

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

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

```
4 df <- read.csv("Amazon.csv", na.strings = c("", "NA")) %>%
5   mutate(
6     ProductID = row_number(),
7     Price = UnitPrice
8   ) %>%
9   select(ProductID, Category, Price, Discount)
10
11 print("--- Original Data ---")
```

```
> library(dplyr)
> library(tidy)
> df <- read.csv("Amazon.csv", na.strings = c("", "NA")) %>%
+   mutate(
+     ProductID = row_number(),
+     Price = UnitPrice
+   ) %>%
+   select(ProductID, Category, Price, Discount)
> print("--- Original Data ---")
[1] "--- Original Data ---"
> print(head(df))
  ProductID Category Price Discount
1         1    Books 106.59    0.00
2         2 Home & Kitchen 251.37    0.05
3         3    Clothing  35.03    0.10
4         4 Home & Kitchen  33.58    0.15
5         5    Clothing 515.64    0.25
6         6    Books 449.73    0.00

> long_df <- df %>%
+   pivot_longer(
+     cols = c(Price, Discount),
+     names_to = "Metric",
+     values_to = "Value"
+   )
> print("--- Long Format ---")
[1] "--- Long Format ---"
> print(head(long_df, 6))
# A tibble: 6 x 4
  ProductID Category Metric Value
  <int> <chr> <chr> <dbl>
1         1    Books Price 107.
2         1    Books Discount 0
3         2 Home & Kitchen Price 251.
```

Environment

Object	Variables	Observations
category_pivot	7	100000
clean_omit	10	500
clean_replace	10	500
data	27	400
data_feb	3	3
data_jan	3	3
data_new_hires	3	2
df	7	100000
df_calc	23	400
df_clean	4	100000

```
4 df <- read.csv("Amazon.csv", na.strings = c("", "NA")) %>%
5   mutate(
6     ProductID = row_number(),
7     Price = UnitPrice
8   ) %>%
9   select(ProductID, Category, Price, Discount)
10
11 print("--- Original Data ---")
```

```
> wide_df <- long_df %>%
+   pivot_wider(
+     names_from = Metric,
+     values_from = Value
+   )
> print("--- wide Format ---")
[1] "--- wide Format ---"
> print(head(wide_df))
# A tibble: 6 x 4
  ProductID Category Price Discount
  <int> <chr> <dbl> <dbl>
1         1    Books 107.      0
2         2 Home & Kitchen 251.    0.05
3         3    Clothing  35.0    0.1
4         4 Home & Kitchen  33.6    0.15
5         5    Clothing 516.    0.25
6         6    Books 450.      0

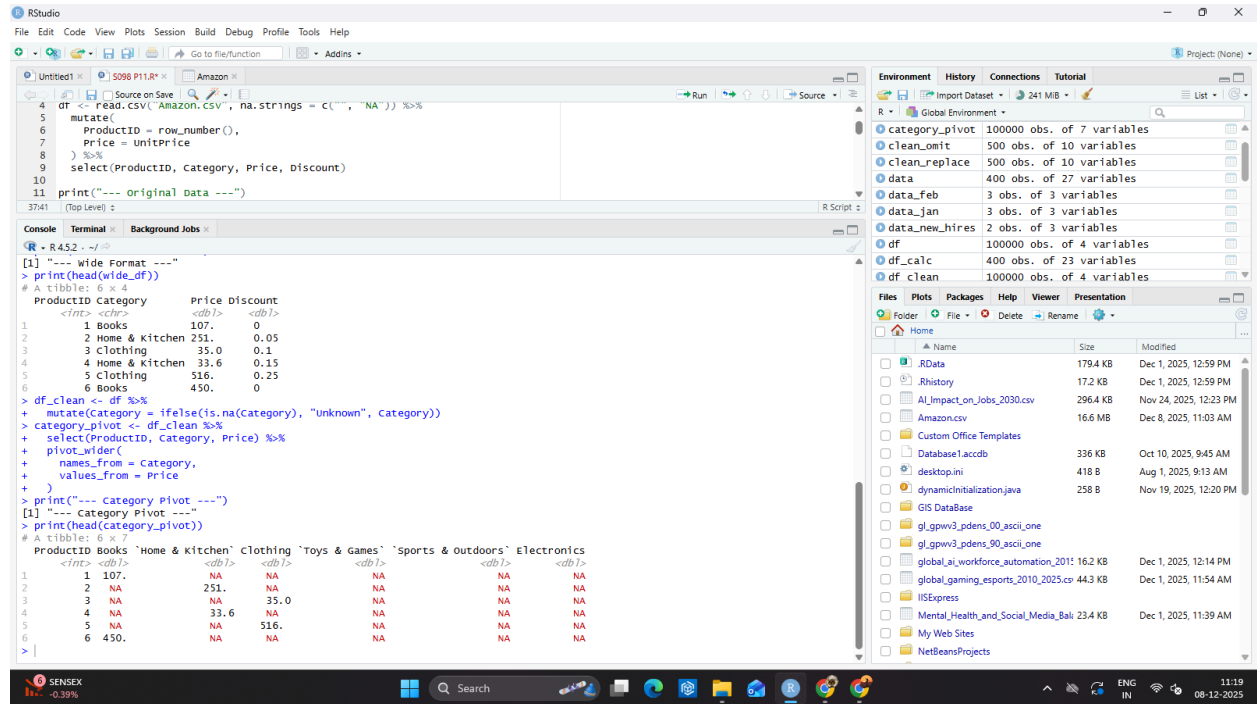
> df_clean <- df %>%
+   mutate(Category = ifelse(is.na(Category), "Unknown", Category))
> category_pivot <- df_clean %>%
+   select(ProductID, Category, Price) %>%
+   pivot_wider(
+     names_from = Category,
+     values_from = Price
+   )
> print("--- Category Pivot ---")
[1] "--- Category Pivot ---"
> print(head(category_pivot))
# A tibble: 6 x 7
```

