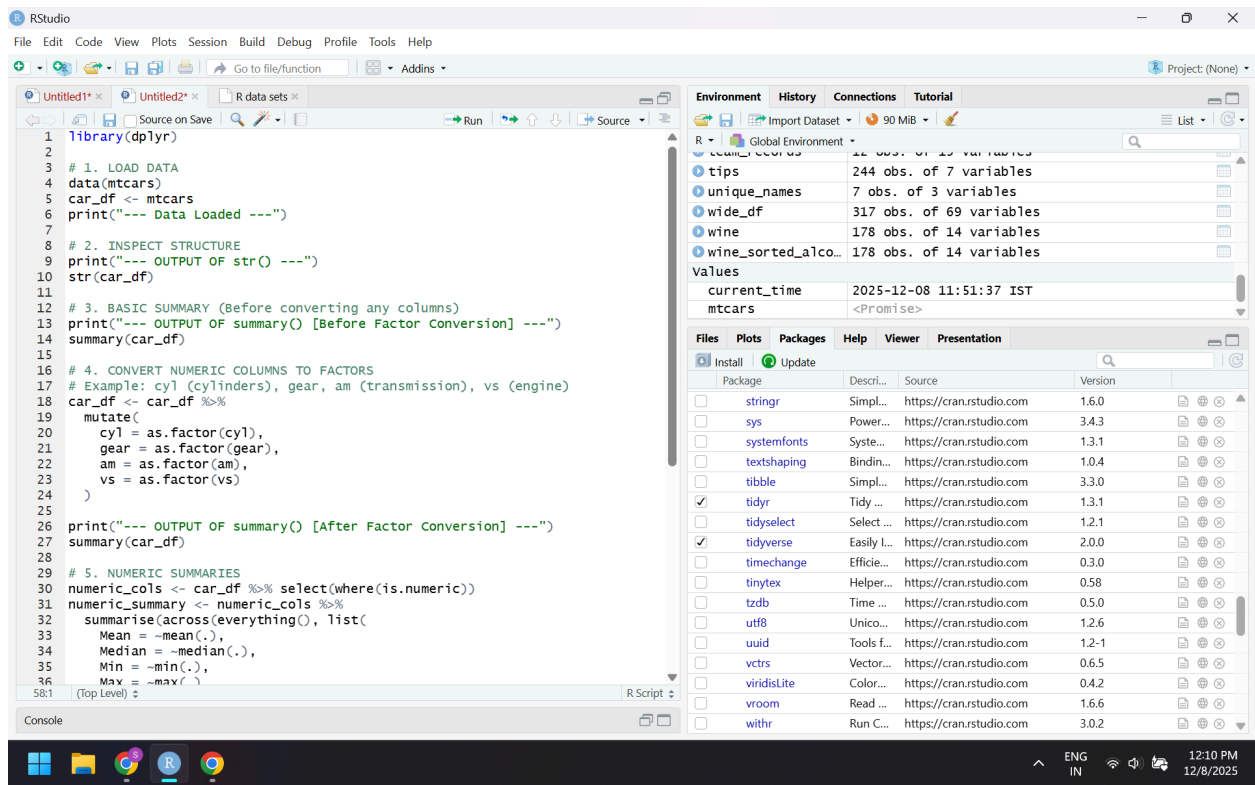


SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE

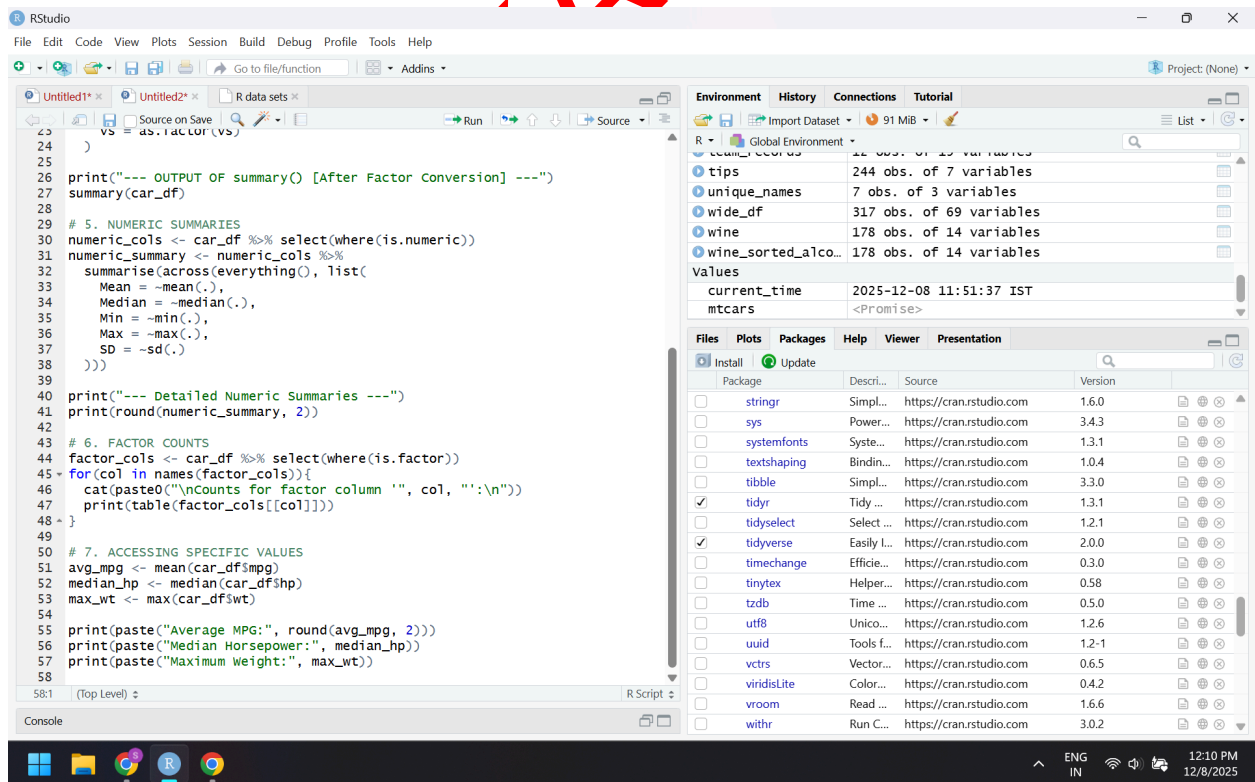
SUBJECT - Data Analysis with SAS / SPSS / R

Aim :- Generating basic summaries using str() or summary() (R).

Input :



```
1 library(dplyr)
2
3 # 1. LOAD DATA
4 data(mtcars)
5 car_df <- mtcars
6 print("--- Data Loaded ---")
7
8 # 2. INSPECT STRUCTURE
9 print("--- OUTPUT OF str() ---")
10 str(car_df)
11
12 # 3. BASIC SUMMARY (Before converting any columns)
13 print("--- OUTPUT OF summary() [Before Factor Conversion] ---")
14 summary(car_df)
15
16 # 4. CONVERT NUMERIC COLUMNS TO FACTORS
17 # Example: cyl (cylinders), gear, am (transmission), vs (engine)
18 car_df <- car_df %>%
19   mutate(
20     cyl = as.factor(cyl),
21     gear = as.factor(gear),
22     am = as.factor(am),
23     vs = as.factor(vs)
24   )
25
26 print("--- OUTPUT OF summary() [After Factor Conversion] ---")
27 summary(car_df)
28
29 # 5. NUMERIC SUMMARIES
30 numeric_cols <- car_df %>% select(where(is.numeric))
31 numeric_summary <- numeric_cols %>%
32   summarise(across(everything(), list(
