

PRACTICAL NO 3

AIM:- Exploring data: View() or print() (R)

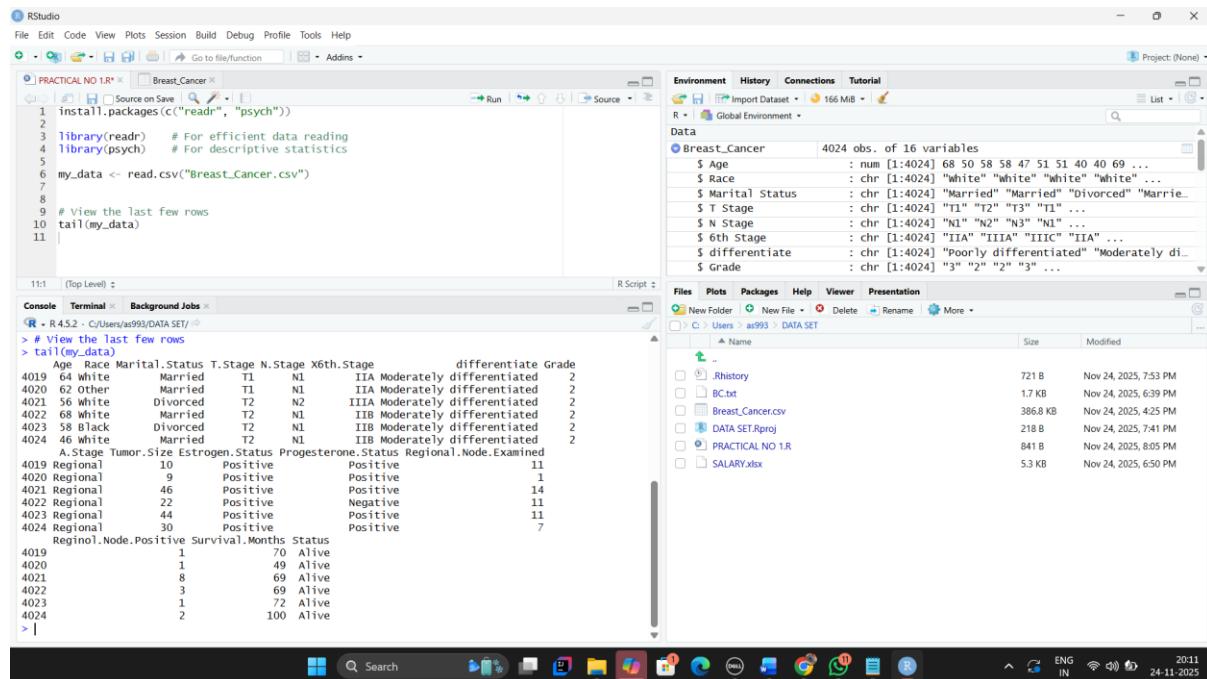
OUTPUT :-

1.The first few rows

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project Bar:** Project: (None)
- Header Bar:** PRACTICAL NO 1.R, Breast_Cancer.R, Go to file/function, Addins.
- Code Editor:** R Script pane containing R code for loading packages and reading the Breast Cancer dataset.
- Environment Tab:** Shows the Breast_Cancer dataset with 4024 observations and 16 variables.
- Data Tab:** Displays the structure of the Breast_Cancer dataset.
- Console Tab:** Shows the R environment (R-4.5.2) and the first few rows of the dataset.
- Plots Tab:** Not visible in the screenshot.
- Packages Tab:** Not visible in the screenshot.
- Help Tab:** Not visible in the screenshot.
- Viewer Tab:** Shows the raw data of the Breast_Cancer dataset.
- Presentation Tab:** Not visible in the screenshot.
- File Explorer:** Shows the local directory structure with files like .History, BC.txt, Breast_Cancer.csv, DATA SET.Rproj, PRACTICAL NO 1.R, and SALARY.xlsx.

