

PRACTICAL NO : 11

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

```

R - R4.5.2 - ~/
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
> library(dplyr)
> library(tidyr)
> # Read data and add an InvoiceIndex (Essential for tracking rows during pivots)
> df <- read.csv("Downloads/SuperMarket Analysis.csv",
+               na.strings = c("", "NA")) %>%
+   mutate(InvoiceIndex = row_number()) %>%
+   select(InvoiceIndex, Product.line, unit.price, Tax.5.)
> print("--- 1. Original wide data ---")
[1] "--- 1. original wide data ---"
> print(head(df))
  InvoiceIndex Product.line unit.price Tax.5.
1           1 Health and beauty 74.69 26.1415
2           2 Electronic accessories 15.28 3.8200
3           3 Home and lifestyle 46.33 16.2155
4           4 Health and beauty 58.22 23.2880
5           5 Sports and travel 86.31 30.2085
6           6 Electronic accessories 85.39 29.8865
> # Combine 'unit.price' and 'Tax.5.' into long format
> long_df <- df %>%
+   pivot_longer(
+     cols = c(unit.price, Tax.5.),
+     names_to = "Metric",
+     values_to = "value"
+   )
> print("--- 2. Long Format (pivot_longer) ---")
[1] "--- 2. Long Format (pivot_longer) ---"
> print(head(long_df, 6))
# A tibble: 6 x 4
  InvoiceIndex Product.line      Metric      value
  <int> <chr> <chr> <dbl>
1           1 Health and beauty unit.price 74.7
2           1 Health and beauty Tax.5. 26.1
3           2 Electronic accessories unit.price 15.3
4           2 Electronic accessories Tax.5. 3.82
5           3 Home and lifestyle unit.price 46.3
6           3 Home and lifestyle Tax.5. 16.2
  
```

```

R - R4.5.2 - ~/
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
> print(head(long_df, 6))
# A tibble: 6 x 4
  InvoiceIndex Product.line      Metric      value
  <int> <chr> <chr> <dbl>
1           1 Health and beauty unit.price 74.7
2           1 Health and beauty Tax.5. 26.1
3           2 Electronic accessories unit.price 15.3
4           2 Electronic accessories Tax.5. 3.82
5           3 Home and lifestyle unit.price 46.3
6           3 Home and lifestyle Tax.5. 16.2
> 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 4
  InvoiceIndex Product.line      unit.price Tax.5.
  <int> <chr> <dbl> <dbl>
1           1 Health and beauty 74.7 26.1
2           2 Electronic accessories 15.3 3.82
3           3 Home and lifestyle 46.3 16.2
4           4 Health and beauty 58.2 23.3
5           5 Sports and travel 86.3 30.2
6           6 Electronic accessories 85.4 29.9
> df_clean <- df %>%
+   mutate(Product.line = ifelse(is.na(Product.line), "Unknown", Product.line))
> category_pivot <- df_clean %>%
+   select(InvoiceIndex, Product.line, unit.price) %>%
+   pivot_wider(
+     names_from = Product.line,
+     values_from = unit.price
+   )
  
```

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains R code for data manipulation using dplyr and tidyr packages. The code includes comments and function calls like `print(head(wide_df))`, `df_clean <- df %>%`, `mutate(Product.line = ifelse(is.na(Product.line), "Unknown", Product.line))`, `category_pivot <- df_clean %>%`, `select(InvoiceIndex, Product.line, unit.price) %>%`, `pivot_wider()`, and `print(head(category_pivot))`.
- Console:** Shows the output of the R code, including a tibble with 6 rows and 4 columns (InvoiceIndex, Product.line, unit.price, Tax.5). The output is formatted with column types: `<int>`, `<chr>`, `<dbl>`, and `<dbl>`.
- Environment:** Lists the objects in the R environment, including `category_pivot` (1000 obs. of 7 variables), `clean_omit` (1338 obs. of 7 variables), `clean_replace` (1338 obs. of 7 variables), `combined_data` (35893 obs. of 2 variables), `data_feb` (8190 obs. of 12 variables), and `data_jan` (1000 obs. of 17 variables).
- Files Panel:** Displays a list of files in the current directory, including `SuperMarket Analysis.csv`, `TCN 1.D.pdf`, `train.csv.zip`, `Unconfirmed 187290.crdownload`, `Unnati prac 1.pdf.odt`, `Unnati prac 1.pdf.pdf`, `Untitled document (1).pdf`, `Untitled document.pdf`, `WhatsApp Unknown 2025-11-18 at 9:41:15 AM`, `WhatsApp Unknown 2025-11-18 at 9:41:15 AM.zip`, `Whols`, `Whols.zip`, `penguins_lter.csv`, and `surveys.csv`.