

# **MVLU COLLEGE**

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

## 10 Creating graphical reports using `ggplot2` (R).

## **OUTPUT:-**

The screenshot shows the RStudio interface with the following details:

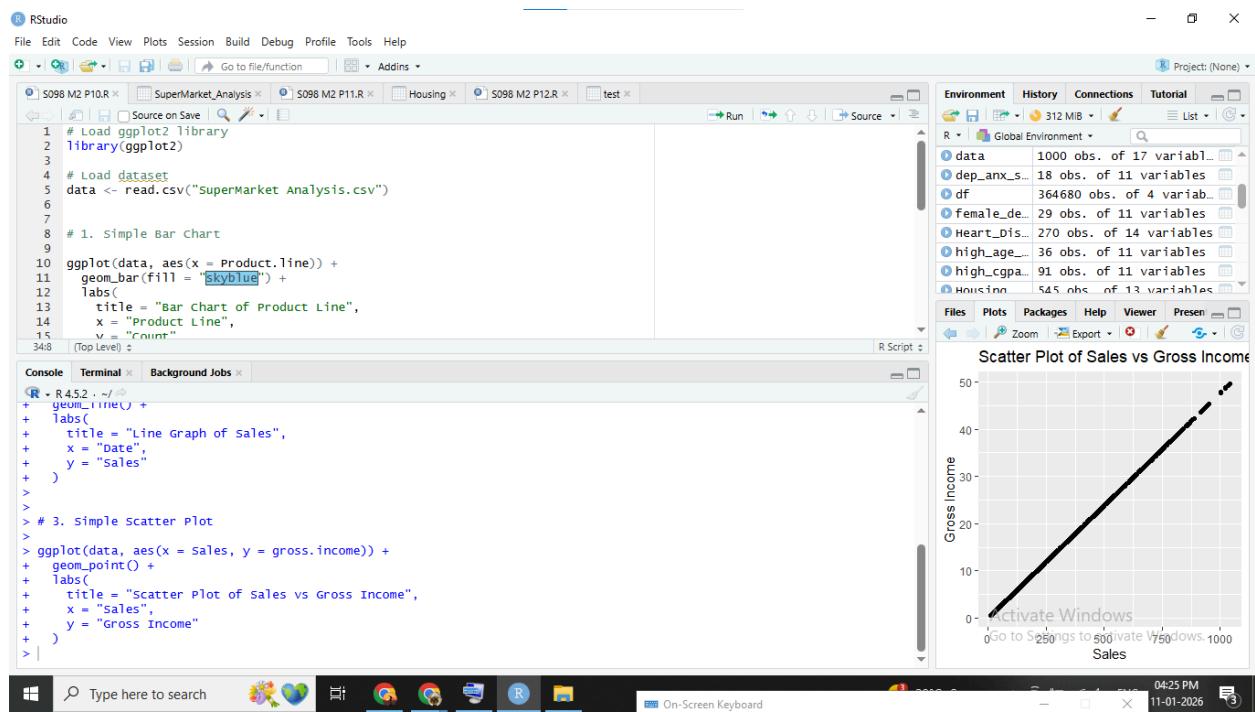
- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project Bar:** S098 M2 P10.R, SuperMarket\_Analysis, S098 M2 P11.R, Housing, S098 M2 P12.R, test.
- Code Editor:** An R script containing code to load ggplot2, read a dataset, and create a bar chart. The code includes comments and annotations.
- Console:** Displays the R session history, showing the same commands as the script.
- Environment:** Shows the global environment with various objects and their dimensions.
- Plots:** A bar chart titled "Bar Chart of Product Line" showing the count of observations for different product lines.
- Bottom Bar:** Taskbar with icons for File Explorer, OneDrive, Edge, Task View, and Start.

The screenshot shows the RStudio interface with the following components:

- Header:** RStudio, File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Go to file/function, Addins.
- Left Panel:** A code editor window titled "S098 M2 P10.R" containing R code for creating a bar chart. The code includes `library(ggplot2)`, `data <- read.csv("SuperMarket Analysis.csv")`, and a ggplot call with `geom\_bar(fill = "skyblue")`.
- Right Panel:** An "Environment" tab showing global variables like `data` (1000 obs.), `dep\_anx...` (18 obs.), and `df` (364680 obs.). Below it is a "Plots" tab displaying a "Line Graph of Sales" with Sales on the y-axis (0 to 1000) and Date on the x-axis.
- Bottom:** A search bar, taskbar icons (Windows, Internet Explorer, Google Chrome, File Explorer, Task View, On-Screen Keyboard), and system status (04:25 PM, 11-01-2026).

