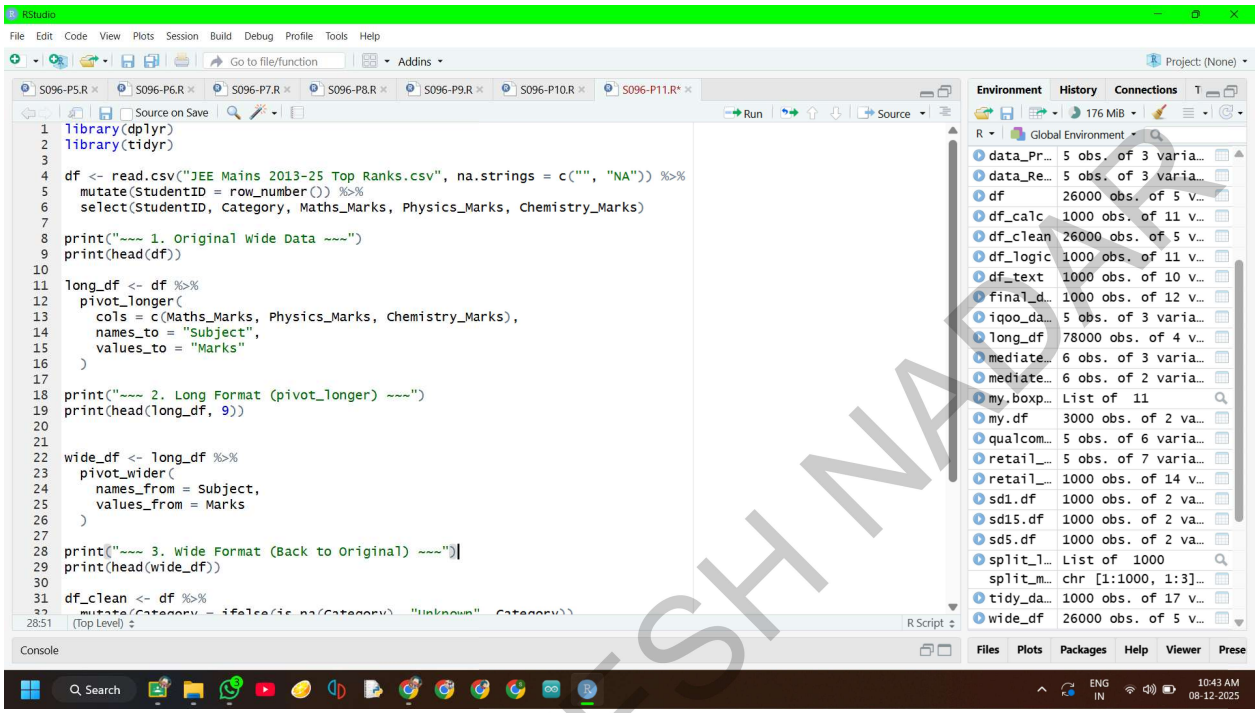
SHETH L.U.J. AND SIR M.V. COLLEGE
DATA ANALYSIS WITH SAS/SPSS/RR

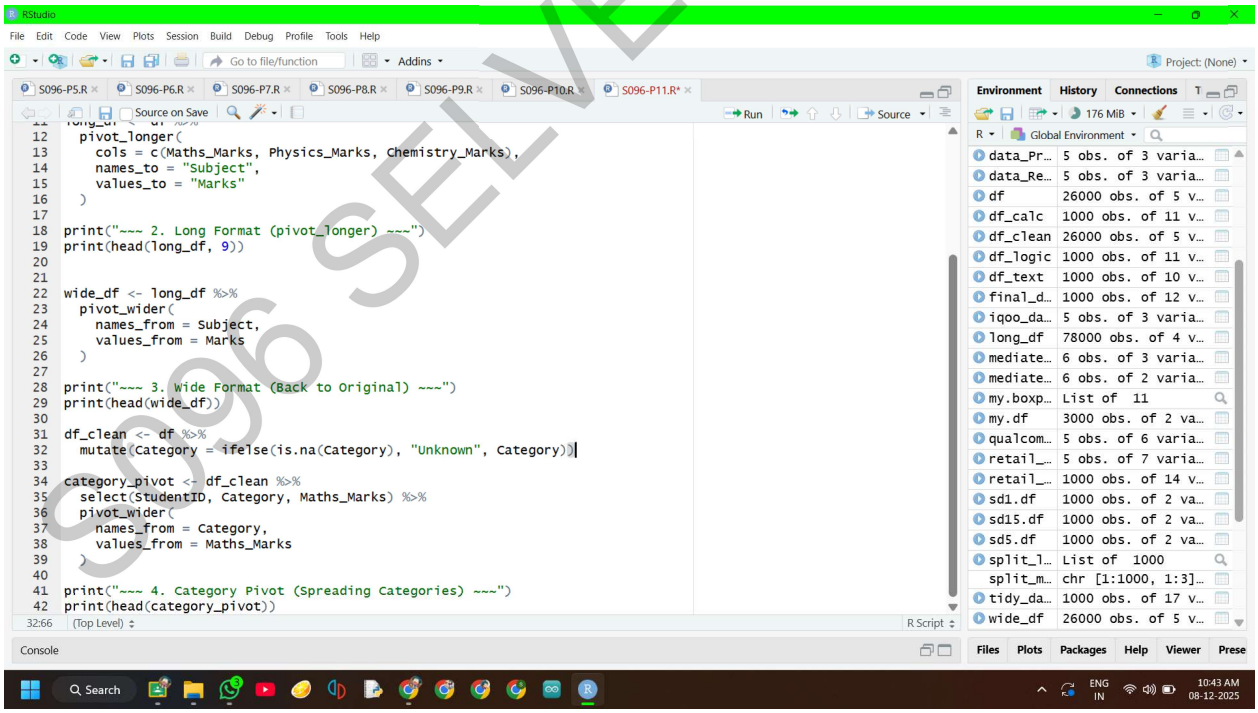
PRACTICAL NO: 11

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

CODE:

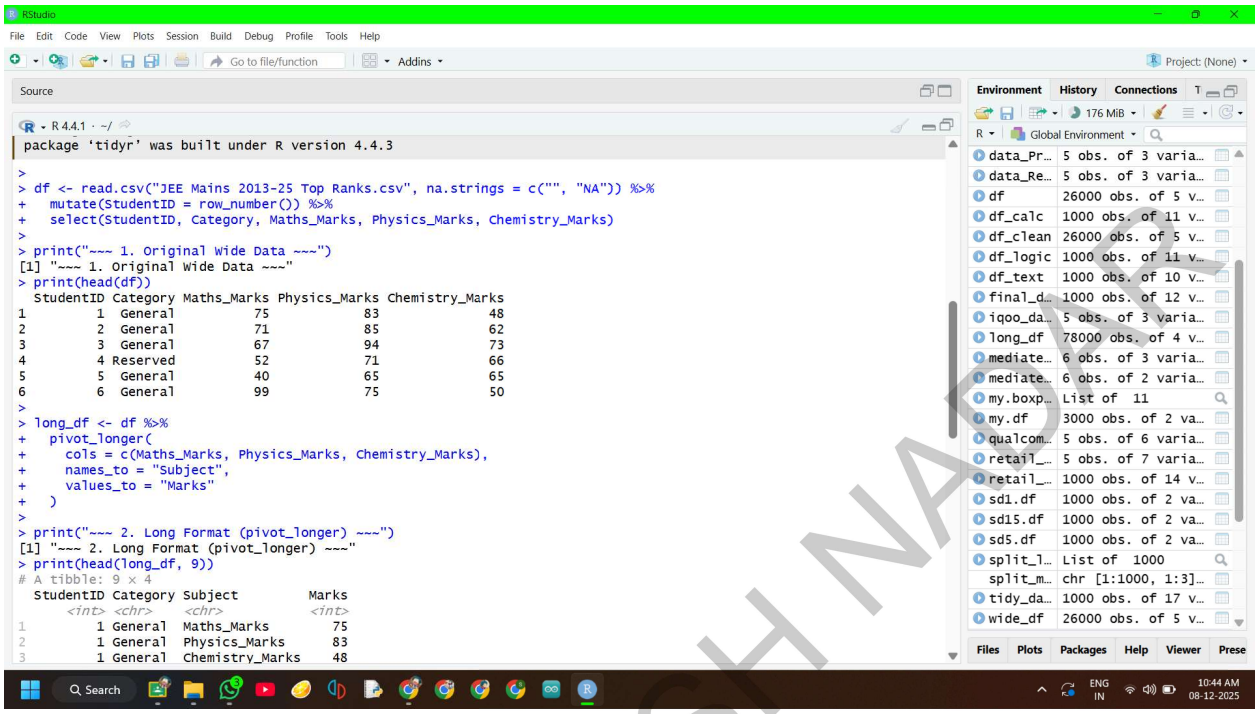


```
library(dplyr)
library(tidyr)

df <- read.csv("JEE Mains 2013-25 Top Ranks.csv", na.strings = c("", "NA")) %>%
  mutate(StudentID = row_number()) %>%
  select(StudentID, Category, Maths_Marks, Physics_Marks, Chemistry_Marks)

print("~~~ 1. Original Wide Data ~~~")
print(head(df))

long_df <- df %>%
  pivot_longer(
    cols = c(Maths_Marks, Physics_Marks, Chemistry_Marks),
    names_to = "Subject",
    values_to = "Marks"
  )

print("~~~ 2. Long Format (pivot_longer) ~~~")
print(head(long_df, 9))


wide_df <- long_df %>%
  pivot_wider(
    names_from = Subject,
    values_from = Marks
  )

print("~~~ 3. Wide Format (Back to Original) ~~~")
print(head(wide_df))

df_clean <- df %>%
```

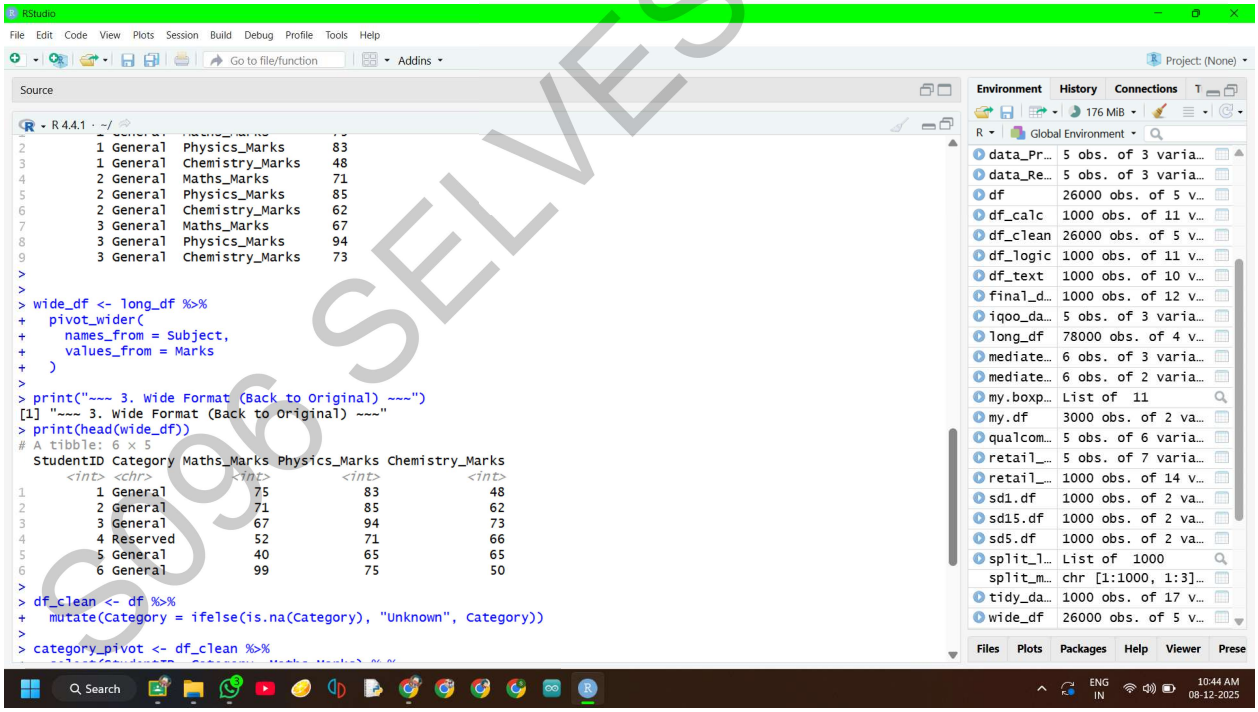


```
  pivot_longer(
    cols = c(Maths_Marks, Physics_Marks, Chemistry_Marks),
