

Prac11 output-

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R4.1.2
6 walmart
> library(dplyr)
> library(tidyrr)
>
> df <- read_csv("Supplement_Sales_Weekly_Expanded.csv", na.strings = c("", "NA")) %>%
+   mutate(ProductID = row_number()) %>%
+   select(ProductID, Category, Price, Discount)
>
> print("--- 1. Original wide Data ---")
[1] "--- 1. Original wide Data ---"
> print(head(df))
  ProductID Category Price Discount
1         1   Protein 31.98    0.03
2         2   Vitamin 42.51    0.04
3         3    Omega 12.91    0.25
4         4   vitamin 16.07    0.08
5         5 Performance 35.47    0.25
6         6 Amino Acid 41.19    0.13
>
> long_df <- df %>%
+   pivot_longer(
+     cols = c(Price, Discount),
+     names_to = "Metric",
+     values_to = "value"
+   )
> print("--- 2. Long Format (pivot_longer) ---")
[1] "--- 2. Long Format (pivot_longer) ---"
> print(head(long_df, 6))
# A tibble: 6 x 4
  ProductID Category Metric    value
  <int>   <chr>   <chr>   <dbl>
1         1 Protein Price    32.0
2         1 Protein Discount  0.03
3         2 Vitamin Price    42.5
4         2 Vitamin Discount  0.04
5         3 Omega Price    12.9
6         3 Omega Discount  0.25
>
> wide_df <- long_df %>%
+   pivot_wider(
+     names_from = Metric,
+     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 4
  ProductID Category Price Discount
  <int>   <chr>   <dbl>   <dbl>
1         1 Protein 32.0    0.03
2         2 Vitamin 42.5    0.04
3         3 Omega 12.9    0.25
4         4 Vitamin 16.1    0.08
5         5 Performance 35.5    0.25
6         6 Amino Acid 41.2    0.13
>
> 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
+   )
```

Environment History Connections Tutorial

Global Environment

- category_pivot 4384 obs. of 11 variables
- df 4384 obs. of 4 variables
- df_clean 4384 obs. of 4 variables
- long_df 8768 obs. of 4 variables
- Supplement_Sal... (1).csv 4384 obs. of 10 variables
- test 4384 obs. of 10 variables
- wide_df 4384 obs. of 4 variables

Files Plots Packages Help Viewer Presentation

Home > practical no-6,7,8,9,10

Name	Size	Modified
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

29°C Sunny

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R4.1.2
1 <int> <chr> <chr> <dbl>
2 1 Protein Price 32.0
3 1 Protein Discount 0.03
4 2 Vitamin Price 42.5
5 2 Vitamin Discount 0.04
6 3 Omega Price 12.9
7 3 Omega Discount 0.25
>
> wide_df <- long_df %>%
+   pivot_wider(
+     names_from = Metric,
+     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 4
  ProductID Category Price Discount
  <int>   <chr>   <dbl>   <dbl>
1         1 Protein 32.0    0.03
2         2 Vitamin 42.5    0.04
3         3 Omega 12.9    0.25
4         4 Vitamin 16.1    0.08
5         5 Performance 35.5    0.25
6         6 Amino Acid 41.2    0.13
>
> 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
+   )
```

Environment History Connections Tutorial

Global Environment

- category_pivot 4384 obs. of 11 variables
- df 4384 obs. of 4 variables
- df_clean 4384 obs. of 4 variables
- long_df 8768 obs. of 4 variables
- Supplement_Sal... (1).csv 4384 obs. of 10 variables
- test 4384 obs. of 10 variables
- wide_df 4384 obs. of 4 variables

