

MVLU COLLEGE

PRACTICAL NO :- 06

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

CODE :-

```
library(dplyr)
```

```
data_main <  
read.csv("C:/Users/itlab/Downloads/S100/Mental_Health_and_Social_Media_Balance_Data  
set.csv")  
names(data_main)  
  
print("--- Original Dataset ---")  
print(head(data_main))
```

2. CREATE TWO SUBSETS FOR MERGING

```
# Dataset A → Basic Info
```

```
data_A <- data_main[, c("User_ID", "Age", "Gender")]
```

```
# Dataset B: Screen, Sleep, Stress, Happiness
```

```
data_B <- data_main[, c("User_ID",  
"Daily_Screen_Time.hrs.",  
"Sleep_Quality.1.10.",  
"Stress_Level.1.10.",  
"Happiness_Index.1.10.")]
```

```
print("--- Dataset A ---")
```

```
print(head(data_A))
```

```
print("--- Dataset B (Mental Health Info) ---")
```

```
print(head(data_B))
```

3. MERGE (Joining Columns)

```
merged_data <- merge(data_A, data_B, by = "User_ID")
```

```
print("--- Merged Dataset ---")
```

```
print(head(merged_data))
```

```
data_new_users <- data.frame(  
  User_ID = c("U151", "U152"),  
  Age = c(29, 33),  
  Gender = c("Female", "Male"),
```

```

Daily_Screen_Time.hrs = c(4.5, 6.2),
Sleep_Quality.1.10 = c(7, 6),
Stress_Level.1.10 = c(5, 7),
Happiness_Index.1.10 = c(8, 7)
)

final_dataset <- bind_rows(merged_data, data_new_users)

print("--- Final Appended Dataset ---")
print(head(final_dataset))

```

OUTPUT :-

The screenshot shows the RStudio interface with the following details:

- Console Tab:** Displays the R session history. It includes code snippets like `library(dplyr)`, `head(data_main)`, and `data_main` data.
- Environment Tab:** Shows the global environment with objects like `data_main` (500 obs. of 10 variables), `data_new_users` (2 obs. of 7 variables), and `final_dataset` (502 obs. of 11 variables).
- File Tab:** Shows tabs for Home, New Folder, New File, Delete, Rename, and More.
- System Status Bar:** Shows system icons, language (ENG), battery level (IN), and date/time (01-12-2025).

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
File Edit Code View Plots Session Build Debug Profile Tools Help
Console Terminal Background Jobs
[R - R 4.52 - ~]
[workspace loaded from ~/.RData]
> 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
> print(head(data_main))
Error: object 'data_main' not found

> data_main <- read.csv("C:/Users/itlab/Downloads/S100/Mental_Health_and_Social_Media_Balance_Dataset.csv")
> print(head(data_main))
#> # Dataset A - Basic Info
#> data_A <- data_main[, c("User_ID", "Age", "Gender")]
#>
#> # Dataset B - Mental health info
#> data_B <- data_main[, c("User_ID", "daily_Screen_Time(hrs)", "Sleep_Quality(1-10)", "Stress_Level(1-10)", "Happiness_Index(1-10)")]
#>
```

User_ID	Age	Gender	Daily_Screen_Time.hrs	Sleep_Quality.1.10	Stress_Level.1.10	Days_without_Social_Media
U001	44	Male	3.1	7	6	2
U002	30	Other	5.1	7	8	5
U003	25	Female	4.4	6	7	1
U004	36	Female	5.7	7	8	1
U005	34	Female	7.0	4	7	5
U006	38	Male	6.6	5	7	4

MVLU COLLEGE

RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Source Terminal Background Jobs
R > R 4.52 - ~/r
Error in [,.data.frame(data_main, , c("User_ID", "daily_Screen_Time(hrs)", :
  undefined columns selected

>
>
> View(data_main)
> View(data_A)
> names(data_main)
[1] "User_ID"          "Age"           "Gender"         "Daily_Screen_Time.hrs."
[5] "Sleep_Quality_1.10." "Stress_Level_1.10." "Days_Without_Social_Media" "Exercise_Frequency.week."
[9] "Social_Media_Platform" "Happiness_Index_1.10."
> # Dataset B: Screen, Sleep, Stress, Happiness
> data_B <- data_main[, c("User_ID",
+                      "daily_Screen_Time.hrs.",
+                      "Sleep_Quality_1.10.",
+                      "Stress_Level_1.10.",
+                      "Happiness_Index_1.10.)]
> print(head(data_A))
User_ID Age Gender
1  U001 44 Male
2  U002 30 Other
3  U003 23 Other
4  U004 36 Female
5  U005 34 Female
6  U006 38 Male
> print(head(data_B))
Error: unexpected symbol in:
print"

> print(head(data_B))
User_ID Daily_Screen_Time.hrs. Sleep_Quality_1.10. Stress_Level_1.10. Happiness_Index_1.10.
1  U001      3.1          7       6        10
2  U002      5.1          7       8        10
3  U003      7.4          6       7        6
4  U004      5.7          7       8        8
5  U005      7.0          4       7        8
6  U006      6.6          5       7        8
> print(head(merged_data))

Error: object 'merged_data' not found

> # 3. MERGE (Joining columns)
> merged_data <- merged(data_A, data_B, by = "User_ID")
> print(head(merged_data))
User_ID Age Gender Daily_Screen_Time.hrs. Sleep_Quality_1.10. Stress_Level_1.10. Happiness_Index_1.10.
1  U001 44 Male      3.1          7       6        10
2  U002 30 Other     5.1          7       8        10
3  U003 23 Other     7.4          6       7        6
4  U004 36 Female    5.7          7       8        8
5  U005 34 Female    7.0          4       7        8
6  U006 38 Male      6.6          5       7        8
> data_new_users <- data.frame(
+   User_ID = c("U151", "U152"),
+   Age = c(29, 33),
+   Gender = c("Female", "Male"),
+   daily_Screen_Time.hrs = c(4.5, 6.2),
+   Sleep_Quality_1.10 = c(7, 6),
+   Stress_Level_1.10 = c(5, 7),
+   Happiness_Index_1.10 = c(8, 7)
+ )
> final_dataset <- bind_rows(merged_data, data_new_users)
> print(head(final_dataset))
User_ID Age Gender Daily_Screen_Time.hrs. Sleep_Quality_1.10. Stress_Level_1.10. Happiness_Index_1.10.
1  U001 44 Male      3.1          7       6        10
2  U002 30 Other     5.1          7       8        10
3  U003 23 Other     7.4          6       7        6
4  U004 36 Female    5.7          7       8        8
5  U005 34 Female    7.0          4       7        8
6  U006 38 Male      6.6          5       7        8
Daily_Screen_Time.hrs sleep_Quality_1.10 Stress_Level_1.10 Happiness_Index_1.10
1 NA NA NA NA
2 NA NA NA NA

