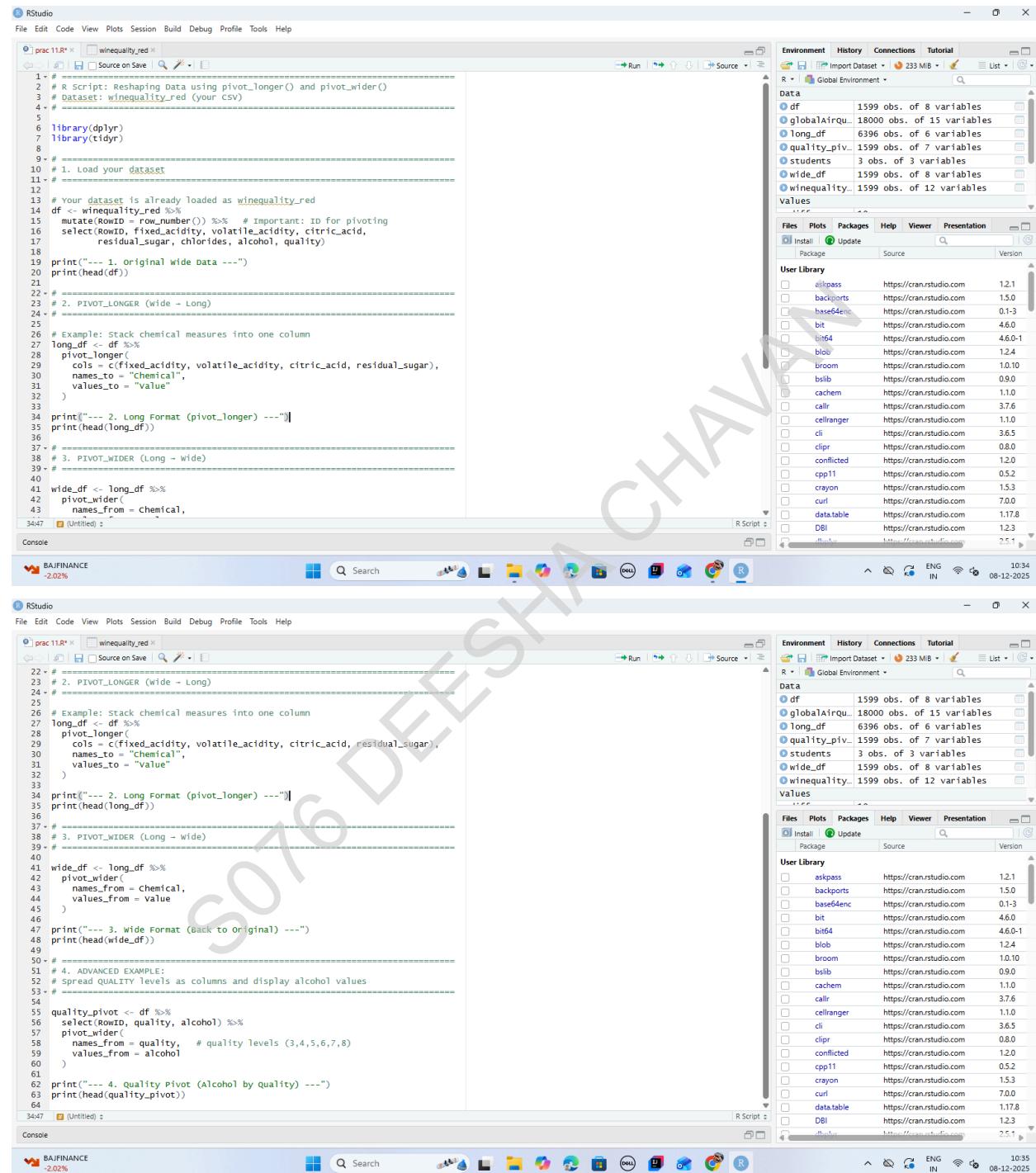


SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT: DATA ANALYSIS WITH R

PRACTICAL NO. 11

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



```

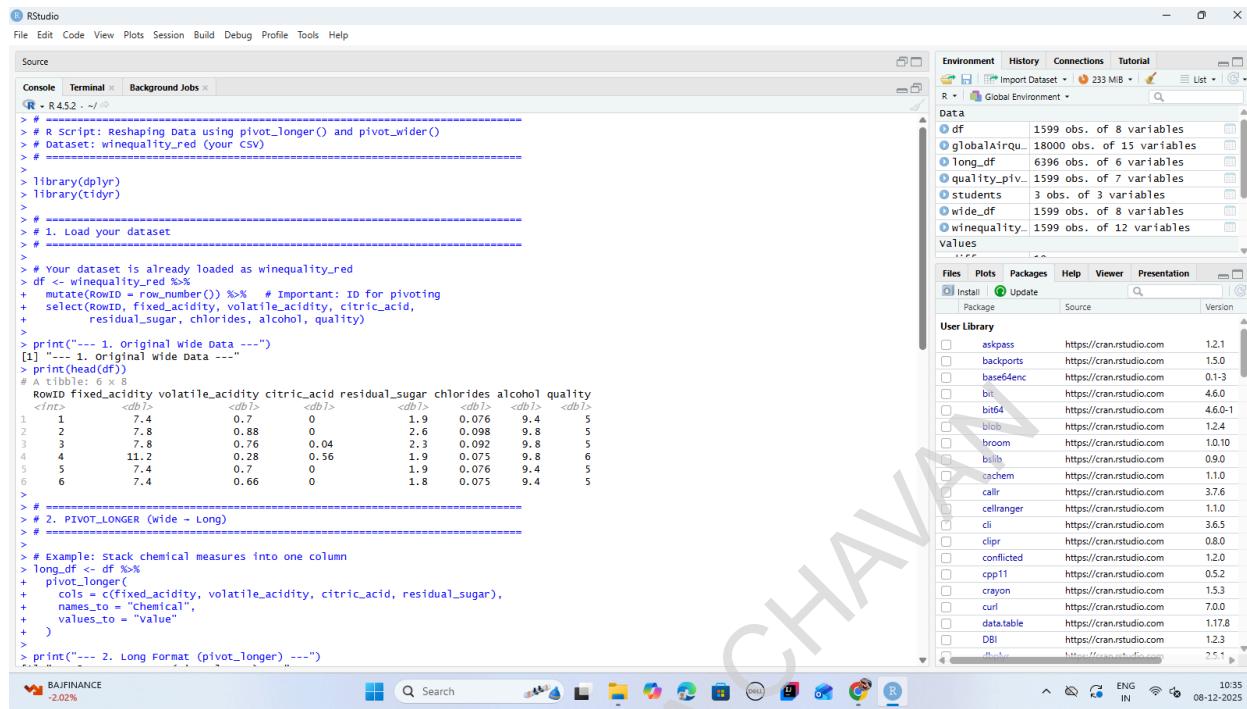
1 # R Script: Reshaping data using pivot_longer() and pivot_wider()
2 # Dataset: winequality_red (your csv)
3 #
4 #
5
6 library(dplyr)
7 library(tidyverse)
8
9 # =====
10 # 1. Load your dataset
11 #
12
13 # Your dataset is already loaded as winequality_red
14 df <- winequality_red %>%
15   mutate(RowID = row_number()) %% # Important: ID for pivoting
16   select(-RowID, fixed_acidity, volatile_acidity, citric_acid,
17         residual_sugar, chlorides, alcohol, quality)
18
19 print("--- 1. Original wide data ---")
20 print(head(df))
21
22 # =====
23 # 2. PIVOT_LONGER (wide -> Long)
24 #
25
26 # Example: Stack chemical measures into one column
27 long_df <- df %>%
28   pivot_longer(
29     cols = c(fixed_acidity, volatile_acidity, citric_acid, residual_sugar),
30     names_to = "chemical",
31     values_to = "value"
32   )
33
34 print("--- 2. Long Format (pivot_longer) ---")
35 print(head(long_df))
36
37 # =====
38 # 3. PIVOT_WIDER (Long -> wide)
39 #
40
41 wide_df <- long_df %>%
42   pivot_wider(
43     names_from = chemical,
44     values_from = value
45   )
46
47 print("--- 3. Wide Format (Back to original) ---")
48 print(head(wide_df))
49
50 # =====
51 # 4. ADVANCED EXAMPLE:
52 # Spread QUALITY Levels as columns and display alcohol values
53 #
54
55 quality_pivot <- df %>%
56   select(-RowID, quality, alcohol) %>%
57   pivot_wider(
58     names_from = quality, # quality levels (3,4,5,6,7,8)
59     values_from = alcohol
60   )
61
62 print("--- 4. quality Pivot (Alcohol by quality) ---")
63 print(head(quality_pivot))
64

```

NAME: DEESHA CHAVAN
 ROLL NO.: S076

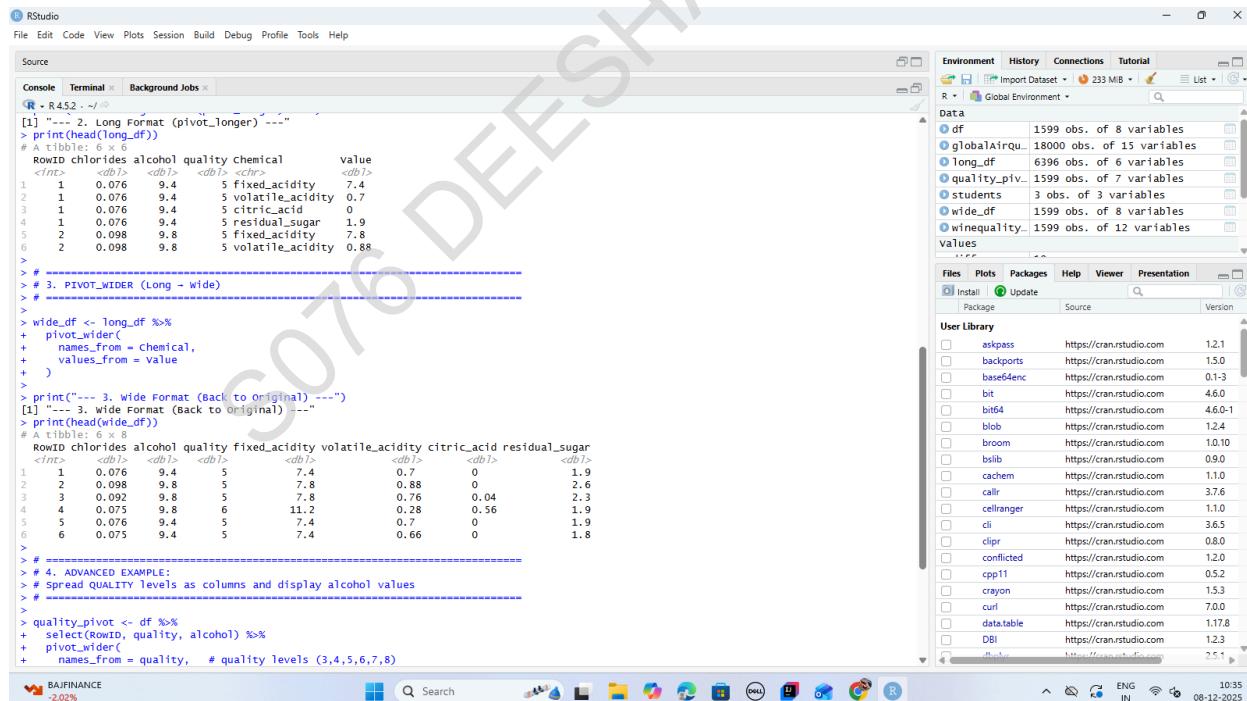
SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT: DATA ANALYSIS WITH R

Output:



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Terminal Background Jobs
[R - R4.52 - ~]
> # R Script: Reshaping Data using pivot_longer() and pivot_wider()
> # Dataset: winequality_red (your CSV)
> #
> #
> library(dplyr)
> library(tidyr)
> #
> # =====
> # 1. Load your dataset
> #
> #
> # Your dataset is already loaded as winequality_red
> df <- winequality_red %>%
+   mutate(rownid = row_number()) %>% # Important: ID for pivoting
+   select(-rowid, fixed_acidity, volatile_acidity, citric_acid,
+         residual_sugar, chlorides, alcohol, quality)
>
> print("--- 1. original wide data ---")
[1] "--- 1. original Wide data ---"
> print(head(df))
#> # tibble: 6 × 8
#> RowID fixed_acidity volatile_acidity citric_acid residual_sugar chlorides alcohol quality
#> <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 1 7.4 0.7 0 1.9 0.076 9.4 5
2 2 7.8 0.88 0 2.6 0.098 9.8 5
3 3 7.8 0.76 0.04 2.3 0.092 9.8 5
4 4 11.2 0.28 0.56 1.9 0.075 9.8 6
5 5 7.4 0.7 0 1.9 0.076 9.4 5
6 6 7.4 0.66 0 1.8 0.075 9.4 5
> #
> # =====
> # 2. PIVOT_LONGER (Wide ~ Long)
> #
> #
> # Example: stack chemical measures into one column
> long_df <- df %>%
+   pivot_longer(
+     cols = c(fixed_acidity, volatile_acidity, citric_acid, residual_sugar),
+     names_to = "chemical",
+     values_to = "value"
+   )
>
> print("--- 2. Long Format (pivot_longer) ---")
  
```



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Terminal Background Jobs
[R - R4.52 - ~]
[1] "--- 2. Long Format (pivot_longer) ---"
> print(head(long_df))
#> # tibble: 6 × 8
#> RowID chlorides alcohol quality chemical value
#> <int> <dbl> <dbl> <dbl> <chr> <dbl>
1 1 0.076 9.4 5 fixed_acidity 7.4
2 1 0.076 9.4 5 volatile_acidity 0.7
3 1 0.076 9.4 5 citric_acid 0
4 1 0.076 9.4 5 residual_sugar 1.9
5 2 0.098 9.8 5 fixed_acidity 7.8
6 2 0.098 9.8 5 volatile_acidity 0.88
>
> #
> # =====
> # 3. PIVOT_WIDER (Long ~ wide)
> #
> #
> wide_df <- long_df %>%
+   pivot_wider(
+     names_from = chemical,
+     values_from = value
+   )
>
> print("--- 3. wide Format (Back to original) ---")
[1] "--- 3. wide Format (Back to original) ---"
> print(head(wide_df))
#> # tibble: 6 × 8
#> RowID chlorides alcohol quality fixed_acidity volatile_acidity citric_acid residual_sugar
#> <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 1 0.076 9.4 5 7.4 0.7 0 1.9
2 2 0.098 9.8 5 7.8 0.88 0 2.6
3 3 0.092 9.8 5 7.8 0.76 0.04 2.3
4 4 0.075 9.8 6 11.2 0.28 0.56 1.9
5 5 0.076 9.4 5 7.4 0.7 0 1.9
6 6 0.075 9.4 5 7.4 0.66 0 1.8
> #
> # =====
> # 4. ADVANCED EXAMPLE:
> # Spread QUALITY levels as columns and display alcohol values
> #
> #
> quality_pivot <- df %>%
+   select(-rowid, quality, alcohol) %>%
+   pivot_wider(
+     names_from = quality, # quality levels (3,4,5,6,7,8)
  
```

NAME: DEESHA CHAVAN
ROLL NO.: S076

**SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT: DATA ANALYSIS WITH R**

The screenshot shows the RStudio interface with the following components:

- Top Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Source Tab:** Contains R code demonstrating various dplyr functions like pivot_wider, pivot_longer, and quality_pivot.
- Environment Tab:** Shows the global environment with objects like `df` (1599 obs. of 8 variables), `globalairqu...` (18000 obs. of 15 variables), and `long_df` (6396 obs. of 6 variables).
- Plots Tab:** Displays a histogram of a variable.
- Packages Tab:** Shows installed packages: `askpass` (1.2.1), `backports` (1.5.0), `base64enc` (0.1-3), and `bit` (4.6.0).
- User Library Tab:** Shows available packages from CRAN.
- Viewer Tab:** Displays the results of the R code execution, including tables and plots.

Code Snippet (from Source tab):

```
R> R.45.2 <- long_df %>%  
> pivot_wider(  
+   names_from = chemical,  
+   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 8  
#>   RowID chlorides alcohol quality fixed_acidity volatile_acidity citric_acid residual_sugar  
#>   <dbl>     <dbl>    <dbl>    <dbl>      <dbl>        <dbl>       <dbl>        <dbl>  
#> 1     1     0.076     9.4      5     7.4      0.7       0.0        1.9  
#> 2     2     0.098     9.8      5     7.8      0.88      0.0        2.6  
#> 3     3     0.092     9.8      5     7.8      0.76      0.04       2.3  
#> 4     4     0.075     9.8      6    11.2      0.28      0.56       1.9  
#> 5     5     0.076     9.4      5     7.4      0.7       0.0        1.9  
#> 6     6     0.075     9.4      5     7.4      0.66      0.0        1.8  
>  
> #-----  
> # 4. ADVANCED EXAMPLE:  
> # Spread QUALITY levels as columns and display alcohol values  
> #-----  
>  
> quality_pivot <- df %>%  
+   select(RowID, quality, alcohol) %>%  
+   pivot_wider(  
+     names_from = quality,  # quality Levels (3,4,5,6,7,8)  
+     values_from = alcohol  
)  
>  
> print("---- 4. quality Pivot (Alcohol by quality) ----")  
[1] "---- 4. quality Pivot (Alcohol by quality) ----"  
> print(head(quality_pivot))  
# A tibble: 6 x 7  
#>   RowID `3`   `6`   `7`   `4`   `8`   `3`  
#>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
#> 1     9.4  NA    NA    NA    NA    NA  
#> 2     9.8  NA    NA    NA    NA    NA  
#> 3     9.8  NA    NA    NA    NA    NA  
#> 4     NA    9.8  NA    NA    NA    NA  
#> 5     9.4  NA    NA    NA    NA    NA  
#> 6     9.4  NA    NA    NA    NA    NA  
>
```

PRACTICAL NO. 12

AIM: Combining datasets vertically (concatenation) using rbind() (R).

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plot, Session, Build, Debug, Profile, Tools, Help.
- Project Bar:** prac_11.R, PRAC_12.R, winequality.rds.
- Code Editor:** A script named `winequality.r` containing R code for vertical concatenation using `rbind()`. The code includes comments for creating a second dataset, ensuring column names match, and performing the vertical concatenation.
- Run Tab:** Run, Source.
- Environment View:** Shows the global environment with various objects and their characteristics.
- Files Tab:** Files, Plots, Packages, Help, Viewer, Presentation.
- User Library:** A list of packages and their URLs.
- Console:** Shows the command `R Script`.
- Bottom Bar:** Includes icons for file operations, search, and system status.

NAME: DEESHA CHAVAN
ROLL NO.: S076

**SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT: DATA ANALYSIS WITH R**

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** PRAC 12.R
- Code Editor:** The script contains R code for concatenating datasets. It starts by loading the dplyr package, then reads a wine dataset from a CSV file. It prints the head of the dataset and then creates a second dataset by simulating another batch of wine data. The code then demonstrates concatenating two datasets (Batch A and Batch B) vertically. Finally, it prints the combined dataset and lists the total number of rows.
- Environment Tab:** Shows the global environment with various objects and their sizes and variable types.
- Files Tab:** Shows the current project files.
- Plots Tab:** No plots are currently displayed.
- Packages Tab:** Shows the installed packages and their versions.
- Help Tab:** Not visible.
- Viewer Tab:** Not visible.
- Presentation Tab:** Not visible.
- Console Tab:** Shows the command history.
- Bottom Bar:** Includes icons for search, file operations, and system status (25°C, ENG IN, battery level).

```
library(dplyr)
# 1. Load Your wine Dataset
wine <- read.csv("D:/5076/winequality_cleaned.csv")
cat("----- Original Wine Dataset -----")
print(head(wine))
# 2. CREATE A SECOND DATASET
# (Simulating another batch of wine data)
# For demonstration, take First 5 rows as Batch A,
# last 5 rows as Batch B
batch_A <- wine[1:5, ]
batch_B <- wine[(nrow(wine) - 4):nrow(wine), ]
# 3. ENSURE COLUMN NAMES MATCH
colnames(batch_A)
colnames(batch_B)
# Both already match because both come from the same dataset
# 4. VERTICAL CONCATENATION
combined_wine <- rbind(batch_A, batch_B)
#----- Combined Dataset (A+B) -----
print(combined_wine)
cat("Total rows:", nrow(combined_wine))
list.files("D:/5076")

```

OUTPUT:

NAME: DEESHA CHAVAN
ROLL NO.: S076

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

SUBJECT: DATA ANALYSIS WITH R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

prac_11.R | PRAC 12.R | winequality_red.r

6 11library(dplyr)

615 (Untitled) :

Console Terminal Background Jobs

```
R + R452: ~/Desktop/winequality_red.r
```

```
> #----- Batch A -----\\n"
> cat("----- Batch A -----\\n")
----- Batch A -----
> print(batch_A)
fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1 7.4 0.70 0.00 1.9 0.076 11 34 0.9978 3.51 0.56
2 7.8 0.88 0.00 2.6 0.098 25 67 0.9968 3.20 0.68
3 7.8 0.76 0.04 2.3 0.092 15 54 0.9970 3.26 0.65
4 11.2 0.28 0.56 1.9 0.075 17 60 0.9980 3.16 0.58
5 7.4 0.66 0.00 1.8 0.075 13 40 0.9978 3.51 0.56
alcohol quality
1 9.4 5
2 9.8 5
3 9.8 5
4 9.8 6
5 9.4 5
> cat("----- Batch B -----\\n")
----- Batch B -----
> print(batch_B)
fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1355 6.8 0.620 0.08 1.9 0.068 28 38 0.99651 3.42 0.82
1356 6.2 0.600 0.08 2.0 0.090 32 44 0.99490 3.45 0.58
1357 5.9 0.550 0.10 2.2 0.062 39 51 0.99512 3.52 0.76
1358 5.9 0.645 0.12 2.0 0.075 32 44 0.99547 3.57 0.71
1359 6.0 0.310 0.47 3.6 0.067 18 42 0.99549 3.39 0.66
alcohol quality
1355 9.5 6
1356 10.5 5
1357 11.2 6
1358 10.2 5
1359 11.0 6
> # ----- 3. ENSURE COLUMN NAMES MATCH
> # -----
> colnames(batch_A)
[1] "fixed.acidity" "volatile.acidity" "citric.acid" "residual.sugar" "chlorides" "density" "pH" "sulphates"
[11] "alcohol" "quality"
```

25°C Sunny Search

Environment History Connections Tutorial

R + Global Environment

- long_df 6396 obs. of 6 variables
- quality_piv_ 1599 obs. of 7 variables
- students 3 obs. of 3 variables
- wide_df 1599 obs. of 8 variables
- wine 1359 obs. of 12 variables
- winequality_ 1599 obs. of 12 variables

values

difference 10

iris <Promise>

Files Plots Packages Help Viewer Presentation

Install Update

User Library

- askpass https://cran.rstudio.com 1.2.1
- backports https://cran.rstudio.com 1.5.0
- base64enc https://cran.rstudio.com 0.1-3
- bit https://cran.rstudio.com 4.6.0
- bit64 https://cran.rstudio.com 4.6.0-1
- blob https://cran.rstudio.com 1.2.4
- broom https://cran.rstudio.com 0.10.10
- bslib https://cran.rstudio.com 0.9.0
- cachet https://cran.rstudio.com 1.1.0
- callr https://cran.rstudio.com 3.7.6
- cellranger https://cran.rstudio.com 1.1.0
- cli https://cran.rstudio.com 3.6.5
- clipr https://cran.rstudio.com 0.8.0
- conflicted https://cran.rstudio.com
- cpp11 https://cran.rstudio.com 0.5.2
- crayon https://cran.rstudio.com 1.5.3
- curl https://cran.rstudio.com 7.0.0
- data.table https://cran.rstudio.com 1.17.8
- DBI https://cran.rstudio.com 1.2.3
- dplyr https://cran.rstudio.com 2.5.1

Snipping Tool

Screenshot copied to clipboard
Automatically saved to screenshots folder.