2. The last few rows



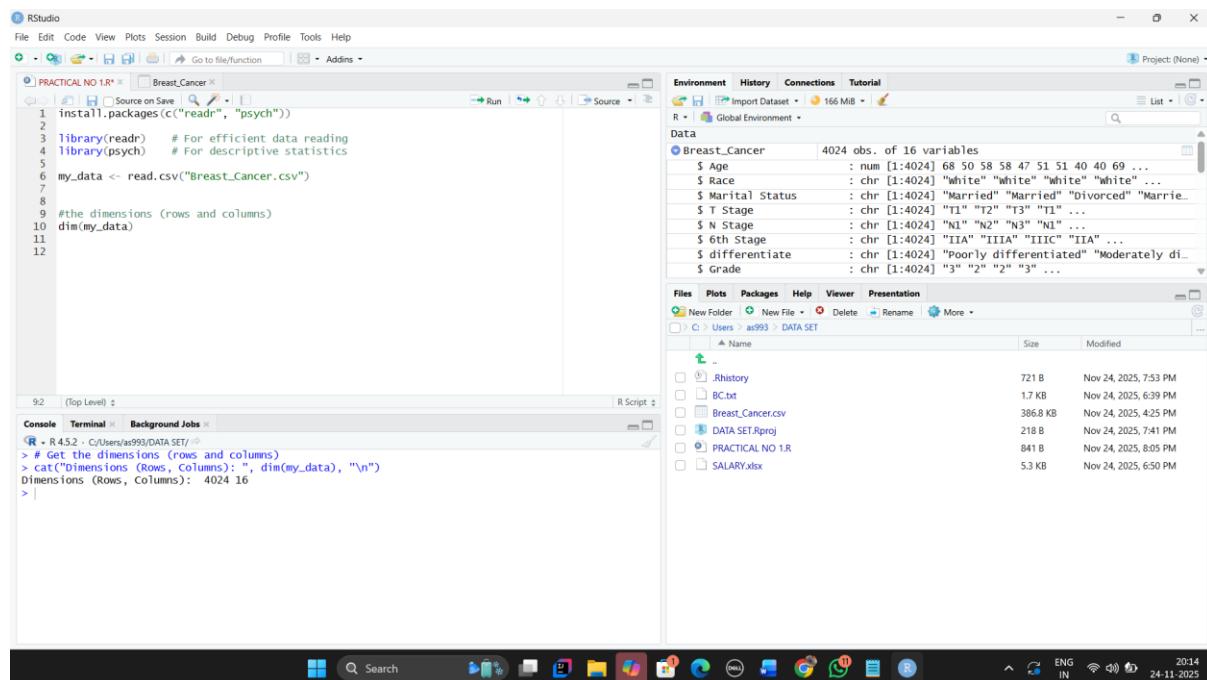
The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project (None).
- Console:**

```
R 4.5.2 · C:/Users/as993/DATA SET/ 
> # View the last few rows
> tail(my_data)
```
- Script Editor:**

```
1 install.packages(c("readr", "psych"))
2
3 library(readr) # For efficient data reading
4 library(psych) # For descriptive statistics
5
6 my_data <- read.csv("Breast_Cancer.csv")
7
8
9 # View the last few rows
10 tail(my_data)
11
```
- Environment View:** Shows the structure of the Breast_Cancer dataset with 4024 observations and 16 variables.
- Files View:** Shows the project directory structure including .Rhistory, BC.txt, Breast_Cancer.csv, DATA SET.Rproj, PRACTICAL NO 1.R, and SALARY.xlsx.
- System Tray:** Shows the date and time as 24-11-2025.