33     Mean = ~mean(.),
34     Median = ~median(.),
35     Min = ~min(.),
36     Max = ~max(.)
37   )))
38
39 print---
```



```
24 )
25
26 print("--- OUTPUT OF summary() [After Factor Conversion] ---")
27 summary(car_df)
28
29 # 5. NUMERIC SUMMARIES
30 numeric_cols <- car_df %>% select(where(is.numeric))
31 numeric_summary <- numeric_cols %>%
32   summarise(across(everything(), list(
33     Mean = ~mean(.),
34     Median = ~median(.),
35     Min = ~min(.),
36     Max = ~max(.),
37     SD = ~sd(.)
38   )))
39
40 print("--- Detailed Numeric Summaries ---")
41 print(round(numeric_summary, 2))
42
43 # 6. FACTOR COUNTS
44 factor_cols <- car_df %>% select(where(is.factor))
45 for(col in names(factor_cols)){
46   cat(paste0("\nCounts for factor column '", col, "':\n"))
47   print(table(factor_cols[[col]]))
48 }
49
50 # 7. ACCESSING SPECIFIC VALUES
51 avg_mpg <- mean(car_df$mpg)
52 median_hp <- median(car_df$hp)
53 max_wt <- max(car_df$wt)
54
55 print(paste("Average MPG:", round(avg_mpg, 2)))
56 print(paste("Median Horsepower:", median_hp))
57 print(paste("Maximum Weight:", max_wt))
58
59 print---
```