Files Plots Packages Help Viewer Presentation

Home > practical no-6,7,8,9,10

Name	Size	Modified
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

29°C Sunny

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Source

```
# A tibble: 6 x 4
  ProductID Category Price Discount
  <int> <chr> <dbl> <dbl>
1 1 Protein 32.0 0.03
2 2 Vitamin 42.5 0.04
3 3 Omega 12.9 0.25
4 4 Vitamin 16.1 0.08
5 5 Performance 35.5 0.25
6 6 Amino Acid 41.2 0.13

> 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("--- 4. Category Pivot (Spreading Categories) ---")
[1] "--- 4. Category Pivot (Spreading Categories) ---"
> print(head(category_pivot))
# A tibble: 6 x 11
  ProductID Protein Vitamin Omega Performance `Amino Acid` Mineral Herbal `Sleep Aid` `Fat Burner`
  <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 1 32.0 NA NA NA NA NA NA NA NA
2 2 NA 42.5 NA NA NA NA NA NA NA
3 3 NA NA 12.9 NA NA NA NA NA NA
4 4 NA 16.1 NA NA NA NA NA NA NA
5 5 NA NA NA 35.5 NA NA NA NA NA
6 6 NA NA NA NA 41.2 NA NA NA NA
# 1 more variable: Hydration <dbl>
```

Environment History Connections Tutorial

R - Global Environment

- category_pivot 4384 obs. of 11 variables
- df 4384 obs. of 4 variables
- df_clean 4384 obs. of 4 variables
- long_df 8768 obs. of 4 variables
- Supplement_Sal... (1).csv 4384 obs. of 10 variables
- test 4384 obs. of 10 variables
- wide_df 4384 obs. of 4 variables

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home practical no-6,7,8,9,10

Name	Size	Modified
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

29°C Sunny

Search

ENG IN 11:43 08-12-2025

Prac12 output-

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Source

```
> combined_data <- rbind(iris_clean, flower_clean)
>
> head(iris_clean)
  FlowerName Height
1 setosa 5.1
2 setosa 4.9
3 setosa 4.7
4 setosa 4.6
5 setosa 5.0
6 setosa 5.4
> head(flower_clean)
  FlowerName Height
1 Artificial NA
2 Artificial NA
3 Artificial NA
4 Artificial NA
5 Artificial NA
6 Artificial NA
> head(combined_data)
  FlowerName Height
1 setosa 5.1
2 setosa 4.9
3 setosa 4.7
4 setosa 4.6
5 setosa 5.0
6 setosa 5.4
>
> nrow(iris_clean)
[1] 150
> nrow(flower_clean)
[1] 358
> nrow(combined_data)
[1] 508
>
```

Environment History Connections Tutorial

R - Global Environment

- drake_df 290 obs. of 5 variables
- flower_clean 358 obs. of 2 variables
- flower_df 358 obs. of 9 variables
- iris 150 obs. of 5 variables
- IRIS 150 obs. of 5 variables
- iris_clean 150 obs. of 2 variables
- iris_df 150 obs. of 5 variables

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home practical no-6,7,8,9,10

Name	Size	Modified
Artificial Flower.csv	22.4 KB	Dec 8, 2025, 12:09 PM
IRIS.csv	4.5 KB	Dec 8, 2025, 12:06 PM
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

SENSEX -0.37%

Search

ENG IN 12:18 08-12-2025

Prac13 output-

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Source

```
library(dplyr)

>
> orders_df <- data.frame(
+   orderID = c(101, 102, 102, 103, 104, 101, 104),
+   Customer = c("jelly", "gum", "chocolate", "milk", "strawberry", "blueberry", "cherry"),
+   Product = c("Laptop", "Phone", "Phone", "Tablet", "Monitor", "Laptop", "Mouse")
+ )
>
> print("Original Dataset")
[1] "Original Dataset"
> print(orders_df)
  orderID Customer Product
1     101    jelly  Laptop
2     102     gum   Phone
3     102 chocolate  Phone
4     103     milk Tablet
5     104 strawberry Monitor
6     101 blueberry Laptop
7     104     cherry Mouse
>
> duplicates_report <- orders_df %>%
+   group_by(orderID, Customer, Product) %>%
+   count() %>%
+   filter(n > 1)
>
> print("Duplicates Report")
[1] "Duplicates Report"
> print(duplicates_report)
# A tibble: 0 x 4
# Groups:   orderID, Customer, Product [0]
# i 4 variables: orderID <dbl>, Customer <chr>, Product <chr>, n <int>
>
> clean_exact <- orders_df %>%
+   distinct()
>
```