3. The dimensions (rows and columns)



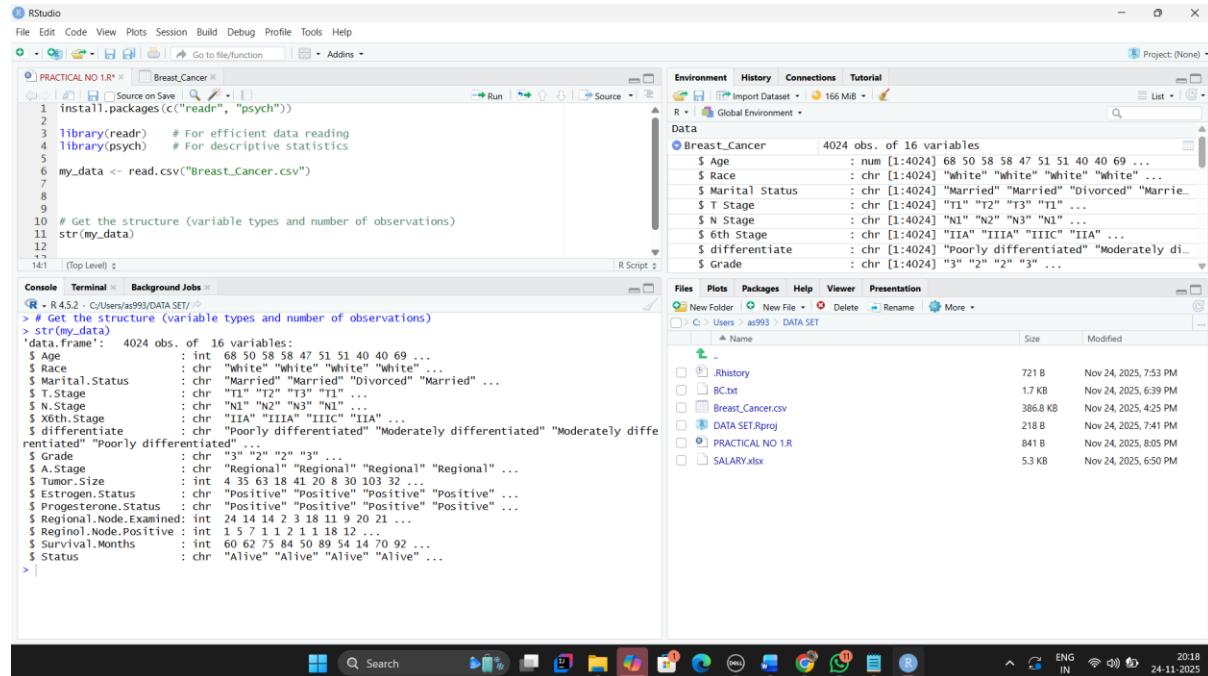
The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project (None).
- Console:**

```
R 4.5.2 · C:/Users/as993/DATA SET/ 
> # the dimensions (rows and columns)
> dim(my_data)
10
11
12
```
- Script Editor:**

```
1 install.packages(c("readr", "psych"))
2
3 library(readr) # For efficient data reading
4 library(psych) # For descriptive statistics
5
6 my_data <- read.csv("Breast_Cancer.csv")
7
8
9 # the dimensions (rows and columns)
10 dim(my_data)
11
12
```
- Environment View:** Shows the structure of the Breast_Cancer dataset with 4024 observations and 16 variables.
- Files View:** Shows the project directory structure including .Rhistory, BC.txt, Breast_Cancer.csv, DATA SET.Rproj, PRACTICAL NO 1.R, and SALARY.xlsx.
- System Tray:** Shows the date and time as 24-11-2025.

4.The structure (variable types and number of observations)



The screenshot shows the RStudio interface with the following details:

- Console:** Displays R code for loading the dataset and viewing its structure.
- Environment:** Shows the global environment with the `Breast_Cancer` dataset loaded.
- Files:** Shows the project structure with files like `BC.txt`, `Breast_Cancer.csv`, `DATA SET.Rproj`, `PRACTICAL NO.1.R`, and `SALARY.xlsx`.
- Taskbar:** Shows the Windows taskbar with various application icons and the date/time: 24-11-2018.