    names_to = "Subject",
    values_to = "Marks"
  )

print("~~~ 2. Long Format (pivot_longer) ~~~")
print(head(long_df, 9))


wide_df <- long_df %>%
  pivot_wider(
    names_from = Subject,
    values_from = Marks
  )

print("~~~ 3. Wide Format (Back to Original) ~~~")
print(head(wide_df))

df_clean <- df %>%
  mutate(Category = ifelse(is.na(Category), "Unknown", Category))

category_pivot <- df_clean %>%
  select(StudentID, Category, Maths_Marks) %>%
  pivot_wider(
    names_from = Category,
    values_from = Maths_Marks
  )

print("~~~ 4. Category Pivot (Spreading Categories) ~~~")
print(head(category_pivot))
```

OUTPUT:



```
package 'tidyr' was built under R version 4.4.3

>
> df <- read.csv("JEE Mains 2013-25 Top Ranks.csv", na.strings = c("", "NA")) %>%
+   mutate(StudentID = row_number()) %>%
+   select(StudentID, Category, Maths_Marks, Physics_Marks, Chemistry_Marks)
>
> print("~~~ 1. Original Wide Data ~~~")
[1] "~~~ 1. Original Wide Data ~~~"
> print(head(df))
  StudentID Category Maths_Marks Physics_Marks Chemistry_Marks
