

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

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

The screenshot shows the RStudio interface with the following content:

```
> library(dplyr)
> data <- read.csv("product_sales_dataset_final.csv")
> print("--- Original Dataset ---")
[1] "--- Original Dataset ---"
> print(head(data))
  Order_ID Order_Date Customer_Name City State Region Country
1      1  08-23-23   Bianca Brown  Jackson Mississippi South United States
2      2  12-20-24   Jared Edwards Grand Rapids Michigan Centre United States
3      3  01-29-24   Susan Valdez Minneapolis Minnesota Centre United States
4      4  11-29-24   Tina Williams Tallahassee Florida South United States
5      5  09-21-23 Catherine Gordon Baltimore Maryland East United States
6      6  10-18-24 Brittany Johnson Boston Massachusetts East United States
  Category Sub_Category Product_Name Quantity Unit_Price Revenue Profit
1 Accessories Small Electronics Phone Case 3 201.01 603.03 221.49
2 Accessories Small Electronics Charging Cable 4 74.30 297.20 97.09
3 Clothing & Apparel Sportswear Nike Air Force 1 1 68.19 68.19 25.47
4 Clothing & Apparel Sportswear Adidas Tracksuit 3 209.64 628.92 231.38
5 Accessories Bags Backpack 1 216.63 216.63 42.46
6 Electronics Home Appliances Instant Pot 1 701.15 701.15 90.63

> data1 <- data[1:100, ]
> data2 <- data[101:200, ]
> print("--- Dataset 1 (First 100 Rows) ---")
[1] "--- Dataset 1 (First 100 Rows) ---"
> print(head(data1))
  Order_ID Order_Date Customer_Name City State Region Country
1      1  08-23-23   Bianca Brown  Jackson Mississippi South United States
2      2  12-20-24   Jared Edwards Grand Rapids Michigan Centre United States
3      3  01-29-24   Susan Valdez Minneapolis Minnesota Centre United States
4      4  11-29-24   Tina Williams Tallahassee Florida South United States
5      5  09-21-23 Catherine Gordon Baltimore Maryland East United States
6      6  10-18-24 Brittany Johnson Boston Massachusetts East United States
  Category Sub_Category Product_Name Quantity Unit_Price Revenue Profit
1 Accessories Small Electronics Phone Case 3 201.01 603.03 221.49
2 Accessories Small Electronics Charging Cable 4 74.30 297.20 97.09
3 Clothing & Apparel Sportswear Nike Air Force 1 1 68.19 68.19 25.47
4 Clothing & Apparel Sportswear Adidas Tracksuit 3 209.64 628.92 231.38
5 Accessories Bags Backpack 1 216.63 216.63 42.46
6 Electronics Home Appliances Instant Pot 1 701.15 701.15 90.63

> print("--- Dataset 2 (Next 100 Rows) ---")
[1] "--- Dataset 2 (Next 100 Rows) ---"
```

The Environment pane on the right shows the following objects:

- combined_data: 200 obs. of 15 variables
- data: 200000 obs. of 14 variables
- data_cat1: 35908 obs. of 14 variables
- data_cat2: 62298 obs. of 14 variables
- data1: 100 obs. of 15 variables
- data2: 100 obs. of 15 variables
- delivery: 18 obs. of 3 variables
- delivery_multi_sort: 18 obs. of 4 variables
- delivery_sorted_dist: 200 obs. of 10 variables

NAME:- CHETAN MANDAVKAR

ROLL NO. S093

SUBJECT:- Data Analysis with SAS / SPSS / R

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

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Source
Console Terminal Background Jobs
R - R4.5.2 - ~/
> print("---- Dataset 2 (Next 100 Rows) ----")
[1] "---- Dataset 2 (Next 100 Rows) ----"
> print(head(data2))
  Order_ID Order_Date Customer_Name City State Region Country
101 101 06-04-24 David Manning Las Vegas Nevada West United States
102 102 09-03-23 Kayla Melendez Savannah Georgia South United States
103 103 07-18-24 Brenda Adkins Jersey City New Jersey East United States
104 104 11-11-24 Stephanie Keller New Orleans Louisiana South United States
105 105 01-11-23 Ellen Barnes Topeka Kansas Centre United States
106 106 11-19-24 Melissa Reed Boise Idaho West United States
  Category Sub_Category Product_Name Quantity Unit_Price Revenue Profit
101 Home & Furniture Kitchenware Instant Pot 6 772.57 4635.42 1536.80
102 Clothing & Apparel Kids Wear Children's Hoodie 2 364.43 728.86 325.52
103 Electronics Laptops Dell XPS 13 3 812.71 2438.13 462.72
104 Clothing & Apparel Women's Wear Old Navy Dress 3 235.05 705.15 298.92
105 Clothing & Apparel Men's Wear Levi's Jeans 2 316.23 632.46 237.00
106 Home & Furniture Bedding Tempur-Pedic Mattress 3 665.51 1996.53 397.07

