

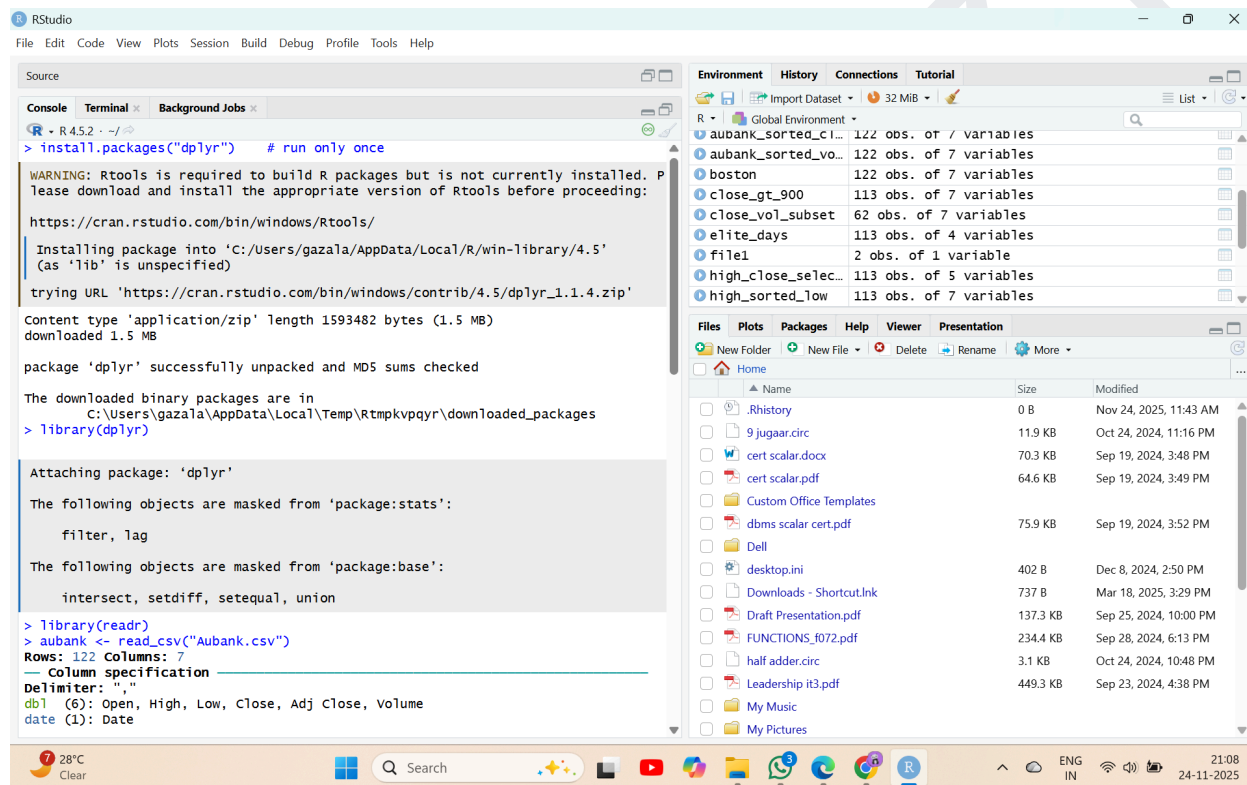
# SHETH L.U.J AND SIR M.V COLLEGE

Subject: Data Analysis with SAS / SPSS /R

## Practical no. 5

Aim: Sorting data using arrange() in R.

Outputs→



The screenshot displays the RStudio environment. The console on the left shows the installation of the 'dplyr' package. The Environment pane on the right lists several data objects, including 'aubank\_sorted\_ci...', 'aubank\_sorted\_vo...', 'boston', 'close\_gt\_900', 'close\_vol\_subset', 'elite\_days', 'file1', 'high\_close\_selec...', and 'high\_sorted\_low'.

```
> install.packages("dplyr") # run only once

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 package 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/dplyr_1.1.4.zip'

Content type 'application/zip' length 1593482 bytes (1.5 MB)
downloaded 1.5 MB

package 'dplyr' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
C:\Users\gazala\AppData\Local\Temp\Rtmpkvpqyr\downloaded_packages
> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':
  filter, lag

The following objects are masked from 'package:base':
  intersect, setdiff, setequal, union

> library(readr)
> aubank <- read_csv("Aubank.csv")
Rows: 122 Columns: 7
  Column specification
Delimiter: ","
dbl (6): Open, High, Low, Close, Adj Close, Volume
date (1): Date
```

Object	Obs	Var
aubank_sorted_ci...	122	7
aubank_sorted_vo...	122	7
boston	122	7
close_gt_900	113	7
close_vol_subset	62	7
elite_days	113	4
file1	2	1
high_close_selec...	113	5
high_sorted_low	113	7

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

Subject: Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface with the following components:

- Source:** Contains R code for data manipulation. The code includes comments and commands for sorting data by 'Close' and 'Volume'.
- Console:** Displays the output of the R code, showing the structure of the data frames and the results of the sorting operations.
- Environment:** Lists the objects in the environment, including 'aubank\_sorted\_close', 'aubank\_sorted\_volume', 'boston', 'close\_gt\_900', 'close\_vol\_subset', 'elite\_days', 'file1', 'high\_close\_selec...', and 'high\_sorted\_low'.
- 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'.

