

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

Subject: Data Analysis with SAS / SPSS /R

## Practical no. 4

Aim: Applying conditional filters subset() or filter() in R.

Outputs→

The screenshot shows the RStudio interface with the following details:

- Console Tab:** Displays R code examples for filtering data based on single conditions (e.g., closing price > 900), multiple conditions (AND, e.g., close > 900 & volume > 1,000,000), and multiple conditions (OR, e.g., high > 930 OR low < 880). It also shows the creation of tibbles and head() functions for subsets.
- Environment Tab:** Shows the global environment with two objects:
  - adj\_summary:** 1 obs. of 4 variables
  - aubank:** 122 obs. of 7 variablesDetails for 'aubank' include:

Variable	Type	Length	Format
\$ Date	num	[1:122]	"2021-01-11" "2021-01-12" ...
\$ Open	num	[1:122]	901 896 909 896 914 ...
\$ High	num	[1:122]	908 909 932 922 ...
\$ Low	num	[1:122]	874 895 884 896 907 ...
\$ Close	num	[1:122]	899 905 892 914 920 ...
\$ Adj Close	num	[1:122]	899 905 892 914 920 ...
\$ Volume	num	[1:122]	564393 472697 730301 1015700 1411827 ...
- File Browser:** Shows a list of files in the 'Home' directory, 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.

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

Subject: Data Analysis with SAS / SPSS /R

RStudio interface showing R code in the console and environment browser.

```
R 4.5.2 : -/-
Days with Open < 900:
> open_lt_900 <- aubank |>
+   filter(open < 900)
> cat("days with Open < 900:", nrow(open_lt_900), "\n")
Days with Open < 900: 7
> summary(open_lt_900$open)
  Min. 1st Qu. Median Mean 3rd Qu. Max.
866.0 882.5 891.0 887.0 895.1 896.4
> # Example 2 - Multiple Conditions (Comma = AND)
> # Days where Adj Close > 910 AND Volume < 600,000
> adj_high_low_vo1 <- aubank |>
+   filter(Adj Close > 910, Volume < 600000)
> cat("days with Adj Close > 910 & volume < 600,000:", nrow(adj_high_low_vo1),
+n")
Days with Adj Close > 910 & Volume < 600,000: 11
> head(adj_high_low_vo1)
# A tibble: 6 × 7
  Date       Open High  Low Close `Adj Close` Volume
  <date>    <dbl> <dbl> <dbl> <dbl>     <dbl>   <dbl>
1 2021-02-15 1110 1115. 1061. 1092.    1092.  591093
2 2021-02-16 1089 1128 1075. 1093.    1093.  455426
3 2021-02-17 1093. 1112. 1081 1098.    1098.  394405
4 2021-02-19 1124. 1132. 1084. 1109.    1109.  553747
5 2021-02-22 1109 1128 1048. 1075.    1075.  462383
6 2021-02-24 1146. 1198. 1140. 1186.    1186.  571815
> # Example 3 - Values in a set (%in%)
> # Days where High is either 908 or 920
> high_specific <- aubank |>
+   filter(High %in% c(908, 920))
> cat("days where High = 908 or 920:", nrow(high_specific), "\n")
Days where High = 908 or 920: 1
> table(high_specific$High)
```

Environment browser showing objects:

- adj\_summary 1 obs. of 4 variables
- aubank 122 obs. of 7 variables
  - \$ Date : Date[1:122], format: "2021-01-11" "2021-01-12" ...
  - \$ Open : num [1:122] 901 896 909 896 914 ...
  - \$ High : num [1:122] 904 908 909 932 922 ...
  - \$ Low : num [1:122] 874 895 884 896 907 ...
  - \$ Close : num [1:122] 899 905 892 914 920 ...
  - \$ Adj Close: num [1:122] 899 905 892 914 920 ...
  - \$ Volume : num [1:122] 564393 472697 730301 1015700 1411827 ...
- .Rhistory 0 B Nov 24, 2025, 11:43 AM
- 9 jugaar.circ 11.9 KB Oct 24, 2024, 11:16 PM
- cert scalar.docx 70.3 KB Sep 19, 2024, 3:48 PM
- cert scalar.pdf 64.6 KB Sep 19, 2024, 3:49 PM
- Custom Office Templates 75.9 KB Sep 19, 2024, 3:52 PM
- Dell 402 B Dec 8, 2024, 2:50 PM
- desktop.ini 737 B Mar 18, 2025, 3:29 PM
- Downloads - Shortcut.lnk 137.3 KB Sep 25, 2024, 10:00 PM
- Draft Presentation.pdf 234.4 KB Sep 28, 2024, 6:13 PM
- FUNCTIONS\_f072.pdf 3.1 KB Oct 24, 2024, 10:48 PM
- half adder.circ 449.3 KB Sep 23, 2024, 4:38 PM
- Leadership i3.pdf 449.3 KB Sep 23, 2024, 4:38 PM
- My Music 64.6 KB Sep 19, 2024, 3:49 PM
- My Pictures 75.9 KB Sep 19, 2024, 3:52 PM