Environment History Connections Tutorial

Global Environment

orders_df	7 obs. of 3 variables	
spotify_clean	100 obs. of 2 variables	
spotify_df	100 obs. of 16 variables	
Supplement_Sal- (1).csv	4384 obs. of 10 variables	
test	4384 obs. of 10 variables	
top.spotify.tr...	100 obs. of 16 variables	
unique_custome...	7 obs. of 3 variables	

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home practical no-6,7,8,9,10

Name	Size	Modified
..		
Artificial Flower.csv	22.4 KB	Dec 8, 2025, 12:09 PM
IRIS.csv	4.5 KB	Dec 8, 2025, 12:06 PM
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

30°C Sunny

Search

ENG IN 12:28 08-12-2025

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Source

```
# A tibble: 0 x 4
# Groups:   orderID, Customer, Product [0]
# i 4 variables: orderID <dbl>, Customer <chr>, Product <chr>, n <int>
>
> clean_exact <- orders_df %>%
+   distinct()
>
> print("Removed Exact Duplicates")
[1] "Removed Exact Duplicates"
> print(clean_exact)
  orderID Customer Product
1     101    jelly  Laptop
2     102     gum   Phone
3     102 chocolate  Phone
4     103     milk Tablet
5     104 strawberry Monitor
6     101 blueberry Laptop
7     104     cherry Mouse
>
> unique_customers <- orders_df %>%
+   distinct(Customer, .keep_all = TRUE)
>
> print("Unique Customers only")
[1] "Unique Customers only"
> print(unique_customers)
  orderID Customer Product
1     101    jelly  Laptop
2     102     gum   Phone
3     102 chocolate  Phone
4     103     milk Tablet
5     104 strawberry Monitor
6     101 blueberry Laptop
7     104     cherry Mouse
>
```

Environment History Connections Tutorial

Global Environment

orders_df	7 obs. of 3 variables	
spotify_clean	100 obs. of 2 variables	
spotify_df	100 obs. of 16 variables	
Supplement_Sal- (1).csv	4384 obs. of 10 variables	
test	4384 obs. of 10 variables	
top.spotify.tr...	100 obs. of 16 variables	
unique_custome...	7 obs. of 3 variables	

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home practical no-6,7,8,9,10

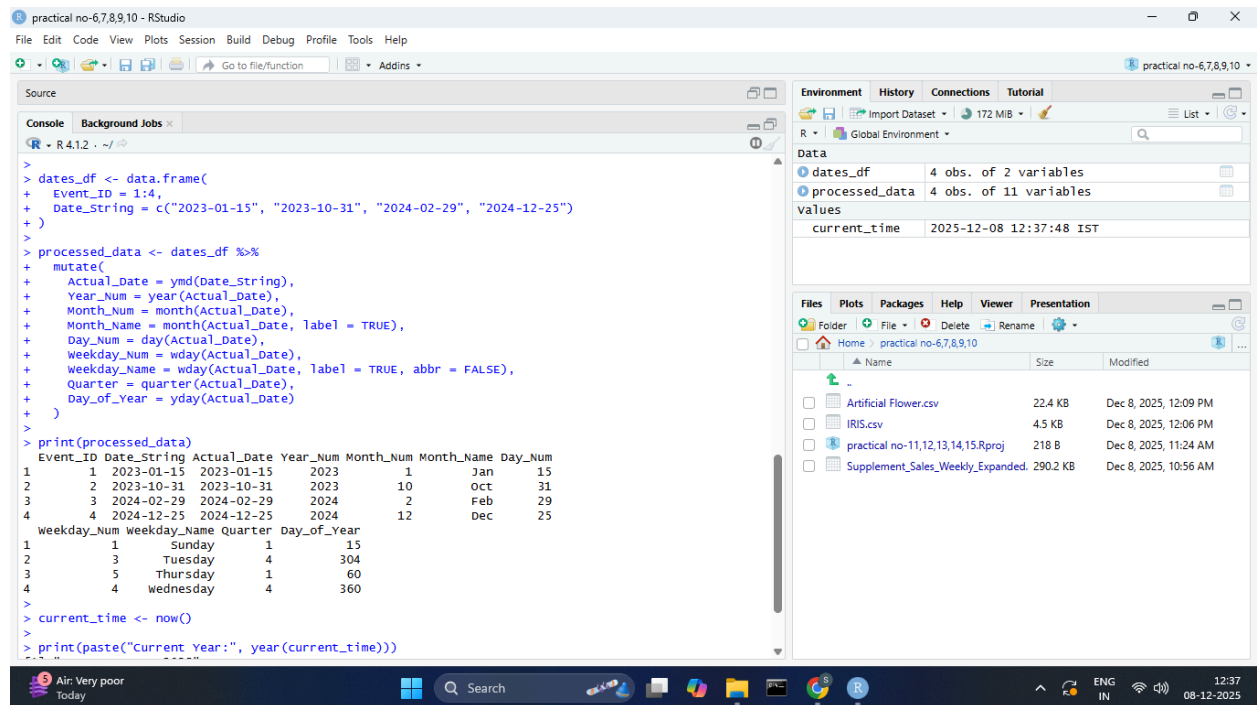
Name	Size	Modified
..		
Artificial Flower.csv	22.4 KB	Dec 8, 2025, 12:09 PM
IRIS.csv	4.5 KB	Dec 8, 2025, 12:06 PM
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

