

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS / R

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

OUTPUT

The first screenshot shows the RStudio interface with the following code in the console:

```
> christmas_df <- read.csv("christmas.csv", na.strings = c("", "NA"))
>
> traffic_df <- read.csv("local_authority_traffic.csv", na.strings = c("", "NA"))
> print("--- Data Structure Before Transformation ---")
[1] "---- Data Structure Before Transformation ----"
> print(names(christmas_df))
[1] "date" "track" "artist" "streams" "day_of_week" "week_of_year"
> print(names(traffic_df))
[1] "local_authority_id" "local_authority_name" "year" "link_length_km"
[5] "link_length_miles" "cars_and_taxis" "all_motor_vehicles"
> christmas_clean <- christmas_df[, c("track", "streams")]
> names(christmas_clean) <- c("Category", "Value")
> traffic_clean <- traffic_df[, c("local_authority_name", "link_length_km")]
> names(traffic_clean) <- c("Category", "Value")
> christmas_clean$Value <- as.numeric(christmas_clean$Value)
> traffic_clean$Value <- as.numeric(traffic_clean$Value)
> combined_data <- rbind(christmas_clean, traffic_clean)
> print("--- Combined Data Summary ---")
[1] "---- Combined Data Summary ----"
> print(paste("Christmas rows:", nrow(christmas_clean)))
[1] "Christmas rows: 376"
> print(paste("Traffic rows:", nrow(traffic_clean)))
[1] "Traffic rows: 5529"
> print(paste("Total rows (Expected):", nrow(christmas_clean) + nrow(traffic_clean)))
[1] "Total rows (Expected): 5905"
> print(paste("Total rows (Actual):", nrow(combined_data)))
[1] "Total rows (Actual): 5905"
> print("---- Preview of Combined Data (Top and Bottom) ----")
[1] "---- Preview of Combined Data (Top and Bottom) ----"
> print(head(combined_data))
```

The second screenshot shows the RStudio interface with the following code in the console:

```
> christmas_clean$Value <- as.numeric(christmas_clean$Value)
> traffic_clean$Value <- as.numeric(traffic_clean$Value)
> combined_data <- rbind(christmas_clean, traffic_clean)
> print("--- Combined Data Summary ---")
[1] "---- Combined Data Summary ----"
> print(paste("Christmas rows:", nrow(christmas_clean)))
[1] "Christmas rows: 376"
> print(paste("Traffic rows:", nrow(traffic_clean)))
[1] "Traffic rows: 5529"
> print(paste("Total rows (Expected):", nrow(christmas_clean) + nrow(traffic_clean)))
[1] "Total rows (Expected): 5905"
> print(paste("Total rows (Actual):", nrow(combined_data)))
[1] "Total rows (Actual): 5905"
> print("---- Preview of Combined Data (Top and Bottom) ----")
[1] "---- Preview of Combined Data (Top and Bottom) ----"
> print(head(combined_data))
```

The Environment pane on the right shows the following data objects:

Object	Size
christmas...	376 obs. of 2 variab1...
christmas...	376 obs. of 6 variab1...
combined...	5905 obs. of 2 variab...
traffic...	5529 obs. of 2 variab...
traffic...	5529 obs. of 7 variab...

The Files pane on the right shows the following files:

Name	Size
.RData	9.9 KB
.Rhistory	5.8 KB
bi.pbix	340 KB
Custom Office Templates	
GIS DataBase	
IISExpress	
My Web Sites	
NetBeansProjects	
Power BI Desktop	
Virtual Machines	
Visual Studio 2022	
local_authority_traffic.csv	327.5 KB
christmas.csv	25.6 KB

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS / R

Screenshot 1: RStudio Data Frame

Category	Value
All I Want for Christmas Is You	3821572
All I Want for Christmas Is You	6130250
Last Christmas	4040197
All I Want for Christmas Is You	11641888
It's Beginning to Look a Lot Like Christmas	4577042
Last Christmas	7133652
Rockin' Around the Christmas Tree	4423222
Santa Tell Me	4221612
All I Want for Christmas Is You	19808428
Feliz Navidad	4780300
It's Beginning to Look a Lot Like Christmas	8172327
Last Christmas	13594807
Rockin' Around the Christmas Tree	6903809
Santa Tell Me	7521462
All I Want for Christmas Is You	21535828
Feliz Navidad	5436759
It's Beginning to Look a Lot Like Christmas	9148642
Last Christmas	14990547
Rockin' Around the Christmas Tree	7888911
Santa Tell Me	8590324