# MVLU COLLEGE

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



# MVLU COLLEGE

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

### 11 Generating histograms and box plots using ggplot2 (R).

#### OUTPUT:-

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source on Save Go to file/function Addins

```

# Load ggplot2
library(ggplot2)
# Load dataset
data <- read.csv("Housing.csv")
# View data
head(data)
  price area bedrooms bathrooms stories mainroad guestroom basement hotwaterheating airconditioning parking
1 13300000 7420 4 2 3 yes no no no yes 2
2 12250000 8960 4 4 4 yes no no no yes 3
3 12250000 9960 3 2 2 yes no yes no no 2
4 12215000 7500 4 2 2 yes no yes no yes 3
5 11410000 7420 4 1 2 yes yes yes no yes 2
6 10850000 7500 3 3 1 yes no yes no yes 2
prefarea furnishingstatus
1 yes furnished
2 no furnished
3 yes semi-furnished
4 yes furnished
5 no furnished
6 yes semi-furnished
# 1. Histogram

```

Console Terminal Background Jobs

R + R 4.5.2 . ~/ ~

Histogram of House Prices

Frequency

Price

Go to Settings to activate Windows.

04:28 PM 11-01-2026

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source on Save Go to file/function Addins

```

# Load ggplot2
library(ggplot2)
# Load dataset
data <- read.csv("Housing.csv")
# View data
# 1. Histogram
ggplot(data, aes(x = price)) +
  geom_histogram(fill = "skyblue", color = "black") +
  labs(
    title = "Histogram of House Prices",
    x = "Price",
    y = "Frequency"
  )
`stat_bin()` using `bins = 30`. Pick better value `binwidth`.
# 2. Box Plot
ggplot(data, aes(y = price)) +
  geom_boxplot(fill = "orange") +
  labs(
    title = "Box Plot of House Prices",
    y = "Price"
  )

```

Environment History Connections Tutorial

R + Global Environment

data 545 obs. of 13 variables

dep\_anx\_subset 18 obs. of 11 variables

df 364680 obs. of 4 variables

female\_dep\_filter 29 obs. of 11 variables

Heart\_Disease\_Predict... 270 obs. of 14 variables

high\_age\_subset 36 obs. of 11 variables

high\_cgpa\_filter 91 obs. of 11 variables

Housing 545 obs. of 13 variables

Files Plots Packages Help Viewer Presentation

Box Plot of House Prices

Price

Activate Windows  
Go to Settings to activate Windows.