> merged_data <- merge(
+   data1,
+   data2,
+   by = "City",
+   suffixes = c("_Part1", "_Part2"))
> print("---- Merged Data (Dataset1 + Dataset2 Columns Added) ----")
[1] "---- Merged Data (Dataset1 + Dataset2 Columns Added) ----"
> print(head(merged_data))
  City Order_ID_Part1 Order_Date_Part1 Customer_Name_Part1 State_Part1 Region_Part1
1 Annapolis 51 01-27-24 Dustin Campbell Maryland East
2 Annapolis 91 11-10-23 Heather Smith Maryland East
3 Annapolis 25 02-18-24 Tricia Price Maryland East
4 Annapolis 88 11-25-24 Anthony Copeland Maryland East
5 Augusta 12 09-09-23 Angela Manning Maine East
6 Augusta 12 09-09-23 Angela Manning Maine East
  Country_Part1 Category_Part1 Sub_Category_Part1 Product_Name_Part1 Quantity_Part1
1 United States Electronics Tablets iPad Pro 2
2 United States Electronics Tablets Samsung Galaxy Tab 1
3 United States Home & Furniture Storage Storage Rack 2
4 United States Electronics Wearables Apple Watch 2
5 United States Accessories Wearable Accessories Watch Strap 1
6 United States Accessories Wearable Accessories Watch Strap 1
  Unit_Price_Part1 Revenue_Part1 Profit_Part1 Order_ID_Part2 Order_Date_Part2
1 576.92 1153.84 105.99 166 01-29-24
2 424.08 424.08 35.63 166 01-29-24
3 618.88 1237.76 158.76 166 01-29-24
4 698.21 1396.42 245.62 166 01-29-24
5 124.92 124.92 24.65 153 09-26-24
6 124.92 124.92 24.65 131 06-01-23
  Customer_Name_Part2 State_Part2 Region_Part2 Country_Part2 Category_Part2 Sub_Category_Part2
1 Thomas Zuniga Maryland East United States Electronics Smartphones
2 Thomas Zuniga Maryland East United States Electronics Smartphones
3 Thomas Zuniga Maryland East United States Electronics Smartphones
4 Thomas Zuniga Maryland East United States Electronics Smartphones
5 Donald Flores Maine East United States Electronics Smartphones
6 Emily Davis Maine East United States Electronics Tablets
  Product_Name_Part2 Quantity_Part2 Unit_Price_Part2 Revenue_Part2 Profit_Part2
1 Apple iPhone 14 1 375.35 375.35 59.87
2 Apple iPhone 14 1 375.35 375.35 59.87
3 Apple iPhone 14 1 375.35 375.35 59.87
4 Apple iPhone 14 1 375.35 375.35 59.87
5 Apple iPhone 14 2 455.82 911.64 114.18
6 iPad Pro 3 143.12 429.36 82.06
```

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Source
Console Terminal Background Jobs
R - R4.5.2 - ~/
> print(head(merged_data))
  City Order_ID_Part1 Order_Date_Part1 Customer_Name_Part1 State_Part1 Region_Part1
1 Annapolis 51 01-27-24 Dustin Campbell Maryland East
2 Annapolis 91 11-10-23 Heather Smith Maryland East
3 Annapolis 25 02-18-24 Tricia Price Maryland East
4 Annapolis 88 11-25-24 Anthony Copeland Maryland East
5 Augusta 12 09-09-23 Angela Manning Maine East
6 Augusta 12 09-09-23 Angela Manning Maine East
  Country_Part1 Category_Part1 Sub_Category_Part1 Product_Name_Part1 Quantity_Part1
1 United States Electronics Tablets iPad Pro 2
2 United States Electronics Tablets Samsung Galaxy Tab 1
3 United States Home & Furniture Storage Storage Rack 2
4 United States Electronics Wearables Apple Watch 2
5 United States Accessories Wearable Accessories Watch Strap 1
6 United States Accessories Wearable Accessories Watch Strap 1
  Unit_Price_Part1 Revenue_Part1 Profit_Part1 Order_ID_Part2 Order_Date_Part2
1 576.92 1153.84 105.99 166 01-29-24
2 424.08 424.08 35.63 166 01-29-24
3 618.88 1237.76 158.76 166 01-29-24
4 698.21 1396.42 245.62 166 01-29-24
5 124.92 124.92 24.65 153 09-26-24
6 124.92 124.92 24.65 131 06-01-23
  Customer_Name_Part2 State_Part2 Region_Part2 Country_Part2 Category_Part2 Sub_Category_Part2
1 Thomas Zuniga Maryland East United States Electronics Smartphones
2 Thomas Zuniga Maryland East United States Electronics Smartphones
3 Thomas Zuniga Maryland East United States Electronics Smartphones
4 Thomas Zuniga Maryland East United States Electronics Smartphones
5 Donald Flores Maine East United States Electronics Smartphones
6 Emily Davis Maine East United States Electronics Tablets
  Product_Name_Part2 Quantity_Part2 Unit_Price_Part2 Revenue_Part2 Profit_Part2
1 Apple iPhone 14 1 375.35 375.35 59.87
2 Apple iPhone 14 1 375.35 375.35 59.87
3 Apple iPhone 14 1 375.35 375.35 59.87
4 Apple iPhone 14 1 375.35 375.35 59.87
5 Apple iPhone 14 2 455.82 911.64 114.18
6 iPad Pro 3 143.12 429.36 82.06

> data1$Part <- "First 100 Rows"
> data2$Part <- "Next 100 Rows"
> combined_data <- bind_rows(data1, data2)
> print("---- Combined Data (Rows Added) ----")
```