Mark-up and share

10:51 08-12-2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

prac_11.R | PRAC 12.R | winequality_red.r

6 11library(dplyr)

615 (Untitled) :

Console Terminal Background Jobs

```
R + R452: ~/
```

```
1359 11.0 6
> # -----
> # 3. ENSURE COLUMN NAMES MATCH
> # -----
> colnames(batch_A)
[1] "fixed.acidity" "volatile.acidity" "citric.acid" "residual.sugar" "chlorides" "density" "pH" "sulphates"
[6] "free.sulfur.dioxide" "total.sulfur.dioxide" "density"
[11] "alcohol" "quality"
> colnames(batch_B)
[1] "fixed.acidity" "volatile.acidity" "citric.acid" "residual.sugar" "chlorides" "density" "pH" "sulphates"
[6] "free.sulfur.dioxide" "total.sulfur.dioxide" "density"
[11] "alcohol" "quality"
> # Both already match because both come from the same dataset
>
> # -----
> # 4. VERTICAL CONCATENATION
> # -----
> combined_wine <- rbind(batch_A, batch_B)
> cat("----- Combined dataset (A+B) -----\\n")
----- Combined dataset (A+B) -----
> print(combined_wine)
fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1 7.4 0.700 0.00 1.9 0.076 11 34 0.99780 3.51 0.56
2 7.8 0.880 0.00 2.6 0.098 25 67 0.99680 3.20 0.68
3 7.8 0.760 0.04 2.3 0.092 15 54 0.99700 3.26 0.65
4 11.2 0.280 0.56 1.9 0.075 17 60 0.99800 3.16 0.58
5 7.4 0.660 0.00 1.8 0.075 13 40 0.99780 3.51 0.56
1355 6.8 0.620 0.08 1.9 0.068 28 38 0.99651 3.42 0.82
1356 6.2 0.600 0.08 2.0 0.090 32 44 0.99490 3.45 0.58
1357 5.9 0.550 0.10 2.2 0.062 39 51 0.99512 3.52 0.76
1358 5.9 0.645 0.12 2.0 0.075 32 44 0.99547 3.57 0.71
1359 6.0 0.310 0.47 3.6 0.067 18 42 0.99549 3.39 0.66
alcohol quality
1 9.4 5
2 9.8 5
3 9.8 5
4 9.8 6
5 9.4 5
> 25°C Sunny Search
```

Environment History Connections Tutorial

R + Global Environment

- long_df 6396 obs. of 6 variables
- quality_piv_ 1599 obs. of 7 variables
- students 3 obs. of 3 variables
- wide_df 1599 obs. of 8 variables
- wine 1359 obs. of 12 variables
- winequality_ 1599 obs. of 12 variables

values

difference 10

iris <Promise>

Files Plots Packages Help Viewer Presentation

Install Update

User Library

- askpass https://cran.rstudio.com 1.2.1
- backports https://cran.rstudio.com 1.5.0
- base64enc https://cran.rstudio.com 0.1-3
- bit https://cran.rstudio.com 4.6.0
- bit64 https://cran.rstudio.com 4.6.0-1
- blob https://cran.rstudio.com 1.2.4
- broom https://cran.rstudio.com 0.10.10
- bslib https://cran.rstudio.com 0.9.0
- cachet https://cran.rstudio.com 1.1.0
- callr https://cran.rstudio.com 3.7.6
- cellranger https://cran.rstudio.com 1.1.0
- cli https://cran.rstudio.com 3.6.5
- clipr https://cran.rstudio.com 0.8.0
- conflicted https://cran.rstudio.com
- cpp11 https://cran.rstudio.com 0.5.2
- crayon https://cran.rstudio.com 1.5.3
- curl https://cran.rstudio.com 7.0.0
- data.table https://cran.rstudio.com 1.17.8
- DBI https://cran.rstudio.com 1.2.3
- dplyr https://cran.rstudio.com 2.5.1

10:51 08-12-2025

NAME: DEESHA CHAVAN
ROLL NO.: S076

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

SUBJECT: DATA ANALYSIS WITH R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

PRAC 12.R winequality_red.r

```

6 library(dplyr)
6:15 (Untitled) : R Script
Console Terminal Background Jobs
R R 4.5.2 : ~
1357 11.2 6
1358 10.2 5
1359 11.0 6
> cat("\nTotal rows:", nrow(combined_wine))
Total rows: 10> list.files("D:/S076")
[1] "cleaned_student_mental_health.csv"
[2] "cleaned_student_mental_health_final.csv"
[3] "D:/S076/3.pkt"
[4] "desktop"
[5] "global air pollution dataset.csv.zip"
[6] "global_air_cleaned.csv"
[7] "Mental_Health_and_Social_Media_Balance_Dataset.csv"
[8] "modified_winequality_red.csv"
[9] "myfirst.js"
[10] "PRAC 10.R"
[11] "PRAC 11.R"
[12] "PRAC 12.R"
[13] "PRAC 3.R"
[14] "prac 4.py"
[15] "PRAC 6.ipynb"
[16] "PRAC 6.R"
[17] "PRAC 7.ipynb"
[18] "PRAC 7.R"
[19] "PRAC 8.ipynb"
[20] "PRAC 8.R"
[21] "prac1.R"
[22] "prac2.py"
[23] "Processed_student_Mental_Health - Processed_Student_Mental_Health.csv"
[24] "S076 PRAC 6.ipynb"
[25] "sales_data - sales_data.csv"
[26] "space_mission1.csv"
[27] "student Mental health - student Mental health (1).csv"
[28] "tip.csv"
[29] "tip_modified.csv"
[30] "winequality-red.csv"
[31] "winequality-cleaned.csv"
[32] "youtube-top-100-songs-2025.csv"
>

```

25°C Sunny

Search

Environment History Connections Tutorial

R Tong_df 6396 obs. of 6 variables

Global Environment

iris 1599 obs. of 12 variables

students 3 obs. of 3 variables

quality_piv_ 1599 obs. of 7 variables

wide_df 1599 obs. of 8 variables

wine 1599 obs. of 12 variables

winequality_ 1599 obs. of 12 variables

values difference 10

User Library

- askpass https://cran.rstudio.com 1.2.1
- backports https://cran.rstudio.com 1.5.0
- base64enc https://cran.rstudio.com 0.1-3
- bit https://cran.rstudio.com 4.6.0
- bit64 https://cran.rstudio.com 4.6.0-1
- blob https://cran.rstudio.com 1.2.4
- broom https://cran.rstudio.com 1.0.10
- bslib https://cran.rstudio.com 0.9.0
- cachem https://cran.rstudio.com 1.1.0
- callr https://cran.rstudio.com 3.7.6
- cellranger https://cran.rstudio.com 1.1.0
- cli https://cran.rstudio.com 3.6.5
- clipr https://cran.rstudio.com 0.8.0
- conflicted https://cran.rstudio.com 1.2.0
- cpp11 https://cran.rstudio.com 0.5.2
- crayon https://cran.rstudio.com 1.5.3
- curl https://cran.rstudio.com 7.0.0
- data.table https://cran.rstudio.com 1.17.8
- DBI https://cran.rstudio.com 1.2.3
- rlang https://cran.rstudio.com 2.5.1

Install Update

Files Plots Packages Help Viewer Presentation

10:52 08-12-2025

PRACTICAL NO. 13

AIM: Identifying and handling duplicates using distinct() (R studio).

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

PRAC 12.R PRAC 13.R winequality_red.r

```

1: # R Script: Identifying and Handling duplicates
2: # dataset: winequality_cleaned.csv (your dataset)
3: #
4: #
5: library(dplyr)
6: #
7: #
8: # -----
9: # 1. LOAD DATA
10: #
11: # use file.choose() to avoid path errors
12: wine <- read.csv(file.choose())
13: 
14: cat("--- 1. Original Dataset ---\n")
15: print(head(wine))
16: 
17: #
18: # 2. IDENTIFY DUPLICATES
19: #
20: # A duplicate row = ALL columns are exactly the same
21: 
22: duplicates_report <- wine %>%
23:   group_by(across(everything())) %>%
24:   count() %>%
25:   filter(n > 1)
26: 
27: cat("--- 2. Duplicate Rows Found ---\n")
28: print(duplicates_report)
29: 
30: #
31: # 3. REMOVE EXACT DUPLICATES
32: #
33: # distinct() removes rows that are identical across ALL columns
34: 
35: clean_wine <- wine %>%
36:   distinct()
37: 
38: cat("--- 3. Dataset After Removing Exact Duplicates ---\n")
39: print(clean_wine)
40: 
41: cat("\nRows before:", nrow(wine))
42: cat("\nRows after removing duplicates:", nrow(clean_wine), "\n")
43: 
44:34 (Untitled) : R Script
Console

```

25°C Sunny

Search

Environment History Connections Tutorial

R Tong_df 6396 obs. of 6 variables

Global Environment

iris 1599 obs. of 12 variables

students 3 obs. of 3 variables

unique_alco_ 65 obs. of 12 variables

wide_df 1599 obs. of 8 variables

wine 1599 obs. of 12 variables

winequality_ 1599 obs. of 12 variables

values difference 10

User Library

- askpass https://cran.rstudio.com 1.2.1
- backports https://cran.rstudio.com 1.5.0
- base64enc https://cran.rstudio.com 0.1-3
- bit https://cran.rstudio.com 4.6.0
- bit64 https://cran.rstudio.com 4.6.0-1
- blob https://cran.rstudio.com 1.2.4
- broom https://cran.rstudio.com 1.0.10
- bslib https://cran.rstudio.com 0.9.0
- cachem https://cran.rstudio.com 1.1.0
- callr https://cran.rstudio.com 3.7.6
- cellranger https://cran.rstudio.com 1.1.0
- cli https://cran.rstudio.com 3.6.5
- clipr https://cran.rstudio.com 0.8.0
- conflicted https://cran.rstudio.com 1.2.0
- cpp11 https://cran.rstudio.com 0.5.2
- crayon https://cran.rstudio.com 1.5.3
- curl https://cran.rstudio.com 7.0.0
- data.table https://cran.rstudio.com 1.17.8
- DBI https://cran.rstudio.com 1.2.3
- rlang https://cran.rstudio.com 2.5.1

Install Update

Files Plots Packages Help Viewer Presentation

10:56 08-12-2025

NAME: DEESHA CHAVAN
ROLL NO.: S076

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

SUBJECT: DATA ANALYSIS WITH R

The screenshot shows the RStudio interface with the following details:

- File Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Source Editor:** A script named "prac 11.R" containing R code for dataset cleaning. The code includes sections for identifying and handling duplicates, removing exact duplicates, and removing duplicates based on specific columns (alcohol). It also prints the head of the cleaned dataset.
- Environment Tab:** Shows objects in the global environment: rong_wine, quality_piv., students, unique_alco., wide_df, wine, and winequality_. All objects are 1599 obs. of 12 variables.
- Files Tab:** Shows installed packages: askpass, backports, base64enc, bit, bit64, blob, broom, bslib, cachem, callr, cellranger, cli, clipr, conflicted, cpp11, crayon, curl, data.table, DBI, and rlang. The DBI package is highlighted.
- Console Tab:** Displays the output of the R code run in the console, including the printed head of the cleaned dataset.
- System Status:** Shows the system is at 25°C and sunny.
- Bottom Bar:** Shows the date (08-12-2025) and time (10:56).

OUTPUT:

The screenshot shows the RStudio interface with the following details:

- File Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Source Editor:** The same "prac 11.R" script as above, showing the full R code and its output.
- Environment Tab:** Shows objects in the global environment: rong_wine, quality_piv., students, unique_alco., wide_df, wine, and winequality_. All objects are 1599 obs. of 12 variables.
- Files Tab:** Shows installed packages: askpass, backports, base64enc, bit, bit64, blob, broom, bslib, cachem, callr, cellranger, cli, clipr, conflicted, cpp11, crayon, curl, data.table, DBI, and rlang. The DBI package is highlighted.
- Console Tab:** Displays the output of the R code run in the console, including the printed head of the cleaned dataset.
- System Status:** Shows the system is at 25°C and sunny.
- Bottom Bar:** Shows the date (08-12-2025) and time (10:57).

