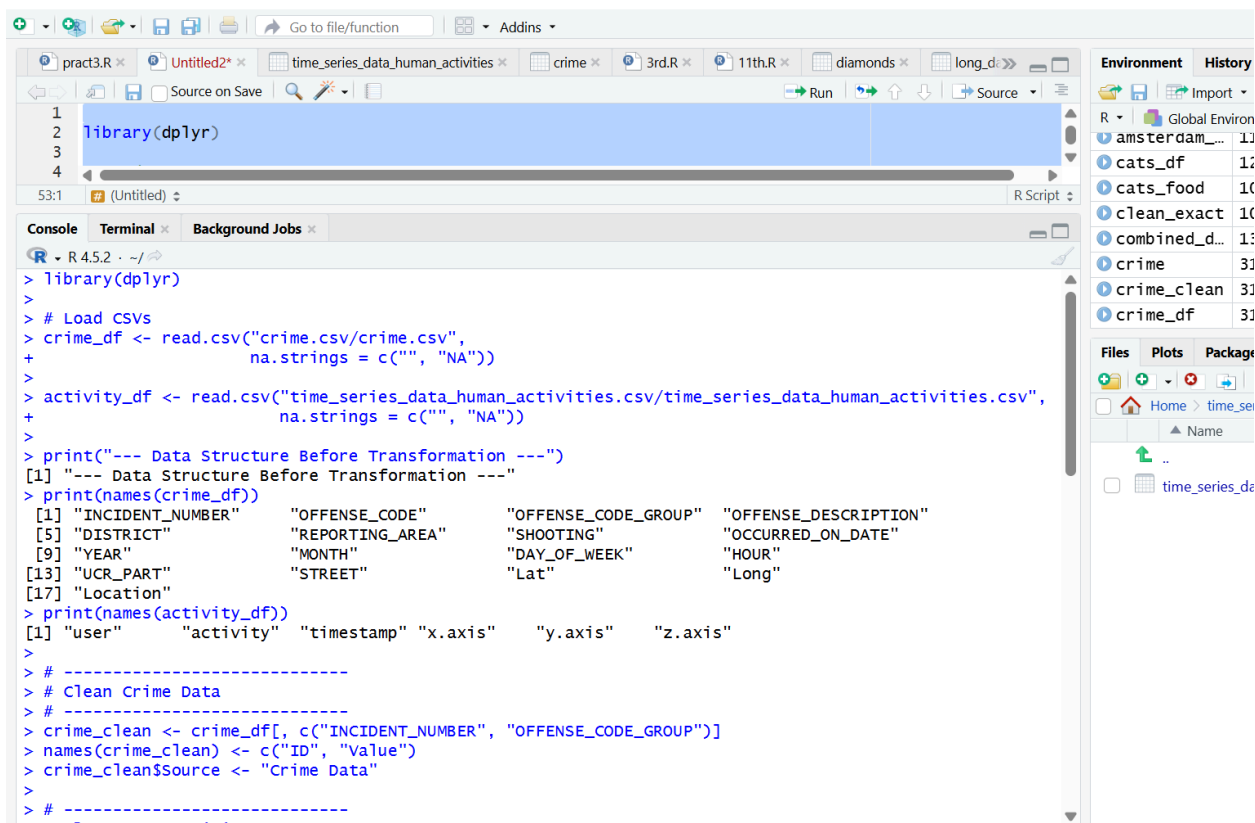


# Sheth I.u.j. And sir m.v. college of arts science and commerce

## Practical no.12th

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



The screenshot displays the R Studio environment. The script editor at the top contains the following R code:

```
1 library(dplyr)
2
3
4
```

The console at the bottom shows the execution of the code, including the loading of dplyr, reading of CSV files, and printing of data structures. The environment pane on the right lists the loaded datasets: crime, crime\_clean, and crime\_df.