Environment

Object	Variables	Observations
category_pivot	7	100000
clean_omit	10	500
clean_replace	10	500
data	27	400
data_feb	3	3
data_jan	3	3
data_new_hires	3	2
df	7	100000
df_calc	23	400
df_clean	4	100000

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R



The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains R code for reading a CSV file, mutating data, and creating a pivot table.
- Console:** Shows the execution of the code, including the output of `print(head(wide_df))` and `print(head(category_pivot))`.
- Environment:** Lists the objects created in the R session, such as `category_pivot`, `clean_omit`, `clean_replace`, `data`, `data_feb`, `data_jan`, `data_new_hires`, `df`, `df_calc`, and `df_clean`.
- Files:** Shows the file explorer with various files and folders, including `Amazon.csv` and `Amazon`.

```
df <- read.csv("Amazon.csv", na.strings = c("", "NA")) %>%
mutate(
  ProductID = row_number(),
  Price = UnitPrice
) %>%
select(ProductID, Category, Price, Discount)
print("--- Original data ---")

[1] "--- Wide Format ---"
> print(head(wide_df))
# A tibble: 6 x 4
  ProductID Category      Price Discount
<int> <chr>      <dbl> <dbl>
1      1 Books        107.         0
2      2 Home & Kitchen 251.        0.05
3      3 Clothing       35.0         0.1
4      4 Home & Kitchen 33.6         0.15
5      5 Clothing       516.        0.25
6      6 Books         450.         0

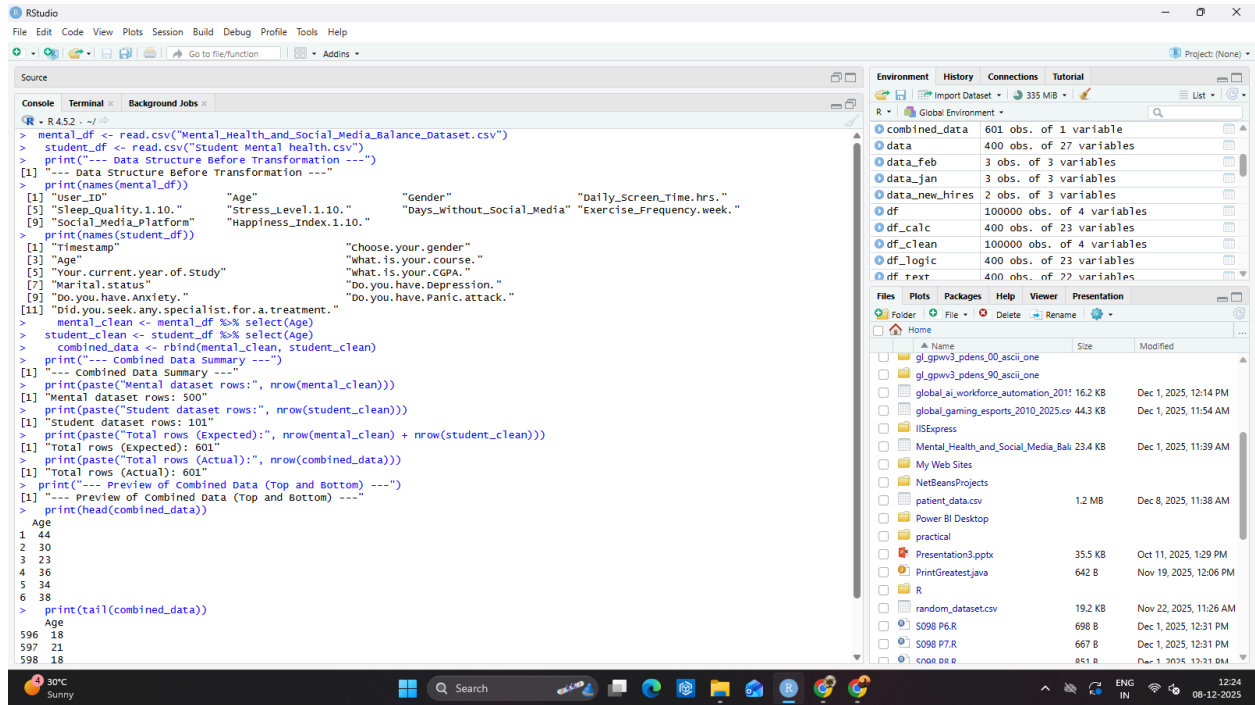
> df_clean <- df %>%
+ mutate(Category = ifelse(is.na(Category), "unknown", Category))
> category_pivot <- df_clean %>%
+ select(ProductID, Category, Price) %>%
+ pivot_wider(
+   names_from = Category,
+   values_from = Price
+ )
> print("--- Category Pivot ---")
[1] "--- Category Pivot ---"
> print(head(category_pivot))
# A tibble: 6 x 7
  ProductID Books `Home & Kitchen` Clothing `Toys & Games` `Sports & outdoors` Electronics
<int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1      1 107. NA NA NA NA NA
2      2 NA 251. NA NA NA NA
3      3 NA NA 35.0 NA NA NA
4      4 NA NA 33.6 NA NA NA
5      5 NA NA 516. NA NA NA
6      6 450. NA NA NA NA NA
```

