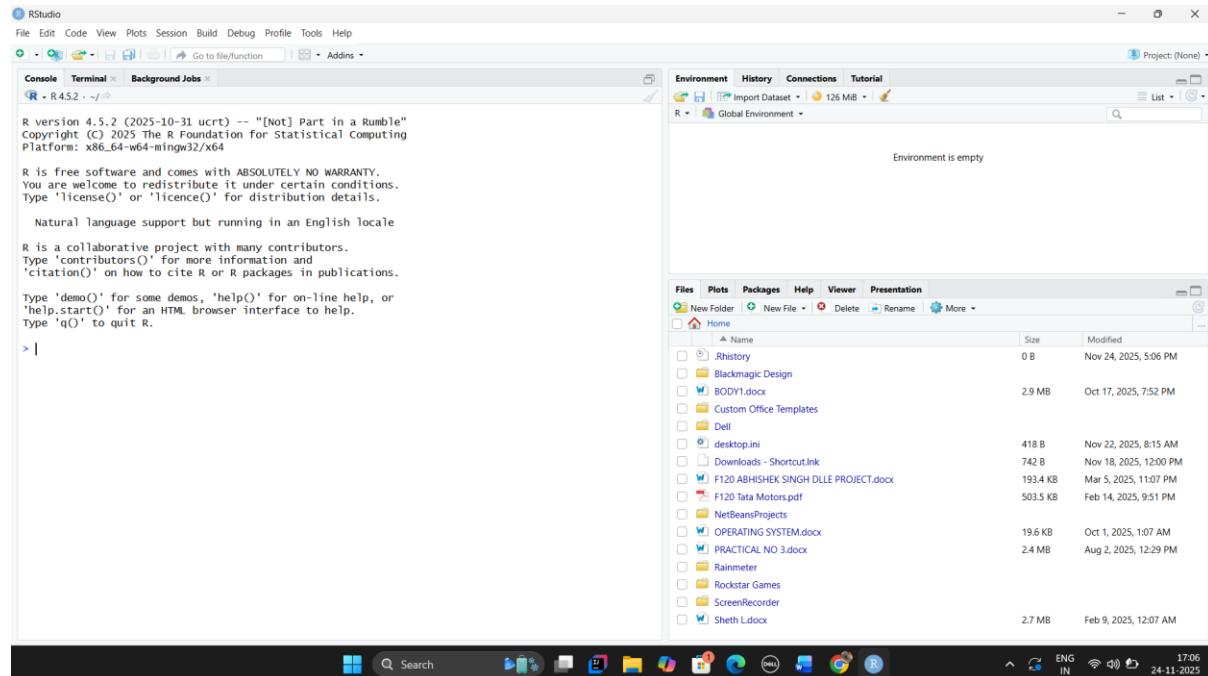


## PRACTICAL NO 2

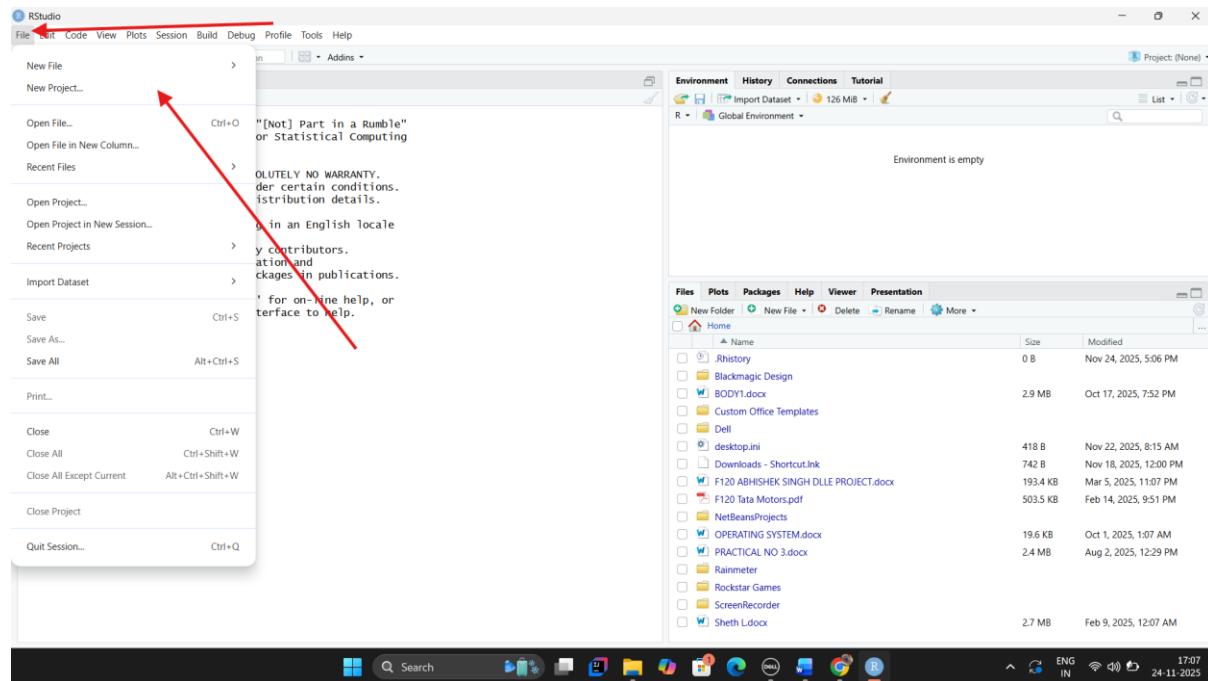
**AIM:-Creating datasets from raw data (text files, CSV files, Excel sheets) and importing data into SAS/SPSS/R.**

