

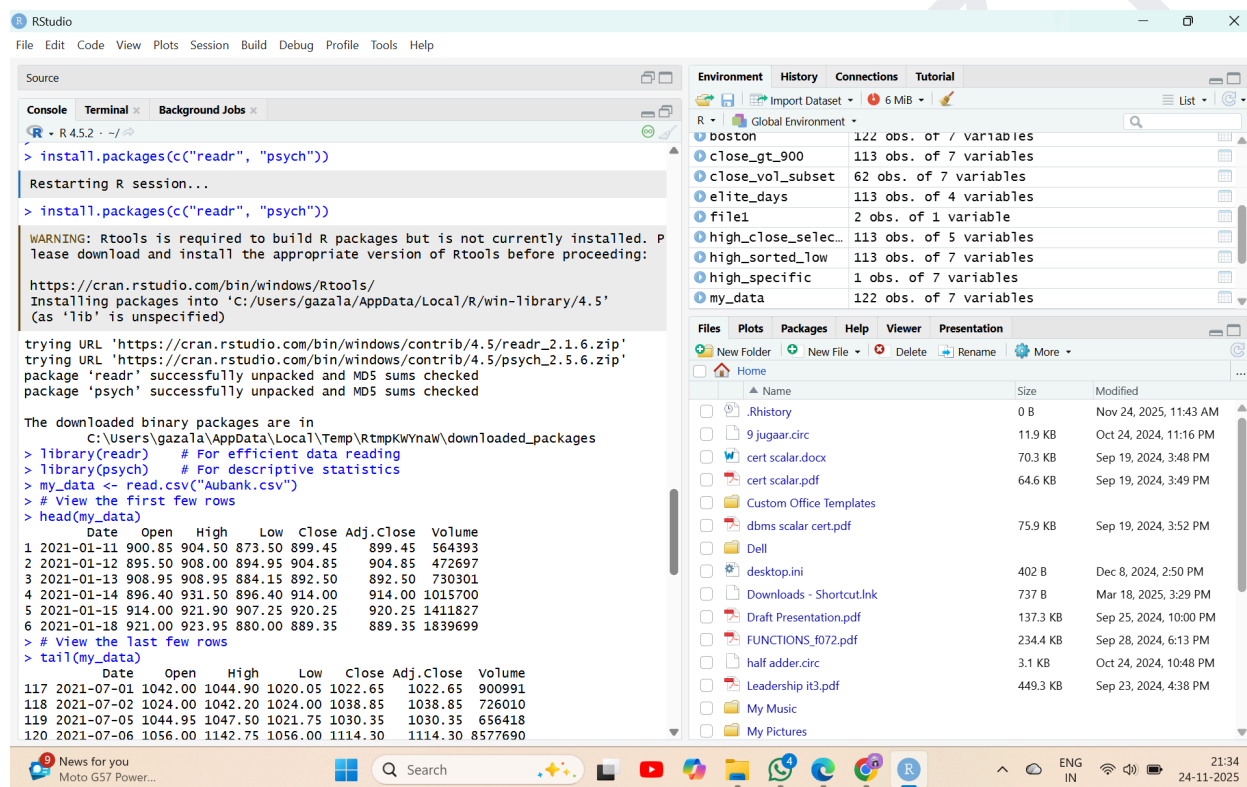
SHETH L.U.J AND SIR M.V COLLEGE

Subject: Data Analysis with SAS / SPSS /R

Practical no. 3

Aim: Exploring data: View() or print() (R).

Outputs→



The screenshot shows the RStudio interface with the following components:

- Source:** The console window shows the installation of 'readr' and 'psych' packages. A warning message indicates that Rtools is required but not installed. The packages are successfully installed into the local library.
- Environment:** The Environment pane shows the global environment with several objects: 'boston' (122 obs. of 7 variables), 'close_gt_900' (113 obs. of 7 variables), 'close_vol_subset' (62 obs. of 7 variables), 'elite_days' (113 obs. of 4 variables), 'file1' (2 obs. of 1 variable), 'high_close_selec...' (113 obs. of 5 variables), 'high_sorted_low' (113 obs. of 7 variables), 'high_specific' (1 obs. of 7 variables), and 'my_data' (122 obs. of 7 variables).
- Files:** The Files pane shows the file explorer with a list of files and folders, including '.Rhistory', '9 jugaar.circ', 'cert scalar.docx', 'cert scalar.pdf', 'Custom Office Templates', 'dbms scalar cert.pdf', 'Dell', 'desktop.ini', 'Downloads - Shortcut.lnk', 'Draft Presentation.pdf', 'FUNCTIONS_f072.pdf', 'half adder.circ', 'Leadership it3.pdf', 'My Music', and 'My Pictures'.

```
> install.packages(c("readr", "psych"))

Restarting R session...

> install.packages(c("readr", "psych"))

WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing packages into 'C:/Users/gazala/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/readr_2.1.6.zip'
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/psych_2.5.6.zip'
package 'readr' successfully unpacked and MD5 sums checked
package 'psych' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
C:\Users\gazala\AppData\Local\Temp\RtmpKwYnaw\downloaded_packages
> library(readr) # For efficient data reading
> library(psych) # For descriptive statistics
> my_data <- read_csv("Aubank.csv")
> # View the first few rows
> head(my_data)
  Date      Open    High    Low   Close Adj.Close Volume
1 2021-01-11 900.85 904.50 873.50 899.45 899.45 564393
2 2021-01-12 895.50 908.00 894.95 904.85 904.85 472697
3 2021-01-13 908.95 908.95 884.15 892.50 892.50 730301
4 2021-01-14 896.40 931.50 896.40 914.00 914.00 1015700
5 2021-01-15 914.00 921.90 907.25 920.25 920.25 1411827
6 2021-01-18 921.00 923.95 880.00 889.35 889.35 1839699
> # View the last few rows
> tail(my_data)
  Date      Open    High    Low   Close Adj.Close Volume
117 2021-07-01 1042.00 1044.90 1020.05 1022.65 1022.65 900991
118 2021-07-02 1024.00 1042.20 1024.00 1038.85 1038.85 726010
119 2021-07-05 1044.95 1047.50 1021.75 1030.35 1030.35 656418
120 2021-07-06 1056.00 1142.75 1056.00 1114.30 1114.30 8577690
```

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

Subject: Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio environment with the following components:

- Source:** Contains the R script code for loading and analyzing the 'my_data' dataset.
- Console:** Displays the output of the R commands, including the dimensions of the dataset (122 rows, 7 columns) and a summary of the data.
- Environment:** Lists the objects in the global environment, including 'boston', 'close_gt_900', 'close_vol_subset', 'elite_days', 'file1', 'high_close_selec...', 'high_sorted_low', 'high_specific', and 'my_data'.
- Files:** Shows the file explorer with various files and folders, including 'Rhistory', '9 jugaar.circ', 'cert scalar.docx', 'cert scalar.pdf', 'Custom Office Templates', 'dbms scalar cert.pdf', 'Dell', 'desktop.ini', 'Downloads - ShortcutLink', 'Draft Presentation.pdf', 'FUNCTIONS_f072.pdf', 'half adder.circ', 'Leadership it3.pdf', 'My Music', and 'My Pictures'.

```
> tail(my_data)
      Date      Open      High      Low      Close Adj.Close Volume
117 2021-07-01 1042.00 1044.90 1020.05 1022.65 1022.65 900991
118 2021-07-02 1024.00 1042.20 1024.00 1038.85 1038.85 726010
119 2021-07-05 1044.95 1047.50 1021.75 1030.35 1030.35 656418
120 2021-07-06 1056.00 1142.75 1056.00 1114.30 1114.30 8577690
121 2021-07-07 1109.00 1133.00 1091.00 1125.00 1125.00 1542754
122 2021-07-08 1115.00 1153.40 1111.75 1141.40 1141.40 1965530
> # Get the dimensions (rows and columns)
> dim(my_data)
[1] 122      7
> # Get the dimensions (rows and columns)
> cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")
Dimensions (Rows, Columns): 122 7
> # Get the structure (variable types and number of observations)
> str(my_data)
'data.frame':   122 obs. of  7 variables:
 $ Date      : chr  "2021-01-11" "2021-01-12" "2021-01-13" "2021-01-14" ...
 $ Open      : num  901 896 909 896 914 ...
 $ High      : num  904 908 909 932 922 ...
 $ Low       : num  874 895 884 896 907 ...
 $ Close     : num  899 905 892 914 920 ...
 $ Adj.Close : num  899 905 892 914 920 ...
 $ Volume    : int  564393 472697 730301 1015700 1411827 1839699 1348639 946234 97
7221 658015 ...
> # See a summary of the dataset
> summary(my_data)
      Date      Open      High      Low
Length:122   Min.   : 866   Min.   : 877.0   Min.   : 842.5
Class:character 1st Qu.: 970   1st Qu.: 989.5   1st Qu.: 958.9
Mode :character Median :1041   Median :1064.4 Median :1020.9
              Mean :1058   Mean :1081.4   Mean :1034.0
              3rd Qu.:1135 3rd Qu.:1158.9 3rd Qu.:1106.7
              Max.   :1280   Max.   :1354.4   Max.   :1242.9
      Close     Adj.Close     Volume
Min.   : 852.7   Min.   : 852.7   Min.   : 258928
1st Qu.: 971.9   1st Qu.: 971.9   1st Qu.: 743190
Median :1040.9   Median :1040.9   Median :1032194
Mean :1058.2     Mean :1058.2     Mean :1449015
3rd Qu.:1139.8   3rd Qu.:1139.8 3rd Qu.:1722323
Max.   :1298.9   Max.   :1298.9   Max.   :8577690
```