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

12 Combining datasets vertically (concatenation) using rbind() (R).

Write code to Combining datasets vertically (concatenation) using rbind() in R studio.



```
R - R 4.5.2 - ~/R
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins

Source
Console Terminal Background Jobs

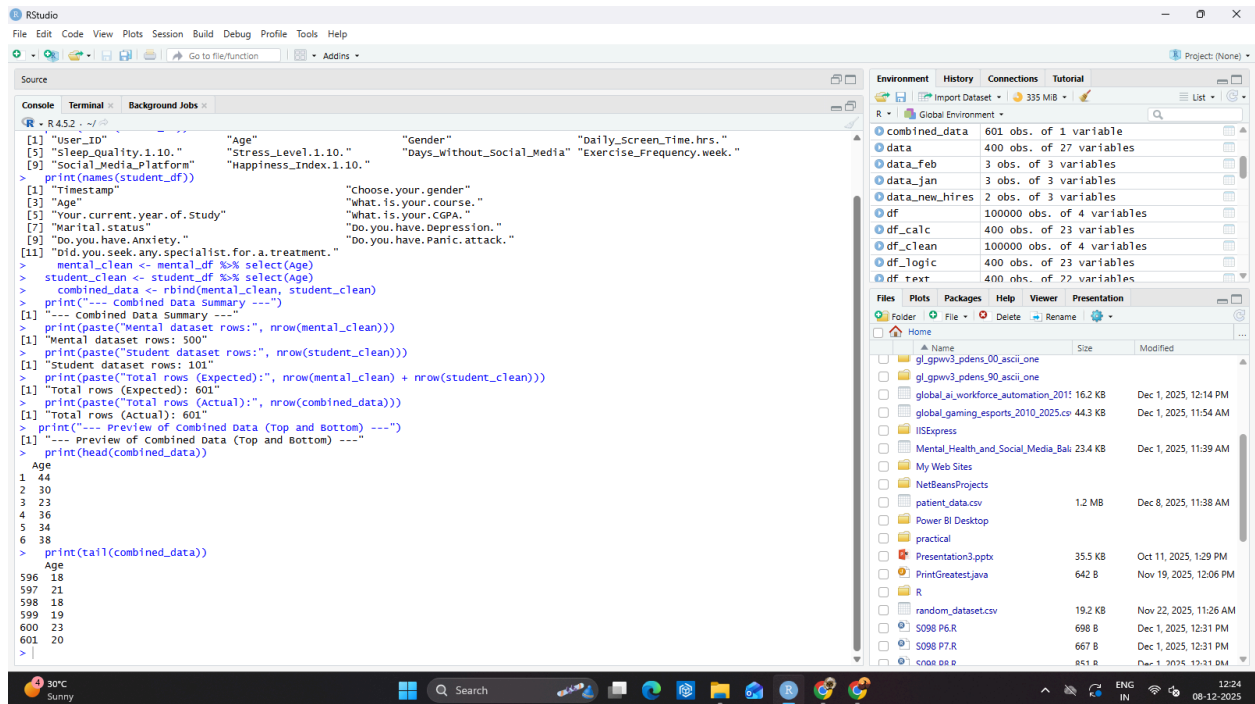
> mental_df <- read.csv("Mental_Health_and_Social_Media_Balance_Dataset.csv")
> student_df <- read.csv("student Mental health.csv")
> print("--- Data Structure Before Transformation ---")
[1] "Data Structure Before Transformation ---"
> print(names(mental_df))
[1] "User_ID" "Age" "Gender" "Daily_Screen_Time.hrs."
[5] "Sleep_Quality.1.10." "Stress_Level.1.10." "Days_Without_Social_Media" "Exercise_Frequency.week."
[9] "Social_Media_Platform" "Happiness_Index.1.10."
> print(names(student_df))
[1] "Timestamp" "Choose.your.gender"
[3] "Age" "What.is.your.course."
[5] "Your.current.year.of.Study" "What.is.your.CGPA."
[7] "Marital.status" "Do.you.have.Depression."
[9] "Do.you.have.Anxiety." "Do.you.have.Panic.attack."
[11] "Did.you.seek.any.specialist.for.a.treatment."
> mental_clean <- mental_df %>% select(Age)
> student_clean <- student_df %>% select(Age)
> combined_data <- rbind(mental_clean, student_clean)
> print("--- Combined Data Summary ---")
[1] "Combined Data Summary ---"
> print(paste("Mental dataset rows:", nrow(mental_clean)))
[1] "Mental dataset rows: 500"
> print(paste("Student dataset rows:", nrow(student_clean)))
[1] "Student dataset rows: 101"
> print(paste("Total rows (Expected):", nrow(mental_clean) + nrow(student_clean)))
[1] "Total rows (Expected): 601"
> print(paste("Total rows (Actual):", nrow(combined_data)))
[1] "Total rows (Actual): 601"
> print("--- Preview of Combined Data (Top and Bottom) ---")
[1] "Preview of Combined Data (Top and Bottom) ---"
> print(head(combined_data))
  Age
1  44
2  30
3  23
4  36
5  34
6  38
> print(tail(combined_data))
  Age
596 18
597 21
598 18
599 19
600 23
601 20
>
```