```

Environment History Connections Tutorial

R > Global Environment

- data_main 500 obs. of 10 variables
- data_new_users 2 obs. of 7 variables
- final_dataset 502 obs. of 11 variables
- high_exam_score 43 obs. of 20 variables
- high_study_high... 727 obs. of 20 variables
- high_study_subset 3063 obs. of 20 variables
- low_sleep_low_mo... 1226 obs. of 20 variables
- mental_health 101 obs. of 11 variables
- merged_data 500 obs. of 7 variables

Files Plots Packages Help Viewer Presentation

Gold 1:05% 10:55 ENG IN 01-12-2025

MVLU COLLEGE

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Console Terminal Background Jobs

```
R > U003      /,0      4      /      8
R > U006      6,6      5      7      8
> print(head(merged_data))
Error: object 'merged_data' not found

> # 3. MERGE (Joining columns)
> merged_data <- merge(data_A, data_B, by = "User_ID")
> print(head(merged_data))
  User_ID Age Gender Daily_Screen_Time.hrs. Sleep_Quality.1.10. Stress_Level.1.10. Happiness_Index.1.10.
1    U001 44   Male       3.1          7          6          10
2    U002 30   Other      5.1          7          8          10
3    U003 23   Other      5.4          6          7          6
4    U004 36 Female      5.7          7          8          8
5    U005 34 Female      7.0          4          7          8
6    U006 38   Male       6.6          5          7          8
> data_new_users <- data.frame(
+   User_ID = c("U151", "U152"),
+   Age = c(29, 33),
+   Gender = c("Female", "Male"),
+   Daily_Screen_Time.hrs <- c(4.5, 6.2),
+   Sleep_Quality.1.10 <- c(6, 6),
+   Stress_Level.1.10 = c(5, 7),
+   Happiness_Index.1.10 = c(8, 7)
+ )
> final_dataset <- bind_rows(merged_data, data_new_users)
> print(head(final_dataset))
  User_ID Age Gender Daily_Screen_Time.hrs. Sleep_Quality.1.10. Stress_Level.1.10. Happiness_Index.1.10.
1    U001 44   Male       3.1          7          6          10
2    U002 30   Other      5.1          7          8          10
3    U003 23   Other      5.4          6          7          6
4    U004 36 Female      5.7          7          8          8
5    U005 34 Female      7.0          4          7          8
6    U006 38   Male       6.6          5          7          8
  Daily_Screen_Time.hrs Sleep_Quality.1.10 Stress_Level.1.10 Happiness_Index.1.10
1           NA          NA          NA          NA
2           NA          NA          NA          NA
3           NA          NA          NA          NA
4           NA          NA          NA          NA
5           NA          NA          NA          NA
6           NA          NA          NA          NA
```

Environment History Connections Tutorial

R > Global Environment

- data_main 500 obs. of 10 variables
- data_new_users 2 obs. of 7 variables
- final_dataset 302 obs. of 11 variables
- high_exam_score 43 obs. of 20 variables
- high_study_high_... 727 obs. of 20 variables
- high_study_subset 3063 obs. of 20 variables
- low_sleep_low_mo... 1226 obs. of 20 variables
- mental_health 101 obs. of 11 variables
- merged_data 500 obs. of 7 variables

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home