### STEP 1:-

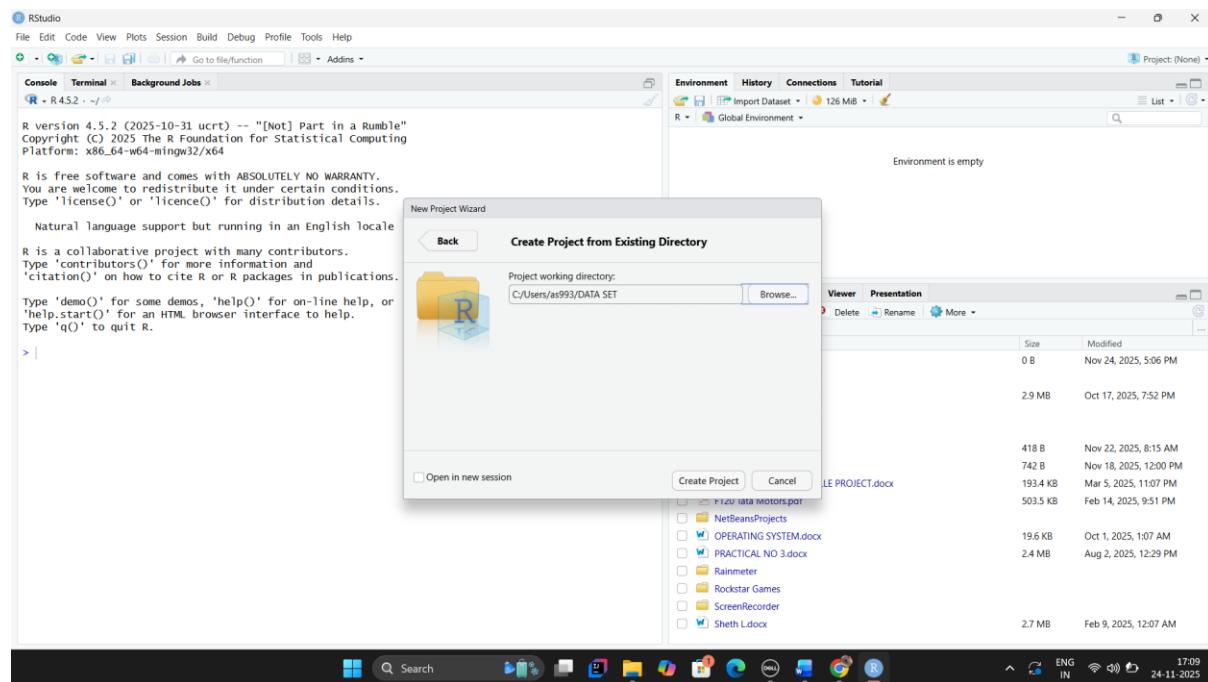
#### OPEN R STUDIO



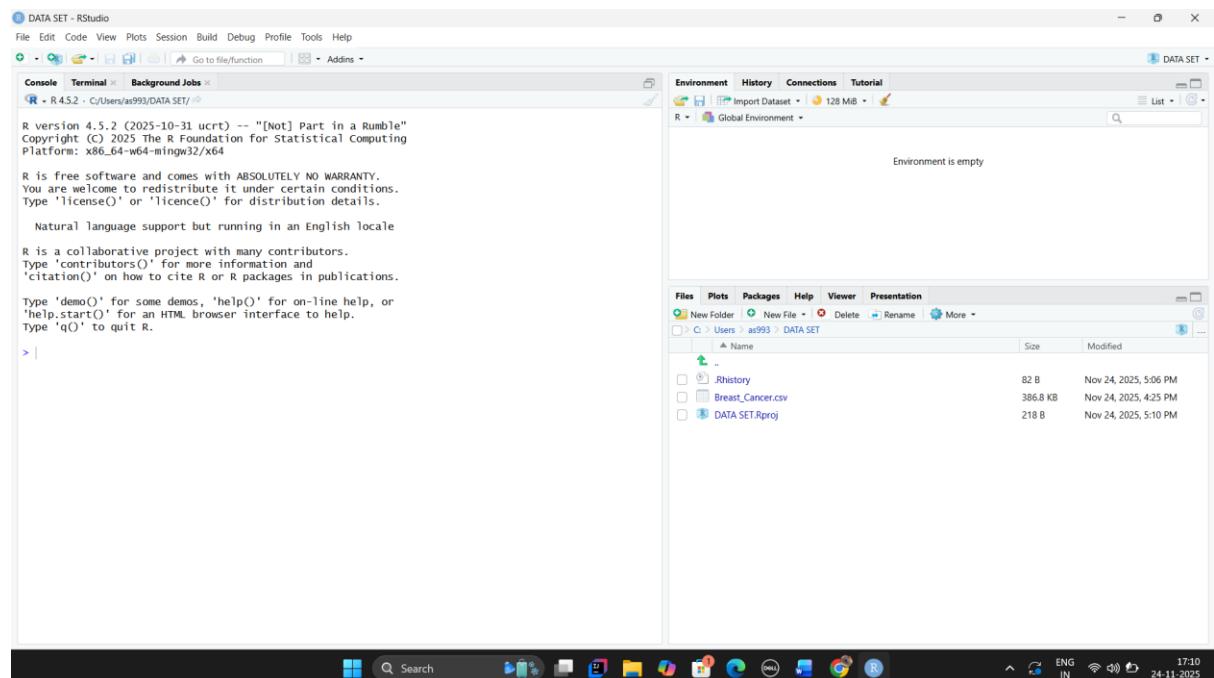