Environment History Connections Tutorial

Object	Class	Size	Modified
combined_data	data.frame	601 obs. of 1 variable	
data	data.frame	400 obs. of 27 variables	
data_feb	data.frame	3 obs. of 3 variables	
data_jan	data.frame	3 obs. of 3 variables	
data_new_hires	data.frame	2 obs. of 3 variables	
df	data.frame	100000 obs. of 4 variables	
df_calc	data.frame	400 obs. of 23 variables	
df_clean	data.frame	100000 obs. of 4 variables	
df_logic	data.frame	400 obs. of 23 variables	
df_text	data.frame	400 obs. of 22 variables	

Files Plots Packages Help Viewer Presentation

Name	Size	Modified
gl_gpwv3_pdens_00_asci_one		
gl_gpwv3_pdens_90_asci_one		
global_ai_workforce_automation_2011	16.2 KB	Dec 1, 2025, 12:14 PM
global_gaming_exports_2010_2025.csv	44.3 KB	Dec 1, 2025, 11:54 AM
IIExpress		
Mental_Health_and_Social_Media_Bali	23.4 KB	Dec 1, 2025, 11:39 AM
My Web Sites		
NetBeansProjects		
patient_data.csv	1.2 MB	Dec 8, 2025, 11:38 AM
Power BI Desktop		
practical		
Presentation3.pptx	35.5 KB	Oct 11, 2025, 1:29 PM
PrintGreatest.java	642 B	Nov 19, 2025, 12:06 PM
R		
random_dataset.csv	19.2 KB	Nov 22, 2025, 11:26 AM
S098 P6.R	698 B	Dec 1, 2025, 12:31 PM
S098 P7.R	667 B	Dec 1, 2025, 12:31 PM
cn08 on e	851 B	Dec 1, 2025, 12:31 PM



```
R - R 4.5.2 - ~/R
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins

Source
Console Terminal Background Jobs

[1] "User_ID" "Age" "Gender" "Daily_Screen_Time.hrs."
[5] "Sleep_Quality.1.10." "Stress_Level.1.10." "Days_Without_Social_Media" "Exercise_Frequency.week."
[9] "Social_Media_Platform" "Happiness_Index.1.10."
> print(names(student_df))
[1] "Timestamp" "Choose.your.gender"
[3] "Age" "What.is.your.course."
[5] "Your.current.year.of.Study" "What.is.your.CGPA."
[7] "Marital.status" "Do.you.have.Depression."
[9] "Do.you.have.Anxiety." "Do.you.have.Panic.attack."
[11] "Did.you.seek.any.specialist.for.a.treatment."
> mental_clean <- mental_df %>% select(Age)
> student_clean <- student_df %>% select(Age)
> combined_data <- rbind(mental_clean, student_clean)
> print("--- Combined Data Summary ---")
[1] "Combined Data Summary ---"
> print(paste("Mental dataset rows:", nrow(mental_clean)))
[1] "Mental dataset rows: 500"
> print(paste("Student dataset rows:", nrow(student_clean)))
[1] "Student dataset rows: 101"
> print(paste("Total rows (Expected):", nrow(mental_clean) + nrow(student_clean)))
[1] "Total rows (Expected): 601"
> print(paste("Total rows (Actual):", nrow(combined_data)))
[1] "Total rows (Actual): 601"
> print("--- Preview of Combined Data (Top and Bottom) ---")
[1] "Preview of Combined Data (Top and Bottom) ---"
> print(head(combined_data))
  Age
1  44
2  30
3  23
4  36
5  34
6  38
> print(tail(combined_data))
  Age
596 18
597 21
598 18
599 19
600 23
601 20
>
```