NAME:- CHETAN MANDAVKAR

ROLL NO. S093

SUBJECT:- Data Analysis with SAS / SPSS / R

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

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains R code for data manipulation:

```
R 4.5.2 ~ />  
2 Thomas Zuniga Maryland East United States Electronics Smartphones  
3 Thomas Zuniga Maryland East United States Electronics Smartphones  
4 Thomas Zuniga Maryland East United States Electronics Smartphones  
5 Donald Flores Maine East United States Electronics Smartphones  
6 Emily Davis Maine East United States Electronics Tablets  
  
Product_Name_Part2 Quantity_Part2 Unit_Price_Part2 Revenue_Part2 Profit_Part2  
1 Apple iPhone 14 1 375.35 375.35 59.87  
2 Apple iPhone 14 1 375.35 375.35 59.87  
3 Apple iPhone 14 1 375.35 375.35 59.87  
4 Apple iPhone 14 1 375.35 375.35 59.87  
5 Apple iPhone 14 2 455.82 911.64 114.18  
6 iPad Pro 3 143.12 429.36 82.06  
  
> data1$Part <- "First 100 Rows"  
> data2$Part <- "Next 100 Rows"  
> combined_data <- bind_rows(data1, data2)  
> print("--- Combined Data (Rows Added) ---")  
[1] "---- Combined Data (Rows Added) ----"  
> print(head(combined_data))  
Order_ID Order_Date Customer_Name City State Region Country  
1 1 08-23-23 Bianca Brown Jackson Mississippi South United States  
2 2 12-20-24 Jared Edwards Grand Rapids Michigan Centre United States  
3 3 01-29-24 Susan Valdez Minneapolis Minnesota Centre United States  
4 4 11-29-24 Tina Williams Tallahassee Florida South United States  
5 5 09-21-23 Catherine Gordon Baltimore Maryland East United States  
6 6 10-18-24 Brittany Johnson Boston Massachusetts East United States  
Category Sub_Category Product_Name Quantity Unit_Price Revenue Profit  
1 Accessories Small Electronics Phone Case 3 201.01 603.03 221.49  
2 Accessories Small Electronics Charging Cable 4 74.30 297.20 97.09  
3 Clothing & Apparel Sportswear Nike Air Force 1 1 68.19 68.19 25.47  
4 Clothing & Apparel Sportswear Adidas Tracksuit 3 209.64 628.92 231.38  
5 Accessories Bags Backpack 1 216.63 216.63 42.46  
6 Electronics Home Appliances Instant Pot 1 701.15 701.15 90.63  
  
1 First 100 Rows  
2 First 100 Rows  
3 First 100 Rows  
4 First 100 Rows  
5 First 100 Rows  
6 First 100 Rows  
>
```
- Environment Pane:** Lists loaded objects:
 - combined_data: 200 obs. of 15 variables
 - data: 200000 obs. of 14 variables
 - data_cat1: 35908 obs. of 14 variables
 - data_cat2: 62298 obs. of 14 variables
 - data1: 100 obs. of 15 variables
 - data2: 100 obs. of 15 variables
 - delivery: 18 obs. of 3 variables
 - delivery_multi_sort: 18 obs. of 4 variables
 - delivery_sorted_dist: 200 obs. of 10 variables
- Files Pane:** Shows the file explorer with various documents and folders.

NAME:- CHETAN MANDAVKAR

ROLL NO. S093

SUBJECT:- Data Analysis with SAS / SPSS /R