NAME: DEESHA CHAVAN
ROLL NO.: S076

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

SUBJECT: DATA ANALYSIS WITH R

R - 4.5.2 - ~

```

> # group_by(everything()) %>% # group by all columns
> count() %% # count occurrences
> filter(n > 1) # keep rows that appear more than once
> cat("---- 2. Duplicate Rows Found ---\n")
--- 2. Duplicate Rows Found ---
> print(duplicates_report)
# A tibble: 220 × 13
# Groups:   fixed.acidity, volatile.acidity, citric.acid, residual.sugar, chlorides, free.sulfur.dioxide, total.sulfur.dioxide, density, pH, sulphates, alcohol, quality [20]
  fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
    <dbl>            <dbl>        <dbl>       <dbl>      <dbl>           <dbl>          <dbl>      <dbl>
1     5.2             0.44        0.09      0.05       0.27          0.92          3.68      0.90
2     5.6             0.5         0.09      0.049      0.17          0.94          3.63      0.63
3     5.6             0.54        0.04      0.049      0.5           0.94          3.72      0.58
4     5.6             0.66        0.0       0.087      0.3           0.94          3.71      0.63
5     5.9             0.61        0.08      0.071      0.16          0.94          3.56      0.77
6     6                0.5         0       0.057      0.15          0.94          3.36      0.45
7     6                0.51        0       0.064      0.4           0.95          3.54      0.93
8     6.1             0.32        0.25      0.071      0.23          0.96          3.42      0.97
9     6.2             0.36        0.24      0.099      0.19          0.97          3.57      0.57
10    6.2             0.56        0.09      0.053      0.24          0.94          3.34      0.6
# i 210 more rows
# i 3 more variables: alcohol <dbl>, quality <int>, n <int>
# i use 'print(n = ...)' to see more rows
>
> # ----- 3. REMOVE EXACT DUPLICATES -----
> # ----- distinct() removes rows that are identical across ALL columns
>
> clean_wine <- wine %>%
>   distinct()
>
> cat("---- 3. Dataset After Removing Exact Duplicates ---\n")
--- 3. dataset After Removing Exact Duplicates ---
> print(clean_wine)
fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1     7.0             0.700        0.00      1.90      0.076      11           0.9978      3.51      0.56
2     7.8             0.880        0.00      2.60      0.098      25           0.9968      3.20      0.68
3     7.8             0.760        0.04      2.30      0.092      15           0.9970      3.26      0.65
4    11.2             0.280        0.56      1.90      0.075      17           0.9980      3.16      0.58
5     7.4             0.660        0.00      1.80      0.075      13           0.9978      3.51      0.56
6     7.9             0.600        0.06      1.60      0.069      15           0.9964      3.30      0.46
7     7.3             0.650        0.00      1.20      0.065      15           0.9946      3.39      0.47

```

R - 4.5.2 - ~

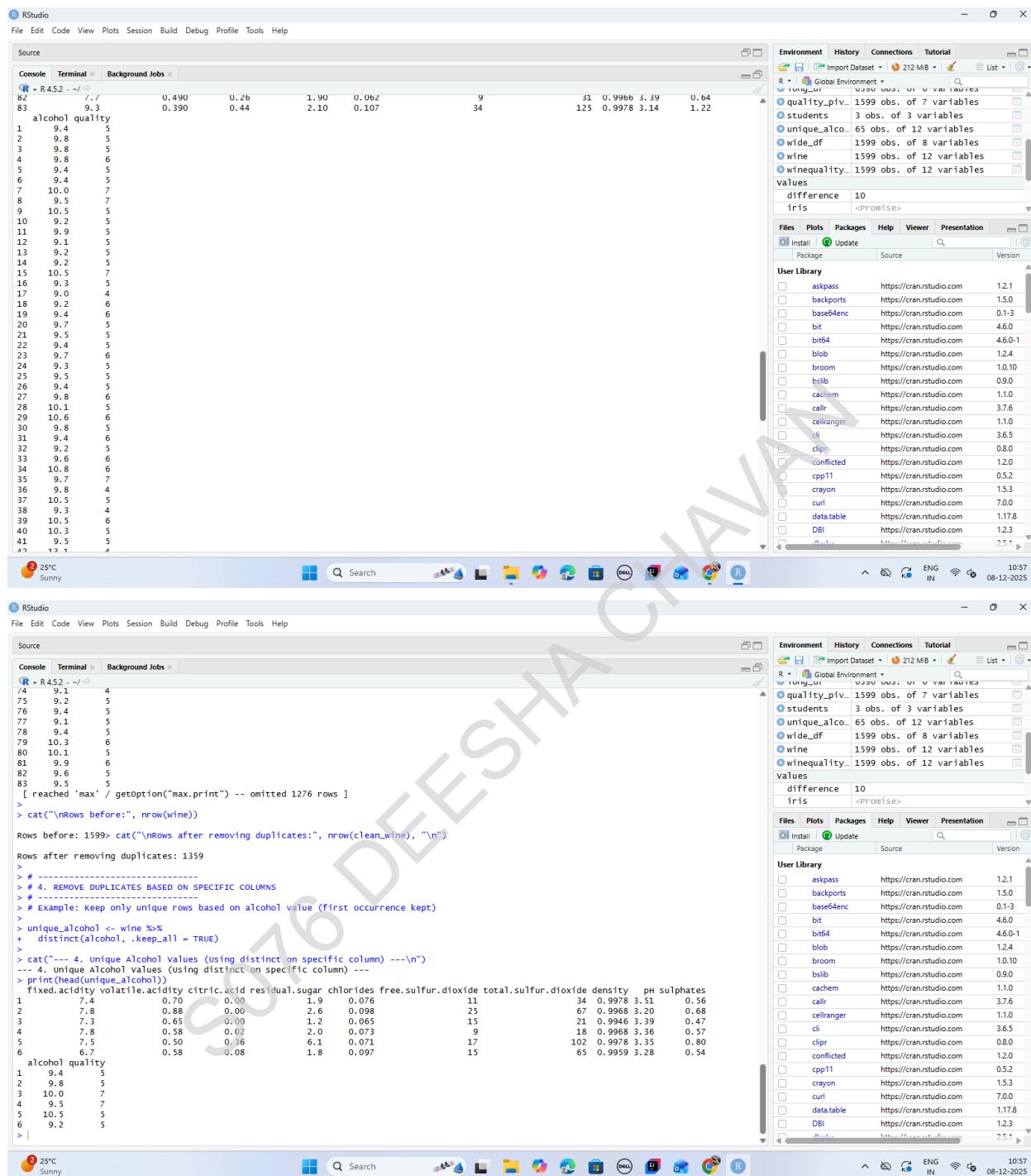
```

39    7.5             0.490        0.20      2.60      0.332      8           0.9979      3.24      0.90
40    8.1             0.600        0.22      2.80      0.069      9           0.9968      3.20      0.60
41    6.8             0.670        0.02      1.80      0.050      5           0.9962      3.48      0.52
42    4.6             0.520        0.15      2.10      0.054      8           0.9934      3.90      0.56
43    7.7             0.935        0.43      2.20      0.114      22          0.9970      3.25      0.73
44    8.7             0.290        0.52      1.60      0.113      12          0.9969      3.25      0.58
45    6.4             0.400        0.23      1.60      0.066      5           0.9958      3.34      0.56
46    5.6             0.310        0.37      1.40      0.074      12          0.9954      3.32      0.58
47    8.8             0.660        0.26      1.70      0.074      4           0.9971      3.15      0.74
48    6.6             0.520        0.04      2.20      0.069      8           0.9971      3.40      0.63
49    6.0             0.500        0.04      2.00      0.068      6           0.9955      3.39      0.64
50    8.6             0.380        0.36      3.00      0.081      30          0.9970      3.20      0.56
51    7.6             0.510        0.15      2.80      0.110      33          0.9955      3.17      0.63
52    7.7             0.620        0.04      3.80      0.084      25          0.9978      3.34      0.53
53   10.2             0.420        0.57      3.40      0.070      4           0.9971      3.04      0.63
54    7.5             0.630        0.12      9.10      0.111      50          0.9983      3.26      0.77
55    7.8             0.590        0.18      2.30      0.076      17          0.9975      3.43      0.59
56    7.3             0.390        0.31      2.40      0.074      9           0.9962      3.41      0.54
57    8.8             0.400        0.40      2.20      0.079      19          0.9962      3.44      0.64
58    7.7             0.690        0.49      1.90      0.155      20          0.9968      3.21      0.71
59    7.5             0.520        0.16      1.90      0.085      12          0.9968      3.38      0.62
60    7.0             0.735        0.05      2.00      0.081      13          0.9966      3.39      0.57
61    7.2             0.725        0.05      4.65      0.086      4           0.9962      3.41      0.39
62    7.5             0.520        0.11      1.50      0.079      11          0.9968      3.42      0.58
63    6.6             0.705        0.07      1.60      0.076      6           0.9962      3.44      0.58
64    9.3             0.320        0.57      2.00      0.074      27          0.9969      3.28      0.79
65    8.0             0.705        0.05      1.90      0.074      8           0.9962      3.34      0.95
66    7.7             0.650        0.08      1.90      0.076      15          0.9962      3.32      0.54
67    7.7             0.670        0.23      2.10      0.088      17          0.9962      3.32      0.54
68    7.7             0.690        0.22      1.90      0.084      18          0.9961      3.31      0.48
69    8.3             0.675        0.26      2.10      0.084      11          0.9976      3.31      0.53
70    9.7             0.320        0.54      2.50      0.094      28          0.9984      3.28      0.82
71    8.8             0.410        0.64      2.20      0.093      9           0.9986      3.54      0.66
72    6.8             0.785        0.00      2.40      0.104      14          0.9966      3.52      0.55
73    6.7             0.750        0.12      2.00      0.086      12          0.9958      3.38      0.52
74    8.3             0.625        0.20      1.50      0.080      27          0.9972      3.16      1.12
75    6.2             0.450        0.09      1.90      0.069      9           0.9962      3.41      0.56
76    7.9             0.430        0.70      1.90      0.164      22          0.9974      3.13      0.58
77    7.4             0.500        0.47      2.00      0.086      21          0.9970      3.36      0.57
78    7.3             0.670        0.26      1.80      0.401      16          0.9969      3.16      1.14
79    6.3             0.300        0.48      1.80      0.069      18          0.9959      3.44      0.78
80    6.9             0.550        0.15      2.20      0.076      19          0.9961      3.41      0.59
81    8.6             0.490        0.28      1.90      0.110      20          0.9972      2.93      1.95
82    7.7             0.490        0.26      1.90      0.062      9           0.9966      3.39      0.64

```

NAME: DEESHA CHAVAN
ROLL NO.: S076

SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT: DATA ANALYSIS WITH R



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
R - R 4.5.2 - ~
82   7.7    0.490   0.2b    1.90   0.062      9      31   0.996b 3.39   0.64
83   9.3    0.390   0.44    2.10   0.107     34     125   0.9978 3.14   1.22
alcohol quality
1   9.4    5
2   9.8    5
3   9.8    5
4   9.8    6
5   9.4    5
6   9.4    5
7   10.0   7
8   9.5    7
9   10.5   5
10  9.2   5
11  9.0   5
12  9.1   5
13  9.2   5
14  9.2   5
15  10.5   7
16  9.3   5
17  9.0   4
18  9.2   6
19  9.4   6
20  9.7   5
21  9.5   5
22  9.4   5
23  9.7   6
24  9.3   5
25  9.5   5
26  9.4   5
27  9.8   6
28  10.1   5
29  10.6   6
30  9.8   5
31  9.4   6
32  9.2   5
33  9.6   6
34  10.8   6
35  9.7   7
36  9.8   4
37  10.3   5
38  9.3   4
39  10.5   6
40  10.3   5
41  9.5   5
42  12.1   4
25°C Sunny Search bit blob broom bslib cachem callr cellranger cli clipr conflicted cpp11 crayon curl data.table DBI httr
Import Dataset - 212 MB List
Global Environment
tung_wine 1599 obs. of 12 variables
quality_piv_ 1599 obs. of 7 variables
students 3 obs. of 3 variables
unique_alco_ 65 obs. of 12 variables
wide_df 1599 obs. of 8 variables
wine 1599 obs. of 12 variables
winequality_ 1599 obs. of 12 variables
values
difference 10
iris <Promise>
User Library
askpass https://cran.rstudio.com 1.2.1
backports https://cran.rstudio.com 1.5.0
base64enc https://cran.rstudio.com 0.1-3
bit https://cran.rstudio.com 4.6.0
bit64 https://cran.rstudio.com 4.6.0-1
blob https://cran.rstudio.com 1.2.4
broom https://cran.rstudio.com 1.0.10
bslib https://cran.rstudio.com 0.9.0
cachem https://cran.rstudio.com 1.1.0
callr https://cran.rstudio.com 3.7.6
cellranger https://cran.rstudio.com 1.1.0
cli https://cran.rstudio.com 3.6.5
clipr https://cran.rstudio.com 0.8.0
conflicted https://cran.rstudio.com 1.2.0
cpp11 https://cran.rstudio.com 0.5.2
crayon https://cran.rstudio.com 1.5.3
curl https://cran.rstudio.com 7.0.0
data.table https://cran.rstudio.com 1.17.8
DBI https://cran.rstudio.com 1.2.3
httr https://cran.rstudio.com 2.5.1
10:57 ENG IN 08-12-2025