Environment History Connections Tutorial

Object	Class	Size	Modified
combined_data	data.frame	601 obs. of 1 variable	
data	data.frame	400 obs. of 27 variables	
data_feb	data.frame	3 obs. of 3 variables	
data_jan	data.frame	3 obs. of 3 variables	
data_new_hires	data.frame	2 obs. of 3 variables	
df	data.frame	100000 obs. of 4 variables	
df_calc	data.frame	400 obs. of 23 variables	
df_clean	data.frame	100000 obs. of 4 variables	
df_logic	data.frame	400 obs. of 23 variables	
df_text	data.frame	400 obs. of 22 variables	

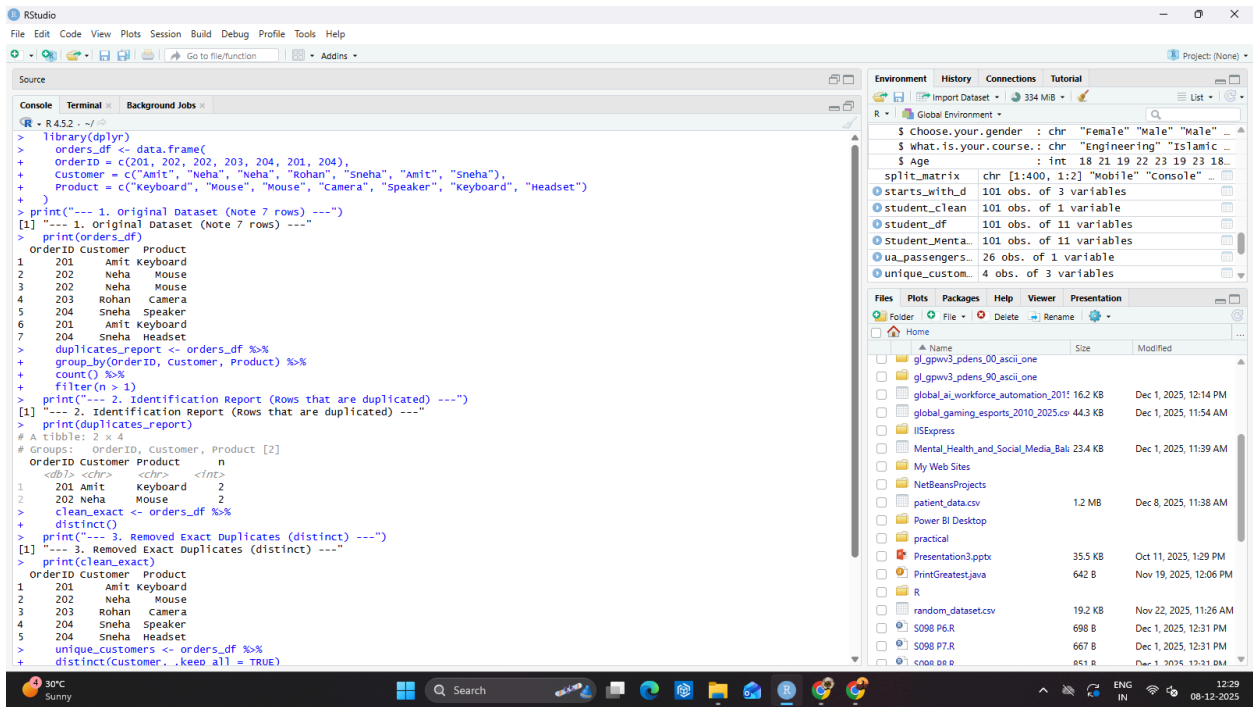
Files Plots Packages Help Viewer Presentation

