

Sheth I.u.j. And sir m.v. college of arts science and commerce

Practical no.11th

Aim: 11Reshaping data using PROC TRANSPOSE (SAS), Restructure Data Wizard (SPSS), and pivot_longer()/pivot_wider() (R).

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
3rd.R x 11th.R x diamonds x long_data x original x wide_data x 13th.R x elevator_traffic_dataset x
Source on Save Run
13
14 names_to = "Variable",
15 values_to = "Value",
16 values_transform = as.character
17
26:24 (Top Level)
Console Terminal Background Jobs
R 4.5.2 ~ /
> library(tidyverse)
> library(dplyr)
>
> data <- read.csv("diamonds.csv/diamonds.csv")
>
> original <- data
> print("Original Data:")
[1] "Original Data:"
> print(original)
  carat    cut    color clarity depth table     x     y     z price
1   0.23  Ideal     E    SI2   61.5   55.0  3.95  3.98  2.43   326
2   0.21  Premium   E    SI1   59.8   61.0  3.89  3.84  2.31   326
3   0.23    Good     E    VS1   56.9   65.0  4.05  4.07  2.31   327
4   0.29  Premium   I    VS2   62.4   58.0  4.20  4.23  2.63   334
5   0.31    Good     J    SI2   63.3   58.0  4.34  4.35  2.75   335
6   0.24 Very Good   J   VVS2   62.8   57.0  3.94  3.96  2.48   336
7   0.24 Very Good   I   VVS1   62.3   57.0  3.95  3.98  2.47   336
8   0.26 Very Good   H    SI1   61.9   55.0  4.07  4.11  2.53   337
9   0.22    Fair     E    VS2   65.1   61.0  3.87  3.78  2.49   337
10  0.23 Very Good   H    VS1   59.4   61.0  4.00  4.05  2.39   338
11  0.30    Good     J    SI1   64.0   55.0  4.25  4.28  2.73   339
12  0.23  Ideal     J    VS1   62.8   56.0  3.93  3.90  2.46   340
13  0.22  Premium   F    SI1   60.4   61.0  3.88  3.84  2.33   342
14  0.31  Ideal     J    SI2   62.2   54.0  4.35  4.37  2.71   344
15  0.20  Premium   E    SI2   60.2   62.0  3.79  3.75  2.27   345
16  0.32  Premium   E     I1   60.9   58.0  4.38  4.42  2.68   345
17  0.30  Ideal     I    SI2   62.0   54.0  4.31  4.34  2.68   348
18  0.30    Good     J    SI1   63.4   54.0  4.23  4.29  2.70   351
```

Sheth I.u.j. And sir m.v. college of arts science and commerce

The screenshot displays an RStudio interface with multiple tabs at the top: 3rd.R, 11th.R, diamonds, long_data, original, wide_data, 13th.R, and elevator_traffic_dataset. The main editor window shows R code for converting a wide dataset to long format using the `pivot_longer()` function. The console output shows the resulting long data structure, which is a tibble with 485,460 rows and 3 columns: `carat`, `Variable`, and `Value`. A sample of the first 10 rows is displayed, showing various diamond attributes like cut, color, clarity, depth, table, x, y, z, and price.

```
13 cols = carat,  
14 names_to = "Variable",  
15 values_to = "Value",  
16 values_transform = as.character  
17 )  
26:24 (Top Level) ↕
```

Console output for long data:

```
> # WIDE -> LONG  
> long_data <- original %>%  
+   pivot_longer(  
+     cols = ~carat,  
+     names_to = "Variable",  
+     values_to = "Value",  
+     values_transform = as.character  
+   )  
>  
> print("Long Data:")  
[1] "Long Data:"  
> print(long_data)  
# A tibble: 485,460 x 3  
  carat Variable Value  
  <dbl> <chr> <chr>  
1 0.23 cut Ideal  
2 0.23 color E  
3 0.23 clarity SI2  
4 0.23 depth 61.5  
5 0.23 table 55  
6 0.23 x 3.95  
7 0.23 y 3.98  
8 0.23 z 2.43  
9 0.23 price 326  
10 0.21 cut Premium  
# i 485,450 more rows  
# i Use `print(n = ...)` to see more rows
```

The second screenshot shows the conversion of the long data back to wide format using the `pivot_wider()` function. The console output shows the resulting wide data structure, which is a tibble with 273 rows and 10 columns: `carat`, `cut`, `color`, `clarity`, `depth`, `table`, `x`, `y`, `z`, and `price`. A sample of the first 10 rows is displayed, showing the reconstructed diamond attributes.