1         1  General          75            83              48
2         2  General          71            85              62
3         3  General          67            94              73
4         4 Reserved          52            71              66
5         5  General          40            65              65
6         6  General          99            75              50
>
> long_df <- df %>%
+   pivot_longer(
+     cols = c(Maths_Marks, Physics_Marks, Chemistry_Marks),
+     names_to = "Subject",
+     values_to = "Marks"
+   )
>
> print("~~~ 2. Long Format (pivot_longer) ~~~")
[1] "~~~ 2. Long Format (pivot_longer) ~~~"
> print(head(long_df, 9))
# A tibble: 9 × 4
  StudentID Category Subject       Marks
      <int> <chr>    <chr>         <int>
1         1 General  Maths_Marks      75
2         1 General  Physics_Marks    83
3         1 General  Chemistry_Marks  48
```



```
2         1 General  Physics_Marks    83
3         1 General  Chemistry_Marks  48
4         2 General  Maths_Marks      71
5         2 General  Physics_Marks    85
6         2 General  Chemistry_Marks  62
7         3 General  Maths_Marks      67
8         3 General  Physics_Marks    94
9         3 General  Chemistry_Marks  73
>
>
> wide_df <- long_df %>%
+   pivot_wider(
+     names_from = Subject,
+     values_from = Marks
+   )
>
> print("~~~ 3. Wide Format (Back to Original) ~~~")
[1] "~~~ 3. Wide Format (Back to Original) ~~~"
> print(head(wide_df))
# A tibble: 6 × 5
  StudentID Category Maths_Marks Physics_Marks Chemistry_Marks
      <int> <chr>          <int>         <int>           <int>
1         1 General           75            83              48
2         2 General           71            85              62
3         3 General           67            94              73
4         4 Reserved          52            71              66
5         5 General           40            65              65
6         6 General           99            75              50
>
> df_clean <- df %>%
+   mutate(Category = ifelse(is.na(Category), "Unknown", Category))
>
> category_pivot <- df_clean %>%
```