30°C Sunny

Search

ENG IN 12:29 08-12-2025

Prac14 output-



practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
> dates_df <- data.frame(
+   Event_ID = 1:4,
+   date_String = c("2023-01-15", "2023-10-31", "2024-02-29", "2024-12-25")
+ )
> 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(processed_data)
  Event_ID Date_String Actual_Date Year_Num Month_Num Month_Name Day_Num
1        1 2023-01-15 2023-01-15    2023         1      Jan     15
2        2 2023-10-31 2023-10-31    2023        10      Oct     31
3        3 2024-02-29 2024-02-29    2024         2      Feb     29
4        4 2024-12-25 2024-12-25    2024        12      Dec     25
  weekday_Num weekday_Name Quarter Day_of_Year
1           1      Sunday      1          15
2           3    Tuesday      4         304
3           5   Thursday      1          60
4           4   Wednesday      4         360
>
> current_time <- now()
> print(paste("Current Year:", year(current_time)))
```

Environment History Connections Tutorial

Data

Object	Size	Variables
dates_df	4 obs. of 2 variables	
processed_data	4 obs. of 11 variables	

Values

current_time
2025-12-08 12:37:48 IST

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

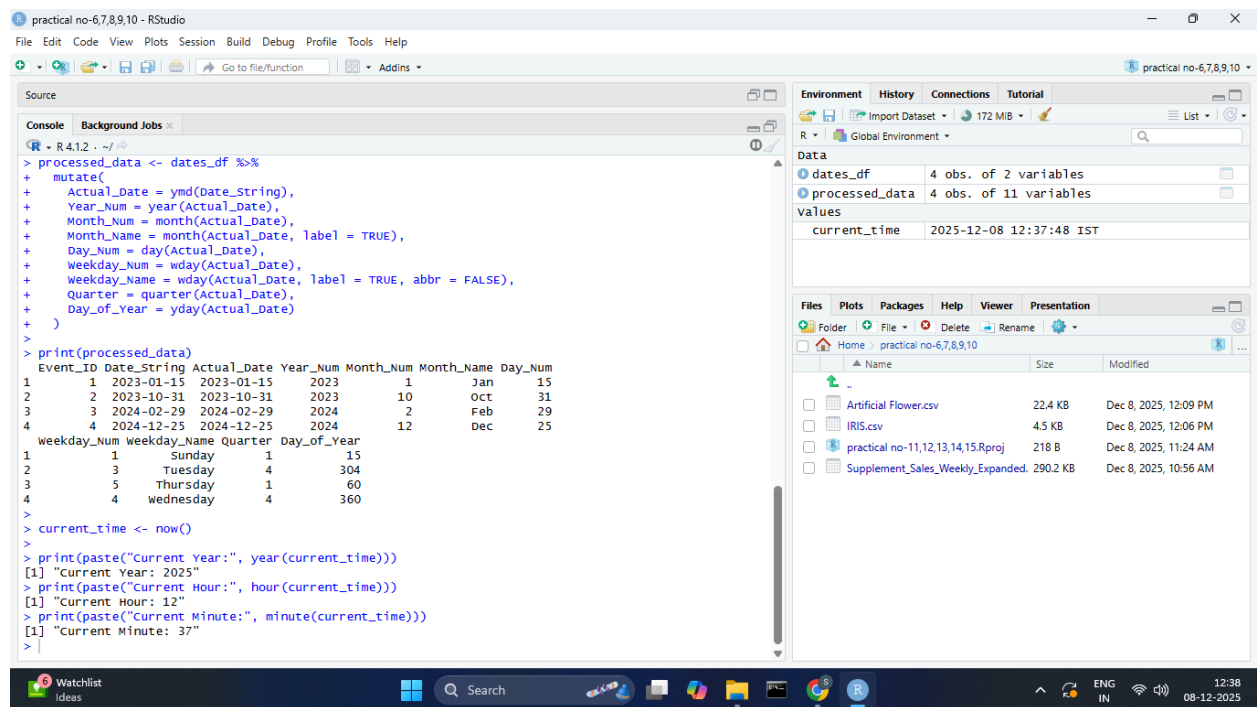
Home practical no-6,7,8,9,10

Name	Size	Modified
Artificial Flower.csv	22.4 KB	Dec 8, 2025, 12:09 PM
IRIS.csv	4.5 KB	Dec 8, 2025, 12:06 PM