RStudio interface showing R code in the console and environment browser.

```
R 4.5.2 : -/-
Days where High = 908 or 920: 1
> table(high_specific$High)

908
1
> # Example 1 - Filter + Select specific columns
> high_close_selected <- aubank |>
+   filter(Close > 900) |>
+   select(Date, High, Low, Close, Volume)
> cat("\nHigh Close days (selected columns):\n")
# A tibble: 6 × 5
  Date       High  Low Close `Adj Close` Volume
  <date>    <dbl> <dbl> <dbl>     <dbl>   <dbl>
1 2021-01-12 908. 895. 905.    472697
2 2021-01-14 932. 896. 914.  1015700
3 2021-01-15 922. 907. 920.  1411827
4 2021-01-20 930. 892. 920.    946234
5 2021-01-21 925. 905. 913.  977221
6 2021-02-01 950. 891. 942.  2308622
> # Example 2 - Sort data by descending Volume
> sorted_by_volume <- aubank |>
+   arrange(desc(Volume))
> cat("\nHighest Volume Days:\n")

Highest Volume Days:
# A tibble: 6 × 7
  Date       Open High  Low Close `Adj Close` Volume
  <date>    <dbl> <dbl> <dbl> <dbl>     <dbl>   <dbl>
1 2021-07-06 1056 1143. 1056 1114.    1114.  8572690
2 2021-03-30 1200. 1354. 1165 1299.    1299.  7550626
3 2021-05-03 949. 950. 914. 924.    924.  6959261
4 2021-04-30 1072 1084. 995. 1004.    1004.  6957375
5 2021-05-05 930. 952. 910. 942.    942.  4578774
6 2021-06-11 996. 1064. 996 1046.    1046.  3552844
```

Environment browser showing objects:

- adj\_summary 1 obs. of 4 variables
- aubank 122 obs. of 7 variables
  - \$ Date : Date[1:122], format: "2021-01-11" "2021-01-12" ...
  - \$ Open : num [1:122] 901 896 909 896 914 ...
  - \$ High : num [1:122] 904 908 909 932 922 ...
  - \$ Low : num [1:122] 874 895 884 896 907 ...
  - \$ Close : num [1:122] 899 905 892 914 920 ...
  - \$ Adj Close: num [1:122] 899 905 892 914 920 ...
  - \$ Volume : num [1:122] 564393 472697 730301 1015700 1411827 ...
- .Rhistory 0 B Nov 24, 2025, 11:43 AM
- 9 jugaar.circ 11.9 KB Oct 24, 2024, 11:16 PM
- cert scalar.docx 70.3 KB Sep 19, 2024, 3:48 PM
- cert scalar.pdf 64.6 KB Sep 19, 2024, 3:49 PM
- Custom Office Templates 75.9 KB Sep 19, 2024, 3:52 PM
- Dell 402 B Dec 8, 2024, 2:50 PM
- desktop.ini 737 B Mar 18, 2025, 3:29 PM
- Downloads - Shortcut.lnk 137.3 KB Sep 25, 2024, 10:00 PM
- Draft Presentation.pdf 234.4 KB Sep 28, 2024, 6:13 PM
- FUNCTIONS\_f072.pdf 3.1 KB Oct 24, 2024, 10:48 PM
- half adder.circ 449.3 KB Sep 23, 2024, 4:38 PM
- Leadership i3.pdf 449.3 KB Sep 23, 2024, 4:38 PM
- My Music 64.6 KB Sep 19, 2024, 3:49 PM
- My Pictures 75.9 KB Sep 19, 2024, 3:52 PM

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

Subject: Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface. The left pane displays the R console with the following code and output:

```
> # Example 3 - Combined: Filter + Sort + Select
> elite_days <- aubank |>
+ filter(High > 920) |>
+ arrange(desc(Close)) |>
+ select(Date, High, Close, Volume)
> cat("\nElite days (High > 920 sorted by Close):\n")
Elite days (High > 920 sorted by Close):
> head(elite_days)
# A tibble: 6 × 4
  Date       High Close  Volume
  <date>     <dbl> <dbl>   <dbl>
1 2021-03-30 1354. 1299. 7650626
2 2021-03-04 1290  1268. 2247523
3 2021-04-01 1278  1268. 1087053
4 2021-03-08 1290  1264. 925121
5 2021-03-05 1294. 1261. 1383020
6 2021-03-12 1275. 1255. 803163
> # Example 4 - Summary for conditions
> adj_summary <- aubank |>
+ filter(Adj Close > 900) |>
+ summarise(
+   count = n(),
+   avg_close = mean(Close),
+   max_close = max(Close),
+   min_close = min(Close)
+ )
> cat("\nsummary of days with Adj Close > 900:\n")
Summary of days with Adj Close > 900:
> print(adj_summary)
# A tibble: 1 × 4
  count avg_close max_close min_close
  <int>    <dbl>    <dbl>    <dbl>
1     113     1072.    1299.    905.
```

The right pane shows the Environment and Global Environment panes. The Global Environment pane lists objects like `adj_summary` and `aubank`. The bottom right corner shows a Mac OS X dock with various application icons.