Name	Size	Modified
gl_gpwv3_pdens_00_asci_one		
gl_gpwv3_pdens_90_asci_one		
global_ai_workforce_automation_2011	16.2 KB	Dec 1, 2025, 12:14 PM
global_gaming_exports_2010_2025.csv	44.3 KB	Dec 1, 2025, 11:54 AM
IIExpress		
Mental_Health_and_Social_Media_Bali	23.4 KB	Dec 1, 2025, 11:39 AM
My Web Sites		
NetBeansProjects		
patient_data.csv	1.2 MB	Dec 8, 2025, 11:38 AM
Power BI Desktop		
practical		
Presentation3.pptx	35.5 KB	Oct 11, 2025, 1:29 PM
PrintGreatest.java	642 B	Nov 19, 2025, 12:06 PM
R		
random_dataset.csv	19.2 KB	Nov 22, 2025, 11:26 AM
S098 P6.R	698 B	Dec 1, 2025, 12:31 PM
S098 P7.R	667 B	Dec 1, 2025, 12:31 PM
cn08 on e	851 B	Dec 1, 2025, 12:31 PM

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 13. Identifying and handling duplicates using distinct() (R studio ).



The screenshot shows the RStudio interface with the following components:

- Console:** Displays the R script execution. The script creates a dataset with 7 rows, identifies duplicates, and removes them using the `distinct()` function. The output shows the original dataset, the duplicates report, and the cleaned dataset.
- Environment:** Shows the global environment with variables like `orders_df`, `duplicates_report`, `clean_exact`, and `unique_customers`.
- Files:** Shows the file explorer with various files and folders.

```
R - R4.5.2 - ~/
> library(dplyr)
> orders_df <- data.frame(
+   orderID = c(201, 202, 202, 203, 204, 201, 204),
+   Customer = c("Amit", "Neha", "Neha", "Rohan", "Sneha", "Amit", "Sneha"),
+   Product = c("Keyboard", "Mouse", "Mouse", "Camera", "Speaker", "Keyboard", "Headset")
+ )
> print("--- 1. Original Dataset (Note 7 rows) ---")
[1] "--- 1. Original Dataset (Note 7 rows) ---"
> print(orders_df)
  orderID Customer Product
1      201    Amit Keyboard
2      202     Neha  Mouse
3      202     Neha  Mouse
4      203     Rohan  Camera
5      204     Sneha Speaker
6      201    Amit Keyboard
7      204     Sneha Headset
> duplicates_report <- orders_df %>%
+   group_by(orderID, Customer, Product) %>%
+   count() %>%
+   filter(n > 1)
> print("--- 2. Identification Report (Rows that are duplicated) ---")
[1] "--- 2. Identification Report (Rows that are duplicated) ---"
# A tibble: 2 x 4
# Groups:   orderID, Customer, Product [2]
  orderID Customer Product     n
  <dbl> <chr>    <chr>    <int>
1     201 Amit      Keyboard     2
2     202 Neha      Mouse      2
> clean_exact <- orders_df %>%
+   distinct()
> print("--- 3. Removed Exact Duplicates (distinct) ---")
[1] "--- 3. Removed Exact Duplicates (distinct) ---"
> print(clean_exact)
  orderID Customer Product
1      201    Amit Keyboard
2      202     Neha  Mouse
3      203     Rohan  Camera
4      204     Sneha Speaker
5      204     Sneha Headset
> unique_customers <- orders_df %>%
+   distinct(Customer, .keep_all = TRUE)
> print("--- 4. Unique Customers only (Partial Duplicates removed) ---")
[1] "--- 4. Unique Customers only (Partial Duplicates removed) ---"
> print(unique_customers)
  orderID Customer Product
1      201    Amit Keyboard
2      202     Neha  Mouse
3      203     Rohan  Camera
4      204     Sneha Speaker
```

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 14. Extracting date components using lubridate:: functions (R).

The screenshot displays the RStudio interface with the following components:

- Source:** Contains R code that uses the `lubridate` package to create a data frame from date strings, process it, and extract date components. The code includes comments and print statements to show the current year, hour, and minute.
- Console:** Shows the execution output, including a data frame with 4 rows and 11 columns, and the current time extraction results.
- Environment:** Lists the loaded packages (`lubridate`, `dplyr`) and the data frame `processed_data` with 100,000 observations and 4 variables.
- Files:** Shows the file explorer with various files and folders, including `global_ai_workforce_automation_2011`, `global_gaming_exports_2010_2025.csv`, and `random_dataset.csv`.

