

MVLU COLLEGE.

PRACTICAL NO :- 12

AIM :- Combining datasets vertically (concatenation) using rbind() (R).

CODE :-

```
library(dplyr)
```

```
iris_df <- read.csv("C:/Users/itlab/Downloads/S100/Iris.csv")
flower_df <- read.csv("C:/Users/itlab/Downloads/S100/flower_dataset.csv")
```

```
print("--- Column Names Before Cleaning ---")
```

```
print(names(iris_df))
```

```
print(names(flower_df))
```

```
iris_clean <- iris_df[, c("Species", "SepalLengthCm")]
```

```
names(iris_clean) <- c("Species", "Height")
```

```
# --- Prepare FLOWER dataset ---
```

```
flower_clean <- flower_df[, c("species", "height_cm")]
```

```
names(flower_clean) <- c("Species", "Height")
```

```
# Convert heights to numeric
```

```
iris_clean$Height <- as.numeric(iris_clean$Height)
```

```
flower_clean$Height <- as.numeric(flower_clean$Height)
```

```
#
```

```
=====
```

```
=====
```

```
# 3. VERTICAL CONCATENATION USING rbind()
```

```
#
```

```
=====
```

```
=====
```

```
combined_data <- rbind(iris_clean, flower_clean)
```

```
# 4. OUTPUT
```

```
print("--- Combined Data Summary ---")
```

```
print(paste("Rows in Iris dataset:", nrow(iris_clean)))
```

```
print(paste("Rows in Flower dataset:", nrow(flower_clean)))
```

```
print(paste("Total Expected Rows:", nrow(iris_clean) + nrow(flower_clean)))
```

```
print(paste("Total Actual Rows:", nrow(combined_data)))
```

MVLU COLLEGE.

```
print("--- Preview of Combined Data ---")
print(head(combined_data))
print(tail(combined_data))
```

This screenshot shows the RStudio interface with the code above running in the console. The environment pane on the right lists various datasets and objects:

- combined_data: 10150 obs. of 2 variables
- df: 497 obs. of 9 variables
- Flower_clean: 10000 obs. of 2 variables
- Flower_df: 10000 obs. of 4 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
- iris_clean: 150 obs. of 2 variables
- iris_df: 150 obs. of 6 variables
- location_pivot: 497 obs. of 10 variables
- long_df: 2485 obs. of 6 variables
- low_sleep_low_mo: 1226 obs. of 20 variables
- Mental_health: 101 obs. of 11 variables
- my_data: 1000 obs. of 14 variables
- sales_data: 1000 obs. of 14 variables
- school_type_filt_: 0 obs. of 20 variables
- Sleep_or_extracu.: 4363 obs. of 20 variables
- student: 6607 obs. of 20 variables
- Student.Mental.h.: 101 obs. of 11 variables
- wide_df: 497 obs. of 9 variables

This screenshot shows the RStudio interface with the code above running in the console. The environment pane on the right lists various datasets and objects, identical to the previous screenshot.