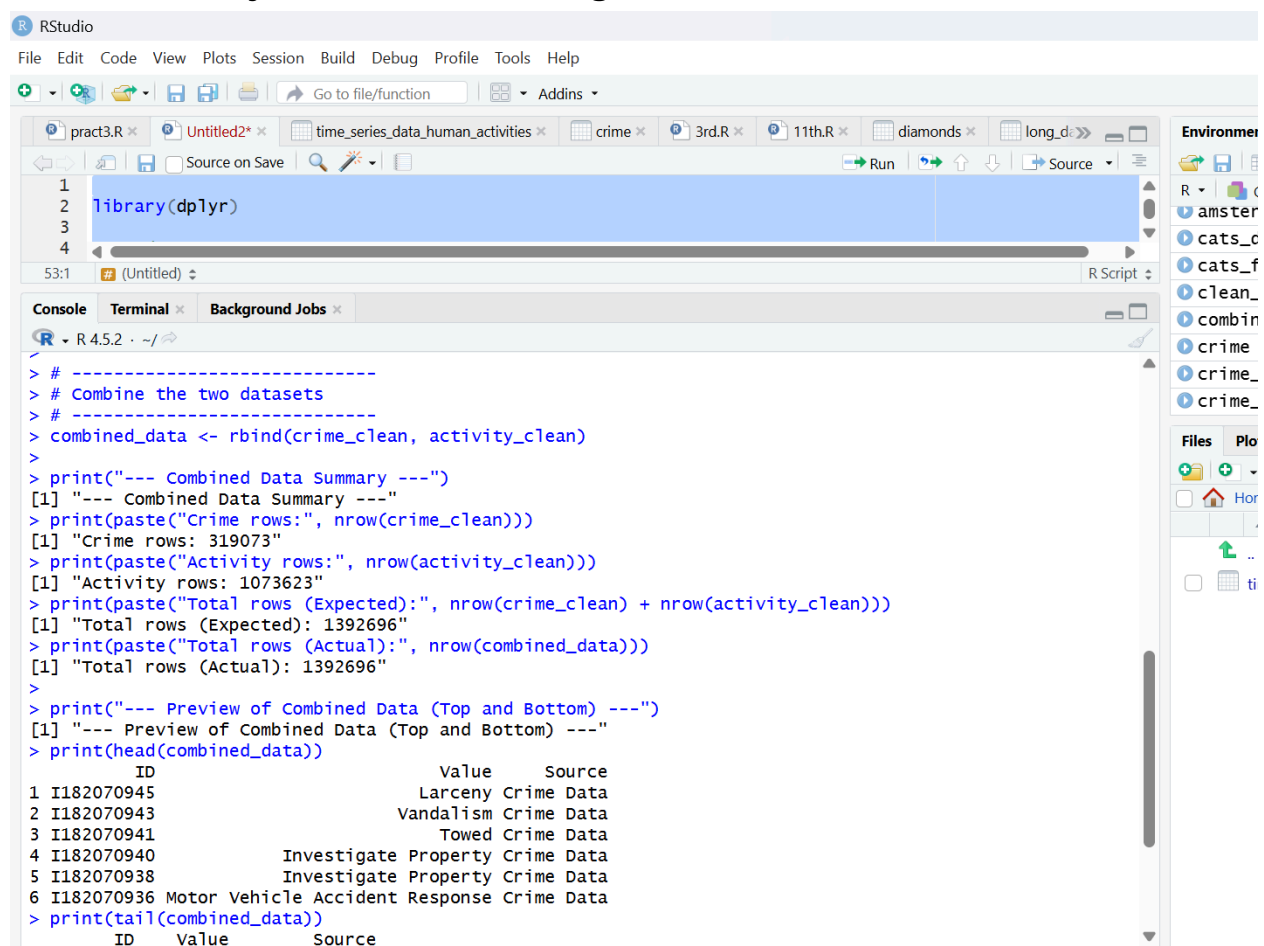
```
> library(dplyr)
>
> # Load CSVs
> crime_df <- read.csv("crime.csv/crime.csv",
+                     na.strings = c("", "NA"))
>
> activity_df <- read.csv("time_series_data_human_activities.csv/time_series_data_human_activities.csv",
+                        na.strings = c("", "NA"))
>
> print("--- 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"
>
> # -----
```

The environment pane on the right shows the following datasets:

Dataset	Size
crime	31
crime_clean	31
crime_df	31

Name: Simran S113

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



The screenshot shows the RStudio environment. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and running code. The main editor window displays an R script with the following code:

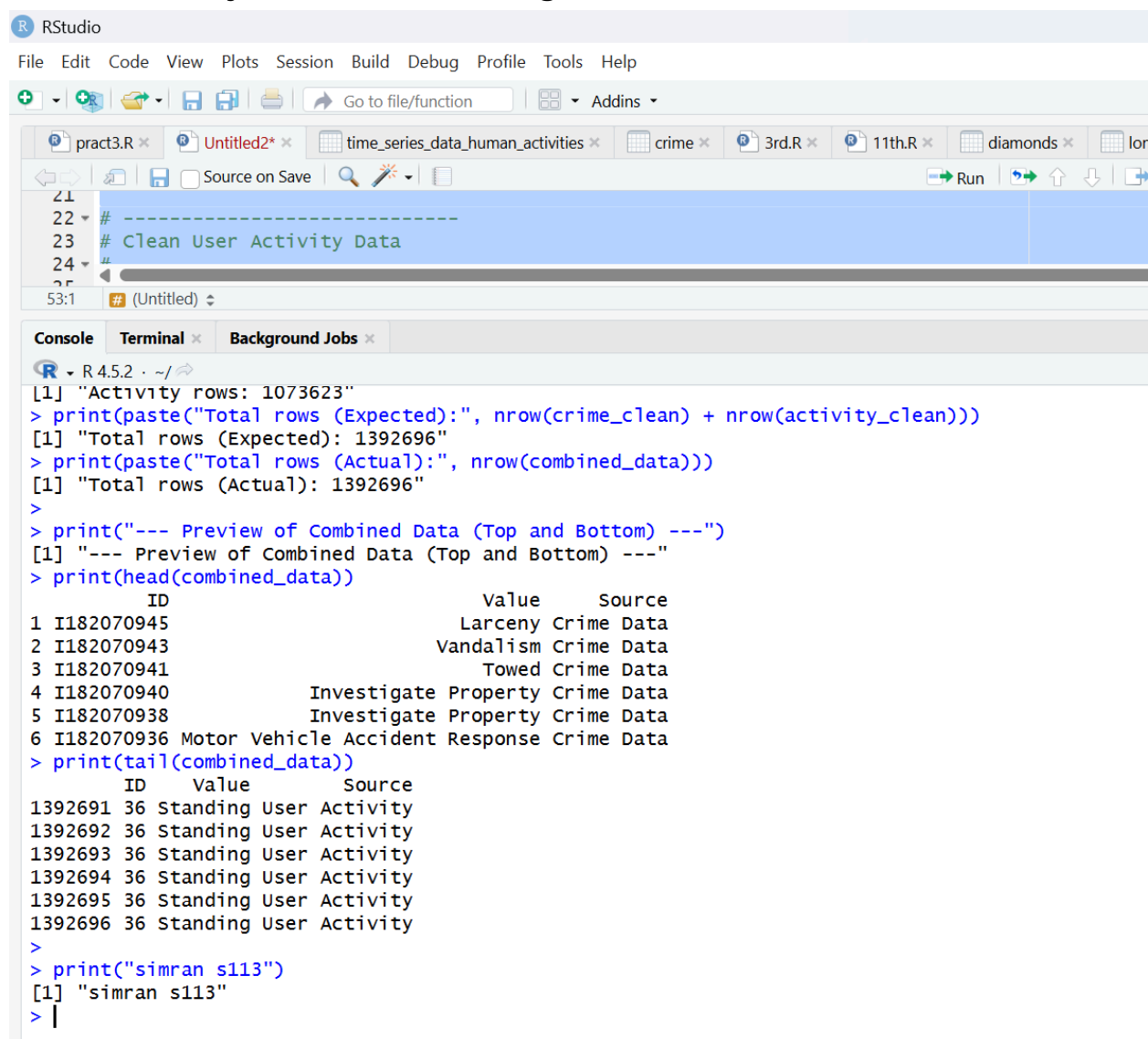
```
1 library(dplyr)
2
3
4
```

The console window shows the output of the script, which includes a summary of the combined data and a preview of the top and bottom rows. The output is as follows:

```
> # -----
> # 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
```

The right-hand side of the RStudio interface shows the Environment pane with a list of objects: R, amster, cats\_c, cats\_f, clean\_, combin, crime, crime\_, and crime\_. The Files and Plots panes are also visible on the right.

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



The screenshot shows the RStudio environment. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a 'Go to file/function' search bar. The file explorer shows several open files: pract3.R, Untitled2\*, time\_series\_data\_human\_activities, crime, 3rd.R, 11th.R, diamonds, and lon. The code editor displays a script with line numbers 21 to 25. Line 22 contains a comment: `# -----`. Line 23 contains a comment: `# Clean User Activity Data`. Line 24 contains a comment: `#`. The console window at the bottom shows the following output:

```
[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
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"
> |
```