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
R - R 4.5.2 - ~
/4 9.1   4
75 9.2   5
76 9.4   5
77 9.1   5
78 9.4   5
79 10.3   6
80 10.1   5
81 9.9   6
82 9.6   5
83 9.5   5
[ reached 'max' / getoption("max.print") -- omitted 1276 rows ]
>
> cat("\nRows before:", nrow(wine))
Rows before: 1599> cat("\nRows after removing duplicates:", nrow(clean_wine), "\n")
Rows after removing duplicates: 1359
>
> # -----
> # 4. REMOVE DUPLICATES BASED ON SPECIFIC COLUMNS
> # -----
> # Example: Keep only unique rows based on alcohol value (first occurrence kept)
>
> unique_alcohol <- wine %>%
+   distinct(alcohol, .keep_all = TRUE)
>
> cat("--- 4. Unique Alcohol values (using distinct on specific column) ---\n")
--- 4. Unique Alcohol values (using distinct on specific column) ---
> print(head(unique_alcohol))
fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1   7.0    0.70    0.025   1.9    0.076    11     20   0.9978 3.12   0.60
2   7.6    0.68    0.00     2.0    0.098    25     57   0.9968 3.00   0.68
3   7.3    0.65    0.00     1.2    0.065    15     21   0.9946 3.39   0.47
4   7.8    0.58    0.02     2.0    0.073      9     18   0.9968 3.36   0.57
5   7.5    0.50    0.36     6.1    0.071    17    102   0.9978 3.35   0.80
6   6.7    0.58    0.08     1.8    0.097     15    65   0.9959 3.28   0.54
alcohol quality
1   9.4    5
2   9.8    5
3   10.0   7
4   9.5    7
5   10.5   5
6   9.2    5
> | 25°C Sunny Search bit blob broom bslib cachem callr cellranger cli clipr conflicted cpp11 crayon curl data.table DBI httr
Import Dataset - 212 MB List
Global Environment
tung_wine 1599 obs. of 12 variables
quality_piv_ 1599 obs. of 7 variables
students 3 obs. of 3 variables
unique_alco_ 65 obs. of 12 variables
wide_df 1599 obs. of 8 variables
wine 1599 obs. of 12 variables
winequality_ 1599 obs. of 12 variables
values
difference 10
iris <Promise>
User Library
askpass https://cran.rstudio.com 1.2.1
backports https://cran.rstudio.com 1.5.0
base64enc https://cran.rstudio.com 0.1-3
bit https://cran.rstudio.com 4.6.0
bit64 https://cran.rstudio.com 4.6.0-1
blob https://cran.rstudio.com 1.2.4
broom https://cran.rstudio.com 1.0.10
bslib https://cran.rstudio.com 0.9.0
cachem https://cran.rstudio.com 1.1.0
callr https://cran.rstudio.com 3.7.6
cellranger https://cran.rstudio.com 1.1.0
cli https://cran.rstudio.com 3.6.5
clipr https://cran.rstudio.com 0.8.0
conflicted https://cran.rstudio.com 1.2.0
cpp11 https://cran.rstudio.com 0.5.2
crayon https://cran.rstudio.com 1.5.3
curl https://cran.rstudio.com 7.0.0
data.table https://cran.rstudio.com 1.17.8
DBI https://cran.rstudio.com 1.2.3
httr https://cran.rstudio.com 2.5.1
10:57 ENG IN 08-12-2025

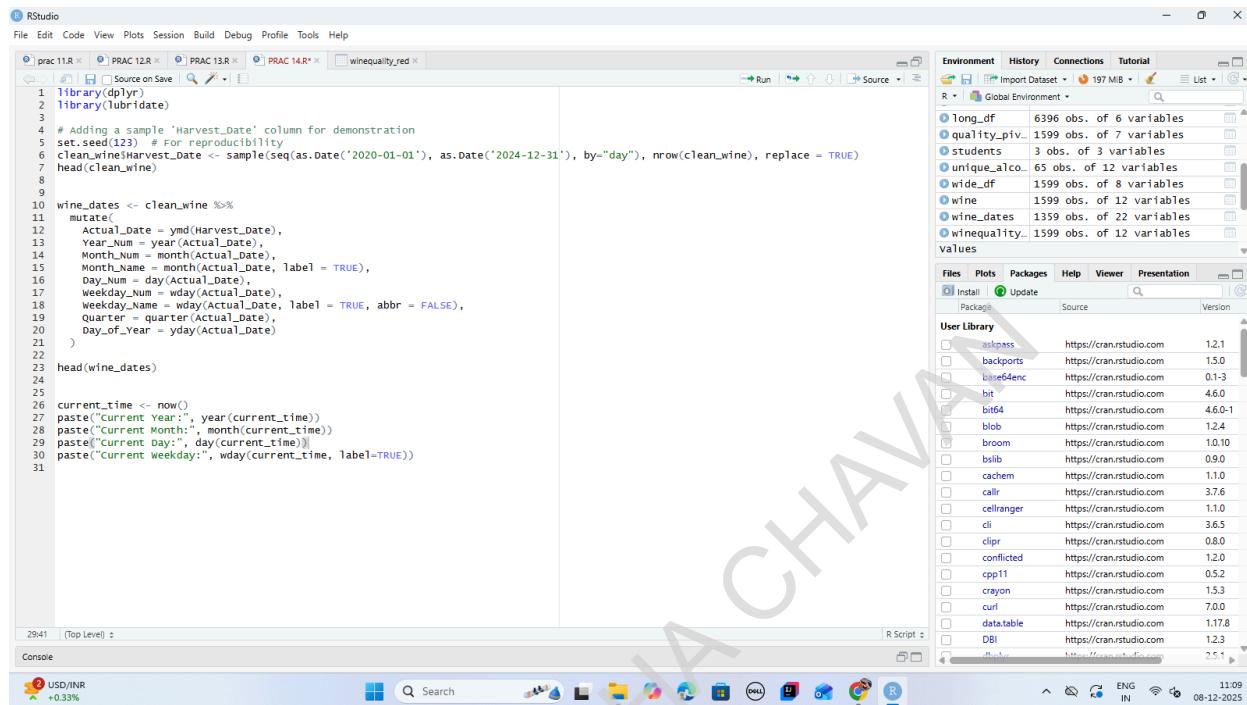
```

NAME: DEESHA CHAVAN
ROLL NO.: S076

SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT: DATA ANALYSIS WITH R

PRACTICAL NO. 14

AIM: Extracting date components using lubridate:: functions (R).



The screenshot shows the RStudio interface. The code editor contains R code for generating date components from a dataset. The environment pane shows various data frames and packages loaded.

```

library(dplyr)
library(lubridate)

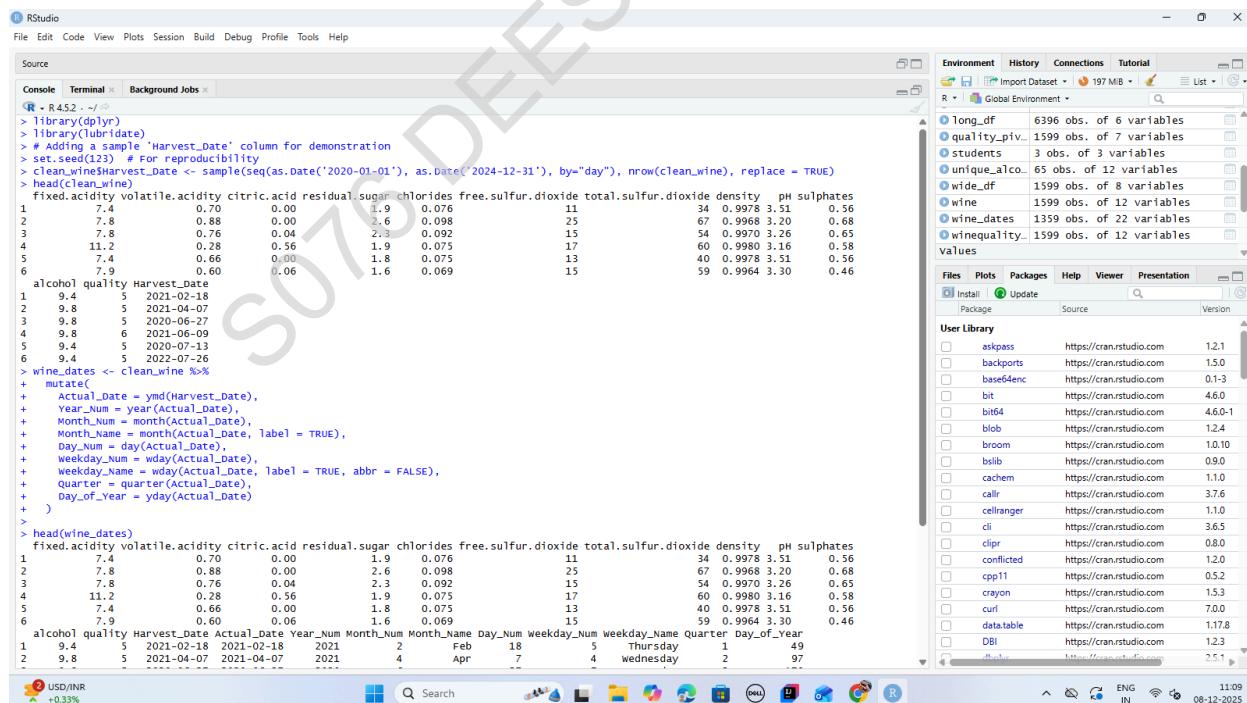
# Adding a sample 'Harvest_Date' column for demonstration
set.seed(123) # For reproducibility
clean_wine$Harvest_Date <- sample(as.Date('2020-01-01'), as.Date('2024-12-31'), by="day"), nrow(clean_wine), replace = TRUE)
head(clean_wine)

wine_dates <- clean_wine %>
  mutate(
    Actual_Date = ymd(Harvest_Date),
    Year_Num = year(Actual_Date),
    Month_Num = month(Actual_Date),
    Month_Name = month(Actual_date, label = TRUE),
    Day_Num = day(Actual_date),
    weekday_Num = wday(Actual_date),
    weekday_Name = wday(Actual_date, label = TRUE, abbr = FALSE),
    quarter = quarter(Actual_date),
    Day_of_Year = yday(Actual_date)
  )
head(wine_dates)

current_time <- now()
paste("Current Year:", year(current_time))
paste("Current Month:", month(current_time))
paste("Current Day:", day(current_time))
paste("Current weekday:", wday(current_time, label=TRUE))

```

OUTPUT:



The screenshot shows the RStudio interface with the output console displaying the results of the R code execution. The environment pane shows the same data frames and packages as the previous screenshot.

```

fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1 7.4 0.70 0.00 1.9 0.076 11 34 0.9978 3.51 0.56
2 7.8 0.88 0.00 2.6 0.098 25 67 0.9968 3.20 0.68
3 7.8 0.76 0.04 2.3 0.092 15 54 0.9970 3.26 0.65
4 11.2 0.28 0.56 1.9 0.075 17 60 0.9980 3.16 0.58
5 7.4 0.66 0.00 1.8 0.075 13 40 0.9978 3.51 0.56
6 7.9 0.60 0.06 1.6 0.069 15 59 0.9964 3.30 0.46

alcohol quality Harvest_Date
1 9.4 5 2021-02-18
2 9.8 5 2021-04-07
3 9.8 5 2020-06-27
4 9.8 6 2021-06-09
5 9.4 5 2020-07-13
6 9.4 5 2022-07-26

wine_dates <- clean_wine %>
  mutate(
    Actual_Date = ymd(Harvest_Date),
    Year_Num = year(Actual_Date),
    Month_Num = month(Actual_Date),
    Month_Name = month(Actual_date, label = TRUE),
    Day_Num = day(Actual_date),
    weekday_Num = wday(Actual_date),
    weekday_Name = wday(Actual_date, label = TRUE, abbr = FALSE),
    quarter = quarter(Actual_date),
    Day_of_Year = yday(Actual_date)
  )
head(wine_dates)

fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1 7.4 0.70 0.00 1.9 0.076 11 34 0.9978 3.51 0.56
2 7.8 0.88 0.00 2.6 0.098 25 67 0.9968 3.20 0.68
3 7.8 0.76 0.04 2.3 0.092 15 54 0.9970 3.26 0.65
4 11.2 0.28 0.56 1.9 0.075 17 60 0.9980 3.16 0.58
5 7.4 0.66 0.00 1.8 0.075 13 40 0.9978 3.51 0.56
6 7.9 0.60 0.06 1.6 0.069 15 59 0.9964 3.30 0.46

