

PRACTICAL NO : 12

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

```

R - R4.5.2 ~ /
> # 1. SETUP: Load necessary libraries
> library(dplyr)
> # 1.1 Load the penguins_lter dataset
> penguins_df <- read.csv("Downloads/penguins_lter.csv", na.strings = c("", "NA"))
> # select relevant columns and rename for consistency
> penguins_clean <- penguins_df %>%
+   select(species, flipper.length.mm) %>%
+   rename(
+     species = species,
+     height = flipper.length.mm
+   )
> penguins_clean$height <- as.numeric(penguins_clean$height)
> # 1.2 Load the survey dataset
> survey_df <- read.csv("Downloads/surveys.csv", na.strings = c("", "NA"))
> # select relevant columns and rename for consistency
> survey_clean <- survey_df %>%
+   select(species_id, hindfoot.length) %>%
+   rename(
+     species = species_id,
+     height = hindfoot.length
+   )
> survey_clean$height <- as.numeric(survey_clean$height)
> combined_data <- rbind(penguins_clean, survey_clean)
> # 3. OUTPUT
> print("--- Combined Data Summary ---")
[1] "--- Combined Data Summary ---"
> print(paste("Penguins rows:", nrow(penguins_clean)))
[1] "Penguins rows: 344"
> print(paste("Survey rows:", nrow(survey_clean)))
[1] "Survey rows: 35549"
> print(paste("Total rows (Expected):", nrow(penguins_clean) + nrow(survey_clean)))
[1] "Total rows (Expected): 35893"
> print(paste("Total rows (Actual):", nrow(combined_data)))
[1] "Total rows (Actual): 35893"
> print(paste("Preview of Combined Data (Top and Bottom) ---"))
[1] "Preview of Combined Data (Top and Bottom) ---"

```

```

R - R4.5.2 ~ /
+   height = hindfoot.length
+ )
> survey_clean$height <- as.numeric(survey_clean$height)
> combined_data <- rbind(penguins_clean, survey_clean)
> # 3. OUTPUT
> print("--- Combined Data Summary ---")
[1] "--- Combined Data Summary ---"
> print(paste("Penguins rows:", nrow(penguins_clean)))
[1] "Penguins rows: 344"
> print(paste("Survey rows:", nrow(survey_clean)))
[1] "Survey rows: 35549"
> print(paste("Total rows (Expected):", nrow(penguins_clean) + nrow(survey_clean)))
[1] "Total rows (Expected): 35893"
> print(paste("Total rows (Actual):", nrow(combined_data)))
[1] "Total rows (Actual): 35893"
> print(paste("Preview of Combined Data (Top and Bottom) ---"))
[1] "Preview of Combined Data (Top and Bottom) ---"
> print(head(combined_data)) # Shows penguins data
  species height
1 Adelle Penguin (Pygoscelis adeliae) 181
2 Adelle Penguin (Pygoscelis adeliae) 186
3 Adelle Penguin (Pygoscelis adeliae) 195
4 Adelle Penguin (Pygoscelis adeliae) NA
5 Adelle Penguin (Pygoscelis adeliae) 193
6 Adelle Penguin (Pygoscelis adeliae) 190
> print(tail(combined_data)) # Shows survey data
  species height
35888 US NA
35889 AH NA
35890 AH NA
35891 RM 15
35892 DO 36
35893 <NA> NA
>

```