-0.4 -0.2 0.0 0.2 0.4

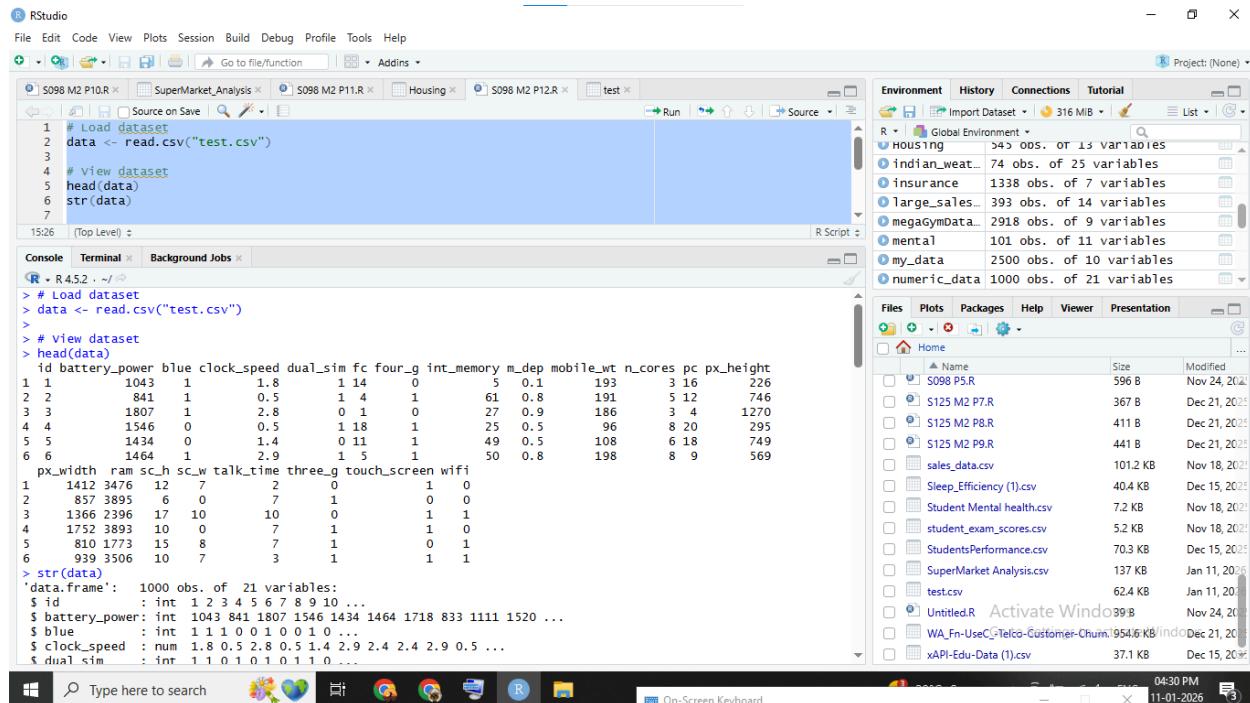
04:28 PM 11-01-2026

# MVLU COLLEGE

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

### 12 Generating correlation matrices using cor() (R).

#### OUTPUT:-



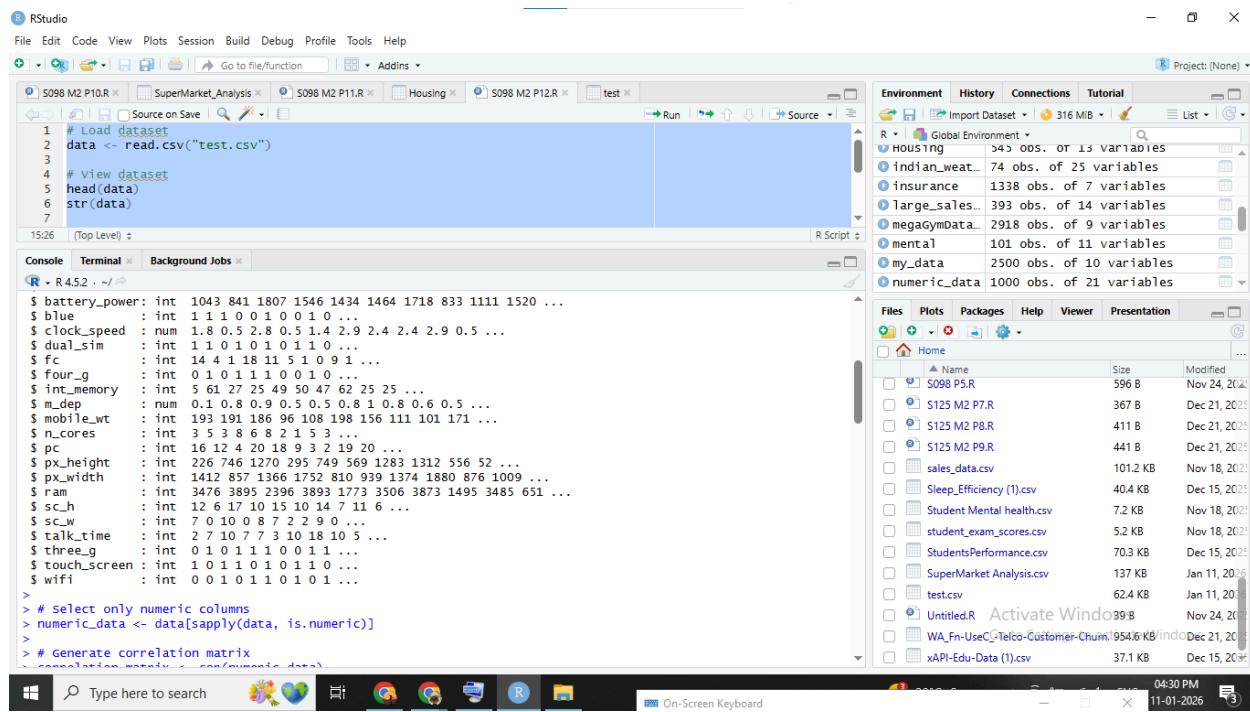
The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project (None) is selected.
- Code Editor:** An R script titled "test.R" is open, containing the following code:

```

1 # Load dataset
2 data <- read.csv("test.csv")
3
4 # View dataset
5 head(data)
6 str(data)

