244 obs. of 7 variables).

The screenshot shows the RStudio interface with the following content:

```
> my_data <- read.csv("C:/Users/mvluc/Downloads/simranpython/tips.csv")
> my_data <- read.csv("C:/Users/mvluc/Downloads/simranpython/tips.csv")
> head(my_data)
  total_bill tip    sex smoker day time size
1    16.99 1.01 Female   No  Sun  Dinner    2
2    10.34 1.66   Male   No  Sun  Dinner    3
3    21.01 3.50   Male   No  Sun  Dinner    3
4    23.68 3.31   Male   No  Sun  Dinner    2
5    24.59 3.61 Female   No  Sun  Dinner    4
6    25.29 4.71   Male   No  Sun  Dinner    4

> tail(my_data)
  total_bill tip    sex smoker day time size
239   35.83 4.67 Female   No  Sat  Dinner    3
240   29.03 5.92   Male   No  Sat  Dinner    3
241   27.18 2.00 Female   Yes  Sat  Dinner    2
242   22.67 2.00   Male   Yes  Sat  Dinner    2
243   17.82 1.75   Male   No  Sat  Dinner    2
244   18.78 3.00 Female   No  Thur Dinner    2

> cat("Dimensions (Rows, Columns): ", paste(dim(my_data), collapse = ", "), "\n")
Dimensions (Rows, Columns):  244, 7

> str(my_data)
'data.frame':   244 obs. of  7 variables:
 $ total_bill: num  17 10.3 21 23.7 24.6 ...
 $ tip       : num  1.01 1.66 3.5 3.31 3.61 4.71 2 3.12 1.96 3.23 ...
 $ sex      : chr  "Female" "male" "male" "male" ...
 $ smoker   : chr  "No" "No" "No" "No" ...
 $ day      : chr  "Sun" "Sun" "Sun" "Sun" ...
 $ time     : chr  "Dinner" "Dinner" "Dinner" "Dinner" ...
 $ size     : int   2 3 3 2 4 4 2 2 2 ...

> summary(my_data)
  total_bill  tip      sex      smoker      day
Min.   : 3.07  Min.   : 1.000  Length:244  Length:244  Length:244
1st Qu.:13.35 1st Qu.: 2.000  Class :character  Class :character  Class :character
Median :17.80 Median : 2.900  Mode  :character  Mode  :character  Mode  :character
Mean   :19.79 Mean   : 2.998
3rd Qu.:24.13 3rd Qu.: 3.562
Max.   :50.81 Max.   :10.000

  time      size
Length:244  Min.   :1.00
Class :character 1st Qu.:2.00
Mode  :character Median :2.00
                Mean  :2.57
                3rd Qu.:3.00
```

The Environment pane on the right shows the loaded objects: my\_data (244 obs. of 7 variables), student\_m (101 obs. of 11 variables), and tips (244 obs. of 7 variables).