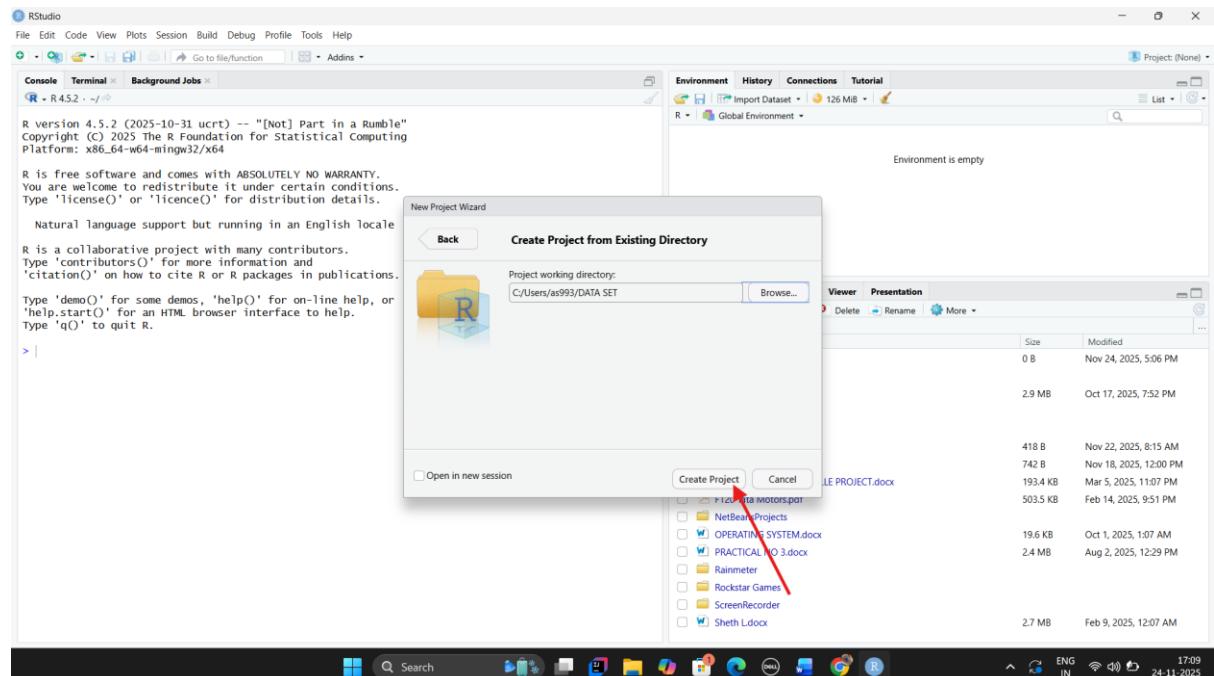
## STEP 2 CREATING A NEW PROJECT



