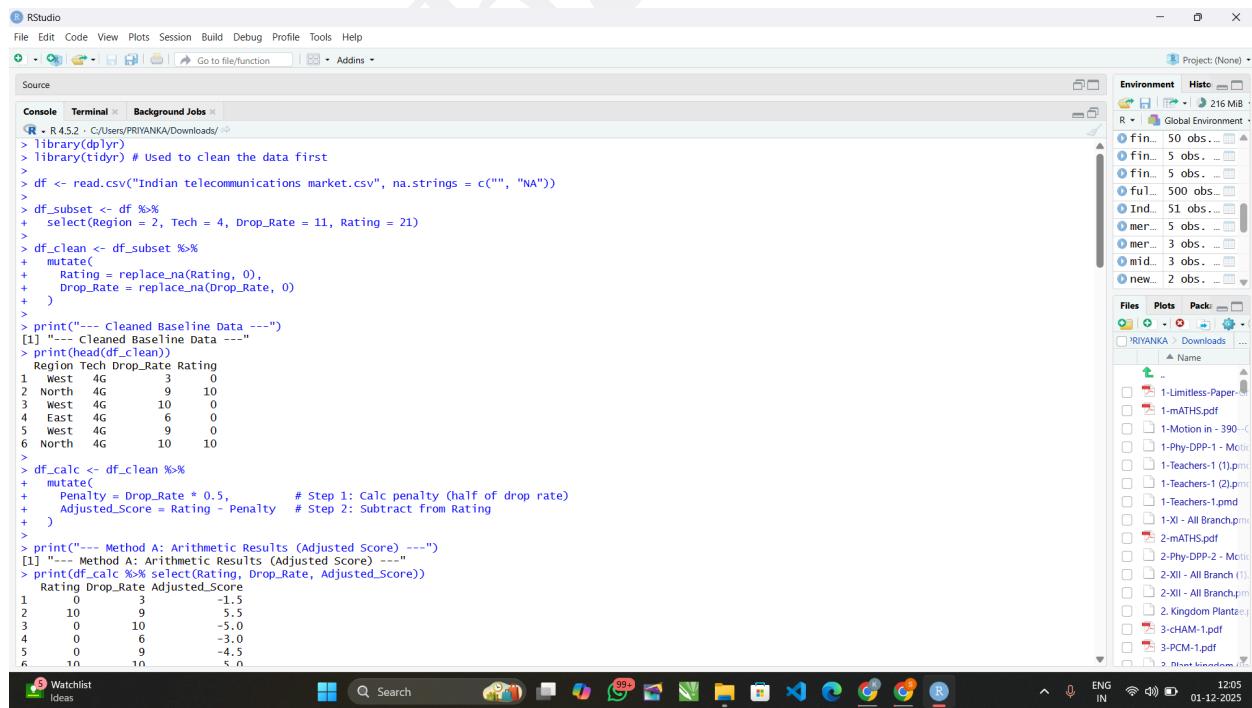


SHETH L.U.J AND SIR M.V. COLLEGE
SUBJECT NAME: DATA ANALYSIS WITH SAS/SPSS/R

Module 1 Practical 10

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

OUTPUT:



The screenshot shows an RStudio interface with the following details:

- Console Tab:** Displays R code for data cleaning and transformation. The code reads a CSV file, removes rows where Region is 2, and then performs various operations like mutate, replace_na, and print statements to show the cleaned data and arithmetic results.
- Environment Tab:** Shows the global environment with objects like fin., ful., Ind., mer., mid., new., and fin..50 obs..
- Files Tab:** Shows a directory structure under 'PRIYANKA > Downloads' containing various PDF files related to mathematics and motion.
- Plots Tab:** Currently empty.
- Packages Tab:** Currently empty.

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
R - R 4.5.2 · C:/users/PRIYANKA/Downloads/...
Go to file/function | Addins |
Source
Console Terminal Background Jobs
> 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

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" "West user on 4G network" "East user on 4G network"
[5] "West user on 4G network" "North user on 4G network"
>
> final_dataset <- df_clean %>%

```

Watchlist Ideas

Search

12:06 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" "West 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 01-12-2025