```
13 cols = carat,  
14 names_to = "Variable",  
15 values_to = "Value",  
16 values_transform = as.character  
17 )  
26:24 (Top Level) ↕
```

Console output for wide data:

```
> print("Wide Data:")  
[1] "Wide Data:"  
> print(wide_data)  
# A tibble: 273 x 10  
  carat cut color clarity depth table x y z price  
  <dbl> <list> <list> <list> <list> <list> <list> <list> <list> <list>  
1 0.23 <chr [293]> <chr [293]> <chr [293]> <chr [293]> <chr> <chr> <chr> <chr> <chr>  
2 0.21 <chr [9]> <chr [9]> <chr [9]> <chr [9]> <chr> <chr> <chr> <chr> <chr>  
3 0.29 <chr [130]> <chr [130]> <chr [130]> <chr [130]> <chr> <chr> <chr> <chr> <chr>  
4 0.31 <chr [2,249]> <chr [2,249]> <chr [2,249]> <chr [2,249]> <chr> <chr> <chr> <chr> <chr>  
5 0.24 <chr [254]> <chr [254]> <chr [254]> <chr [254]> <chr> <chr> <chr> <chr> <chr>  
6 0.26 <chr [253]> <chr [253]> <chr [253]> <chr [253]> <chr> <chr> <chr> <chr> <chr>  
7 0.22 <chr [5]> <chr [5]> <chr [5]> <chr [5]> <chr> <chr> <chr> <chr> <chr>  
8 0.3 <chr [2,604]> <chr [2,604]> <chr [2,604]> <chr [2,604]> <chr> <chr> <chr> <chr> <chr>  
9 0.2 <chr [12]> <chr [12]> <chr [12]> <chr [12]> <chr> <chr> <chr> <chr> <chr>  
10 0.32 <chr [1,840]> <chr [1,840]> <chr [1,840]> <chr [1,840]> <chr> <chr> <chr> <chr> <chr>  
# i 263 more rows  
# i Use `print(n = ...)` to see more rows  
>  
> # CREATE A SMALL REPORT  
> # -----  
> report <- original %>%  
+   summarise(  
+     total_rows = n(),  
+     avg_carat = mean(carat, na.rm = TRUE),  
+     min_carat = min(carat, na.rm = TRUE),  
+   )
```

On the right side of the RStudio interface, the 'Data' pane shows the following objects and their sizes:

Object	Size
data	53
long_data	48
original	53
report	1
wide_data	27

The 'Files' pane on the right shows a list of files and folders, including .Rhistory, 11th.R, 13th.R, 14th.R.pdf, 14th.R, 15th.R, apache-maven, apache-maven, apache-tomca, apache-tomca, Arduino, College_Marks, and Default.rdp.

Sheth I.u.j. And sir m.v. college of arts science and commerce

```
Source on Save
14 names_to = "Variable",
15 values_to = "value",
16 values_transform = as.character
17 )
26:24 (Top Level) R Script

Console Terminal Background Jobs
R 4.5.2 ~
2 0.21 <chr [9]> <chr [9]> <chr [9]> <chr [9]> <chr> <chr> <chr> <chr> <chr>
3 0.29 <chr [130]> <chr [130]> <chr [130]> <chr [130]> <chr> <chr> <chr> <chr> <chr>
4 0.31 <chr [2,249]> <chr [2,249]> <chr [2,249]> <chr [2,249]> <chr> <chr> <chr> <chr> <chr>
5 0.24 <chr [254]> <chr [254]> <chr [254]> <chr [254]> <chr> <chr> <chr> <chr> <chr>
6 0.26 <chr [253]> <chr [253]> <chr [253]> <chr [253]> <chr> <chr> <chr> <chr> <chr>
7 0.22 <chr [5]> <chr [5]> <chr [5]> <chr [5]> <chr> <chr> <chr> <chr> <chr>
8 0.3 <chr [2,604]> <chr [2,604]> <chr [2,604]> <chr [2,604]> <chr> <chr> <chr> <chr> <chr>
9 0.2 <chr [12]> <chr [12]> <chr [12]> <chr [12]> <chr> <chr> <chr> <chr> <chr>
10 0.32 <chr [1,840]> <chr [1,840]> <chr [1,840]> <chr [1,840]> <chr> <chr> <chr> <chr> <chr>
# i 263 more rows
# i Use `print(n = ...)` to see more rows
>
> # CREATE A SMALL REPORT
> # -----
> report <- original %>%
+ summarise(
+   total_rows = n(),
+   avg_carat = mean(carat, na.rm = TRUE),
+   min_carat = min(carat, na.rm = TRUE),
+   max_carat = max(carat, na.rm = TRUE)
+ )
>
> print("Small Report:")
[1] "Small Report:"
> print(report)
total_rows avg_carat min_carat max_carat
1 53940 0.7979397 0.2 5.01
> |
```