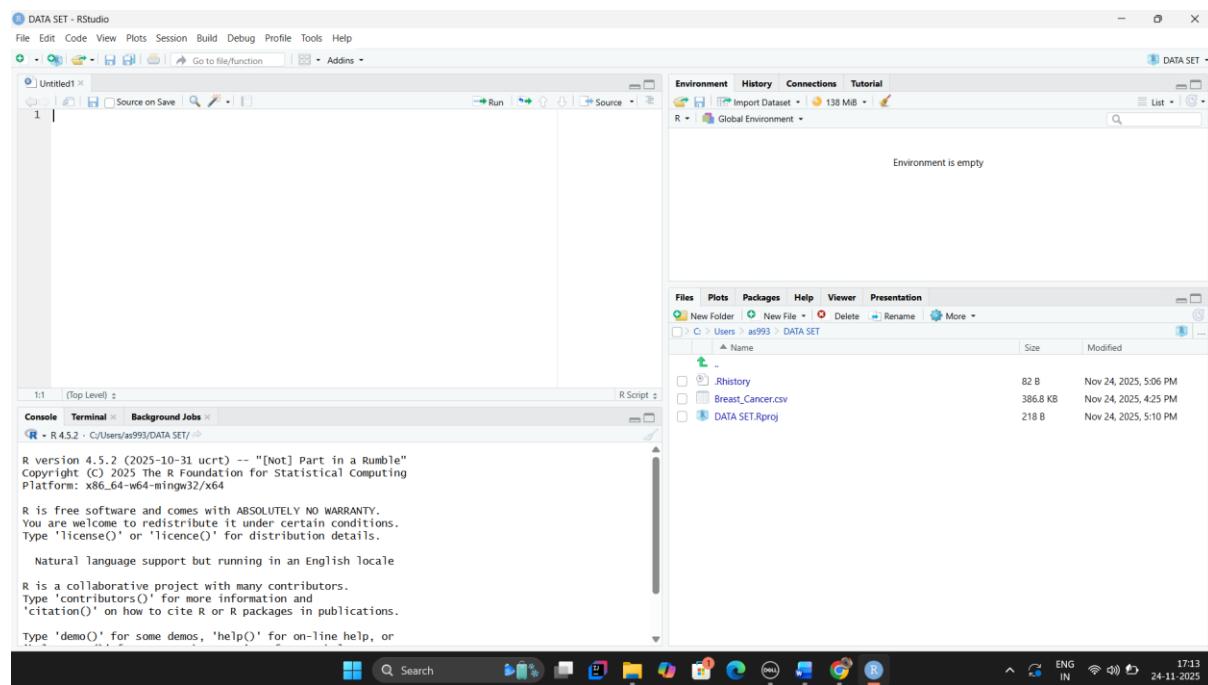
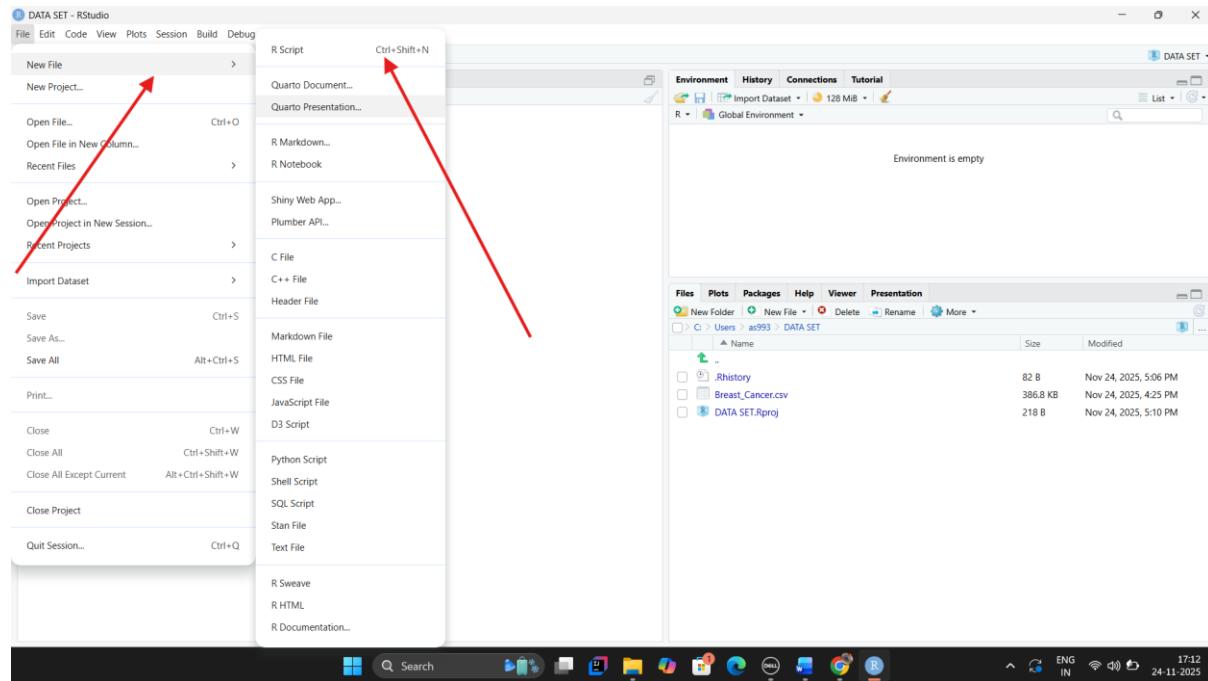
## STEP 3 CHOOSE A DIRECTORY FOR PROJECT