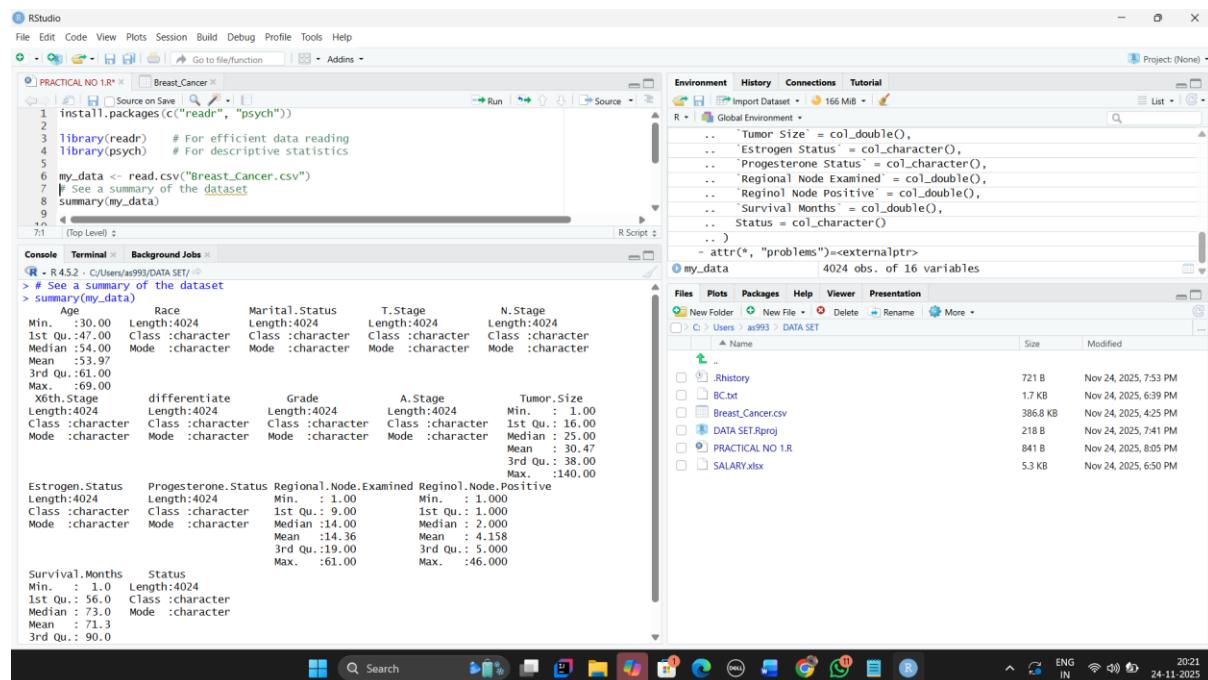
```

install.packages(c("readr", "psych"))
library(readr) # For efficient data reading
library(psych) # For descriptive statistics
my_data <- read.csv("Breast_Cancer.csv")
# Get the structure (variable types and number of observations)
summary(my_data)

'data.frame': 4024 obs. of 16 variables:
 $ Age : int 68 50 58 58 47 51 51 40 40 69 ...
 $ Race : chr "white" "white" "white" "white" ...
 $ Marital.Status: chr "Married" "Married" "Divorced" "Married" ...
 $ T.Stage : chr "T1" "T2" "T3" "T1" ...
 $ N.Stage: chr "N1" "N2" "N3" "N1" ...
 $ X6th.Stage: chr "IIA" "IIIA" "IIIC" "IIA" ...
 $ differentiate: chr "Poorly differentiated" "Moderately differentiated" "Moderately differentiated" "Poorly differentiated"
 $ Grade : chr "3" "2" "3" ...
 $ A.Stage : chr "Regional" "Regional" "Regional" "Regional" ...
 $ Tumor.Size : int 4 35 63 18 41 20 8 30 103 32 ...
 $ Estrogen.Status: chr "Positive" "Positive" "Positive" "Positive" ...
 $ Progesterone.Status: chr "Positive" "Positive" "Positive" "Positive" ...
 $ Regional.Node.Examined: int 24 14 14 2 3 18 11 9 20 21 ...
 $ Reginol.Node.Positive : int 1 5 7 1 1 2 1 1 18 12 ...
 $ Survival.Months : int 60 62 75 84 50 89 54 14 70 92 ...
 $ Status : chr "Alive" "Alive" "Alive" "Alive" ...

