

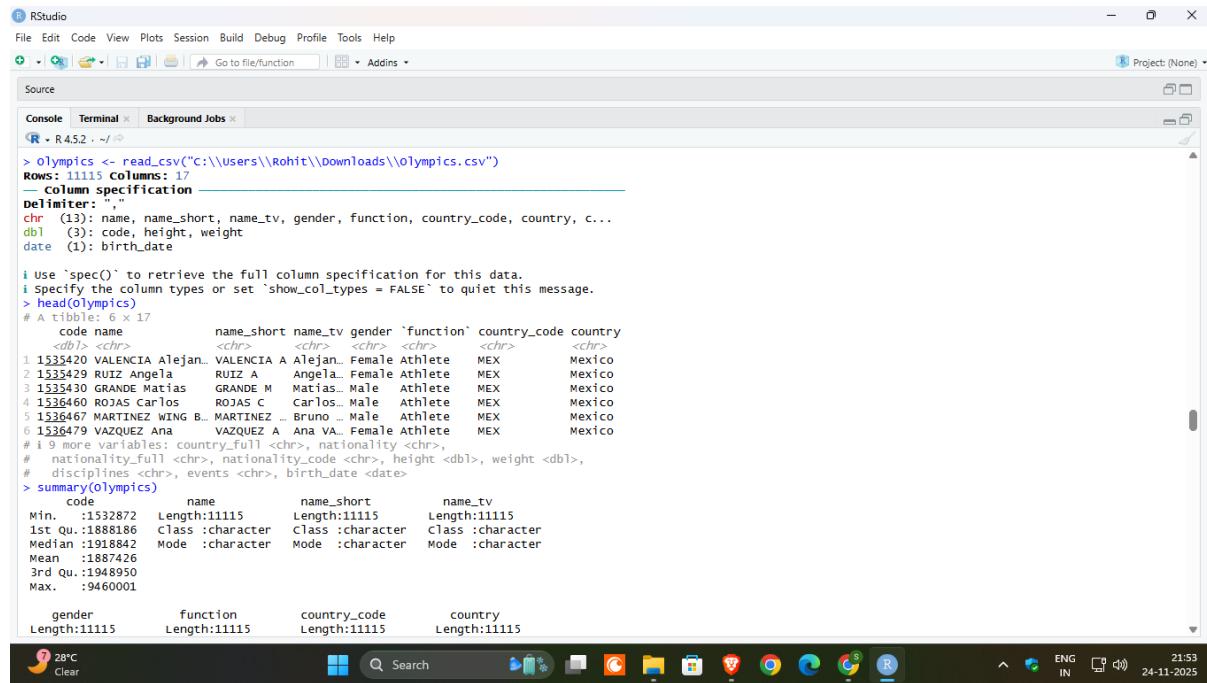
SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT :- DATA ANALYSIS WITH SAS / SPSS / R

