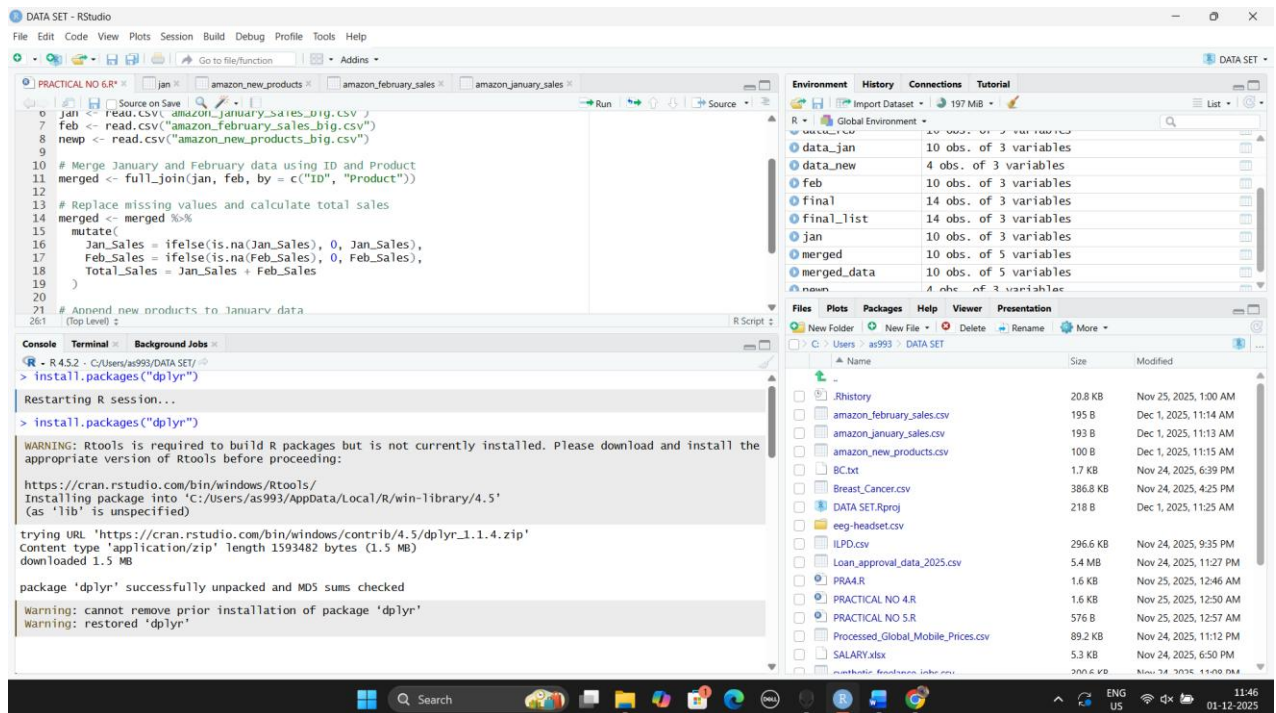


## PRACTICAL NO 6

AIM:- Combining and appending datasets using merge() or bind\_rows() in R.