```

5.Summary of the dataset



The screenshot shows the RStudio interface with the following details:

- Console:** Displays R code for loading the dataset and viewing its summary statistics.
- Environment:** Shows the global environment with the `my_data` dataset loaded.
- Files:** Shows the project structure with files like `BC.txt`, `Breast_Cancer.csv`, `DATA SET.Rproj`, `PRACTICAL NO.1.R`, and `SALARY.xlsx`.
- Taskbar:** Shows the Windows taskbar with various application icons and the date/time: 24-11-2018.

```

install.packages(c("readr", "psych"))
library(readr) # For efficient data reading
library(psych) # For descriptive statistics
my_data <- read.csv("Breast_Cancer.csv")
# See a summary of the dataset
summary(my_data)

Age          Race        Marital.Status   T.Stage      N.Stage
Min. :40.00  Length:4024  Length:4024  Length:4024  Length:4024
1st Qu.:47.00 Class :character  Class :character  Class :character  Class :character
Median :54.00 Mode  :character  Mode  :character  Mode  :character  Mode  :character
Mean   :53.97
3rd Qu.:61.00
Max.  :69.00

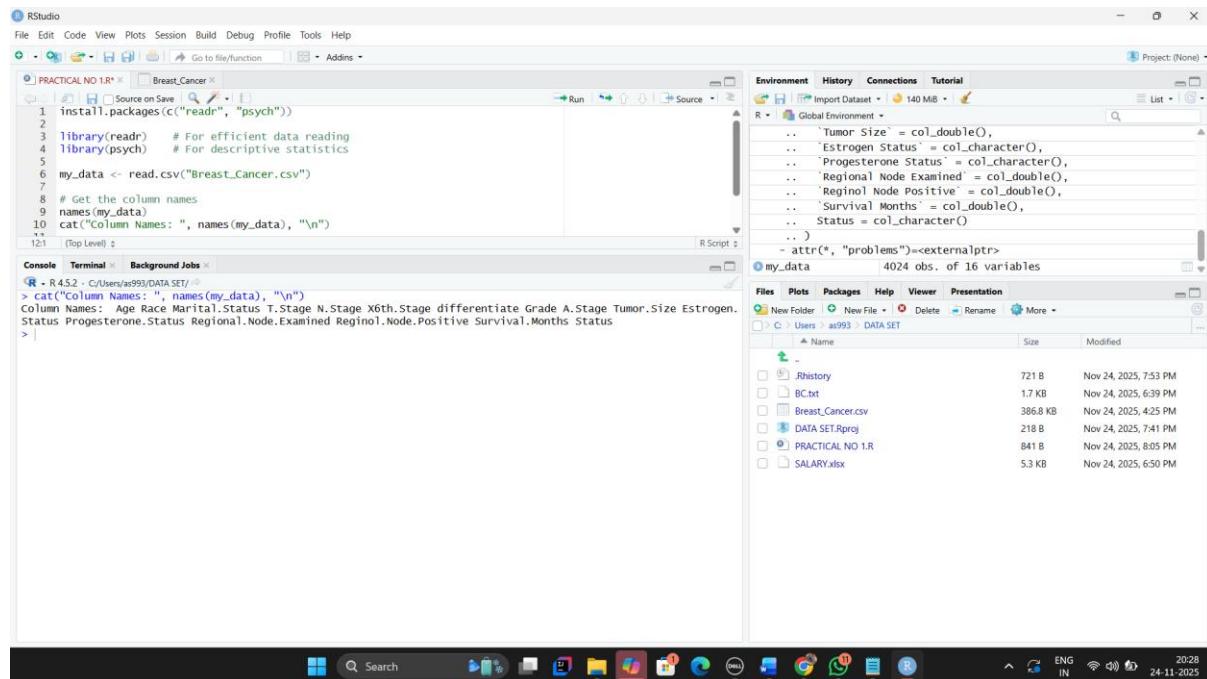
X6th.Stage    differentiate    Grade       A.Stage      Tumor.Size
Length:4024  Length:4024  Length:4024  Length:4024  Min.   : 1.00
Class :character  Class :character  Class :character  Class :character  1st Qu.: 16.00
Mode  :character  Mode  :character  Mode  :character  Mode  :character  Median : 25.00
                           Mean   : 30.47
                           3rd Qu.: 38.00
                           Max.  :140.00

Estrogen.Status Progesterone.Status Regional.Node.Examined Reginol.Node.Positive
Length:4024  Length:4024  Length:4024  Length:4024
Class :character  Class :character  Class :character  Class :character
Mode  :character  Mode  :character  Mode  :character  Mode  :character
                           Min.   : 1.00
                           1st Qu.: 9.00
                           Median :14.00
                           Mean   :14.36
                           3rd Qu.:19.00
                           Max.  :61.00

Survival.Months Status
Min.   : 1.00  Length:4024
1st Qu.: 56.00  Class :character
Median : 73.00  Mode  :character
Mean   : 71.3   Max.  :46.00
3rd Qu.: 90.00