```
- Console:** The output of the R code is displayed, showing the first few rows of the dataset "test.csv". The columns include battery\_power, blue, clock\_speed, dual\_sim, fc, four\_g, int\_memory, m\_dep, mobile\_wt, n\_cores, pc, px\_height, px\_width, ram, sc\_h, sc\_w, talk\_time, three\_g, touch\_screen, and wifi.
- Environment:** A sidebar showing the global environment with various datasets loaded, such as "Housing", "indian\_weat...", "insurance", "large\_sales...", "megagymdata", "mental", "my\_data", and "numeric\_data".
- Files:** A sidebar showing the file structure and files in the project, including "S098 M2 P10.R", "SuperMarket\_Analysis", "S098 M2 P11.R", "Housing", "S098 M2 P12.R", and "test.csv".
- Bottom Bar:** Includes icons for search, file operations, and an on-screen keyboard.



The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project:** Project (None) is selected.
- Code Editor:** An R script titled "test.R" is open, containing the following code:

```

1 # Load dataset
2 data <- read.csv("test.csv")
3
4 # View dataset
5 head(data)
6 str(data)

```
- Console:** The output of the R code is displayed, showing the first few rows of the dataset "test.csv". The columns include battery\_power, blue, clock\_speed, dual\_sim, fc, four\_g, int\_memory, m\_dep, mobile\_wt, n\_cores, pc, px\_height, px\_width, ram, sc\_h, sc\_w, talk\_time, three\_g, touch\_screen, and wifi.
- Environment:** A sidebar showing the global environment with various datasets loaded, such as "Housing", "indian\_weat...", "insurance", "large\_sales...", "megagymdata", "mental", "my\_data", and "numeric\_data".
- Files:** A sidebar showing the file structure and files in the project, including "S098 M2 P10.R", "SuperMarket\_Analysis", "S098 M2 P11.R", "Housing", "S098 M2 P12.R", and "test.csv".
- Bottom Bar:** Includes icons for search, file operations, and an on-screen keyboard.

# MVLU COLLEGE

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

RStudio interface showing a script editor with R code and a file browser on the right.

```

1 # Load dataset
2 data <- read.csv("test.csv")
3
4 # View dataset
5 head(data)
6 str(data)
7