alcohol quality Harvest_Date Actual_Date Year_Num Month_Num Month_Name Day_Num weekday_Num weekday_Name quarter Day_of_Year
1 9.4 5 2021-02-18 2021-02-18 2021 2 Feb 18 5 Thursday 1 49
2 9.8 5 2021-04-07 2021-04-07 2021 4 Apr 7 4 Wednesday 2 97

```

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

SUBJECT: DATA ANALYSIS WITH R

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help

Source
R - R 4.5.2 - ~/ ...
Console Terminal Background Jobs
[R - R 4.5.2 - ~/ ...]
1 9.4      5 2021-02-18
2 9.8      5 2021-04-07
3 9.8      5 2020-06-27
4 9.8      6 2021-06-09
5 9.4      5 2020-07-13
6 9.4      5 2022-07-26
> wine_dates <- clear_wine %>
+ mutate(
+   Actual_Date = ymd(Harvest_Date),
+   Year_Num = year(Actual_Date),
+   Month_Num = month(Actual_Date),
+   Month_Name = month(Actual_Date, label = TRUE),
+   Day_Num = day(Actual_Date),
+   Weekday_Num = wday(Actual_Date),
+   Weekday_Name = wday(Actual_Date, label = TRUE, abbr = FALSE),
+   Quarter = quarter(Actual_Date),
+   Day_of_Year = yday(Actual_Date)
+ )
>
> head(wine_dates)
fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide density pH sulphates
1 7.4      0.70 0.00 1.9 0.076 11 34 0.988 3.51 0.56
2 7.8      0.88 0.00 2.6 0.098 25 57 0.989 3.0 0.68
3 7.8      0.76 0.04 2.3 0.092 15 54 0.9970 3.26 0.65
4 11.2     0.28 0.56 1.9 0.075 17 60 0.9980 3.16 0.58
5 7.4      0.66 0.00 1.8 0.075 13 40 0.9978 3.51 0.56
6 7.9      0.60 0.06 1.6 0.069 15 59 0.9964 3.30 0.46
alcohol quality Harvest_Date Actual_Date Year_Num Month_Num Month_Name Day_Num Weekday_Num weekday_Name Quarter Day_of_Year
1 9.4      5 2021-02-18 2021-02-18 2021 2 Feb 18 5 Thursday 1 49
2 9.8      5 2021-04-07 2021-04-07 2021 4 Apr 7 4 wednesday 2 97
3 9.8      5 2020-06-27 2020-06-27 2020 6 Jun 27 7 Saturday 2 179
4 9.6      5 2021-06-09 2021-06-09 2021 6 Jun 9 4 wednesday 2 160
5 9.4      5 2020-07-13 2020-07-13 2020 7 Jul 13 2 Monday 3 195
6 9.4      5 2022-07-26 2022-07-26 2022 7 Jul 26 3 Tuesday 3 207
> current_time <- now()
> paste("Current Year:", year(current_time))
[1] "Current Year: 2025"
> paste("Current Month:", month(current_time))
[1] "Current Month: 12"
> paste("Current Day:", day(current_time))
[1] "Current Day: 8"
> paste("Current Weekday:", wday(current_time, label=TRUE))
[1] "Current weekday: Mon"
> |
> |

```

USD/INR +0.33% 11:09 ENG IN 08-12-2025

PRACTICAL NO. 15

AIM: Generating basic summaries using str() or summary() (R).