```

6.The column names



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Source Save Go to file/function Run Source
PRACTICAL NO 1.R Breast_Cancer
1 install.packages(c("readr", "psych"))
2
3 library(readr) # For efficient data reading
4 library(psych) # For descriptive statistics
5
6 my_data <- read.csv("Breast_Cancer.csv")
7
8 # Get the column names
9 names(my_data)
10 cat("Column Names: ", names(my_data), "\n")
11
12 [Top Level] R Script
Console Terminal Background Jobs
R - R 4.5.2 - C:/Users/as993/DATA SET/
> cat("Column Names: ", names(my_data), "\n")
Column Names: Age Race Marital.Status T.Stage N.Stage X6th.Stage differentiate Grade A.Stage Tumor.Size Estrogen.Status Progesterone.Status Regional.Node.Examined Reginol.Node.Positive Survival.Months Status
>

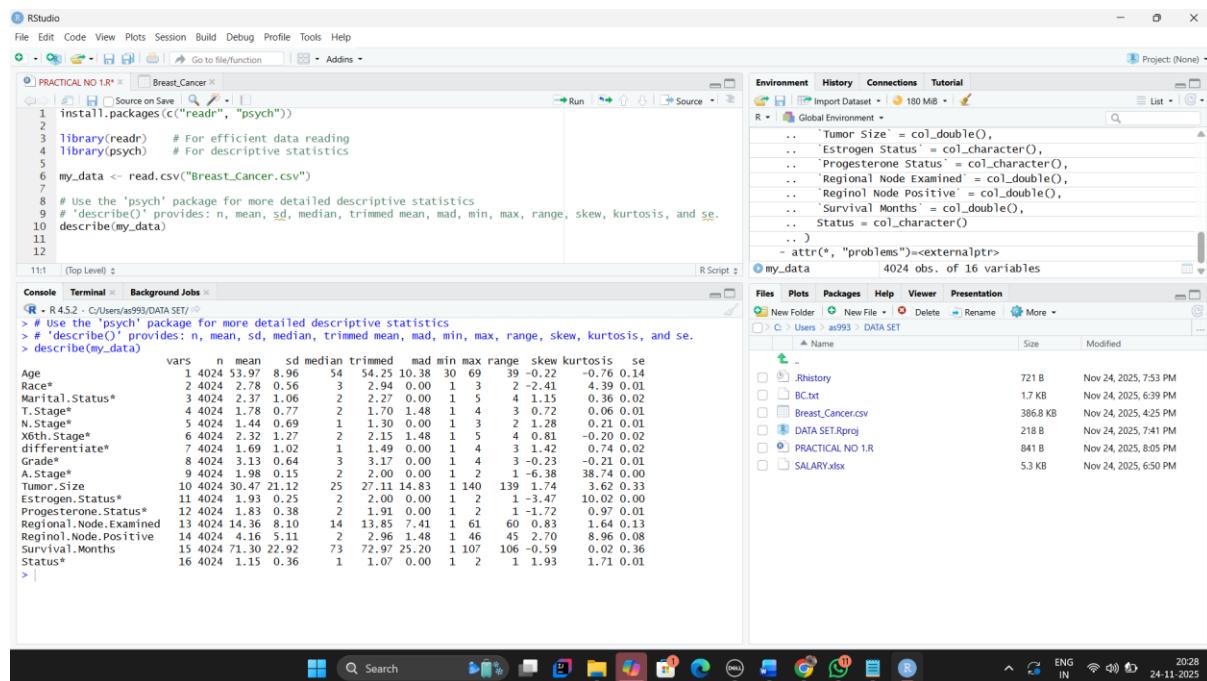
```

Environment History Connections Tutorial
R Global Environment
my_data 4024 obs. of 16 variables
Files Plots Packages Help Viewer Presentation
C:/Users/as993/DATA SET

Name	Size	Modified
RHistory	721 B	Nov 24, 2025, 7:53 PM
BC.txt	1.7 KB	Nov 24, 2025, 6:39 PM
Breast_Cancer.csv	386.8 KB	Nov 24, 2025, 4:25 PM
DATA SET.Rproj	218 B	Nov 24, 2025, 7:41 PM
PRACTICAL NO 1.R	841 B	Nov 24, 2025, 8:05 PM
SALARY.xlsx	5.3 KB	Nov 24, 2025, 6:50 PM

7. 'psych' package for more detailed descriptive statistics

'describe()' provides: n, mean, sd, median, trimmed mean, mad, min, max, range, skew, kurtosis, and se.



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Source Save Go to file/function Run Source
PRACTICAL NO 1.R Breast_Cancer
1 install.packages(c("readr", "psych"))
2
3 library(readr) # For efficient data reading
4 library(psych) # For descriptive statistics
5
6 my_data <- read.csv("Breast_Cancer.csv")
7