**R Code:**

```
R - R 4.52 - ~/
> library(lubridate)
> library(dplyr)
> dates_df <- data.frame(
+   Event_ID = 1:4,
+   Date_String = c("2022-05-10", "2023-08-21", "2024-11-30", "2025-03-18")
+ )
> processed_data <- dates_df %>%
+   mutate(
+     Actual_Date = ymd(Date_String),
+     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)
+   )
> print("---- Data with Extracted Date Components ----")
[1] "---- Data with Extracted Date Components ----"
[1] print(processed_data)
Event_ID Date_String Actual_Date Year_Num Month_Num Month_Name Day_Num Weekday_Num Weekday_Name Quarter Day_of_Year
1 1 2022-05-10 2022-05-10 2022 5 May 10 3 Tuesday 2 130
2 2 2023-08-21 2023-08-21 2023 8 Aug 21 2 Monday 3 233
3 3 2024-11-30 2024-11-30 2024 11 Nov 30 7 Saturday 4 335
4 4 2025-03-18 2025-03-18 2025 3 Mar 18 3 Tuesday 1 77
> current_time <- now()
> print("---- Current Time Extraction ----")
[1] "---- Current Time Extraction ----"
> print(paste("Current Year:", year(current_time)))
[1] "Current Year: 2025"
> print(paste("Current Hour:", hour(current_time)))
[1] "Current Hour: 12"
> print(paste("Current Minute:", minute(current_time)))
[1] "Current Minute: 32"
>
```

**Console Output:**

```
[1] "---- Data with Extracted Date Components ----"
[1] print(processed_data)
Event_ID Date_String Actual_Date Year_Num Month_Num Month_Name Day_Num Weekday_Num Weekday_Name Quarter Day_of_Year
1 1 2022-05-10 2022-05-10 2022 5 May 10 3 Tuesday 2 130
2 2 2023-08-21 2023-08-21 2023 8 Aug 21 2 Monday 3 233
3 3 2024-11-30 2024-11-30 2024 11 Nov 30 7 Saturday 4 335
4 4 2025-03-18 2025-03-18 2025 3 Mar 18 3 Tuesday 1 77
> current_time <- now()
> print("---- Current Time Extraction ----")
[1] "---- Current Time Extraction ----"
> print(paste("Current Year:", year(current_time)))
[1] "Current Year: 2025"
> print(paste("Current Hour:", hour(current_time)))
[1] "Current Hour: 12"
> print(paste("Current Minute:", minute(current_time)))
[1] "Current Minute: 32"
>
```

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

### 15. Generating basic summaries using str() or summary() (R).

The screenshot shows the RStudio interface with the following code in the console:

```
R - R 4.5.2 ~ />  
> retail_df <- data.frame(  
+   ID = 1:6,  
+   Category = c("Grocery", "Electronics", "Clothing", "Grocery", "Sports", "Electronics"),  
+   Price = c(120.00, 999.00, 450.75, NA, 800.00, 1500.00),  
+   In_Stock = c(TRUE, FALSE, TRUE, TRUE, FALSE, TRUE),  
+   Rating = c(4.0, 4.8, 3.9, 4.3, 4.6, 4.9)  
+ )  
> print("--- Data Loaded ---")  
[1] --- Data Loaded ---  
> print("--- OUTPUT of str() ---")  
[1] --- OUTPUT of str() ---  
> str(retail_df)  
'data.frame': 6 obs. of 5 variables:  
 $ ID      : int  1 2 3 4 5 6  
 $ Category: chr  "Grocery" "Electronics" "Clothing" "Grocery" ...  
 $ Price   : num  120.00 999.00 450.75 NA 800.00 1500.00  
 $ In_Stock: logi  TRUE FALSE TRUE TRUE FALSE TRUE  
 $ Rating  : num  4.0 4.8 3.9 4.3 4.6 4.9  
> print("--- OUTPUT of summary() [Before Factor Conversion] ---")  
[1] --- OUTPUT of summary() [Before Factor Conversion] ---  
> summary(retail_df)  
      ID      Category      Price      In_Stock      Rating  
Min.   :1.00   Length:6   Min.   :120.0   Mode :logical   Min.   :3.900  
1st Qu.:2.25   Class :character 1st Qu.: 450.8   FALSE:2   1st Qu.:4.075  
Median :3.50   Mode :character  Median : 800.0   TRUE :4      Median :4.450  
Mean   :3.50                Mean   : 774.0   Mean   :4.417  
3rd Qu.:4.75                3rd Qu.: 999.0   3rd Qu.:4.750  
Max.   :6.00                Max.   :1500.0   Max.   :4.900  
      NA's :1  
> retail_df$Category <- as.factor(retail_df$Category)  
> print("--- OUTPUT of summary() [After Factor Conversion] ---")  
[1] --- OUTPUT of summary() [After Factor Conversion] ---  
> summary(retail_df)  
      ID      Category      Price      In_Stock      Rating  
Min.   :1.00   Clothing :1   Min.   :120.0   Mode :logical   Min.   :3.900  
1st Qu.:2.25   Electronics:2 1st Qu.: 450.8   FALSE:2   1st Qu.:4.075  
Median :3.50   Grocery   :2   Median : 800.0   TRUE :4      Median :4.450  
Mean   :3.50   Sports    :1   Mean   : 774.0   Mean   :4.417  
3rd Qu.:4.75                3rd Qu.: 999.0   3rd Qu.:4.750  
Max.   :6.00                Max.   :1500.0   Max.   :4.900  
      NA's :1  
> avg_rating <- mean(retail_df$Rating)  
> max_price <- max(retail_df$Price, na.rm = TRUE)
```