## STEP 4 CLICK ON CREATE PROJECT



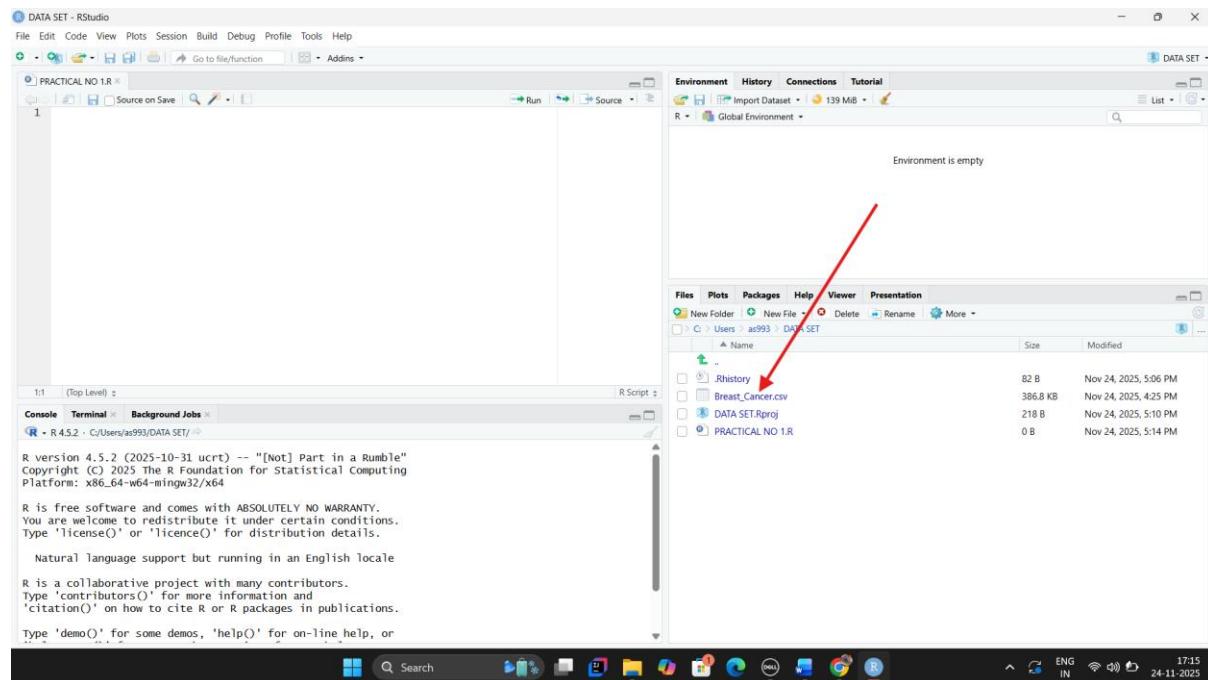
## STEP 5 CREATING A NEW FILE



## STEP 6 SAVE THE FILE