8 # Use the 'psych' package for more detailed descriptive statistics
9 # 'describe()' provides: n, mean, sd, median, trimmed mean, mad, min, max, range, skew, kurtosis, and se.
10 describe(my_data)
11
12 [Top Level] R Script
Console Terminal Background Jobs
R - R 4.5.2 - C:/Users/as993/DATA SET/
> # Use the 'psych' package for more detailed descriptive statistics
> # 'describe()' provides: n, mean, sd, median, trimmed mean, mad, min, max, range, skew, kurtosis, and se.
> describe(my_data)
   vars   n   mean      sd median trimmed   mad   min   max range skew kurtosis    se
Age      1 4024 53.97 8.96  54  54.25 10.38 30  69   39 -0.22 -0.76 0.14
Race*    2 4024 2.78 0.56   3  2.94 0.00  1  3   2 -2.41  4.39 0.01
Marital.Status* 3 4024 2.37 1.06   2  2.27 0.00  1  5   4  1.15  0.36 0.02
T.Stage* 4 4024 0.77 0.77   2  1.70 1.48  1  4   3  0.72  0.06 0.01
N.Stage* 5 4024 1.44 0.69   1  1.00 0.00  3  2  1.26  0.00 0.01
X6th.Stage* 6 4024 2.32 1.27   2  2.15 1.48  1  5   4  0.81 -0.20 0.02
differentiate* 7 4024 1.69 1.02   1  1.49 0.00  1  4   3  1.42  0.74 0.02
Grade*   8 4024 3.13 0.64   3  3.17 0.00  1  4   3 -0.23 -0.21 0.01
A.Stage* 9 4024 1.98 0.15   2  2.00 0.00  1  2   1 -6.38  38.74 0.00
Tumor.Size 10 4024 30.47 21.12  25  27.11 14.83  1 140  138  1.74  3.62 0.33
Estrogen.Status* 11 4024 1.93 0.25   2  2.00 0.00  1  2   1 -3.47  10.02 0.00
Progesterone.Status* 12 4024 1.83 0.38   2  1.91 0.00  1  2   1 -1.72  0.97 0.01
Regional.Node.Examined 13 4024 14.36 8.10  14  13.85 7.41  1  61   60  0.83  1.64 0.13
Reginol.Node.Positive 14 4024 4.16 5.11   2  2.96 1.48  1  46   45  2.70  8.96 0.08
Survival.Months 15 4024 71.30 22.92  73  72.97 25.20  1 107  106 -0.59  0.02 0.36
Status*   16 4024 1.15 0.36   1  1.07 0.00  1  2   1  1.93  1.71 0.01

```