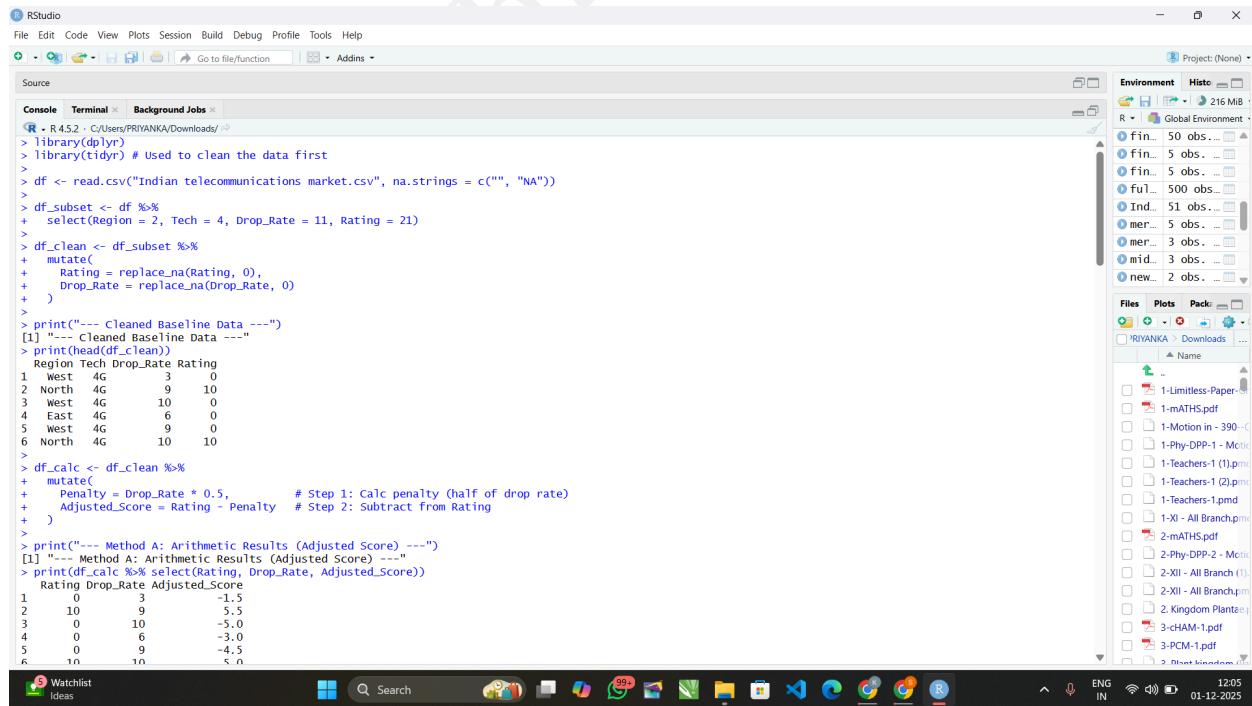


**SHETH L.U.J AND SIR M.V. COLLEGE**  
**SUBJECT NAME: DATA ANALYSIS WITH SAS/SPSS/R**

## Module 1 Practical 10

**Aim:** Creating new variables using transformations and calculations in R. import dataset.

### OUTPUT:



The screenshot shows the RStudio interface with the following R code in the Console tab:

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Environment History Project: (None)
Console Terminal Background Jobs
R - R 4.5.2 · C:/users/PRIYANKA/Downloads/...
> library(dplyr)
> library(tidyverse) # used to clean the data first
>
> df <- read.csv("Indian telecommunications market.csv", na.strings = c("", "NA"))
>
> df_subset <- df %>%
+   select(Region = 2, Tech = 4, Drop_Rate = 11, Rating = 21)
>
> df_clean <- df_subset %>%
+   mutate(
+     Rating = replace_na(Rating, 0),
+     Drop_Rate = replace_na(Drop_Rate, 0)
+   )
>
> print("---- Cleaned Baseline Data ----")
[1] "---- Cleaned Baseline Data ---"
> print(head(df_clean))
  Region Tech Drop_Rate Rating
1  West    4G       3      0
2 North   4G       9     10
3  West    4G      10      0
4  East    4G       6      0
5  West    4G       9      0
6 North   4G      10     10
>
> df_calc <- df_clean %>%
+   mutate(
+     Penalty = Drop_Rate * 0.5,      # Step 1: Calc penalty (half of drop rate)
+     Adjusted_Score = Rating - Penalty # Step 2: Subtract from Rating
+   )
>
> print("---- Method A: Arithmetic Results (Adjusted Score) ----")
[1] "---- Method A: Arithmetic Results (Adjusted Score) ---"
> print(df_calc %>% select(Rating, Drop_Rate, Adjusted_Score))
  Rating Drop_Rate Adjusted_Score
1      0       3        -1.5
2     10      9        -5.0
3      0      10        -5.0
4      0       6        -3.0
5      0       9        -4.5
6     10      10         5.0
```

**SHETH L.U.J AND SIR M.V. COLLEGE**  
**SUBJECT NAME: DATA ANALYSIS WITH SAS/SPSS/R**

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The left sidebar has tabs for Source, Console, Terminal, and Background Jobs, with the Console tab active. The main area displays the following CSV data:

5	6	7	8
10	10	0	0
10	0	4	5
11	9	7	7
12	0	0	0
13	8	8	8
14	0	9	9
15	0	3	3
16	8	2	2
17	9	9	9
18	0	4	4
19	0	0	0
20	0	8	8
21	0	9	9
22	9	9	9
23	0	8	8
24	0	8	8
25	0	3	3
26	0	6	6
27	8	8	8
28	6	6	6
29	7	6	6
30	8	4	4
31	10	10	10
32	0	7	7
33	0	7	7
34	9	9	9
35	0	9	9
36	0	6	6
37	0	8	8
38	0	1	1
39	0	9	9
40	7	7	7
41	5	5	5
42	0	9	9
43	7	8	8
44	0	3	3
45	n	<	<

The right sidebar shows the Project Environment, listing files like fin\_50.pdf, fin\_5.pdf, ful\_500.pdf, Ind\_51.pdf, mer\_5.pdf, mer\_3.pdf, mid\_3.pdf, new\_2.pdf, and 1-Limitless-Paper.pdf. The Files, Plots, and Packets tabs are also visible.

# SHETH L.U.J AND SIR M.V. COLLEGE

## SUBJECT NAME: DATA ANALYSIS WITH SAS/SPSS/R

RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
[R - R 4.5.2 - C:/users/PRIYANKA/Downloads/]
23 0 Average/Low 8 Unstable
24 0 Average/Low 8 Unstable
25 0 Average/Low 3 Stable
26 0 Average/Low 6 Stable
27 8 Average/Low 8 Unstable
28 6 Average/Low 6 Stable
29 7 Average/Low 6 Stable
30 8 Average/Low 4 Stable
31 10 High 10 Unstable
32 0 Average/Low 7 Stable
33 0 Average/Low 7 Stable
34 9 High 9 Unstable
35 0 Average/Low 9 Unstable
36 0 Average/Low 6 Stable
37 0 Average/Low 8 Unstable
38 0 Average/Low 1 Stable
39 0 Average/Low 9 Unstable
40 7 Average/Low 7 Stable
41 5 Average/Low 5 Stable
42 0 Average/Low 9 Unstable
43 7 Average/Low 8 Unstable
44 0 Average/Low 3 Stable
45 0 Average/Low 5 Stable
46 0 Average/Low 3 Stable
47 0 Average/Low 8 Unstable
48 0 Average/Low 8 Unstable
49 0 Average/Low 9 Unstable
50 9 High 7 Stable
>
> df_text <- df_clean %>%
+   mutate(
+     # paste connects strings with a space by default
+     user_Profile = paste(Region, "user on", Tech, "network")
+   )
> print("---- Method C: Text Transformation ----")
[1] "---- Method C: Text Transformation ----"
> print(head(df_text$user_Profile))
[1] "West user on 4G network" "North user on 4G network" "East user on 4G network"
[5] "West user on 4G network" "North user on 4G network"
>
> final_dataset <- df_clean %>%
+   filter(Region != "North")

```

Watchlist Ideas

Search

12:06 ENG IN 01-12-2025

RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
[R - R 4.5.2 - C:/users/PRIYANKA/Downloads/]
41 5 Average/Low 5 Stable
42 0 Average/Low 9 Unstable
43 7 Average/Low 8 Unstable
44 0 Average/Low 3 Stable
45 0 Average/Low 5 Stable
46 0 Average/Low 3 Stable
47 0 Average/Low 8 Unstable
48 0 Average/Low 8 Unstable
49 0 Average/Low 9 Unstable
50 9 High 7 Stable
>
> df_text <- df_clean %>%
+   mutate(
+     # paste connects strings with a space by default
+     user_Profile = paste(Region, "user on", Tech, "network")
+   )
> print("---- Method C: Text Transformation ----")
[1] "---- Method C: Text Transformation ----"
> print(head(df_text$user_Profile))
[1] "West user on 4G network" "North user on 4G network" "East user on 4G network"
[5] "West user on 4G network" "North user on 4G network"
>
> final_dataset <- df_clean %>%
+   mutate(
+     Adjusted_Score = Rating - (Drop_Rate * 0.5),
+     Is_Top_Tier = ifelse(Adjusted_Score > 7, TRUE, FALSE),
+     Status_Report = paste0("Region: ", Region, " / Score: ", round(Adjusted_Score, 1))
+   )
>
> print("---- Final Combined Dataset ----")
[1] "---- Final Combined Dataset ----"
> print(head(final_dataset))
Region Tech Drop_Rate Rating Adjusted_Score Is_Top_Tier Status_Report
1 West 4G 3 0 -1.5 FALSE Region: West / Score: -1.5
2 North 4G 9 10 5.5 FALSE Region: North / Score: 5.5
3 West 4G 10 0 -5.0 FALSE Region: West / Score: -5
4 East 4G 6 0 -3.0 FALSE Region: East / Score: -3
5 West 4G 9 0 -4.5 FALSE Region: West / Score: -4.5
6 North 4G 10 10 5.0 FALSE Region: North / Score: 5

```

Watchlist Ideas

Search

12:06 ENG IN 01-12-2025