

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

PRACTICAL NO:4

AIM: Applying conditional filters subset() or filter() in R

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Console Terminal Background Jobs
> install.packages("dplyr")
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/IT-03/documents/R/win-library/4.1'
(as 'lib' is unspecified)
also installing the dependency 'vctrs'

There are binary versions available but the source versions are later:
  binary source needs_compilation
vctrs 0.6.1 0.6.5      TRUE
dplyr 1.1.2 1.1.4      TRUE

Binaries will be installed

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.1/vctrs_0.6.1.zip'
Content type 'application/zip' length 1581321 bytes (1.5 MB)
downloaded 1.5 MB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.1/dplyr_1.1.2.zip'
Content type 'application/zip' length 1646641 bytes (1.6 MB)
downloaded 1.6 MB

package 'vctrs' successfully unpacked and MD5 sums checked
Warning: cannot remove prior installation of package 'vctrs'
Warning: restored 'vctrs'
package 'dplyr' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:/Users/IT-03/AppData/Local/Temp/RtmpCSHKBC/downloaded_packages
Warning message:
In file.copy(savedcopy, lib, recursive = TRUE) :

```

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Console Terminal Background Jobs
> 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

Warning message:
package 'dplyr' was built under R version 4.1.3

> library(readr) # For efficient reading
> # Load your Housing.csv file
> housing <- read_csv("My Web Sites/Housing.csv")
Rows: 545 Columns: 13
  Column specification:
  Delimiter: ","
  chr (7): mainroad, guestroom, basement, hotwaterheating, airconditioning, pr...
  dbl (6): price, area, bedrooms, bathrooms, stories, parking

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
> # Quick look at the data structure
> head(housing)
# A tibble: 6 x 13
  price area bedrooms bathrooms stories mainroad guestroom basement
  <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <chr> <chr>
```