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help

winequality_red PRAC 12.R PRAC 13.R PRAC 14.R PRAC 15.R prac 11.R
Source Source on Save Run Environment History Connections Tutorial
[R - R 4.5.2 - ~/ ...]
1 # R Script: Generating Basic Summaries
2 # Functions: str() and summary()
3 # Dataset: Wine Quality (Red wine)
4 #
5 #
6 #
7 #
8 # 1. SETUP: Load wine dataset
9 #
10 #
11 # NOTE: Use file.choose() if the csv is not in your working directory
12 # wine_df <- read.csv(file.choose())
13 #
14 wine_df <- read.csv("D:/S076/winequality-red.csv") # -- use correct path or file.choose()
15 #
16 print("--- wine Data Loaded ---")
17 #
18 #
19 #
20 # 2. USING str() (Structure)
21 #
22 #
23 #
24 # Purpose: Compactly display the internal structure of the R object.
25 # What it tells you:
26 # - The class of the frame
27 # - Number of observations (rows) and variables (columns)
28 # - The data type of each column (num, int, chr)
29 # - The first few values
30 #
31 print("--- OUTPUT OF str() ---")
32 str(wine_df)
33 #
34 #
35 #
36 # 3. USING summary() (Statistical Summary)
37 #
38 #
39 # Purpose: detailed summary statistics for each column.
40 # What it tells you:
41 # - Numeric cols: Min, 1st Qu, Median, Mean, 3rd Qu, Max
42 # - Integer cols: Same statistical summary
43 # - Character cols: Length and class
44 #
45 #
46 #
47 #
48 #
49 #
50 #
51 #
52 #
53 #
54 #
55 #
56 #
57 #
58 #
59 #
60 #
61 #
62 #
63 #
64 #
65 #
66 #
67 #
68 #
69 #
70 #
71 #
72 #
73 #
74 #
75 #
76 #
77 #
78 #
79 #
80 #
81 #
82 #
83 #
84 #
85 #
86 #
87 #
88 #
89 #
90 #
91 #
92 #
93 #
94 #
95 #
96 #
97 #
98 #
99 #
100 #
101 #
102 #
103 #
104 #
105 #
106 #
107 #
108 #
109 #
110 #
111 #
112 #
113 #
114 #
115 #
116 #
117 #
118 #
119 #
120 #
121 #
122 #
123 #
124 #
125 #
126 #
127 #
128 #
129 #
130 #
131 #
132 #
133 #
134 #
135 #
136 #
137 #
138 #
139 #
140 #
141 #
142 #
143 #
144 #
145 #
146 #
147 #
148 #
149 #
150 #
151 #
152 #
153 #
154 #
155 #
156 #
157 #
158 #
159 #
160 #
161 #
162 #
163 #
164 #
165 #
166 #
167 #
168 #
169 #
170 #
171 #
172 #
173 #
174 #
175 #
176 #
177 #
178 #
179 #
180 #
181 #
182 #
183 #
184 #
185 #
186 #
187 #
188 #
189 #
190 #
191 #
192 #
193 #
194 #
195 #
196 #
197 #
198 #
199 #
200 #
201 #
202 #
203 #
204 #
205 #
206 #
207 #
208 #
209 #
210 #
211 #
212 #
213 #
214 #
215 #
216 #
217 #
218 #
219 #
220 #
221 #
222 #
223 #
224 #
225 #
226 #
227 #
228 #
229 #
230 #
231 #
232 #
233 #
234 #
235 #
236 #
237 #
238 #
239 #
240 #
241 #
242 #
243 #
244 #
245 #
246 #
247 #
248 #
249 #
250 #
251 #
252 #
253 #
254 #
255 #
256 #
257 #
258 #
259 #
2510 #
2511 #
2512 #
2513 #
2514 #
2515 #
2516 #
2517 #
2518 #
2519 #
2520 #
2521 #
2522 #
2523 #
2524 #
2525 #
2526 #
2527 #
2528 #
2529 #
2530 #
2531 #
2532 #
2533 #
2534 #
2535 #
2536 #
2537 #
2538 #
2539 #
2540 #
2541 #
2542 #
2543 #
2544 #
2545 #
2546 #
2547 #
2548 #
2549 #
2550 #
2551 #
2552 #
2553 #
2554 #
2555 #
2556 #
2557 #
2558 #
2559 #
2560 #
2561 #
2562 #
2563 #
2564 #
2565 #
2566 #
2567 #
2568 #
2569 #
2570 #
2571 #
2572 #
2573 #
2574 #
2575 #
2576 #
2577 #
2578 #
2579 #
2580 #
2581 #
2582 #
2583 #
2584 #
2585 #
2586 #
2587 #
2588 #
2589 #
2590 #
2591 #
2592 #
2593 #
2594 #
2595 #
2596 #
2597 #
2598 #
2599 #
25900 #
25901 #
25902 #
25903 #
25904 #
25905 #
25906 #
25907 #
25908 #
25909 #
25910 #
25911 #
25912 #
25913 #
25914 #
25915 #
25916 #
25917 #
25918 #
25919 #
25920 #
25921 #
25922 #
25923 #
25924 #
25925 #
25926 #
25927 #
25928 #
25929 #
25930 #
25931 #
25932 #
25933 #
25934 #
25935 #
25936 #
25937 #
25938 #
25939 #
25940 #
25941 #
25942 #
25943 #
25944 #
25945 #
25946 #
25947 #
25948 #
25949 #
25950 #
25951 #
25952 #
25953 #
25954 #
25955 #
25956 #
25957 #
25958 #
25959 #
25960 #
25961 #
25962 #
25963 #
25964 #
25965 #
25966 #
25967 #
25968 #
25969 #
25970 #
25971 #
25972 #
25973 #
25974 #
25975 #
25976 #
25977 #
25978 #
25979 #
25980 #
25981 #
25982 #
25983 #
25984 #
25985 #
25986 #
25987 #
25988 #
25989 #
25990 #
25991 #
25992 #
25993 #
25994 #
25995 #
25996 #
25997 #
25998 #
25999 #
259000 #
259001 #
259002 #
259003 #
259004 #
259005 #
259006 #
259007 #
259008 #
259009 #
259010 #
259011 #
259012 #
259013 #
259014 #
259015 #
259016 #
259017 #
259018 #
259019 #
259020 #
259021 #
259022 #
259023 #
259024 #
259025 #
259026 #
259027 #
259028 #
259029 #
259030 #
259031 #
259032 #
259033 #
259034 #
259035 #
259036 #
259037 #
259038 #
259039 #
259040 #
259041 #
259042 #
259043 #
259044 #
259045 #
259046 #
259047 #
259048 #
259049 #
259050 #
259051 #
259052 #
259053 #
259054 #
259055 #
259056 #
259057 #
259058 #
259059 #
259060 #
259061 #
259062 #
259063 #
259064 #
259065 #
259066 #
259067 #
259068 #
259069 #
259070 #
259071 #
259072 #
259073 #
259074 #
259075 #
259076 #
259077 #
259078 #
259079 #
259080 #
259081 #
259082 #
259083 #
259084 #
259085 #
259086 #
259087 #
259088 #
259089 #
259090 #
259091 #
259092 #
259093 #
259094 #
259095 #
259096 #
259097 #
259098 #
259099 #
2590100 #
2590101 #
2590102 #
2590103 #
2590104 #
2590105 #
2590106 #
2590107 #
2590108 #
2590109 #
2590110 #
2590111 #
2590112 #
2590113 #
2590114 #
2590115 #
2590116 #
2590117 #
2590118 #
2590119 #
2590120 #
2590121 #
2590122 #
2590123 #
2590124 #
2590125 #
2590126 #
2590127 #
2590128 #
2590129 #
2590130 #
2590131 #
2590132 #
2590133 #
2590134 #
2590135 #
2590136 #
2590137 #
2590138 #
2590139 #
2590140 #
2590141 #
2590142 #
2590143 #
2590144 #
2590145 #
2590146 #
2590147 #
2590148 #
2590149 #
2590150 #
2590151 #
2590152 #
2590153 #
2590154 #
2590155 #
2590156 #
2590157 #
2590158 #
2590159 #
2590160 #
2590161 #
2590162 #
2590163 #
2590164 #
2590165 #
2590166 #
2590167 #
2590168 #
2590169 #
2590170 #
2590171 #
2590172 #
2590173 #
2590174 #
2590175 #
2590176 #
2590177 #
2590178 #
2590179 #
2590180 #
2590181 #
2590182 #
2590183 #
2590184 #
2590185 #
2590186 #
2590187 #
2590188 #
2590189 #
2590190 #
2590191 #
2590192 #
2590193 #
2590194 #
2590195 #
2590196 #
2590197 #
2590198 #
2590199 #
2590200 #
2590201 #
2590202 #
2590203 #
2590204 #
2590205 #
2590206 #
2590207 #
2590208 #
2590209 #
2590210 #
2590211 #
2590212 #
2590213 #
2590214 #
2590215 #
2590216 #
2590217 #
2590218 #
2590219 #
2590220 #
2590221 #
2590222 #
2590223 #
2590224 #
2590225 #
2590226 #
2590227 #
2590228 #
2590229 #
2590230 #
2590231 #
2590232 #
2590233 #
2590234 #
2590235 #
2590236 #
2590237 #
2590238 #
2590239 #
2590240 #
2590241 #
2590242 #
2590243 #
2590244 #
2590245 #
2590246 #
2590247 #
2590248 #
2590249 #
2590250 #
2590251 #
2590252 #
2590253 #
2590254 #
2590255 #
2590256 #
2590257 #
2590258 #
2590259 #
2590260 #
2590261 #
2590262 #
2590263 #
2590264 #
2590265 #
2590266 #
2590267 #
2590268 #
2590269 #
2590270 #
2590271 #
2590272 #
2590273 #
2590274 #
2590275 #
2590276 #
2590277 #
2590278 #
2590279 #
2590280 #
2590281 #
2590282 #
2590283 #
2590284 #
2590285 #
2590286 #
2590287 #
2590288 #
2590289 #
2590290 #
2590291 #
2590292 #
2590293 #
2590294 #
2590295 #
2590296 #
2590297 #
2590298 #
2590299 #
2590300 #
2590301 #
2590302 #
2590303 #
2590304 #
2590305 #
2590306 #
2590307 #
2590308 #
2590309 #
2590310 #
2590311 #
2590312 #
2590313 #
2590314 #
2590315 #
2590316 #
2590317 #
2590318 #
2590319 #
2590320 #
2590321 #
2590322 #
2590323 #
2590324 #
2590325 #
2590326 #
2590327 #
2590328 #
2590329 #
2590330 #
2590331 #
2590332 #
2590333 #
2590334 #
2590335 #
2590336 #
2590337 #
2590338 #
2590339 #
2590340 #
2590341 #
2590342 #
2590343 #
2590344 #
2590345 #
2590346 #
2590347 #
2590348 #
2590349 #
2590350 #
2590351 #
2590352 #
2590353 #
2590354 #
2590355 #
2590356 #
2590357 #
2590358 #
2590359 #
2590360 #
2590361 #
2590362 #
2590363 #
2590364 #
2590365 #
2590366 #
2590367 #
2590368 #
2590369 #
2590370 #
2590371 #
2590372 #
2590373 #
2590374 #
2590375 #
2590376 #
2590377 #
2590378 #
2590379 #
2590380 #
2590381 #
2590382 #
2590383 #
2590384 #
2590385 #
2590386 #
2590387 #
2590388 #
2590389 #
2590390 #
2590391 #
2590392 #
2590393 #
2590394 #
2590395 #
2590396 #
2590397 #
2590398 #
2590399 #
2590400 #
2590401 #
2590402 #
2590403 #
2590404 #
2590405 #
2590406 #
2590407 #
2590408 #
2590409 #
2590410 #
2590411 #
2590412 #
2590413 #
2590414 #
2590415 #
2590416 #
2590417 #
2590418 #
2590419 #
2590420 #
2590421 #
2590422 #
2590423 #
2590424 #
2590425 #
2590426 #
2590427 #
2590428 #
2590429 #
2590430 #
2590431 #
2590432 #
2590433 #
2590434 #
2590435 #
2590436 #
2590437 #
2590438 #
2590439 #
2590440 #
2590441 #
2590442 #
2590443 #
2590444 #
2590445 #
2590446 #
2590447 #
2590448 #
2590449 #
2590450 #
2590451 #
2590452 #
2590453 #
2590454 #
2590455 #
2590456 #
2590457 #
2590458 #
2590459 #
2590460 #
2590461 #
2590462 #
2590463 #
2590464 #
2590465 #
2590466 #
2590467 #
2590468 #
2590469 #
2590470 #
2590471 #
2590472 #
2590473 #
2590474 #
2590475 #
2590476 #
2590477 #
2590478 #
2590479 #
2590480 #
2590481 #
2590482 #
2590483 #
2590484 #
2590485 #
2590486 #
2590487 #
2590488 #
2590489 #
2590490 #
2590491 #
2590492 #
2590493 #
2590494 #
2590495 #
2590496 #
2590497 #
2590498 #
2590499 #
2590500 #
2590501 #
2590502 #
2590503 #
2590504 #
2590505 #
2590506 #
2590507 #
2590508 #
2590509 #
2590510 #
2590511 #
2590512 #
2590513 #
2590514 #
2590515 #
2590516 #
2590517 #
2590518 #
2590519 #
2590520 #
2590521 #
2590522 #
2590523 #
2590524 #
2590525 #
2590526 #
2590527 #
2590528 #
2590529 #
2590530 #
2590531 #
2590532 #
2590533 #
2590534 #
2590535 #
2590536 #
2590537 #
2590538 #
2590539 #
2590540 #
2590541 #
2590542 #
2590543 #
2590544 #
2590545 #
2590546 #
2590547 #
2590548 #
2590549 #
2590550 #
2590551 #
2590552 #
2590553 #
2590554 #
2590555 #
2590556 #
2590557 #
2590558 #
2590559 #
2590560 #
2590561 #
2590562 #
2590563 #
2590564 #
2590565 #
2590566 #
2590567 #
2590568 #
2590569 #
2590570 #
2590571 #
2590572 #
2590573 #
2590574 #
2590575 #
2590576 #
2590577 #
2590578 #
2590579 #
2590580 #
2590581 #
2590582 #
2590583 #
2590584 #
2590585 #
2590586 #
2590587 #
2590588 #
2590589 #
2590590 #
2590591 #
2590592 #
2590593 #
2590594 #
2590595 #
2590596 #
2590597 #
2590598 #
2590599 #
25905900 #
25905901 #
25905902 #
25905903 #
25905904 #
25905905 #
25905906 #
25905907 #
25905908 #
25905909 #
25905910 #
25905911 #
25905912 #
25905913 #
25905914 #
25905915 #
25905916 #
25905917 #
25905918 #
25905919 #
25905920 #
25905921 #
25905922 #
25905923 #
25905924 #
25905925 #
25905926 #
25905927 #
25905928 #
25905929 #
25905930 #
25905931 #
25905932 #
25905933 #
25905934 #
25905935 #
25905936 #
25905937 #
25905938 #
25905939 #
25905940 #
25905941 #
25905942 #
25905943 #
25905944 #
25905945 #
25905946 #
25905947 #
25905948 #
25905949 #
25905950 #
25905951 #
25905952 #
25905953 #
25905954 #
25905955 #
25905956 #
25905957 #
25905958 #
25905959 #
259059500 #
259059501 #
259059502 #
259059503 #
259059504 #
259059505 #
259059506 #
259059507 #
259059508 #
259059509 #
259059510 #
259059511 #
259059512 #
259059513 #
259059514 #
259059515 #
259059516 #
259059517 #
259059518 #
259059519 #
259059520 #
259059521 #
259059522 #
259059523 #
259059524 #
259059525 #
259059526 #
259059527 #
259059528 #
259059529 #
259059530 #
259059531 #
259059532 #
259059533 #
259059534 #
259059535 #
259059536 #
259059537 #
259059538 #
259059539 #
259059540 #
259059541 #
259059542 #
259059543 #
259059544 #
259059545 #
259059546 #
259059547 #
259059548 #
259059549 #
259059550 #
259059551 #
259059552 #
259059553 #
259059554 #
259059555 #
259059556 #
259059557 #
259059558 #
259059559 #
259059560 #
259059561 #
259059562 #
259059563 #
259059564 #
259059565 #
259059566 #
259059567 #
259059568 #
259059569 #
259059570 #
259059571 #
259059572 #
259059573 #
259059574 #
259059575 #
259059576 #
259059577 #
259059578 #
259059579 #
259059580 #
259059581 #
259059582 #
259059583 #
259059584 #
259059585 #
259059586 #
259059587 #
259059588 #
259059589 #
259059590 #
259059591 #
259059592 #
259059593 #
259059594 #
259059595 #
259059596 #
259059597 #
259059598 #
259059599 #
2590595000 #
2590595001 #
2590595002 #
2590595003 #
2590595004 #
2590595005 #
2590595006 #
2590595007 #
2590595008 #
2590595009 #
2590595010 #
2590595011 #
2590595012 #
2590595013 #
2590595014 #
2590595015 #
2590595016 #
2590595017 #
2590595018 #
2590595019 #
2590595020 #
2590595021 #
2590595022 #
2590595023 #
2590595024 #
2590595025 #
2590595026 #
2590595027 #
2590595028 #
2590595029 #
2590595030 #
2590595031 #
2590595032 #
2590595033 #
2590595034 #
2590595035 #
2590595036 #
2590595037 #
2590595038 #
2590595039 #
2590595040 #
2590595041 #
2590595042 #
2590595043 #
2590595044 #
2590595045 #
2590595046 #
2590595047 #
2590595048 #
2590595049 #
2590595050 #
2590595051 #
2590595052 #
2590595053 #
2590595054 #
2590595055 #
2590595056 #
2590595057 #
2590595058 #
2590595059 #
2590595060 #
2590595061 #
2590595062 #
2590595063 #
2590595064 #
2590595065 #
2590595066 #
2590595067 #
2590595068 #
2590595069 #
2590595070 #
2590595071 #
2590595072 #
2590595073 #
2590595074 #
2590595075 #
2590595076 #
2590595077 #
2590595078 #
2590595079 #
2590595080 #
2590595081 #
2590595082 #
2590595083 #
2590595084 #
2590595085 #
2590595086 #
2590595087 #
2590595088 #
2590595089 #
2590595090 #
2590595091 #
2590595092 #
2590595093 #
2590595094 #
2590595095 #
2590595096 #
2590595097 #
2590595098 #
2590595099 #
2590595100 #
2590595101 #
2590595102 #
2590595103 #
2590595104 #
2590595105 #
2590595106 #
2590595107 #
2590595108 #
2590595109 #
2590595110 #
2590595111 #
2590595112 #
2590595113 #
2590595114 #
2590595115 #
2590595116 #
2590595117 #
2590595118 #
2590595119 #
2590595120 #
2590595121 #
2590595122 #
2590595123 #
2590595124 #
2590595125 #
2590595126 #
2590595127 #
2590595128 #
2590595129 #
2590595130 #
2590595131 #
2590595132 #
2590595133 #
259
```

**SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT: DATA ANALYSIS WITH R**

The screenshot shows the RStudio interface. The left pane contains a script editor with R code for wine quality analysis. The right pane shows the Global Environment viewer with a list of objects and their details. A large watermark 'DataCamp' is visible across the bottom of the screen.

```
winequality.rd | PRAC 12.R | PRAC 13.R | PRAC 14.R | PRAC 15.R | prac 11.R
```