```

Console output:

```

> correlation_matrix <- cor(numeric_data)
>
> # Display correlation matrix
> print(correlation_matrix)
      id battery_power    bluet   clock_speed dual_sim     fc
id  1.000000000 -0.0215114289  0.0004644277  0.0359170888 -0.0027208943  0.016933840
battery_power -0.0215114289  1.0000000000 -0.0466096315 -0.0390754297 -0.0611708963 -0.007846110
bluet   0.0004644277 -0.0466096315  1.0000000000  0.0347539970 -0.0111000991 -0.056063410
clock_speed 0.0359170888 -0.0390754297  0.0347539970  1.0000000000 -0.0124226539  0.010127218
dual_sim  -0.0027208943 -0.0611708963 -0.0111000991 -0.0124226539  1.0000000000  0.057606192
fc    0.0169338400 -0.0078461109 -0.0560634097  0.0101272178  0.0576061919  1.000000000
four_g   0.0309206471 -0.0425200594 -0.0011689397  0.0246648231  0.0249068154  0.032832204
int_memory -0.0140234070  0.0037509249 -0.0124163277 -0.0304868400 -0.0121578310 -0.006564568
m_dep    -0.0027944074 -0.0090648939  0.0181319438  0.0169946882  0.0217602256  0.020859314
mobile_wt  -0.0075411165 -0.047645578  0.0235133686 -0.0141069129 -0.0017341884  0.018352697
n_cores   0.0159345009  0.0275731558  0.0028282090 -0.0122469187 -0.0031290360  0.020827702
pc      0.0019685860  0.0128467464 -0.0252471248  0.0474691494  0.0739364180  0.659338498
px_height 0.0250561723 -0.0486468811 -0.0588100983  0.0172774418  0.0068419402 -0.017982400
px_width  -0.0121377560 -0.0533653977 -0.0320535827  0.0705848351  0.0156095870  0.030549905
ram     0.0434421522 -0.0232656392  0.0575699860 -0.0064696637  0.0481710716 -0.0151996670
sc_h    -0.0119719018 -0.0566562209  0.0127796584 -0.0395030496  0.0062952549  0.045157825
sc_w    -0.0029181267 -0.0239047827  0.0042230804 -0.0271276160 -0.00260464597  0.006115145
talk_time 0.0308072019 -0.0155460124 -0.0139952096 -0.0787968458  0.0043902542 -0.051548197
three_g  0.0495709892  0.0151416155  0.0135298590 -0.0214064414  0.0006895838 -0.011212045
touch_screen 0.0397679064 -0.0101378408 -0.0600307436  0.0618927630  0.0340196691  0.015467061

```

RStudio interface showing a script editor with R code and a file browser on the right.

```

1 # Load dataset
2 data <- read.csv("test.csv")
3
4 # View dataset
5 head(data)
6 str(data)
7

```

Console output:

```

> correlation_matrix <- cor(numeric_data)
>
> # Display correlation matrix
> print(correlation_matrix)
      id battery_power    bluet   clock_speed dual_sim     fc
id  1.000000000 -0.00366434123  0.0255675984 -0.0485931165  0.0315453283 -0.0603727273
battery_power -0.00366434123  0.0309206471 -0.0140234070 -0.0027944047 -0.0075411165 -0.015934501  0.001968586 -0.025056172
bluet   0.0255675984 -0.00366434123  1.0000000000 -0.0407646578  0.025731558  0.0128467468  0.048646881
clock_speed 0.0019685860 -0.00366434123 -0.0027944047  1.0000000000  0.0235133680 -0.025247123 -0.058810098
dual_sim  -0.0250561723 -0.0075411165 -0.0027944047 -0.0027944047  1.0000000000  0.0122469194  0.0474691494 -0.0122469194
fc    0.0159345009 -0.015934501  0.001968586 -0.00366434123 -0.0027944047 -0.0027944047  1.0000000000  0.0062952549
four_g   0.0019685860 -0.00366434123 -0.0027944047 -0.0027944047 -0.0027944047 -0.0027944047 -0.0027944047  1.000000000
int_memory -0.00374884991  1.0000000000 -0.004386083 -0.0104473507  0.021601083  0.022682215 -0.009564431
m_dep    -0.0148062274 -0.004386083  1.0000000000 -0.0419938426  0.010062154  0.012662551  0.062558682
mobile_wt  -0.0005086964 -0.010447351 -0.041993843  1.0000000000 -0.038907544  0.027342597  0.011157044
n_cores   0.0667157390  0.021601083  0.010062154 -0.0389075443  1.0000000000  0.014375511  0.054432727
pc      0.0376690186  0.02682215  0.012662551  0.0273425965  0.014375511  1.0000000000  0.028909812
px_height 0.0336552474 -0.009564431  0.062558682  0.0111570445 -0.054432727  0.028909812  1.0000000000
px_width  -0.0365445743 -0.0038765353  0.034861054 -0.0145768285 -0.059388199  0.056397187  0.517649808
ram     0.0308209628 -0.007107274  0.018348503  0.0287859346 -0.042750444 -0.045987256  0.027945304
sc_h    -0.0150869147 -0.0092492935 -0.026159507 -0.0220531896 -0.034057185  0.019737958  0.011161867
sc_w    -0.0008931884  0.024521033 -0.023392943  0.0221475038  0.012830091 -0.006315969  0.043486313
talk_time 0.0136919001  0.023759083  0.024124410 -0.0217044320 -0.005640231  0.038434209  0.052382641
three_g  0.5535282792 -0.015922391 -0.029278273  0.0062618805  0.050935894 -0.015202382 -0.01124604
touch_screen -0.0100033817  0.022185887  0.040253625  0.0445253109 -0.0166115579  0.021995877 -0.019636564
wifii  -0.0356515451  0.011859684 -0.039705365  0.0697616072 -0.007255598 -0.054955025 -0.012459113
id    -0.0121377560 -0.0434421522 -0.011971902  0.0029181267  0.030807202  0.0495709892  0.03976791

