

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

SUBJECT : Data Analysis with SAS / SPSS / R

AIM : Reshaping data using pivot_longer() and pivot_wider() (R).

OUTPUT

The first screenshot shows the RStudio interface with the following console output:

```
R - R 4.5.1 - ~/R
> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':
  filter, lag

The following objects are masked from 'package:base':
  intersect, setdiff, setequal, union

Warning message:
package 'dplyr' was built under R version 4.5.2

> library(tidyr)

Warning message:
package 'tidyr' was built under R version 4.5.2

> df <- read.csv("local_authority_traffic.csv", na.strings = c("", "NA")) %>%
+   mutate(RowID = row_number())
> print("--- 1. Original Wide Data ---")
> print(head(df))
  local_authority_id local_authority_name year link_length_km link_length_miles cars_and_taxis
1                45      Aberdeenshire 2019      6273.378      3898.10      1537817161
2                107      Lambeth 2019      377.000      234.26      397710911
3                172  Newcastle upon Tyne 2019      984.992      612.05      1006029025
4                 93    Tower Hamlets 2019      287.327      178.54      465293697
5                 158    St. Helens 2019      740.220      459.95      728175389
```

The second screenshot shows the RStudio interface with the following console output:

```
R - R 4.5.1 - ~/R
> print("--- 1. Original Wide Data ---")
> print(head(df))
  local_authority_id local_authority_name year link_length_km link_length_miles cars_and_taxis
1                45      Aberdeenshire 2019      6273.378      3898.10      1537817161
2                107      Lambeth 2019      377.000      234.26      397710911
3                172  Newcastle upon Tyne 2019      984.992      612.05      1006029025
4                 93    Tower Hamlets 2019      287.327      178.54      465293697
5                 158    St. Helens 2019      740.220      459.95      728175389
6                 69  Worcestershire 2019      4273.420      2655.38      3556865608

all_motor_vehicles RowID
1      2055244624      1
2      547123377      2
3      1209012058      3
4      630938804      4
5      930780139      5
6      4546565782      6

> long_df <- df %>%
+   pivot_longer(
+     cols = c(link_length_km, link_length_miles, cars_and_taxis, all_motor_vehicles),
+     names_to = "Metric",
+     values_to = "Value"
+   )
> print("--- 2. Long Format (pivot_longer) ---")
> print(head(long_df, 10))
# A tibble: 10 x 6
  local_authority_id local_authority_name year RowID Metric Value
  <int> <chr> <int> <int> <chr> <dbl>
1      45 Aberdeenshire 2019 1 link_length_km 6273.
2      45 Aberdeenshire 2019 1 link_length_miles 3898.
3      45 Aberdeenshire 2019 1 cars_and_taxis 1537817161
4      45 Aberdeenshire 2019 1 all_motor_vehicles 2055244624
5      107 Lambeth 2019 2 link_length_km 377
```

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

SUBJECT : Data Analysis with SAS / SPSS / R

The image displays two screenshots of the RStudio interface, showing the execution of R code for data manipulation and analysis.

Top Screenshot:

- Source:** The R script shows the following code:

```
R - R 4.5.1 - ~/r
107 Lambeth 2019 2 all_motor_vehicles 547123377.
172 Newcastle upon Tyne 2019 3 link_length_km 985.
172 Newcastle upon Tyne 2019 3 link_length_miles 612.
> wide_df <- long_df %>%
+ pivot_wider(
+ names_from = Metric,
+ values_from = Value
+ )
> print("--- 3. Wide Format (Back to Original) ---")
[1] "--- 3. Wide Format (Back to Original) ---"
> print(head(wide_df))
# A tibble: 6 x 8
  local_authority_id local_authority_name year RowID link_length_km link_length_miles cars_and_taxis
  <int> <chr> <int> <int> <dbl> <dbl> <dbl>
1 45 Aberdeenshire 2019 1 6273. 3898. 1537812161
2 107 Lambeth 2019 2 377. 234. 397210911.
3 172 Newcastle upon Tyne 2019 3 985. 612. 1006029025
4 93 Tower Hamlets 2019 4 287. 179. 465293697.
5 158 St. Helens 2019 5 740. 460. 728125389.
6 69 Worcestershire 2019 6 4273. 2655. 3556865608
```
- Console:** The output shows the execution of the code, including the print statement and the head of the wide format data frame.
- Environment:** The environment pane shows the loaded data frames: `authorit...` (5529 obs. of 210 varia...), `df` (5529 obs. of 8 variabl...), `df_calc` (100 obs. of 12 variabl...), `df_clean` (100 obs. of 11 variabl...), `df_logic` (100 obs. of 13 variabl...), and `df_text` (100 obs. of 12 variabl...).

