

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

SUBJECT: Data Analysis with R

Aim: Combining and appending datasets using merge() or bind_rows() in R.

The top screenshot shows the RStudio interface with the following data in the Environment pane:

data_feb
3 obs. of 3 variables

data_jan
3 obs. of 3 variables

data_new_hires
2 obs. of 3 variables

final_list
5 obs. of 3 variables

merged_data
3 obs. of 4 variables

The bottom screenshot shows the R console output for the following code:

```
> library(dplyr) # Load the library for bind_rows
> # Dataset 1: Sales in January
> data_jan <- data.frame(
+   ID = c(1, 2, 3),
+   Name = c("Amit", "Piyush", "Rahul"),
+   Jan_Sales = c(120, 180, 250)
+ )
> # Dataset 2: Sales in February
> data_feb <- data.frame(
+   ID = c(1, 2, 3),
+   Name = c("Amit", "Piyush", "Rahul"),
+   Feb_Sales = c(140, 200, 270)
+ )
> # Dataset 3: New Employees (For appending example)
> data_new_hires <- data.frame(
+   ID = c(4, 5),
+   Name = c("Sneha", "Rohan"),
+   Jan_Sales = c(90, 110)
+ )
> print("--- Data January ---")
[1] "--- Data January ---"
> print(data_jan)
  ID Name Jan_Sales
1  1 Amit       120
2  2 Piyush      180
3  3 Rahul      250
> print("--- Data February ---")
[1] "--- Data February ---"
> print(data_feb)
  ID Name Feb_Sales
1  1 Amit       140
2  2 Piyush      200
3  3 Rahul      270
> # 2. MERGE (Joining Columns)
> merged_data <- merge(data_jan, data_feb, by = c("ID", "Name"))
> print("--- Merged data (Columns Added) ---")
[1] "--- Merged Data (Columns Added) ---"
> print(merged_data)
  ID Name Jan_Sales Feb_Sales
1  1 Amit       120       140
2  2 Piyush      180       200
3  3 Rahul      250       270
> # 3. APPEND (Stacking Rows)
> final_list <- bind_rows(data_jan, data_new_hires)
> print("--- Appended Data (Rows Added) ---")
[1] "--- Appended Data (Rows Added) ---"
> print(final_list)
  ID Name Jan_Sales
1  1 Amit       120
2  2 Piyush      180
3  3 Rahul      250
4  4 Sneha        90
5  5 Rohan       110
> View(data_feb)
> View(data_jan)
```

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

SUBJECT: Data Analysis with R

The image displays two screenshots of the R Studio environment, illustrating data manipulation steps in R.

Top Screenshot:

- Environment Pane:** Shows data objects: `data_feb` (3 obs. of 3 variables), `data_jan` (3 obs. of 3 variables), `data_new_hires` (2 obs. of 3 variables), `final_list` (5 obs. of 3 variables), and `merged_data` (3 obs. of 4 variables).
- Files Pane:** Shows a file explorer view of the user's home directory.
- Console:** Displays the following R code and output:

```
> print(data_feb)
  ID Name Jan_Sales
1  1  Amit      140
2  2 Piyush     200
3  3  Rahul     270

> # 2. MERGE (Joining columns)
> # -----
> merged_data <- merge(data_jan, data_feb, by = c("ID", "Name"))
> print("---- Merged Data (Columns Added) ----")
[1] "---- Merged Data (Columns Added) ----"
> print(merged_data)
  ID Name Jan_Sales Feb_Sales
1  1  Amit      120      140
2  2 Piyush     180      200
3  3  Rahul     250      270

> # 3. APPEND (Stacking Rows)
> # -----
> final_list <- bind_rows(data_jan, data_new_hires)
> print("---- Appended Data (Rows Added) ----")
[1] "---- Appended Data (Rows Added) ----"
> print(final_list)
  ID Name Jan_Sales
1  1  Amit      120
2  2 Piyush     180
3  3  Rahul     250
4  4  Sneha      90
5  5  Rohan     110

> View(data_feb)
> View(data_jan)
> View(data_new_hires)
>
```

Bottom Screenshot:

- Environment Pane:** Shows the same data objects as the top screenshot.
- Files Pane:** Shows the same file explorer view.
- Console:** Displays the same R code and output as the top screenshot.
- Taskbar:** A Windows taskbar is visible at the bottom, showing the system clock as 11:06 on 01-12-2025.

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

SUBJECT: Data Analysis with R

The screenshot displays the R Studio environment with several open files: `S103 6th practice.R`, `data_feb`, `data_jan`, `data_new_hires`, `final_list`, and `merged_data`. The main editor shows a data frame with 3 rows and 4 columns:

ID	Name	Jan_Sales	Feb_Sales
1	Amit	120	140
2	Piyush	180	200
3	Rahul	250	270

The console shows the following R code and output:

```
> # 3. APPEND (Stacking Rows)
> #
> final_list <- bind_rows(data_jan, data_new_hires)
> print("---- Appended Data (Rows Added) ----")
[1] "---- Appended Data (Rows Added) ----"
> print(final_list)
  ID Name Jan_Sales
1  1 Amit       120
2  2 Piyush      180
3  3 Rahul      250
4  4 Sneha       90
5  5 Rohan      110
> view(data_feb)
> view(data_jan)
> view(data_new_hires)
> view(data_new_hires)
> view(final_list)
> view(merged_data)
>
```

The Environment pane on the right lists the following objects:

- `data_feb`: 3 obs. of 3 variables
- `data_jan`: 3 obs. of 3 variables
- `data_new_hires`: 2 obs. of 3 variables
- `final_list`: 5 obs. of 3 variables
- `merged_data`: 3 obs. of 4 variables

The Files pane on the right shows a list of files in the current directory, including `disease_diagnosis.csv`, `GIS DataBase`, `ISExpress`, `loan_approval.csv`, `My Music`, `My Pictures`, `My Videos`, `My Web Sites`, `NetBeansProjects`, `Power BI Desktop`, `S103 Cn 1st.pdf`, `scale for DS`, `Sound Recordings`, `T050_AnjaliT`, `vgsales.csv`, `Virtual Machines`, and `Visual Studio 18`.