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

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R 4.1.2 - ~
> head(housing)
# A tibble: 6 x 13
  price area bedrooms bathrooms stories mainroad guestroom basement
  <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <chr>
1 13300000 2420     4      2     3 yes    no     no
2 12250000 8960     4      4     4 yes    no     no
3 12250000 9960     3      2     2 yes    no     yes
4 12215000 2500     4      2     2 yes    no     yes
5 11410000 2420     4      1     2 yes    yes   yes
6 10850000 2500     3      3     1 yes    no     yes
# i 5 more variables: hotwaterheating <chr>, airconditioning <chr>,
# parking <dbl>, prefarea <chr>, furnishingstatus <chr>
> high_price_subset <- subset(housing, price > 10000000)
> cat("Number of high-price houses (price > 10M):", nrow(high_price_subset), "\n")
Number of high-price houses (price > 10M): 8
> summary(high_price_subset$price)
  Min. 1st Qu. Median Mean 3rd Qu. Max.
10150000 10675000 11812500 11571875 12250000 13300000
> large_highprice_subset <- subset(housing, price > 10000000 & area > 8000)
> cat("Number of large, high-price houses:", nrow(large_highprice_subset), "\n")
Number of large, high-price houses: 4
> head(large_highprice_subset)
# A tibble: 4 x 13
  price area bedrooms bathrooms stories mainroad guestroom basement
  <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <chr>
1 12250000 8960     4      4     4 yes    no     no
2 12250000 9960     3      2     2 yes    no     yes
3 10150000 8580     4      3     4 yes    no     no
4 10150000 16200    5      3     2 yes    no     no
# i 5 more variables: hotwaterheating <chr>, airconditioning <chr>,
# parking <dbl>, prefarea <chr>, furnishingstatus <chr>
> special_houses_subset <- subset(housing, guestroom == "yes" | airconditioning == "ye
s")
> cat("Number of special houses (guestroom OR AC):", nrow(special_houses_subset), "\n")
Number of special houses (guestroom OR AC): 225
```

Environment History Connections Tutorial

Data

- furnished_subset 367 obs. of 13 variables
- high_price_subset 8 obs. of 13 variables
- housing 545 obs. of 13 variables
- Housing 545 obs. of 13 variables
- insurance 1338 obs. of 7 variables
- large_highprice_subset 4 obs. of 13 variables
- mainroad_basement 168 obs. of 13 variables

Files Plots Package Help Viewer Presentation

New Folder New File Delete Rename More

Home My Web Sites

Name	Size	Modified
insurance.csv	543 KB	Nov 24, 2025, 11:39 AM
WebSite1		
Housing.csv	293 KB	Nov 18, 2025, 11:51 AM

Windows Type here to search 31°C Sunny 12:58 24-11-2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R - R 4.1.2 - ~
> head(special_houses_subset)
# A tibble: 6 x 13
  price area bedrooms bathrooms stories mainroad guestroom basement
  <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <chr>
1 13300000 2420     4      2     3 yes    no     no
2 12250000 8960     4      4     4 yes    no     no
3 12250000 2500     4      2     2 yes    no     yes
4 11410000 2420     4      1     2 yes    yes   yes
5 10850000 2500     3      3     1 yes    no     yes
6 10150000 8580     4      3     4 yes    no     no
# i 5 more variables: hotwaterheating <chr>, airconditioning <chr>,
# parking <dbl>, prefarea <chr>, furnishingstatus <chr>
> many_bathrooms <- housing[>
+ filter(bathrooms > 3)
> cat("Number of houses with > 3 bathrooms:", nrow(many_bathrooms), "\n")
Number of houses with > 3 bathrooms: 1
> summary(many_bathrooms$bathrooms)
  Min. 1st Qu. Median Mean 3rd Qu. Max.
        4        4        4        4        4
> mainroad_basement <- housing[>
+ filter(mainroad == "yes", basement == "yes")
> cat("Houses on main road with basement:", nrow(mainroad_basement), "\n")
Houses on main road with basement: 168
> head(mainroad_basement)
# A tibble: 6 x 13
  price area bedrooms bathrooms stories mainroad guestroom basement
  <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <chr>
1 12250000 9960     3      2     2 yes    no     yes
2 12215000 2500     4      2     2 yes    no     yes
3 11410000 2420     4      1     2 yes    yes   yes
4 10850000 2500     3      3     1 yes    no     yes
5 9870000 8100     4      1     2 yes    yes   yes
6 9800000 13200    3      1     2 yes    no     yes
# i 5 more variables: hotwaterheating <chr>, airconditioning <chr>,
# parking <dbl>, prefarea <chr>, furnishingstatus <chr>
```

Environment History Connections Tutorial

Data

- furnished_subset 367 obs. of 13 variables
- high_price_subset 8 obs. of 13 variables
- housing 545 obs. of 13 variables
- Housing 545 obs. of 13 variables
- insurance 1338 obs. of 7 variables
- large_highprice_subset 4 obs. of 13 variables
- mainroad_basement 168 obs. of 13 variables

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home My Web Sites

Name	Size	Modified
insurance.csv	543 KB	Nov 24, 2025, 11:39 AM
WebSite1		
Housing.csv	293 KB	Nov 18, 2025, 11:51 AM

Windows Type here to search 31°C Sunny 12:58 24-11-2025

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

The screenshot shows the RStudio interface with the following details:

- Console Tab:** Displays R code and its output. The code filters a dataset to find houses with more than 3 bathrooms, identifies houses on main roads with basements, and counts furnished and semi-furnished houses.
- Data View:** Shows a table of house features for 6 rows. The columns include price, area, bedrooms, bathrooms, stories, mainroad, guestroom, and basement.
- Environment Tab:** Lists various datasets and objects in the global environment, such as `furnished_subset`, `high_price_subset`, and `insurance`.
- Files Tab:** Shows files in the current directory: `insurance.csv`, `WebSite1`, and `Housing.csv`.
- System Status Bar:** Shows the date (24-11-2025), time (12:58), weather (31°C Sunny), and battery status.