

## Aim: Reshaping data using pivot\_longer() and pivot\_wider() (R).

**Screenshot 1: Initial Data Loading and Conversion**

```

library(dplyr)
library(tidyverse)

# Read the dataset
df <- read.csv("C:/Users/mvluc/downloads/survey lung cancer.csv",
              na.strings = c("", "NA"))

# Print the original data (selected columns)
print("1) original Data (Selected columns) ~~~")
df_selected <- df %>%
  select(GENDER, AGE, SMOKING, ANXIETY, PEER_PRESSURE)

# Print the head of the selected data
print(head(df_selected))

# CONVERT WIDE TO LONG FORMAT
long_df <- df_selected %>%
  pivot_longer(
    cols = c(SMOKING, ANXIETY, PEER_PRESSURE),
    names_to = "Factor",
    values_to = "Status"
  )

# Print the long format data
print("2) Long Format Data ~~~")
print(head(long_df))
  
```

**Screenshot 2: Conversion to Wide Format and Summary Report**

```

# Print the wide format data
print("3) Wide Format Data ~~~")
print(head(wide_df))

# CREATE A SMALL REPORT
report_df <- df_selected %>%
  group_by(GENDER) %>%
  summarise(
    Avg_Age = mean(AGE, na.rm = TRUE),
    Smoking_Count = sum(SMOKING == 2, na.rm = TRUE),
    Anxiety_Count = sum(ANXIETY == 2, na.rm = TRUE),
    Peer_Pressure_Count = sum(PEER_PRESSURE == 2, na.rm = TRUE)
  )

# Print the health behavior report
print("4) Health Behavior Report ~~~")
print(head(report_df))
  
```

The screenshots show the RStudio interface with the console, source editor, and environment pane. The environment pane lists the data frames created: category, df, df\_full, df\_selected, long\_df, report\_df, survey\_lu, and wide\_df. The source editor shows the R code being executed, and the console shows the output of the commands.

YUKTA SONAWANE S120

R PROGRAMMING PRACTICAL NO: 11