### OUTPUT:-



```
DATA SET - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
PRACTICAL NO 6.R | jan | amazon_new_products | amazon_february_sales | amazon_january_sales
0 jan <- read.csv("amazon_january_sales_big.csv")
7 feb <- read.csv("amazon_february_sales_big.csv")
8 newp <- read.csv("amazon_new_products_big.csv")
9
10 # Merge January and February data using ID and Product
11 merged <- full_join(jan, feb, by = c("ID", "Product"))
12
13 # Replace missing values and calculate total sales
14 merged <- merged %>%
15 mutate(
16   Jan_Sales = ifelse(is.na(Jan_Sales), 0, Jan_Sales),
17   Feb_Sales = ifelse(is.na(Feb_Sales), 0, Feb_Sales),
18   Total_Sales = Jan_Sales + Feb_Sales
19 )
20
21 # Append new products to January data
261 (Top Level) |

Console | Terminal | Background Jobs
R - R 4.5.2 - C:/Users/as993/DATA SET/
> install.packages("dplyr")

Restarting R session...

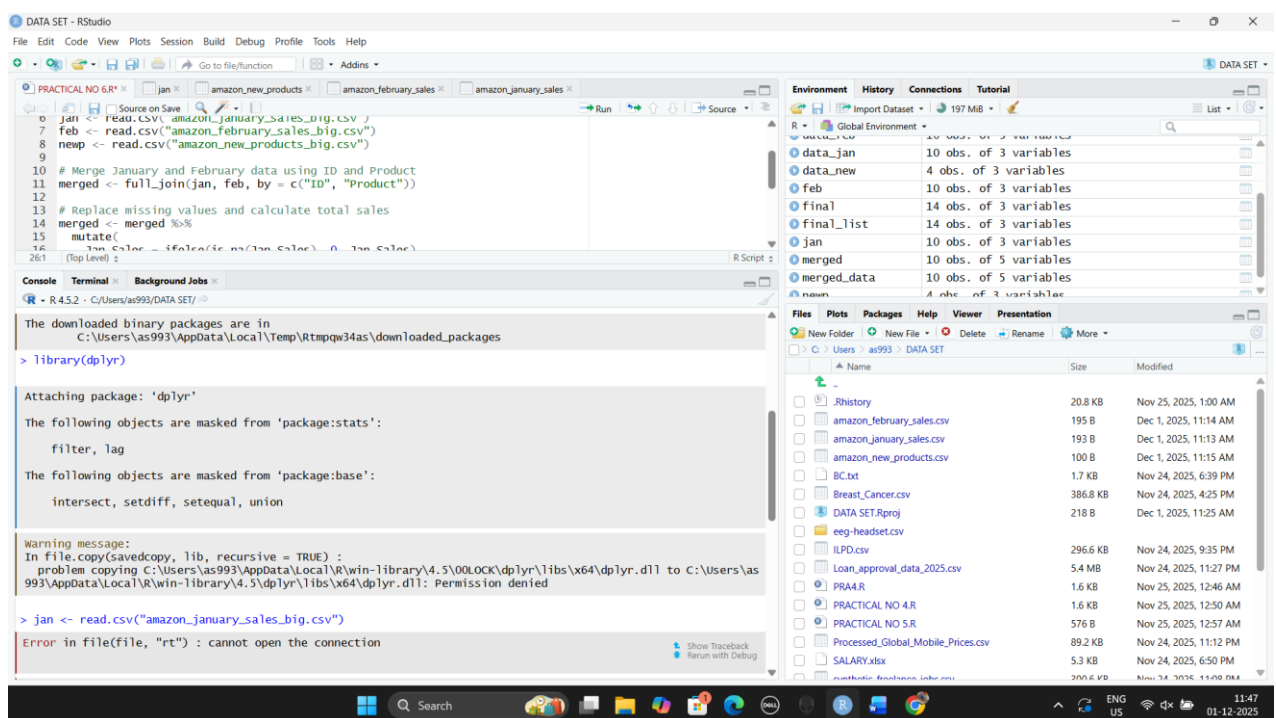
> install.packages("dplyr")

WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the
appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/as993/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/dplyr_1.1.4.zip'
Content type 'application/zip' length 1593482 bytes (1.5 MB)
downloaded 1.5 MB

package 'dplyr' successfully unpacked and MD5 sums checked
Warning: cannot remove prior installation of package 'dplyr'
Warning: restored 'dplyr'
```



```
DATA SET - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
PRACTICAL NO 6.R | jan | amazon_new_products | amazon_february_sales | amazon_january_sales
0 jan <- read.csv("amazon_january_sales_big.csv")
7 feb <- read.csv("amazon_february_sales_big.csv")
8 newp <- read.csv("amazon_new_products_big.csv")
9
10 # Merge January and February data using ID and Product
11 merged <- full_join(jan, feb, by = c("ID", "Product"))
12
13 # Replace missing values and calculate total sales
14 merged <- merged %>%
15 mutate(
16   Jan_Sales = ifelse(is.na(Jan_Sales), 0, Jan_Sales),
17   Feb_Sales = ifelse(is.na(Feb_Sales), 0, Feb_Sales),
18   Total_Sales = Jan_Sales + Feb_Sales
19 )
20
21 # Append new products to January data
261 (Top Level) |

Console | Terminal | Background Jobs
R - R 4.5.2 - C:/Users/as993/DATA SET/
> library(dplyr)

The downloaded binary packages are in
C:/Users/as993/AppData/Local/Temp/Rtmpq34as/downloaded_packages

> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':
  filter, lag

The following objects are masked from 'package:base':
  intersect, setdiff, setequal, union

Warning message:
In file.copy(savedcopy, lib, recursive = TRUE) :
  problem copying C:/Users/as993/AppData/Local/R/win-library/4.5/00LOCK/dplyr/lib/x64/dplyr.dll to C:/Users/as993/AppData/Local/R/win-library/4.5/dplyr/lib/x64/dplyr.dll: Permission denied

> jan <- read.csv("amazon_january_sales_big.csv")

Error in file(file, "rt") : cannot open the connection
```