AND WRITE THIS CODE ALONG NAME WITH YOUR DATA SET

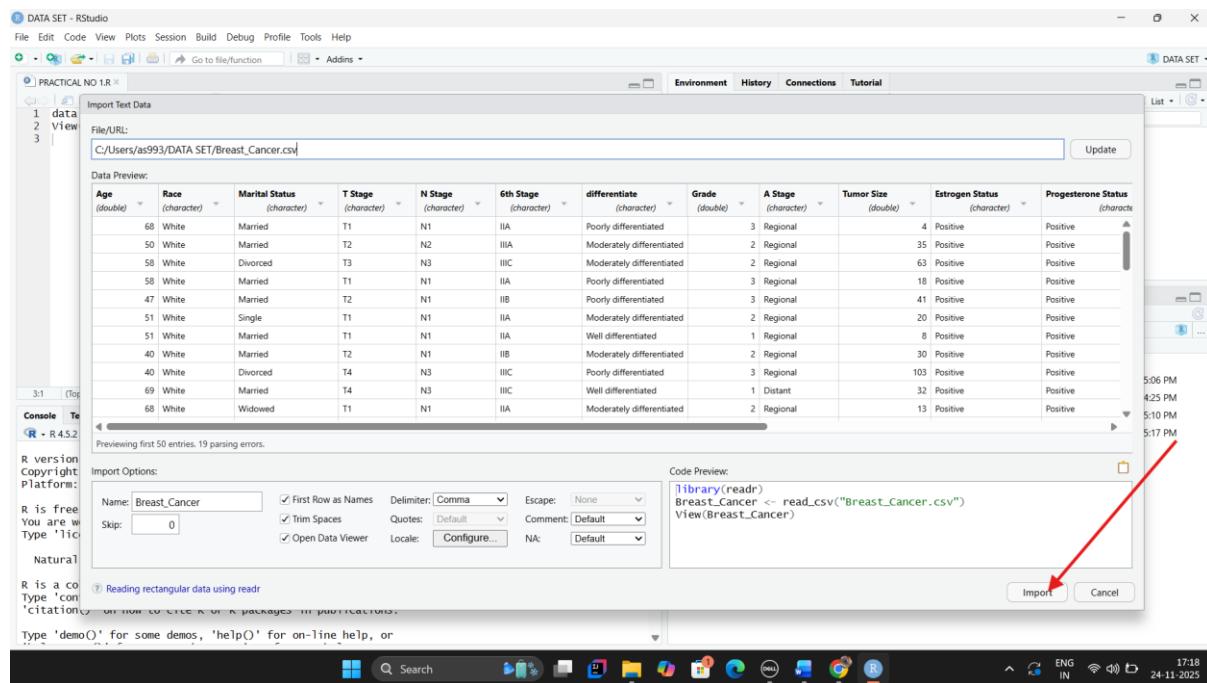
EX:- `data <- read.csv("Breast_Cancer.csv")`