Name - Mithil Kadam
Roll No - S083

SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE
SUBJECT - Data Analysis with SAS / SPSS / R

Output :

The screenshot displays the RStudio interface with the following components:

- Console:** Shows R code for loading the `airquality` dataset, parsing date components, and printing the results. The output shows the first 10 rows of the `airquality` dataset with extracted date components.
- Environment:** Lists the loaded datasets: `tips` (244 obs. of 7 variables), `unique_names` (7 obs. of 3 variables), `wide_df` (317 obs. of 69 variables), `wine` (178 obs. of 14 variables), and `wine_sorted_alco...` (178 obs. of 14 variables).
- Files:** Shows the file explorer with the `mtcars` dataset loaded.
- Packages:** Lists installed and available packages, including `stringr`, `sys`, `systemfonts`, `textshaping`, `tibble`, `tidyr`, `tidyselect`, `tidyverse`, `timechange`, `tinytex`, `tzdb`, `utf8`, `uuid`, `vcdrs`, `viridisLite`, `vroom`, and `withr`.

The console output shows the first 10 rows of the `airquality` dataset with extracted date components:

	ozone	Solar.R	Wind	Temp	Month	Day	Date	Year_Num	Month_Num	Month_Name
1	41	190	7.4	67	5	1	1973-05-01	1973	5	May
2	36	118	8.0	72	5	2	1973-05-02	1973	5	May
3	12	149	12.6	74	5	3	1973-05-03	1973	5	May
4	18	313	11.5	62	5	4	1973-05-04	1973	5	May
5	NA	NA	14.3	56	5	5	1973-05-05	1973	5	May
6	28	NA	14.9	66	5	6	1973-05-06	1973	5	May
7	23	299	8.6	65	5	7	1973-05-07	1973	5	May
8	19	99	13.8	59	5	8	1973-05-08	1973	5	May
9	8	19	20.1	61	5	9	1973-05-09	1973	5	May
10	NA	194	8.6	69	5	10	1973-05-10	1973	5	May

The console output also shows the extracted date components for the first 10 rows:

Day_Num	Weekday_Num	Weekday_Name	Quarter	Day_of_Year
1	3	Tuesday	2	121
2	4	Wednesday	2	122
3	5	Thursday	2	123
4	6	Friday	2	124
5	7	Saturday	2	125
6	1	Sunday	2	126
7	2	Monday	2	127
8	3	Tuesday	2	128
9	4	Wednesday	2	129
10	5	Thursday	2	130

Name - Mithil Kadam
Roll No - S083

SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE
SUBJECT - Data Analysis with SAS / SPSS / R

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains R code for data manipulation and time extraction. The code includes a data frame with columns for Day_Num, Weekday_Num, Weekday_Name, Quarter, and Day_of_Year. It also includes a system date extraction script.
- Console:** Shows the output of the R code, including the current time extraction results and the data frame structure.
- Environment:** Lists the loaded packages and their versions, including stringr, sys, systemfonts, textshaping, tibble, tidyr, tidyselect, tidyverse, timechange, tinytex, tzdb, utf8, uuid, vctrs, viridisLite, vroom, and withr.
- Files:** Shows the file explorer with a list of files and folders.
- Plots:** Shows the plot viewer with a list of plots.
- Packages:** Shows the package manager with a list of installed and available packages.
- Help:** Shows the help viewer with a list of help topics.
- Viewer:** Shows the viewer pane with a list of viewer items.
- Presentation:** Shows the presentation pane with a list of presentation items.

The R code in the Source Editor is as follows:

```
> # 3. SYSTEM DATE: Extract components from current timestamp
> current_time <- now()
> print("--- Current Time Extraction ---")
[1] "--- Current Time Extraction ---"
> print(paste("Current Year:", year(current_time)))
[1] "Current Year: 2025"
> print(paste("Current Month:", month(current_time)))
[1] "Current Month: 12"
> print(paste("Current Day:", day(current_time)))
[1] "Current Day: 8"
> print(paste("Current Hour:", hour(current_time)))
[1] "Current Hour: 11"
> print(paste("Current Minute:", minute(current_time)))
[1] "Current Minute: 51"
>
> data("mtcars")
> |
```

Name - Mithil Kadam
Roll No - S083