Showing 1 to 21 of 5,905 entries. 2 total columns

Screenshot 2: RStudio Data Frame

date	track	artist	streams	day_of_week	week_of_year
16-11-2017	All I Want for Christmas Is You	Mariah Carey	3821572	Thursday	46
23-11-2017	All I Want for Christmas Is You	Mariah Carey	6130250	Thursday	47
23-11-2017	Last Christmas	Wham!	4040197	Thursday	47
30-11-2017	All I Want for Christmas Is You	Mariah Carey	11641888	Thursday	48
30-11-2017	It's Beginning to Look a Lot Like Christmas	Michael Bublé	4577042	Thursday	48
6-12-2017	Last Christmas	Wham!	7133652	Thursday	48
7-12-2017	Rockin' Around the Christmas Tree	Brenda Lee	4423222	Thursday	48
8-12-2017	Santa Tell Me	Ariana Grande	4221612	Thursday	48
9-12-2017	All I Want for Christmas Is You	Mariah Carey	19808428	Thursday	49
10-12-2017	Feliz Navidad	José Feliciano	4780300	Thursday	49
11-12-2017	It's Beginning to Look a Lot Like Christmas	Michael Bublé	8172327	Thursday	49
12-12-2017	Last Christmas	Wham!	13594807	Thursday	49
13-12-2017	Rockin' Around the Christmas Tree	Brenda Lee	6903809	Thursday	49
14-12-2017	Santa Tell Me	Ariana Grande	7521462	Thursday	49
14-12-2017	All I Want for Christmas Is You	Mariah Carey	21535828	Thursday	50
16-12-2017	Feliz Navidad	José Feliciano	5436759	Thursday	50
14-12-2017	It's Beginning to Look a Lot Like Christmas	Michael Bublé	9148642	Thursday	50
18-12-2017	Last Christmas	Wham!	14990547	Thursday	50
14-12-2017	Rockin' Around the Christmas Tree	Brenda Lee	7888911	Thursday	50
14-12-2017	Santa Tell Me	Ariana Grande	8590324	Thursday	50

Showing 1 to 21 of 376 entries. 6 total columns

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS / R

RStudio interface showing a data frame with columns: local_authority_id, local_authority_name, year, link_length_km, link_length_miles, cars_and_taxis, and all_motor_vehicles. The data is sorted by local_authority_id. The console shows the data frame structure.

	local_authority_id	local_authority_name	year	link_length_km	link_length_miles	cars_and_taxis	all_motor_vehicles
1	45	Aberdeenshire	2019	6273.378	3898.10	1537817161.0	2055244624
2	107	Lambeth	2019	377.000	234.26	397710910.8	547123377
3	172	Newcastle upon Tyne	2019	984.992	612.05	1006029025.0	1209012058
4	93	Tower Hamlets	2019	287.327	178.54	465293697.1	630938804
5	158	St. Helens	2019	740.220	459.95	728175389.4	930780139
6	69	Worcestershire	2019	4273.420	2655.38	3556865608.0	4546565782
7	104	Lewisham	2019	447.188	277.87	467457465.5	613517557
8	145	Camden	2019	279.342	173.58	218491001.6	297009025
9	169	Kingston upon Hull. City of	2019	772.361	479.92	827160827.6	1043436141
10	56	Stockport	2019	1004.974	624.46	1161305176.0	1412753855
11	186	Bedford	2019	930.400	578.12	722473805.8	917378332
12	8	Swansea	2019	1167.030	725.16	983067119.3	1200222980
13	5	Somerset	2019	6807.584	4230.04	3685156283.0	4704198951
14	19	Cardiff	2019	1093.979	679.77	1734110249.0	2107567897
15	197	Kirklees	2019	1909.033	1186.22	1518187481.0	1959431313
16	115	Bath and North East Somerset	2019	1079.621	670.85	628532401.9	780570733
17	180	Bracknell Forest	2019	474.714	294.97	413645269.3	497682894
18	30	Falkirk	2019	970.375	602.96	781875898.7	1015092641
19	78	Hertfordshire	2019	5026.716	3123.46	6335184262.0	8088341294
20	184	Torbay	2019	530.393	329.57	366969775.4	446634659

Showing 1 to 21 of 5,529 entries. 7 total columns

Console