`View(data)`

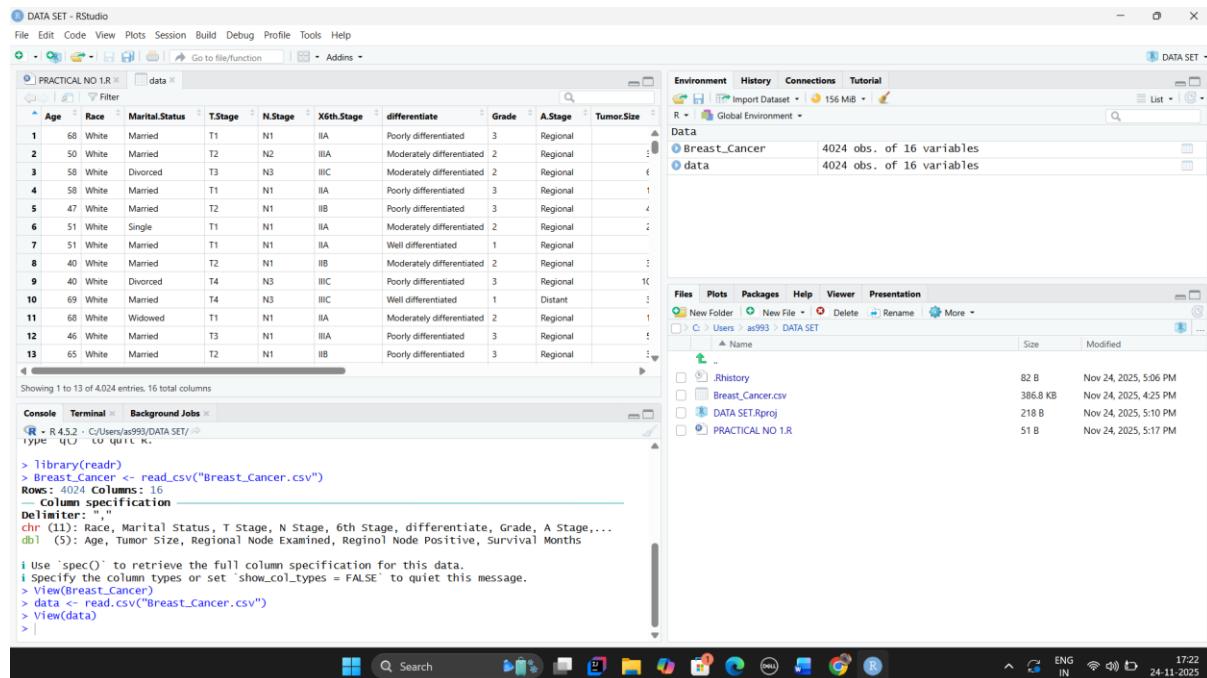


# Sheth L.U.J. College of Arts And Sir M.V. College of Science and Commerce

## Data Analysis with SAS / SPSS / R



## DONE WE IMPORTED COMMA SEPERATED VALUE



# Sheth L.U.J. College of Arts And Sir M.V. College of Science and Commerce

## Data Analysis with SAS / SPSS / R

### TXT FILE

The screenshot shows the RStudio interface. In the top-left pane, there is a code editor with the following content:

```

1. Definition
2. Breast Cancer is a disease in which cells in the breast grow uncontrollably.
It usually begins in the milk ducts or lobules.
5. It is the most common cancer among women worldwide.
6.
7. 2. Causes
8. Genetic factors (BRCA1, BRCA2 mutations)
10. Increasing age, especially after 40
13. Hormonal factors (early menstruation, late menopause)
14. Lifestyle factors (obesity, alcohol, lack of exercise)
16. Radiation exposure
19. 3. Symptoms
20. Lump or thickening in the breast
23. Change in breast shape or size
24. Nipple pain or unusual discharge
26. Dimpling of skin (orange-peel appearance)
28. Swelling under the arm (lymph nodes)

```

In the top-right pane, the "Environment" tab shows the global environment with two objects:

- Breast\_Cancer: 4024 obs. of 16 variables
- data: 52 obs. of 1 variable

The bottom-right pane shows the file browser with the following files listed:

Name	Size	Modified
.Rhistory	82 B	Nov 24, 2025, 5:06 PM
BC.txt	1.7 KB	Nov 24, 2025, 6:39 PM
