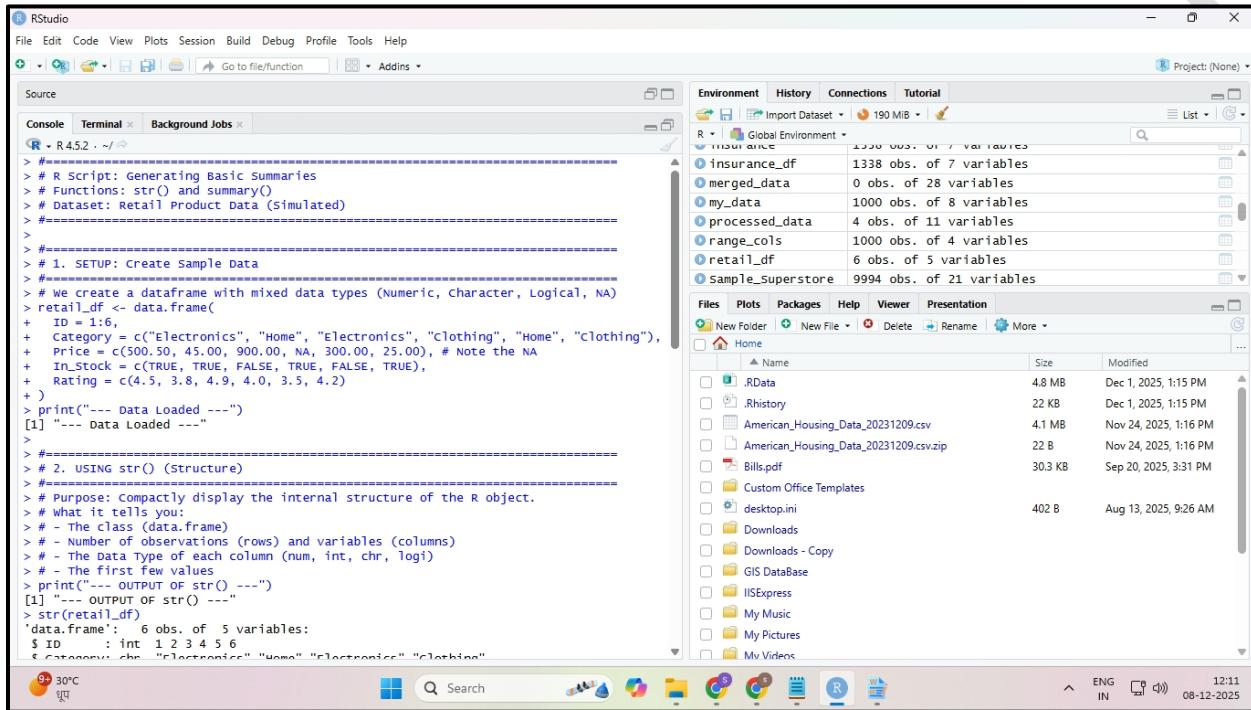


Sheth L.U.J. College of Arts & Sir M.V. College Of Science & Commerce  
SUBJECT NAME: Data Analysis with SAS / SPSS/R