```

# MVLU COLLEGE

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

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source on Save | Go to file/function | Addins | Project: (None) | Environment History Connections Tutorial

```
1 # Load dataset
2 data <- read.csv("test.csv")
3
4 # View dataset
5 head(data)
6 str(data)
```

Console Terminal Background Jobs

R - R 4.5.2 - ~

```
battery_power -0.033635398 -0.0323656392 -0.055665221 -0.0239047827 0.015546012 0.0315141655 -0.01013784
blue -0.0320353583 0.0575699860 0.0127779658 0.0042230804 -0.031995210 0.0135298590 -0.06003074
clock_speed -0.0705843835 -0.0064966367 -0.039503050 -0.0271376160 -0.078796846 -0.0214064414 0.06189276
dual_sim 0.015609587 0.0481710716 0.006295255 -0.0020644597 0.004390254 0.0006298588 0.03401967
fc -0.03054905 -0.0519966705 0.045157825 0.0061151445 -0.051458197 -0.0111210450 0.01546706
four_g 0.036544574 0.0308209628 -0.015086915 -0.0008931884 0.013691909 0.5535282792 -0.01000338
int_memory -0.003876535 -0.0071072737 -0.009249235 0.0245210335 0.023759083 -0.0159223913 0.02218589
m_dep 0.034861058 0.0183485033 -0.026159507 -0.0233929427 0.024124410 -0.0292782727 0.04025363
mobile_wt -0.014576829 0.0287859346 -0.0202053190 0.0221475038 -0.021704432 0.0062618805 0.04452531
n_cores -0.059388199 -0.0427504441 -0.034057185 0.0128300914 -0.005640231 0.0509358945 -0.01661558
pc 0.056397187 -0.0459872564 0.019737958 -0.0063159642 -0.038434209 -0.0152203818 0.02199588
px_height 0.517649803 0.0279453038 0.011161867 0.0434863135 0.052382641 -0.0111246035 -0.01963656
px_width 1.000000000 -0.0263782762 -0.022610052 0.0046921000 0.053422954 0.0194385698 -0.03998625
ram -0.026378276 1.0000000000 0.022894404 0.0306779596 -0.003418840 0.0297124608 0.04365416
sc_h -0.022610052 0.0228944079 1.000000000 0.497154333 0.026061586 -0.0195283321 -0.01366232
sc_w 0.004692102 0.0306779596 0.497154331 1.0000000000 0.036990199 0.0011559393 -0.04766466
talk_time 0.053442294 -0.0034188404 0.026061586 0.0639901986 1.000000000 -0.0018050743 0.02893599
three_g 0.019438570 0.0297124608 -0.019528332 0.0011559393 -0.001805074 1.0000000000 0.000000000
touch_screen -0.039986247 -0.0436541613 -0.013662322 -0.0476646559 0.028935995 0.0000000000 1.000000000
wififi -0.073996910 -0.0319042189 -0.002994426 0.0154750685 0.016709900 -0.0246454405 -0.02600255
> |
```

On-Screen Keyboard

04:31 PM 11-01-2026

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source on Save | Go to file/function | Addins | Project: (None) | Environment History Connections Tutorial

```
1 # Load dataset
2 data <- read.csv("test.csv")
3
4 # View dataset
5 head(data)
6 str(data)
```

Console Terminal Background Jobs

R - R 4.5.2 - ~

```
wififi -0.039986247 -0.0436541613 -0.013662322 -0.0476646559 0.028935995 0.0000000000 1.000000000
wififi -0.073996910 -0.0319042189 -0.002994426 0.0154750685 0.016709900 -0.0246454405 -0.02600255
> |
```

On-Screen Keyboard

04:31 PM 11-01-2026