The Environment pane on the right shows the objects created: `ua_passengers...` (26 obs. of 1 variable), `unique_custom...` (4 obs. of 3 variables), `unique_patien...` (20000 obs. of 10 variables), and `wide_df` (100000 obs. of 4 variables). The Files pane shows the project structure.

The screenshot shows the RStudio interface with the following code in the console:

```
R - R 4.5.2 ~ />  
> print("--- Data Loaded ---")  
[1] --- Data Loaded ---  
> print("--- OUTPUT of str() ---")  
[1] --- OUTPUT of str() ---  
> str(retail_df)  
'data.frame': 6 obs. of 5 variables:  
 $ ID      : int  1 2 3 4 5 6  
 $ Category: chr  "Grocery" "Electronics" "Clothing" "Grocery" ...  
 $ Price   : num  120.00 999.00 450.75 NA 800.00 1500.00  
 $ In_Stock: logi  TRUE FALSE TRUE TRUE FALSE TRUE  
 $ Rating  : num  4.0 4.8 3.9 4.3 4.6 4.9  
> print("--- OUTPUT of summary() [Before Factor Conversion] ---")  
[1] --- OUTPUT of summary() [Before Factor Conversion] ---  
> summary(retail_df)  
      ID      Category      Price      In_Stock      Rating  
Min.   :1.00   Length:6   Min.   :120.0   Mode :logical   Min.   :3.900  
1st Qu.:2.25   Electronics:2 1st Qu.: 450.8   FALSE:2   1st Qu.:4.075  
Median :3.50   Grocery   :2   Median : 800.0   TRUE :4      Median :4.450  
Mean   :3.50   Sports    :1   Mean   : 774.0   Mean   :4.417  
3rd Qu.:4.75                3rd Qu.: 999.0   3rd Qu.:4.750  
Max.   :6.00                Max.   :1500.0   Max.   :4.900  
      NA's :1  
> retail_df$Category <- as.factor(retail_df$Category)  
> print("--- OUTPUT of summary() [After Factor Conversion] ---")  
[1] --- OUTPUT of summary() [After Factor Conversion] ---  
> summary(retail_df)  
      ID      Category      Price      In_Stock      Rating  
Min.   :1.00   Clothing :1   Min.   :120.0   Mode :logical   Min.   :3.900  
1st Qu.:2.25   Electronics:2 1st Qu.: 450.8   FALSE:2   1st Qu.:4.075  
Median :3.50   Grocery   :2   Median : 800.0   TRUE :4      Median :4.450  
Mean   :3.50   Sports    :1   Mean   : 774.0   Mean   :4.417  
3rd Qu.:4.75                3rd Qu.: 999.0   3rd Qu.:4.750  
Max.   :6.00                Max.   :1500.0   Max.   :4.900  
      NA's :1  
> avg_rating <- mean(retail_df$Rating)  
> max_price <- max(retail_df$Price, na.rm = TRUE)  
> print(paste("Average Rating:", avg_rating))  
[1] "Average Rating: 4.41666666666667"  
> print(paste("Highest Price:", max_price))  
[1] "Highest Price: 1500"
```

The Environment pane on the right shows the objects created: `ua_passengers...` (26 obs. of 1 variable), `unique_custom...` (4 obs. of 3 variables), `unique_patien...` (20000 obs. of 10 variables), and `wide_df` (100000 obs. of 4 variables). The Files pane shows the project structure.