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

SUBJECT NAME: **Data Analysis with SAS / SPSS /R**

Practical No. 11

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

Output:

```
> print("--- 1. Original Wide Data ---")
[1] "--- 1. Original Wide Data ---"
> print(head(df))
             datetime     symbol  open  high   low close volume
1 2025-07-31 22:00:00 NYMEX:CL1! 69.35 69.46 69.22 69.44    449
2 2025-07-31 22:15:00 NYMEX:CL1! 69.43 69.43 69.34 69.35     91
3 2025-07-31 22:30:00 NYMEX:CL1! 69.35 69.38 69.33 69.38     44
4 2025-07-31 22:45:00 NYMEX:CL1! 69.38 69.43 69.38 69.42     45
5 2025-07-31 23:00:00 NYMEX:CL1! 69.40 69.43 69.38 69.39     71
6 2025-07-31 23:15:00 NYMEX:CL1! 69.38 69.39 69.35 69.38     61
> # Combine 'open', 'high', 'low', 'close' into a single 'Price' column
> # This is useful for plotting multiple price lines on one graph in ggplot2
> long_df <- df %>%
+   pivot_longer(
+     cols = c(open, high, low, close),
+     names_to = "PriceType",  # This will contain "open", "high", etc.
+     values_to = "Price"      # This will contain the numeric values
+   )
> print("--- 1. Original Wide Data ---")
[1] "--- 1. Original Wide Data ---"
> print(head(df))
             datetime     symbol  open  high   low close volume
1 2025-07-31 22:00:00 NYMEX:CL1! 69.35 69.46 69.22 69.44    449
2 2025-07-31 22:15:00 NYMEX:CL1! 69.43 69.43 69.34 69.35     91
3 2025-07-31 22:30:00 NYMEX:CL1! 69.35 69.38 69.33 69.38     44
4 2025-07-31 22:45:00 NYMEX:CL1! 69.38 69.43 69.38 69.42     45
5 2025-07-31 23:00:00 NYMEX:CL1! 69.40 69.43 69.38 69.39     71
6 2025-07-31 23:15:00 NYMEX:CL1! 69.38 69.39 69.35 69.38     61
> # Combine 'open', 'high', 'low', 'close' into a single 'Price' column
> # This is useful for plotting multiple price lines on one graph in ggplot2
> long_df <- df %>%
+   pivot_longer(
+     cols = c(open, high, low, close),
+     names_to = "PriceType",  # This will contain "open", "high", etc.
+     values_to = "Price"      # This will contain the numeric values
+   )
```

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

SUBJECT NAME: **Data Analysis with SAS / SPSS /R**

Practical No. 11

```
> print(head(long_df, 10))
# A tibble: 10 x 5
   datetime            symbol     volume PriceType Price
   <chr>               <chr>       <dbl> <chr>     <dbl>
 1 2025-07-31 22:00:00 NYMEX:CL1!    449 open       69.4
 2 2025-07-31 22:00:00 NYMEX:CL1!    449 high       69.5
 3 2025-07-31 22:00:00 NYMEX:CL1!    449 low        69.2
 4 2025-07-31 22:00:00 NYMEX:CL1!    449 close      69.4
 5 2025-07-31 22:15:00 NYMEX:CL1!     91 open       69.4
 6 2025-07-31 22:15:00 NYMEX:CL1!     91 high       69.4
 7 2025-07-31 22:15:00 NYMEX:CL1!     91 low        69.3
 8 2025-07-31 22:15:00 NYMEX:CL1!     91 close      69.4
 9 2025-07-31 22:30:00 NYMEX:CL1!     44 open       69.4
10 2025-07-31 22:30:00 NYMEX:CL1!     44 high       69.4
> # Converting back to the original OHLC format
> wide_df <- long_df %>%
+   pivot_wider(
+     names_from = PriceType,
+     values_from = Price
+   )
> print("--- 3. Wide Format (Back to Original) ---")
[1] "--- 3. Wide Format (Back to Original) ---"
> print(head(wide_df))
# A tibble: 6 x 7
  datetime            symbol     volume  open  high   low close
  <chr>               <chr>       <dbl> <dbl> <dbl> <dbl> <dbl>
1 2025-07-31 22:00:00 NYMEX:CL1!    449  69.4  69.5  69.2  69.4
2 2025-07-31 22:15:00 NYMEX:CL1!     91  69.4  69.4  69.3  69.4
3 2025-07-31 22:30:00 NYMEX:CL1!     44  69.4  69.4  69.3  69.4
4 2025-07-31 22:45:00 NYMEX:CL1!     45  69.4  69.4  69.4  69.4
5 2025-07-31 23:00:00 NYMEX:CL1!     71  69.4  69.4  69.4  69.4
6 2025-07-31 23:15:00 NYMEX:CL1!     61  69.4  69.4  69.4  69.4
> # We isolate just the closing price for comparison
> category_pivot <- df %>%
+   select(datetime, symbol, close) %>%
+   pivot_wider(
+     names_from = symbol,
+     values_from = close
+   )
```

```
+     values_from = close
+   )
> print("--- 4. Symbol Pivot (Spreading Symbols) ---")
[1] "--- 4. Symbol Pivot (Spreading Symbols) ---"
> print(head(category_pivot))
# A tibble: 6 x 2
  datetime            `NYMEX:CL1!`
  <chr>                      <dbl>
1 2025-07-31 22:00:00         69.4
2 2025-07-31 22:15:00         69.4
3 2025-07-31 22:30:00         69.4
4 2025-07-31 22:45:00         69.4
5 2025-07-31 23:00:00         69.4
6 2025-07-31 23:15:00         69.4
```