The screenshot shows the RStudio environment with the following components:

- Source:** Contains the R script code for loading and analyzing the 'my_data' dataset.
- Console:** Displays the output of the R commands, including the dimensions of the dataset (122 rows, 7 columns) and a summary of the data.
- Environment:** Lists the objects in the global environment, including 'boston', 'close_gt_900', 'close_vol_subset', 'elite_days', 'file1', 'high_close_selec...', 'high_sorted_low', 'high_specific', and 'my_data'.
- Files:** Shows the file explorer with various files and folders, including 'Rhistory', '9 jugaar.circ', 'cert scalar.docx', 'cert scalar.pdf', 'Custom Office Templates', 'dbms scalar cert.pdf', 'Dell', 'desktop.ini', 'Downloads - ShortcutLink', 'Draft Presentation.pdf', 'FUNCTIONS_f072.pdf', 'half adder.circ', 'Leadership it3.pdf', 'My Music', and 'My Pictures'.

```
> # Get the column names
> names(my_data)
[1] "Date" "Open" "High" "Low" "Close" "Adj.Close"
[7] "Volume"
> cat("Column Names: ", names(my_data), "\n")
Column Names: Date Open High Low Close Adj.Close Volume
> # Use the 'psych' package for more detailed descriptive statistics
> # 'describe()' provides: n, mean, sd, median, trimmed mean, mad, min, max, range, skew, kurtosis, and se.
> describe(my_data)
      vars  n      mean      sd      median      trimmed      mad
Date*    1 122      61.50      35.36      61.50      61.50      45.22
Open     2 122     1058.32     111.53     1041.00     1053.96     116.31
High     3 122     1081.44     118.02     1064.45     1077.67     129.62
Low      4 122     1034.00     104.24     1020.90     1030.33     103.71
Close    5 122     1058.17     111.31     1040.90     1054.55     119.35
Adj.Close 6 122     1058.17     111.31     1040.90     1054.55     119.35
Volume   7 122    1449015.01    1341884.78    1032193.50    1196407.27    551765.90
      min      max      range      skew      kurtosis      se
Date*    1.00      122.00      121.00      0.00      -1.23      3.20
Open     866.00    1280.00     414.00      0.36     -0.91     10.10
High     877.00    1354.40     477.40      0.30     -0.93     10.68
Low      842.50    1242.90     400.40      0.32     -0.84      9.44
Close    852.65    1298.90     446.25      0.29     -0.89     10.08
Adj.Close 852.65    1298.90     446.25      0.29     -0.89     10.08
Volume   258928.00 8577690.00 8318762.00      3.28     12.35    121488.54
```