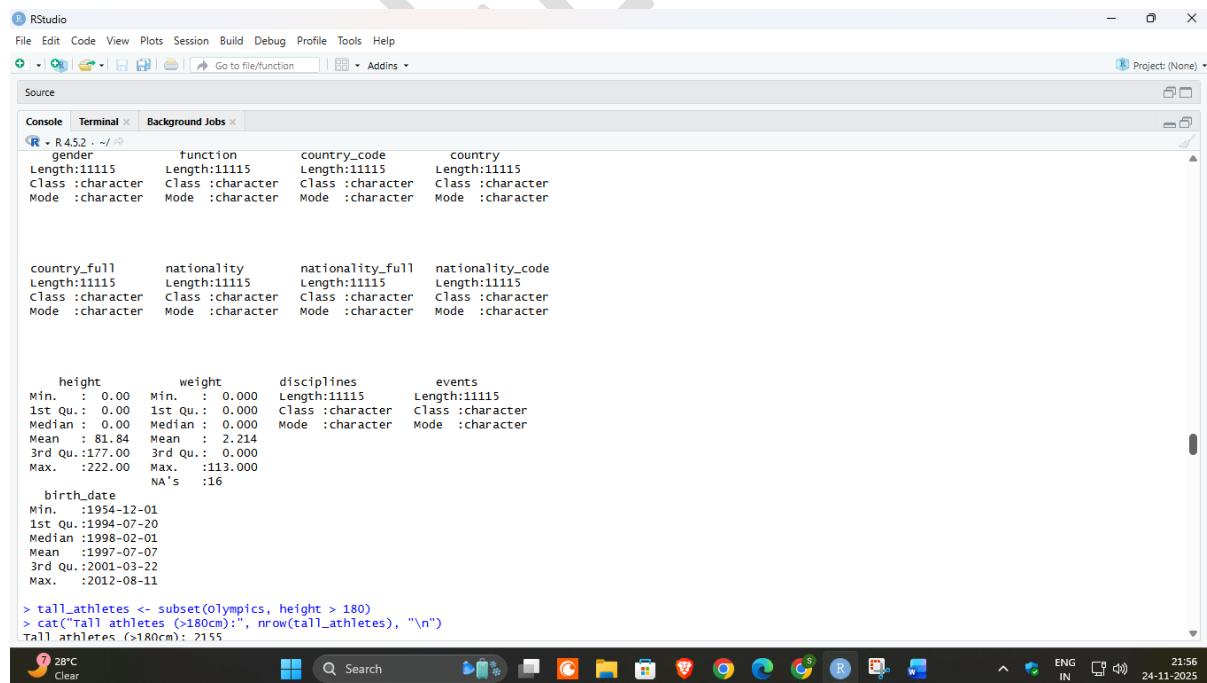
PRACTICAL – 4

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

OUTPUT:-



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ - Go to file/function Addins Project: (None)
Source Console Terminal Background Jobs
R v 4.5.2 . ~/~>
> Olympics <- read_csv("C:\\users\\Rohit\\Downloads\\olympics.csv")
Rows: 11115 Columns: 17
--- Column specification ---
Delimiter: ","
chr (13): name, name_short, name_tv, gender, function, country_code, country, c...
dbl (3): code, height, weight
date (1): birth_date
# 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.
> head(Olympics)
# A tibble: 6 × 17
   code      name    name_short  name_tv gender  function  country_code country
   <dbl>     <chr>    <chr>       <chr>   <chr>    <chr>      <chr>
1 1535420 VALENCIA Alejan. VALENCIA A Alejan.. Female Athlete MEX Mexico
2 1535420 RUIZ Angelia RUIZ A Angelia.. Female Athlete MEX Mexico
3 1535430 GRANDE Matias GRANDE M Matias.. Male Athlete MEX Mexico
4 1536460 ROJAS Carlos ROJAS C Carlos.. Male Athlete MEX Mexico
5 1536467 MARTINEZ WING B. MARTINEZ .. Bruno .. Male Athlete MEX Mexico
6 1536479 VAZQUEZ Ana VAZQUEZ A Ana VA.. Female Athlete MEX Mexico
# i 9 more variables: country_full <chr>, nationality <chr>,
# nationality_full <chr>, nationality_code <chr>, height <dbl>, weight <dbl>,
# disciplines <chr>, events <chr>, birth_date <date>
> summary(Olympics)
   code      name    name_short  name_tv gender  function  country_code country
   Min. :1532872 Length:11115  Length:11115  Length:11115
  1st Qu.:1888186 Class :character Class :character Class :character
  Median :1918842 Mode  :character Mode  :character Mode  :character
  Mean   :1887426
  3rd Qu.:1948950
  Max.  :9460001
   gender      function  country_code  country
   Length:11115  Length:11115  Length:11115  Length:11115
  
```



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ - Go to file/function Addins Project: (None)
Source Console Terminal Background Jobs
R v 4.5.2 . ~/~>
> gender      function  country_code  country
Length:11115  Length:11115  Length:11115  Length:11115
Class :character Class :character Class :character Class :character
Mode  :character Mode  :character Mode  :character Mode  :character

country_full  nationality  nationality_full  nationality_code
Length:11115  Length:11115  Length:11115  Length:11115
Class :character Class :character Class :character Class :character
Mode  :character Mode  :character Mode  :character Mode  :character

height      weight      disciplines      events
Min. : 0.00  Min. : 0.000  Length:11115  Length:11115
1st Qu.: 0.00  1st Qu.: 0.000  Class :character Class :character
Median : 0.00  Median : 0.000  Mode  :character Mode  :character
Mean   : 81.84  Mean   : 2.214
3rd Qu.:177.00  3rd Qu.: 0.000
Max.  :222.00  Max.  :113.000
NA's   :16

birth_date
Min. :1954-12-01
1st Qu.:1994-07-20
Median :1998-02-01
Mean   :1997-07-07
3rd Qu.:2001-03-22
Max.  :2012-08-11

> tall_athletes <- subset(Olympics, height > 180)
> cat("Tall athletes (>180cm):", nrow(tall_athletes), "\n")
Tall athletes (>180cm): 2155
  
```

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

SUBJECT :- DATA ANALYSIS WITH SAS / SPSS / R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

