

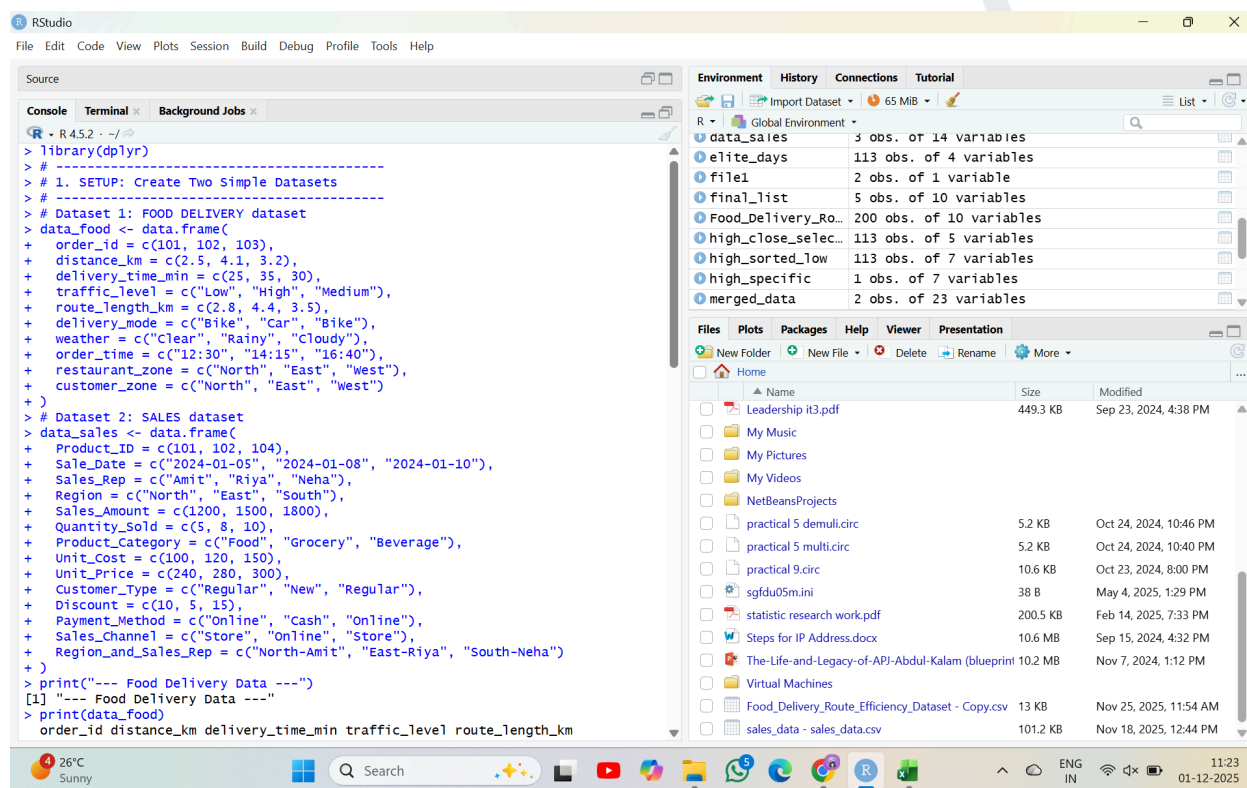
SHETH L.U.J AND SIR M.V COLLEGE

Subject: Data Analysis with SAS / SPSS /R

Practical no. 6

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

Outputs→



```
> library(dplyr)
> # -----
> # 1. SETUP: Create Two Simple Datasets
> # -----
> # Dataset 1: FOOD DELIVERY dataset
> data_food <- data.frame(
+   order_id = c(101, 102, 103),
+   distance_km = c(2.5, 4.1, 3.2),
+   delivery_time_min = c(25, 35, 30),
+   traffic_level = c("Low", "High", "Medium"),
+   route_length_km = c(2.8, 4.4, 3.5),
+   delivery_mode = c("Bike", "Car", "Bike"),
+   weather = c("Clear", "Rainy", "Cloudy"),
+   order_time = c("12:30", "14:15", "16:40"),
+   restaurant_zone = c("North", "East", "West"),
+   customer_zone = c("North", "East", "West")
+ )
> # Dataset 2: SALES dataset
> data_sales <- data.frame(
+   Product_ID = c(101, 102, 104),
+   Sale_Date = c("2024-01-05", "2024-01-08", "2024-01-10"),
+   Sales_Rep = c("Amit", "Riya", "Neha"),
+   Region = c("North", "East", "South"),
+   Sales_Amount = c(1200, 1500, 1800),
+   Quantity_Sold = c(5, 8, 10),
+   Product_Category = c("Food", "Grocery", "Beverage"),
+   Unit_Cost = c(100, 120, 150),
+   Unit_Price = c(240, 280, 300),
+   Customer_Type = c("Regular", "New", "Regular"),
+   Discount = c(10, 5, 15),
+   Payment_Method = c("Online", "Cash", "Online"),
+   Sales_Channel = c("Store", "Online", "Store"),
+   Region_and_Sales_Rep = c("North-Amit", "East-Riya", "South-Neha")
+ )
> print("--- Food Delivery Data ---")
[1] "--- Food Delivery Data ---"
> print(data_food)
  order_id distance_km delivery_time_min traffic_level route_length_km
1       101         2.5             25           Low         2.8
2       102         4.1             35          High         4.4
3       103         3.2             30        Medium         3.5
```