Breast_Cancer.csv	368.8 KB	Nov 24, 2025, 4:25 PM
DATA SET.Rproj	218 B	Nov 24, 2025, 5:10 PM
PRACTICAL NO 1.R	158 B	Nov 24, 2025, 6:47 PM

### XLSX FILE

The screenshot shows the RStudio interface. In the top-left pane, there is a code editor with the following content:

```

> data <- read.table("BC.txt",
+   header = TRUE,
+   sep = "\t",
+   stringsAsFactors = FALSE)
>
>
>
>
>
>

```

In the top-right pane, the "File/URL:" field contains the path: C:/Users/as993/DATA SET/SALARY.xlsx. The "Import Excel Data" button is highlighted.

The bottom-right pane shows the file browser with the following files listed:

Name	Size	Modified
.Rhistory	82 B	Nov 24, 2025, 5:06 PM
BC.txt	1.7 KB	Nov 24, 2025, 6:39 PM
Breast_Cancer.csv	368.8 KB	Nov 24, 2025, 4:25 PM
DATA SET.Rproj	218 B	Nov 24, 2025, 5:10 PM
PRACTICAL NO 1.R	158 B	Nov 24, 2025, 6:47 PM

# Sheth L.U.J. College of Arts And Sir M.V. College of Science and Commerce

## Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface. The left pane displays a data frame titled "SALARY" with columns "ID", "Name", and "Salary (INR)". The data contains 20 entries. The right pane shows the "Data" browser with datasets like "Breast\_Cancer", "data", and "SALARY". The bottom pane shows the R console with the following code:

```
R > library(readxl)
> SALARY <- read_excel("SALARY.xlsx")
> View(SALARY)
```

The taskbar at the bottom includes icons for various applications like File Explorer, Edge, and Google Chrome.