

# Sheth I.u.j. And sir m.v. college of arts science and commerce

## Practical no.12th

Aim: Combining datasets vertically (concatenation) using SET statement (SAS), Merge Files - Add Cases (SPSS), and rbind() (R).

The screenshot shows the RStudio interface with the following details:

- Top Bar:** Includes icons for file operations (New, Open, Save, etc.), a search bar labeled "Go to file/function", and a dropdown menu "Addins".
- Left Sidebar:** Shows a list of open files: "pract3.R", "Untitled2\*", "time\_series\_data\_human\_activities", "crime", "3rd.R", "11th.R", "diamonds", "long\_d...".
- Console Tab:** Displays the R session history:

```
R 4.5.2 --/ 
> library(dplyr)
>
> # Load CSVs
> crime_df <- read.csv("crime.csv/crime.csv",
+                      na.strings = c("", "NA"))
>
> activity_df <- read.csv("time_series_data_human_activities.csv/time_series_data_human_activities.csv",
+                          na.strings = c("", "NA"))
>
> print("--- Data Structure Before Transformation ---")
[1] "--- Data Structure Before Transformation ---"
> print(names(crime_df))
[1] "INCIDENT_NUMBER"      "OFFENSE_CODE"           "OFFENSE_CODE_GROUP"   "OFFENSE_DESCRIPTION"
[5] "DISTRICT"              "REPORTING_AREA"        "SHOOTING"             "OCCURRED_ON_DATE"
[9] "YEAR"                  "MONTH"                 "DAY_OF_WEEK"          "HOUR"
[13] "UCR_PART"              "STREET"                "Lat"                  "Long"
[17] "Location"
> print(names(activity_df))
[1] "user"      "activity"    "timestamp"  "x.axis"     "y.axis"     "z.axis"
>
> # -----
> # Clean Crime Data
> # -----
> crime_clean <- crime_df[, c("INCIDENT_NUMBER", "OFFENSE_CODE_GROUP")]
> names(crime_clean) <- c("ID", "Value")
> crime_clean$source <- "Crime Data"
>
> # -----
```
- Environment Tab:** Shows global variables: "Global Environ", "amsterdam\_...", "cats\_df", "cats\_food", "clean\_exact", "combined\_d...", "crime", "crime\_clean", "crime\_df".
- Files Tab:** Shows the current project structure: "Home > time\_se...".

# Sheth I.u.j. And sir m.v. college of arts science and commerce

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Go to file/function, Run, Source.
- Environment Browser:** Shows a list of objects including amster, cats\_c, cats\_f, clean\_, combin, crime, crime\_, crime\_, and crime\_. A sidebar also lists Files and Plots.
- Code Editor:** A script named "pract3.R" is open, showing the following R code:

```
1 library(dplyr)
2
3
4
53:1 # (Untitled) ◊
```

```
> # -----
> # Combine the two datasets
> # -----
> combined_data <- rbind(crime_clean, activity_clean)
>
> print("---- Combined Data Summary ---")
[1] "---- Combined Data Summary ---"
> print(paste("Crime rows:", nrow(crime_clean)))
[1] "Crime rows: 319073"
> print(paste("Activity rows:", nrow(activity_clean)))
[1] "Activity rows: 1073623"
> print(paste("Total rows (Expected):", nrow(crime_clean) + nrow(activity_clean)))
[1] "Total rows (Expected): 1392696"
> print(paste("Total rows (Actual):", nrow(combined_data)))
[1] "Total rows (Actual): 1392696"
>
> print("---- Preview of Combined Data (Top and Bottom) ---")
[1] "---- Preview of Combined Data (Top and Bottom) ---"
> print(head(combined_data))
   ID          Value     Source
1 I182070945      Larceny Crime Data
2 I182070943      Vandalism Crime Data
3 I182070941      Towed Crime Data
4 I182070940 Investigate Property Crime Data
5 I182070938 Investigate Property Crime Data
6 I182070936 Motor Vehicle Accident Response Crime Data
> print(tail(combined_data))
   ID          Value     Source
```

# Sheth I.u.j. And sir m.v. college of arts science and commerce

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Includes icons for New, Open, Save, Run, and Addins.
- Project Explorer:** Shows multiple files: pract3.R, Untitled2\*, time\_series\_data\_human\_activities, crime, 3rd.R, 11th.R, diamonds, and lon.
- Code Editor:** Displays R code. Lines 21-25 are visible, starting with "# Clean User Activity Data". Line 53:1 shows "# (Untitled)".
- Console Tab:** Shows the R session output. It includes:
  - [1] "Activity rows: 1073623"
  - > print(paste("Total rows (Expected):", nrow(crime\_clean) + nrow(activity\_clean)))  
[1] "Total rows (Expected): 1392696"
  - > print(paste("Total rows (Actual):", nrow(combined\_data)))  
[1] "Total rows (Actual): 1392696"
  - >
  - > print("--- Preview of Combined Data (Top and Bottom) ---")  
[1] "--- Preview of Combined Data (Top and Bottom) ---"
  - > print(head(combined\_data))

ID	Value	Source
I182070945	Larceny	Crime Data
I182070943	Vandalism	Crime Data
I182070941	Towed	Crime Data
I182070940	Investigate	Property Crime Data
I182070938	Investigate	Property Crime Data
I182070936	Motor Vehicle Accident Response	Crime Data
  - > print(tail(combined\_data))

ID	Value	Source
1392691	36 Standing	User Activity
1392692	36 Standing	User Activity
1392693	36 Standing	User Activity
1392694	36 Standing	User Activity
1392695	36 Standing	User Activity
1392696	36 Standing	User Activity
  - >
  - > print("simran s113")  
[1] "simran s113"
  - > |