The Environment pane shows the following datasets:

Dataset	Observations	Variables
data_sales	3	14
elite_days	113	4
file1	2	1
final_list	5	10
Food_Delivery_Ro...	200	10
high_close_selec...	113	5
high_sorted_low	113	7
high_specific	1	7
merged_data	2	23

SHETH L.U.J AND SIR M.V COLLEGE

Subject: Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface with the following components:

- Source:** R script code for loading and merging data.
- Console:** Output of the R script, showing data frames and their dimensions.
- Environment:** List of loaded datasets and their dimensions.
- Files:** File explorer showing the project directory.

Script Code:

```
R - R4.5.2 - ~/
+ Sales_Channel = c("Store", "Online", "Store"),
+ Region_and_Sales_Rep = c("North-Amit", "East-Riya", "South-Neha")
+ )
> print("--- Food Delivery Data ---")
[1] "--- Food Delivery Data ---"
> print(data_food)
  order_id distance_km delivery_time_min traffic_level route_length_km
1      101         2.5             25             Low             2.8
2      102         4.1             35             High             4.4
3      103         3.2             30             Medium           3.5
  delivery_mode weather order_time restaurant_zone customer_zone
1         Bike   Clear    12:30             North             North
2         Car    Rainy    14:15             East              East
3         Bike   Cloudy    16:40             West              West
> print("--- Sales Data ---")
[1] "--- Sales Data ---"
> print(data_sales)
  Product_ID Sale_Date Sales_Rep Region Sales_Amount Quantity_Sold
1      101 2024-01-05      Amit North      1200             5
2      102 2024-01-08      Riya East       1500             8
3      104 2024-01-10      Neha South      1800            10
  Product_Category Unit_Cost Unit_Price Customer_Type Discount Payment_Method
1         Food       100       240       Regular       10       Online
2      Grocery       120       280         New         5         Cash
3      Beverage       150       300       Regular       15       Online
  Sales_Channel Region_and_Sales_Rep
1         Store      North-Amit
2      Online      East-Riya
3         Store      South-Neha
> merged_data <- merge(
+   data_food,
+   data_sales,
+   by.x = "order_id",
+   by.y = "Product_ID"
+ )
> print("--- Merged Data (Food + Sales) ---")
[1] "--- Merged Data (Food + Sales) ---"
> print(merged_data)
```

Environment Panel:

Dataset	Obs	Var
data_sales	3	14
elite_days	113	4
file1	2	1
final_list	5	10
Food_Delivery_Ro...	200	10
high_close_selec...	113	5
high_sorted_low	113	7
high_specific	1	7
merged_data	2	23

The screenshot shows the RStudio interface with the following components:

- Source:** R script code for binding rows and printing the final list.
- Console:** Output of the R script, showing the final merged data frame.
- Environment:** List of loaded datasets and their dimensions.
- Files:** File explorer showing the project directory.

Script Code:

```
R - R4.5.2 - ~/
+ by.x = "order_id",
+ by.y = "Product_ID"
+ )
> print("--- Merged Data (Food + Sales) ---")
[1] "--- Merged Data (Food + Sales) ---"
> print(merged_data)
  order_id distance_km delivery_time_min traffic_level route_length_km
1      101         2.5             25             Low             2.8
2      102         4.1             35             High             4.4
3      103         3.2             30             Medium           3.5
4      104         5.2             40             High             5.5
5      105         3.5             28             Low             3.8
  delivery_mode weather order_time restaurant_zone customer_zone
1         Bike   Clear    12:30             North             North
2         Car    Rainy    14:15             East              East
3         Bike   Cloudy    16:40             West              West
4         Car    Rainy    18:20             South             South
5         Bike   Clear    19:05             North             North
> final_list <- bind_rows(data_food, new_food_orders)
> print("--- Appended Data (Food + New Orders) ---")
[1] "--- Appended Data (Food + New Orders) ---"
> print(final_list)
```

Environment Panel:

Dataset	Obs	Var
data_sales	3	14
elite_days	113	4
file1	2	1
final_list	5	10
Food_Delivery_Ro...	200	10
high_close_selec...	113	5
high_sorted_low	113	7
high_specific	1	7
merged_data	2	23