Bottom Screenshot:

- Source:** The R script continues with the following code:

```
# i 1 more variable: all_motor_vehicles <dbl>
> authority_pivot <- df %>%
+ select(RowID, local_authority_name, link_length_km) %>%
+ pivot_wider(
+ names_from = local_authority_name,
+ values_from = link_length_km
+ )
> print("--- 4. Local Authority Pivot (Columns = Authority Names) ---")
[1] "--- 4. Local Authority Pivot (Columns = Authority Names) ---"
> print(head(authority_pivot))
# A tibble: 6 x 210
  RowID Aberdeenshire Lambeth 'Newcastle upon Tyne' 'Tower Hamlets' 'St. Helens' Worcestershire Lewisham
  <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 1 6273. NA NA NA NA NA NA
2 2 NA 377 NA NA NA NA NA
3 3 NA NA 985. NA NA NA NA NA
4 4 NA NA 287. NA NA NA NA NA
5 5 NA NA 740. NA NA 740. NA NA
6 6 NA NA 4273. NA NA NA 4273. NA
```
- Console:** The output shows the execution of the code, including the print statement and the head of the authority pivot data frame.
- Environment:** The environment pane shows the loaded data frames: `authorit...` (5529 obs. of 210 varia...), `df` (5529 obs. of 8 variabl...), `df_calc` (100 obs. of 12 variabl...), `df_clean` (100 obs. of 11 variabl...), `df_logic` (100 obs. of 13 variabl...), and `df_text` (100 obs. of 12 variabl...).

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

SUBJECT : Data Analysis with SAS / SPSS / R

The first screenshot shows the RStudio interface with a data frame named 'authority_pivot' containing 8 columns: local_authority_id, local_authority_name, year, link_length_km, link_length_miles, cars_and_taxis, all_motor_vehicles, and RowID. The data is displayed in a table with 18 rows. The console shows the command `View(authority_pivot)`.

The second screenshot shows the same data frame after being pivoted into a long format. The columns are now: local_authority_id, local_authority_name, year, RowID, Metric, and Value. The data is displayed in a table with 19 rows. The console shows the command `View(authority_pivot)`.

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

SUBJECT : Data Analysis with SAS / SPSS / R

RStudio interface showing a data frame 'authority_pivot' with 5529 observations and 8 variables. The variables are: local_authority_id, local_authority_name, year, RowID, link_length_km, link_length_miles, cars_and_taxis, and all_motor_vehicles. The console shows the command: `# i use 'colnames()' to see all variable names` and `> View(authority_pivot)`. The environment pane shows the data frame with 5529 observations and 8 variables.

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

Showing 1 to 18 of 5,529 entries. 8 total columns

Console: `# i use 'colnames()' to see all variable names`
`> View(authority_pivot)`

Environment: R, Global Environment, 165 MiB

Data: authority_pivot (5529 obs. of 210 variab...), df (5529 obs. of 8 variab...), df_calc (100 obs. of 12 variab...), df_clean (100 obs. of 11 variab...), df_logic (100 obs. of 13 variab...), df_text (100 obs. of 12 variab...)

Files: Home, RData (9.9 KB)

Snipping Tool: Screenshot copied to clipboard, Automatically saved to screenshots folder, Mark-up and share