The code in the Source pane includes the following examples:

```
# Example 1 - Sort by Close (ascending)
> aubank_sorted_close <- aubank |>
+ arrange(Close)
> head(aubank_sorted_close, 5)
# A tibble: 5 x 7
  Date       Open High Low Close `Adj Close` Volume
<date>     <dbl> <dbl> <dbl> <dbl> <dbl>    <dbl>
1 2021-01-27  885  891.  842.  853.  565921
2 2021-01-28  866  877  845  870.  306074
3 2021-01-29  880  888.  863.  872.  1127886
4 2021-01-25  902  902.  880  885.  258928
5 2021-01-18  921  924.  880  889.  1839699

# Example 2 - Sort by Volume (descending)
> aubank_sorted_volume_desc <- aubank |>
+ arrange(desc(Volume))
> head(aubank_sorted_volume_desc, 5)
# A tibble: 5 x 7
  Date       Open High Low Close `Adj Close` Volume
<date>     <dbl> <dbl> <dbl> <dbl> <dbl>    <dbl>
1 2021-07-06 1056 1143. 1056 1114.  8527690
2 2021-03-30 1200. 1354. 1165 1299.  7650626
3 2021-05-03  949  950  914.  924.   6959261
4 2021-04-30 1072 1084  995. 1004.  6952375
5 2021-05-05  930  952.  910  942.   4578774

# Example 3 - Sort by Date (ascending) then Close (descending)
> aubank_multi_sort <- aubank |>
+ arrange(Date, desc(Close))
> head(aubank_multi_sort, 10)
# A tibble: 10 x 7
  Date       Open High Low Close `Adj Close` Volume
<date>     <dbl> <dbl> <dbl> <dbl> <dbl>    <dbl>
1 2021-01-11  901.  904.  874.  899.   564393
2 2021-01-12  896.  908  895.  905.   472697
3 2021-01-13  909.  909.  884.  892.   730301
4 2021-01-14  896.  932.  886.  914   1015700
5 2021-01-15  914  922.  907.  920.   1411827
6 2021-01-18  921  924.  880  889.   1839699
7 2021-01-19  895.  909.  894  899.   1348639
8 2021-01-20  905.  930.  892.  920.   946234
9 2021-01-21  922.  925.  905  913.   972221
10 2021-01-22  914.  914.  881.  897.   658015
```

The screenshot shows the RStudio interface with the following components:

- Source:** Contains R code for data manipulation. The code includes comments and commands for sorting data by 'Date' and 'Close', and filtering data based on 'High' price.
- Console:** Displays the output of the R code, showing the structure of the data frames and the results of the sorting and filtering operations.
- Environment:** Lists the objects in the environment, including 'aubank\_sorted\_close', 'aubank\_sorted\_volume', 'boston', 'close\_gt\_900', 'close\_vol\_subset', 'elite\_days', 'file1', 'high\_close\_selec...', and 'high\_sorted\_low'.
- 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'.

The code in the Source pane includes the following examples:

```
# Example 3 - Sort by Date (ascending) then Close (descending)
> aubank_multi_sort <- aubank |>
+ arrange(Date, desc(Close))
> head(aubank_multi_sort, 10)
# A tibble: 10 x 7
  Date       Open High Low Close `Adj Close` Volume
<date>     <dbl> <dbl> <dbl> <dbl> <dbl>    <dbl>
1 2021-01-11  901.  904.  874.  899.   564393
2 2021-01-12  896.  908  895.  905.   472697
3 2021-01-13  909.  909.  884.  892.   730301
4 2021-01-14  896.  932.  886.  914   1015700
5 2021-01-15  914  922.  907.  920.   1411827
6 2021-01-18  921  924.  880  889.   1839699
7 2021-01-19  895.  909.  894  899.   1348639
8 2021-01-20  905.  930.  892.  920.   946234
9 2021-01-21  922.  925.  905  913.   972221
10 2021-01-22  914.  914.  881.  897.   658015

# Example 4 - Filter + Sort
> high_sorted_low <- aubank |>
+ filter(High > 920) |> # filter first
+ arrange(Low)         # then sort
> cat("Top 5 high-price days sorted by Low:\n")
> print(high_sorted_low |> select(Date, High, Low, Close) |> head(5))
# A tibble: 5 x 4
  Date       High Low Close
<date>     <dbl> <dbl> <dbl>
1 2021-01-18  924.  880  889.
2 2021-02-01  950  891  942.
3 2021-01-20  930.  892.  920.
4 2021-01-14  932.  896.  914
5 2021-01-21  925.  905  913.
```

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

Subject: Data Analysis with SAS / SPSS /R

gazala S071