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded	290.2 KB	Dec 8, 2025, 10:56 AM

Air: Very poor
Today

Search

ENG IN 12:37 08-12-2025



practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
> 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(processed_data)
  Event_ID Date_String Actual_Date Year_Num Month_Num Month_Name Day_Num
1        1 2023-01-15 2023-01-15    2023         1      Jan     15
2        2 2023-10-31 2023-10-31    2023        10      Oct     31
3        3 2024-02-29 2024-02-29    2024         2      Feb     29
4        4 2024-12-25 2024-12-25    2024        12      Dec     25
  weekday_Num weekday_Name Quarter Day_of_Year
1           1      Sunday      1          15
2           3    Tuesday      4         304
3           5   Thursday      1          60
4           4   Wednesday      4         360
>
> current_time <- now()
> 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: 37"
```

Environment History Connections Tutorial

Data

Object	Size	Variables
dates_df	4 obs. of 2 variables	
processed_data	4 obs. of 11 variables	

Values

current_time
2025-12-08 12:37:48 IST

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home practical no-6,7,8,9,10

Name	Size	Modified
Artificial Flower.csv	22.4 KB	Dec 8, 2025, 12:09 PM
IRIS.csv	4.5 KB	Dec 8, 2025, 12:06 PM
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded	290.2 KB	Dec 8, 2025, 10:56 AM