```
1 # Load wine quality dataset
2 library(tidyverse)
3
4 # Load wine quality dataset
5 wine_wq <- read.csv("winequality.csv")
6
7 # View first few rows of the dataset
8 head(wine_wq)
```

Environment History Connections Tutorial

R + Global Environment

Data

- batch_A 5 obs. of 12 variables
- batch_B 5 obs. of 12 variables
- combined_wq... 10 obs. of 12 variables
- globalairqu... 18000 obs. of 15 variables
- globalairqu... 6 obs. of 4 variables
- wine 1359 obs. of 12 variables
- wine_df 1599 obs. of 12 variables
- winequality_... 1599 obs. of 12 variables

Files Plots Packages Help Viewer Presentation

Install Update Source Version

User Library

- askpass https://cran.rstudio.com 1.2.1
- backports https://cran.rstudio.com 1.5.0
- base64enc https://cran.rstudio.com 0.1.3
- bit https://cran.rstudio.com 4.6.0
- bit64 https://cran.rstudio.com 4.6.0-1
- blob https://cran.rstudio.com 1.2.4
- broom https://cran.rstudio.com 1.0.10
- bitlib https://cran.rstudio.com 0.9.0
- cachem https://cran.rstudio.com 1.1.0
- callr https://cran.rstudio.com 3.7.6
- ceilranger https://cran.rstudio.com 1.1.0
- cli https://cran.rstudio.com 3.6.5
- clipr https://cran.rstudio.com 0.8.0
- conflicted https://cran.rstudio.com 1.2.0
- cpp11 https://cran.rstudio.com 0.5.2
- crayon https://cran.rstudio.com 1.5.3
- curl https://cran.rstudio.com 7.0.0
- data.table https://cran.rstudio.com 1.17.8
- DBI https://cran.rstudio.com 1.2.3
- httr https://cran.rstudio.com 1.1.1

Script

Console

27°C Sunny

Search

ENG IN

08-12 11:58

Output:

The screenshot shows an RStudio interface with the following details:

- File Edit Code View Plots Session Build Debug Profile Tools Help**
- Source** tab selected.
- Console** tab active, showing R code for generating basic summaries, loading the wine dataset, and printing its structure.
- Data** pane on the right lists various datasets:
 - batch_A: 5 obs. of 12 variables
 - batch_B: 5 obs. of 12 variables
 - combined_wine: 10 obs. of 12 variables
 - globalairquality: 18000 obs. of 15 variables
 - retail_df: 6 obs. of 4 variables
 - wine: 1359 obs. of 12 variables
 - wine_df: 1599 obs. of 12 variables
 - winequality: 1599 obs. of 12 variables
- Files Plots Packages Help Viewer Presentation** tabs at the bottom of the RStudio window.
- User Library** pane on the right lists packages and their URLs, such as askpass, backports, base64enc, bit, blob, broom, bslib, cachem, callr, cellranger, cli, clipr, conflicted, cpp11, crayon, curl, data.table, DBI, etc.
- System tray icons** include battery level (27% Sunny), search bar, and system status indicators.

NAME: DEESHA CHAVAN
ROLL NO.: S076

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

SUBJECT: DATA ANALYSIS WITH R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R 4.5.2 - ~
```

```
$ sulphates      : num  0.56 0.68 0.65 0.58 0.56 0.56 0.46 0.47 0.57 0.8 ...
$ alcohol        : num  9.4 9.8 9.8 9.4 9.4 9.4 10 9.5 10.5 ...
$ quality        : int  5 5 5 6 5 5 5 7 7 5 ...
>
> # =====
> # 3. USING summary() (Statistical summary)
> # =====
>
> # Purpose: detailed summary statistics for each column.
> # What it tells you:
> # = Numeric Cols: Min, 1st QU, Median, Mean, 3rd QU, Max
> # - Integer Cols: Same statistical summary
> # - Character Cols: Length and Class
>
> print("---- OUTPUT OF summary() [original data] ----")
[1] "---- OUTPUT OF summary() [original data] ----"
> summary(wine_df)

fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide
Min. : 4.60   Min. :0.1200  Min. :0.0000  Min. :0.900  Min. :0.01200  Min. : 1.00  Min. : 6.00
1st Qu.: 7.10  1st Qu.:0.3900  1st Qu.:0.0900  1st Qu.: 1.900  1st Qu.:0.07000  1st Qu.: 7.00  1st Qu.:22.00
Median : 7.90  Median :0.5200  Median :0.2600  Median : 2.200  Median :0.07900  Median :14.00  Median :38.00
Mean   : 8.32  Mean   :0.5278  Mean   :0.271  Mean   : 2.530  Mean   :0.08747  Mean   :15.87  Mean   :46.47
3rd Qu.: 9.20  3rd Qu.:0.6400  3rd Qu.:0.420  3rd Qu.: 2.600  3rd Qu.:0.09000  3rd Qu.:21.00  3rd Qu.:62.00
Max.   :15.90  Max.   :1.5800  Max.   :1.0000  Max.   :15.500  Max.   :0.61100  Max.   :72.00  Max.   :289.00

density    pH    sulphates    alcohol    quality
Min. :0.9901  Min. :3.2740  Min. :0.3300  Min. : 8.40  Min. : 1.0000
1st Qu.:0.9956  1st Qu.:3.210  1st Qu.:0.5500  1st Qu.: 9.50  1st Qu.:15.000
Median :0.9968  Median :3.310  Median :0.6200  Median :10.20  Median :16.000
Mean   :0.9967  Mean   :3.311  Mean   :0.6581  Mean   :10.42  Mean   :15.636
3rd Qu.:0.9978  3rd Qu.:3.400  3rd Qu.:0.7300  3rd Qu.:11.10  3rd Qu.:16.000
Max.   :1.0037  Max.   :4.010  Max.   :2.0000  Max.   :14.90  Max.   :8.000
>
> # =====
> # 4. IMPROVING summary() WITH FACTORS
> # =====
>
> # 'quality' is often treated as a categorical variable
> # Converting it to factor shows counts per quality score
>
> wine_df$quality <- as.factor(wine_df$quality)
>
> print("---- OUTPUT OF summary() [After Factor Conversion] ----")
[1] "---- OUTPUT OF summary() [After Factor Conversion] ----"
```

27°C Sunny

Windows Search

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R 4.5.2 - ~
```

```
>
> # =====
> # 4. IMPROVING summary() WITH FACTORS
> # =====
>
> # 'quality' is often treated as a categorical variable
> # Converting it to factor shows counts per quality score
>
> wine_df$quality <- as.factor(wine_df$quality)
>
> print("---- OUTPUT OF summary() [After Factor Conversion] ----")
[1] "---- OUTPUT OF summary() [After Factor Conversion] ----"
> summary(wine_df)

fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide
Min. : 4.60   Min. :0.1200  Min. :0.0000  Min. :0.900  Min. :0.01200  Min. : 1.00  Min. : 6.00
1st Qu.: 7.10  1st Qu.:0.3900  1st Qu.:0.0900  1st Qu.: 1.900  1st Qu.:0.07000  1st Qu.: 7.00  1st Qu.:22.00
Median : 7.90  Median :0.5200  Median :0.2600  Median : 2.200  Median :0.07900  Median :14.00  Median :38.00
Mean   : 8.32  Mean   :0.5278  Mean   :0.271  Mean   : 2.530  Mean   :0.08747  Mean   :15.87  Mean   :46.47
3rd Qu.: 9.20  3rd Qu.:0.6400  3rd Qu.:0.420  3rd Qu.: 2.600  3rd Qu.:0.09000  3rd Qu.:21.00  3rd Qu.:62.00
Max.   :15.90  Max.   :1.5800  Max.   :1.0000  Max.   :15.500  Max.   :0.61100  Max.   :72.00  Max.   :289.00

density    pH    sulphates    alcohol    quality
Min. :0.9901  Min. :3.2740  Min. :0.3300  Min. : 8.40  Min. : 1.0000
1st Qu.:0.9956  1st Qu.:3.210  1st Qu.:0.5500  1st Qu.: 9.50  1st Qu.:15.000
Median :0.9968  Median :3.310  Median :0.6200  Median :10.20  Median :16.000
Mean   :0.9967  Mean   :3.311  Mean   :0.6581  Mean   :10.42  Mean   :15.636
3rd Qu.:0.9978  3rd Qu.:3.400  3rd Qu.:0.7300  3rd Qu.:11.10  3rd Qu.:16.000
Max.   :1.0037  Max.   :4.010  Max.   :2.0000  Max.   :14.90  Max.   :8.000
>
> # =====
> # 5. Accessing Specific Summaries
> # =====
>
> # Sometimes you want only particular metrics, not the whole summary block.
>
> avg_alcohol <- mean(wine_df$alcohol)
> max_sulphates <- max(wine_df$sulphates)
>
> print(paste("Average Alcohol:", avg_alcohol))
[1] "Average Alcohol: 10.422983144465"
> print(paste("Highest Sulphates:", max_sulphates))
[1] "Highest Sulphates: 2"
> |
```

27°C Sunny

Windows Search

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R 4.5.2 - ~
```

```
>
> # =====
> # 4. IMPROVING summary() WITH FACTORS
> # =====
>
> # 'quality' is often treated as a categorical variable
> # Converting it to factor shows counts per quality score
>
> wine_df$quality <- as.factor(wine_df$quality)
>
> print("---- OUTPUT OF summary() [After Factor Conversion] ----")
[1] "---- OUTPUT OF summary() [After Factor Conversion] ----"
> summary(wine_df)

fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide
Min. : 4.60   Min. :0.1200  Min. :0.0000  Min. :0.900  Min. :0.01200  Min. : 1.00  Min. : 6.00
1st Qu.: 7.10  1st Qu.:0.3900  1st Qu.:0.0900  1st Qu.: 1.900  1st Qu.:0.07000  1st Qu.: 7.00  1st Qu.:22.00
Median : 7.90  Median :0.5200  Median :0.2600  Median : 2.200  Median :0.07900  Median :14.00  Median :38.00
Mean   : 8.32  Mean   :0.5278  Mean   :0.271  Mean   : 2.530  Mean   :0.08747  Mean   :15.87  Mean   :46.47
3rd Qu.: 9.20  3rd Qu.:0.6400  3rd Qu.:0.420  3rd Qu.: 2.600  3rd Qu.:0.09000  3rd Qu.:21.00  3rd Qu.:62.00
Max.   :15.90  Max.   :1.5800  Max.   :1.0000  Max.   :15.500  Max.   :0.61100  Max.   :72.00  Max.   :289.00

density    pH    sulphates    alcohol    quality
Min. :0.9901  Min. :3.2740  Min. :0.3300  Min. : 8.40  Min. : 1.0000
1st Qu.:0.9956  1st Qu.:3.210  1st Qu.:0.5500  1st Qu.: 9.50  1st Qu.:15.000
Median :0.9968  Median :3.310  Median :0.6200  Median :10.20  Median :16.000
Mean   :0.9967  Mean   :3.311  Mean   :0.6581  Mean   :10.42  Mean   :15.636
3rd Qu.:0.9978  3rd Qu.:3.400  3rd Qu.:0.7300  3rd Qu.:11.10  3rd Qu.:16.000
Max.   :1.0037  Max.   :4.010  Max.   :2.0000  Max.   :14.90  Max.   :8.000
>
> # =====
> # 5. Accessing Specific Summaries
> # =====
>
> # Sometimes you want only particular metrics, not the whole summary block.
>
> avg_alcohol <- mean(wine_df$alcohol)
> max_sulphates <- max(wine_df$sulphates)
>
> print(paste("Average Alcohol:", avg_alcohol))
[1] "Average Alcohol: 10.422983144465"
> print(paste("Highest Sulphates:", max_sulphates))
[1] "Highest Sulphates: 2"
> |
```

27°C Sunny

Windows Search

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R 4.5.2 - ~
```

```
>
> # =====
> # 4. IMPROVING summary() WITH FACTORS
> # =====
>
> # 'quality' is often treated as a categorical variable
> # Converting it to factor shows counts per quality score
>
> wine_df$quality <- as.factor(wine_df$quality)
>
> print("---- OUTPUT OF summary() [After Factor Conversion] ----")
[1] "---- OUTPUT OF summary() [After Factor Conversion] ----"
> summary(wine_df)

fixed.acidity volatile.acidity citric.acid residual.sugar chlorides free.sulfur.dioxide total.sulfur.dioxide
Min. : 4.60   Min. :0.1200  Min. :0.0000  Min. :0.900  Min. :0.01200  Min. : 1.00  Min. : 6.00
1st Qu.: 7.10  1st Qu.:0.3900  1st Qu.:0.0900  1st Qu.: 1.900  1st Qu.:0.07000  1st Qu.: 7.00  1st Qu.:22.00
Median : 7.90  Median :0.5200  Median :0.2600  Median : 2.200  Median :0.07900  Median :14.00  Median :38.00
Mean   : 8.32  Mean   :0.5278  Mean   :0.271  Mean   : 2.530  Mean   :0.08747  Mean   :15.87  Mean   :46.47
3rd Qu.: 9.20  3rd Qu.:0.6400  3rd Qu.:0.420  3rd Qu.: 2.600  3rd Qu.:0.09000  3rd Qu.:21.00  3rd Qu.:62.00
Max.   :15.90  Max.   :1.5800  Max.   :1.0000  Max.   :15.500  Max.   :0.61100  Max.   :72.00  Max.   :289.00

density    pH    sulphates    alcohol    quality
Min. :0.9901  Min. :3.2740  Min. :0.3300  Min. : 8.40  Min. : 1.0000
1st Qu.:0.9956  1st Qu.:3.210  1st Qu.:0.5500  1st Qu.: 9.50  1st Qu.:15.000
Median :0.9968  Median :3.310  Median :0.6200  Median :10.20  Median :16.000
Mean   :0.9967  Mean   :3.311  Mean   :0.6581  Mean   :10.42  Mean   :15.636
3rd Qu.:0.9978  3rd Qu.:3.400  3rd Qu.:0.7300  3rd Qu.:11.10  3rd Qu.:16.000
Max.   :1.0037  Max.   :4.010  Max.   :2.0000  Max.   :14.90  Max.   :8.000
>
> # =====
> # 5. Accessing Specific Summaries
> # =====
>
> # Sometimes you want only particular metrics, not the whole summary block.
>
> avg_alcohol <- mean(wine_df$alcohol)
> max_sulphates <- max(wine_df$sulphates)
>
> print(paste("Average Alcohol:", avg_alcohol))
[1] "Average Alcohol: 10.422983144465"
> print(paste("Highest Sulphates:", max_sulphates))
[1] "Highest Sulphates: 2"
> |
```

27°C Sunny

Windows Search

NAME: DEESHA CHAVAN
ROLL NO.: S076