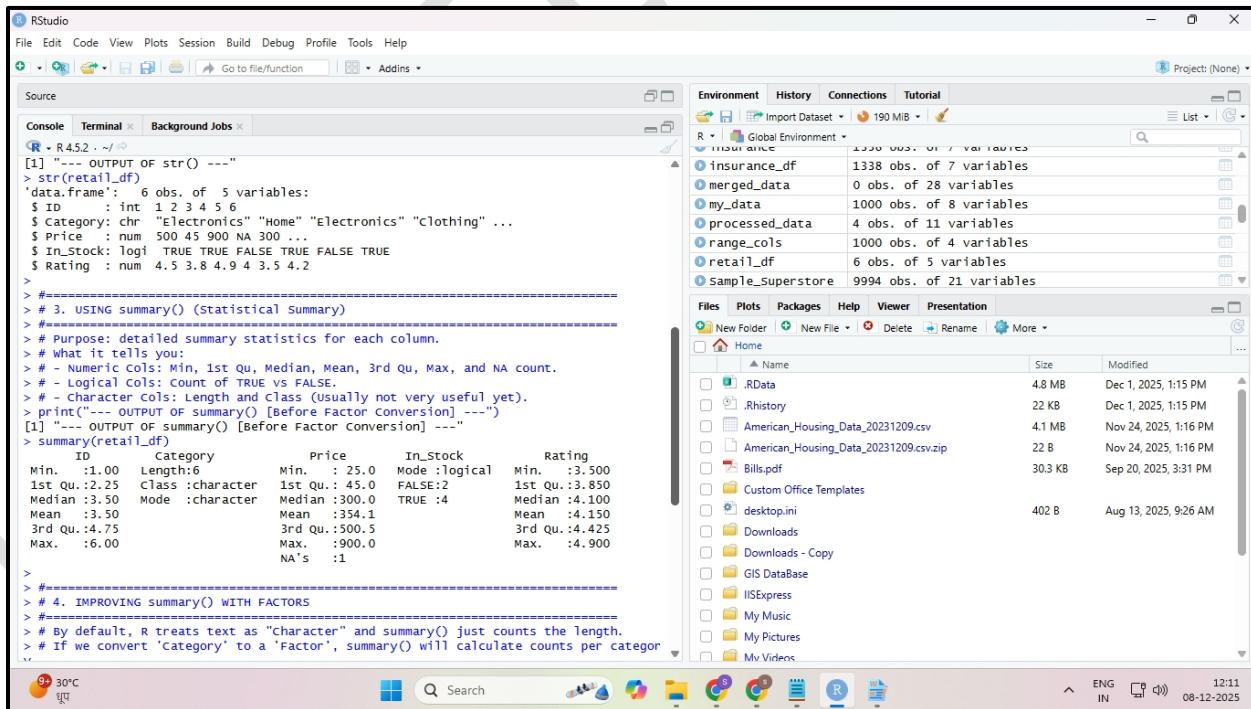
**PRACTICAL NO : 15**

**AIM :** Generating basic summaries using str() or summary() (R).



The screenshot shows the RStudio interface with the following details:

- Console:** Displays R code for generating basic summaries, including the creation of a simulated retail product dataset and its structure.
- Environment:** Shows the global environment with various data frames like insurance\_df, merged\_data, my\_data, processed\_data, range\_cols, retail\_df, and Sample\_Superstore.
- File Explorer:** Shows files in the current directory, including .RData, .Rhistory, American\_Housing\_Data\_20231209.csv, American\_Housing\_Data\_20231209.csv.zip, Bills.pdf, Custom Office Templates, desktop.ini, Downloads, Downloads - Copy, GIS DataBase, IISExpress, My Music, My Pictures, and Mv Videos.
- Bottom Bar:** Includes system icons for battery, signal, and date/time (08-12-2025, 12:11).



The screenshot shows the RStudio interface with the following details:

- Console:** Displays R code for generating basic summaries, including the use of str() and summary() functions on the retail\_df dataset.
- Environment:** Shows the global environment with various data frames.
- File Explorer:** Shows files in the current directory.
- Bottom Bar:** Includes system icons and date/time (08-12-2025, 12:11).

Sheth L.U.J. College of Arts & Sir M.V. College Of Science & Commerce

SUBJECT NAME: Data Analysis with SAS / SPSS/R

The screenshot shows the RStudio environment. The left pane displays the R console with the following code and output:

```
NA's :1
> # 4. IMPROVING summary() WITH FACTORS
> # By default, R treats text as "character" and summary() just counts the length.
> # If we convert 'Category' to a 'Factor', summary() will calculate counts per category
y.
> retail_df$Category <- as.factor(retail_df$Category)
> print("--- OUTPUT OF summary() [After Factor conversion] ---")
[1] "--- OUTPUT OF summary() [After Factor conversion] ---"
> # Notice how 'Category' now shows: Clothing:2, Electronics:2, Home:2
> summary(retail_df)
   ID      Category    Price  In_Stock Rating
Min. :1.00  Clothing :2  Min. :25.0 Mode :logical  Min. :3.500
1st Qu.:2.25 Electronics:2  1st Qu.:45.0 FALSE:2   1st Qu.:3.850
Median :3.50  Home     :2  Median :300.0 TRUE :4    Median :4.100
Mean   :3.50          Mean  :354.1          Mean  :4.150
3rd Qu.:4.75         3rd Qu.:500.5          3rd Qu.:4.425
Max.  :6.00          Max.  :900.0           Max.  :4.900
NA's   :1
```

The right pane shows the Global Environment and File Explorer. The Global Environment lists several datasets:

- insurance\_df: 1338 obs. of 7 variables
- merged\_data: 0 obs. of 28 variables
- my\_data: 1000 obs. of 8 variables
- processed\_data: 4 obs. of 11 variables
- range\_cols: 1000 obs. of 4 variables
- retail\_df: 6 obs. of 5 variables
- Sample\_Superstore: 9994 obs. of 21 variables

The File Explorer shows the local directory structure:

- .RData
- .History
- American\_Housing\_Data\_20231209.csv
- American\_Housing\_Data\_20231209.csv.zip
- Bills.pdf
- Custom Office Templates
- desktop.ini
- Downloads
- Downloads - Copy
- GIS DataBase
- IISExpress
- My Music
- My Pictures
- Mv Videos

At the bottom, the taskbar shows various application icons and the system clock indicates 12:12 on 08-12-2025.