Watchlist
Ideas

Search

ENG IN 12:38 08-12-2025

Prac15 output-

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Source

```
R - R4.1.2 - ~/ /  
> retail_df <- data.frame(  
+ ID = 1:6,  
+ Category = c("appliances", "Home", "Electronics", "gadgets", "lawn", "Clothing"),  
+ Price = c(500.50, 45.00, 900.00, NA, 300.00, 25.00),  
+ In_Stock = c(TRUE, TRUE, FALSE, TRUE, FALSE, TRUE),  
+ Rating = c(4.5, 3.8, 4.9, 4.0, 3.5, 4.2)  
+ )  
> print("Data Loaded")  
[1] "Data Loaded"  
> str(retail_df)  
'data.frame': 6 obs. of 5 variables:  
 $ ID      : int  1 2 3 4 5 6  
 $ Category: chr  "appliances" "Home" "Electronics" "gadgets" ...  
 $ Price   : num  500 45 900 NA 300 ...  
 $ In_Stock: logi  TRUE TRUE FALSE TRUE FALSE TRUE  
 $ Rating  : num  4.5 3.8 4.9 4.0 3.5 4.2  
>  
> print("Summary Before Factor Conversion")  
[1] "Summary Before Factor conversion"  
> summary(retail_df)  
      ID      Category      Price      In_Stock  
Min.   :1.00 Length:6      Min.   : 25.0 Mode :logical  
1st Qu.:2.25 Class :character 1st Qu.: 45.0 FALSE:2  
Median :3.50 Mode  :character Median :300.0 TRUE :4  
Mean   :3.50              Mean  :354.1  
3rd Qu.:4.75              3rd Qu.:500.5  
Max.   :6.00              Max.   :900.0  
              NA's   :1  
  
      Rating  
Min.   :3.500  
1st Qu.:3.850  
Median :4.100  
Mean   :4.150  
3rd Qu.:4.425  
Max.   :4.900
```

Environment History Connections Tutorial

R - Global Environment

Data

- dates_df 4 obs. of 2 variables
- processed_data 4 obs. of 11 variables
- retail_df 6 obs. of 5 variables

values

- avg_rating 4.15
- current_time 2025-12-08 12:37:48 IST
- max_price 900

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home practical no-6,7,8,9,10

Name	Size	Modified
Artificial Flower.csv	22.4 KB	Dec 8, 2025, 12:09 PM
IRIS.csv	4.5 KB	Dec 8, 2025, 12:06 PM
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

30°C Sunny

Search

ENG IN 12:43 08-12-2025

practical no-6,7,8,9,10 - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Source

```
R - R4.1.2 - ~/ /  
3rd Qu.:4.75      3rd Qu.:500.5  
Max.   :6.00      Max.   :900.0  
              NA's   :1  
  
      Rating  
Min.   :3.500  
1st Qu.:3.850  
Median :4.100  
Mean   :4.150  
3rd Qu.:4.425  
Max.   :4.900  
  
>  
> retail_df$Category <- as.factor(retail_df$Category)  
>  
> print("Summary After Factor conversion")  
[1] "Summary After Factor conversion"  
> summary(retail_df)  
      ID      Category      Price      In_Stock      Rating  
Min.   :1.00 appliances :1      Min.   : 25.0 Mode :logical Min.   :3.500  
1st Qu.:2.25 Clothing  :1      1st Qu.: 45.0 FALSE:2      1st Qu.:3.850  
Median :3.50 Electronics:1      Median :300.0 TRUE :4      Median :4.100  
Mean   :3.50 gadgets  :1      Mean  :354.1              Mean  :4.150  
3rd Qu.:4.75 Home    :1      3rd Qu.:500.5           3rd Qu.:4.425  
Max.   :6.00 lawn    :1      Max.   :900.0           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: 4.15"  
> print(paste("Highest Price:", max_price))  
[1] "Highest Price: 900"  
>
```

Environment History Connections Tutorial

R - Global Environment

Data

- dates_df 4 obs. of 2 variables
- processed_data 4 obs. of 11 variables
- retail_df 6 obs. of 5 variables

values

- avg_rating 4.15
- current_time 2025-12-08 12:37:48 IST
- max_price 900

Files Plots Packages Help Viewer Presentation

Folder File Delete Rename

Home practical no-6,7,8,9,10

Name	Size	Modified
Artificial Flower.csv	22.4 KB	Dec 8, 2025, 12:09 PM
IRIS.csv	4.5 KB	Dec 8, 2025, 12:06 PM
practical no-11,12,13,14,15.Rproj	218 B	Dec 8, 2025, 11:24 AM
Supplement_Sales_Weekly_Expanded.	290.2 KB	Dec 8, 2025, 10:56 AM

30°C Sunny

Search

ENG IN 12:44 08-12-2025