# 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 script editor contains the following code:

```
0 jan <- read.csv("amazon_january_sales_big.csv")
7 feb <- read.csv("amazon_february_sales_big.csv")
8 newp <- read.csv("amazon_new_products_big.csv")
9
10 # Merge January and February data using ID and Product
11 merged <- full_join(jan, feb, by = c("ID", "Product"))
12
13 # Replace missing values and calculate total sales
14 merged <- merged %>%
15   mutate(
16     Jan_Sales = ifelse(is.na(Jan_Sales), 0, Jan_Sales)
17   )
18
19 # Merge new products with the merged data
20 newp <- read.csv("amazon_new_products_big.csv")
21 newp <- newp %>%
22   mutate(
23     Jan_Sales = ifelse(is.na(Jan_Sales), 0, Jan_Sales)
24   )
25
26 # Final merge
27 final <- full_join(merged, newp, by = c("ID", "Product"))
28
29 # Write to CSV
30 write.csv(merged, "merged_output.csv", row.names = FALSE)
31 write.csv(final, "final_output.csv", row.names = FALSE)
```

The console shows the following errors:

```
In addition: Warning message:
In file(file, "rt") :
cannot open file 'amazon_january_sales_big.csv': No such file or directory

> feb <- read.csv("amazon_february_sales_big.csv")
Error in file(file, "rt") : cannot open the connection

In addition: Warning message:
In file(file, "rt") :
cannot open file 'amazon_february_sales_big.csv': No such file or directory

> newp <- read.csv("amazon_new_products_big.csv")
Error in file(file, "rt") : cannot open the connection

In addition: Warning message:
In file(file, "rt") :
cannot open file 'amazon_new_products_big.csv': No such file or directory
```

The Environment pane on the right shows the following objects:

Object	Class	Attributes
data_jan	data.frame	10 obs. of 3 variables
data_new	data.frame	4 obs. of 3 variables
feb	data.frame	10 obs. of 3 variables
final	data.frame	14 obs. of 3 variables
final_list	data.frame	14 obs. of 3 variables
jan	data.frame	10 obs. of 3 variables
merged	data.frame	10 obs. of 5 variables
merged_data	data.frame	10 obs. of 5 variables
newp	data.frame	4 obs. of 3 variables

The screenshot shows the RStudio interface with the completed script. The script editor contains the following code:

```
0 jan <- read.csv("amazon_january_sales_big.csv")
7 feb <- read.csv("amazon_february_sales_big.csv")
8 newp <- read.csv("amazon_new_products_big.csv")
9
10 # Merge January and February data using ID and Product
11 merged <- full_join(jan, feb, by = c("ID", "Product"))
12
13 # Replace missing values and calculate total sales
14 merged <- merged %>%
15   mutate(
16     Jan_Sales = ifelse(is.na(Jan_Sales), 0, Jan_Sales),
17     Feb_Sales = ifelse(is.na(Feb_Sales), 0, Feb_Sales),
18     Total_Sales = Jan_Sales + Feb_Sales
19   )
20
21 # Merge new products with the merged data
22 newp <- read.csv("amazon_new_products_big.csv")
23 newp <- newp %>%
24   mutate(
25     Jan_Sales = ifelse(is.na(Jan_Sales), 0, Jan_Sales)
26   )
27
28 # Final merge
29 final <- bind_rows(jan, newp)
30
31 # Write to CSV
32 write.csv(merged, "merged_output.csv", row.names = FALSE)
33 write.csv(final, "final_output.csv", row.names = FALSE)
```

The console shows the following errors:

```
In addition: Warning message:
In file(file, "rt") :
cannot open file 'amazon_february_sales_big.csv': No such file or directory

> newp <- read.csv("amazon_new_products_big.csv")
Error in file(file, "rt") : cannot open the connection

In addition: Warning message:
In file(file, "rt") :
cannot open file 'amazon_new_products_big.csv': No such file or directory
```

The Environment pane on the right shows the following objects:

Object	Class	Attributes
data_jan	data.frame	10 obs. of 3 variables
data_new	data.frame	4 obs. of 3 variables
feb	data.frame	10 obs. of 3 variables
final	data.frame	14 obs. of 3 variables
final_list	data.frame	14 obs. of 3 variables
jan	data.frame	10 obs. of 3 variables
merged	data.frame	10 obs. of 5 variables
merged_data	data.frame	10 obs. of 5 variables
newp	data.frame	4 obs. of 3 variables

ABHISHEK DINESH SINGH  
S116

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

## Data Analysis with SAS / SPSS /R

ID	Product	Jan_Sales	Feb_Sales	Total_Sales
1	Headphon	120	150	270
2	USB Cable	80	90	170
3	Laptop Sta	45	60	105
4	Wireless M	60	75	135
5	Keyboard	110	125	235
6	Webcam	95	110	205
7	Power Bar	150	180	330
8	Smartwatc	130	140	270
9	Bluetooth	175	190	365
10	Phone Cas	90	100	190

ID	Product	Jan_Sales
1	Headphon	120
2	USB Cable	80
3	Laptop Sta	45
4	Wireless M	60
5	Keyboard	110
6	Webcam	95
7	Power Bar	150
8	Smartwatc	130
9	Bluetooth	175
10	Phone Cas	90
11	Gaming Ct	200
12	VR Headse	175
13	Charging C	85
14	LED Ring L	95