Console Terminal Background Jobs

```
> tall_athletes <- subset(Olympics, height > 180)
> cat("tall_athletes (>180cm):", nrow(tall_athletes), "\n")
tall_athletes (>180cm): 2155
> summary(tall_athletes$height)
   Min. 1st Qu. Median  Mean 3rd Qu. Max.
181.0 185.0 189.0 190.1 194.0 222.0
> tall_heavy_athletes <- subset(Olympics, height > 180 & weight > 80)
> cat("tall & Heavy athletes:", nrow(tall_heavy_athletes), "\n")
tall & Heavy athletes: 89
> head(tall_heavy_athletes)
# A tibble: 6 × 17
  code name      name_short name_tv gender `function` country_code country
  <dbl> <chr>      <chr>     <chr>    <chr>      <chr>
1 1539950 McNulty Harry McNULTY H Harry .. Male Athlete IRL Ireland
2 1539951 Mollen Bryan MOLLEN B Bryan .. Male Athlete IRL Ireland
3 1539952 Kennedy Terry KENNEDY T Terry .. Male Athlete IRL Ireland
4 1540534 Vera Feldi... VERA FELD... Santia... Male Athlete ARG Argentina
5 1540536 OSADCUK Matias OSADCUK M Matias.. Male Athlete ARG Argentina
6 1540537 Alvarez Santiago ALVAREZ S Santia... Male Athlete ARG Argentina
# i 9 more variables: country_full <chr>, nationality <chr>,
#   nationality_full <chr>, nationality_code <chr>, height <dbl>, weight <dbl>,
#   disciplines <chr>, events <chr>, birth_date <date>
> female_or_very_tall <- subset(Olympics, gender == "Female" | height > 190)
> cat("Female OR >190 cm:", nrow(female_or_very_tall), "\n")
Female OR >190 cm: 6243
> head(female_or_very_tall)
# A tibble: 6 × 17
  code name      name_short name_tv gender `function` country_code country
  <dbl> <chr>      <chr>     <chr>    <chr>      <chr>
1 1535420 Valencia Alejand... VALENCIA A Alejan... Female Athlete MEX Mexico
2 1535429 Ruiz Angel... RUIZ A Angel... Female Athlete MEX Mexico
3 1536479 Vazquez Ana VAZQUEZ A Ana VA... Female Athlete MEX Mexico
4 1536632 Fallah Mobina FALLAH M Mobina... Female Athlete IRI IR Iran
5 1538449 AMAistroae Mad... AMAISTROA M Madali... Female Athlete ROU Romania
# i 9 more variables: country_full <chr>, nationality <chr>,
#   nationality_full <chr>, nationality_code <chr>, height <dbl>, weight <dbl>,
#   disciplines <chr>, events <chr>, birth_date <date>
```

28°C Clear 21:56 24-11-2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

Console Terminal Background Jobs

```
> mex_athletes <- Olympics[1]
+ filter(country_code == "MEX")
> cat("Mexican Athletes:", nrow(mex_athletes), "\n")
Mexican Athletes: 100
> heavy_male <- Olympics[1]
+ filter(gender == "Male", weight > 85)
> cat("Heavy Male Athletes:", nrow(heavy_male), "\n")
Heavy Male Athletes: 95
> head(heavy_male)
# A tibble: 6 × 17
  code name      name_short name_tv gender `function` country_code country
  <dbl> <chr>      <chr>     <chr>    <chr>      <chr>
1 1539950 McNulty Harry McNULTY H Harry .. Male Athlete IRL Ireland
2 1539951 Mollen Bryan MOLLEN B Bryan .. Male Athlete IRL Ireland
3 1540536 OSADCUK Matias OSADCUK M Matias.. Male Athlete ARG Argentina
4 1540537 Alvarez Santiago ALVAREZ S Santia... Male Athlete ARG Argentina
5 1540538 Schulz German SCHULZ G German.. Male Athlete ARG Argentina
6 1540603 Elizalde Tomas ELIZALDE T Tomas .. Male Athlete ARG Argentina
# i 9 more variables: country_full <chr>, nationality <chr>,
#   nationality_full <chr>, nationality_code <chr>, height <dbl>, weight <dbl>,
#   disciplines <chr>, events <chr>, birth_date <date>
> selected_countries <- Olympics[1]
+ filter(country_code %in% c("MEX", "USA", "IND"))
> cat("Athletes from selected countries:", nrow(selected_countries), "\n")
Athletes from selected countries: 840
> table(selected_countries$country_code)
```

28°C Clear 21:56 24-11-2025

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

SUBJECT :- DATA ANALYSIS WITH SAS / SPSS / R



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ Go to file/function Addins Project: (None)
Source
Console Terminal Background Jobs
R > R 4.5.2 - ~
1 1539950 McNULTY Harry McNULTY H Harry ... Male Athlete IRL Ireland
2 1539951 MOLLEN Bryan MOLLEN B Bryan ... Male Athlete IRL Ireland
3 1540536 OSADCUK Matias OSADCUK M Matias... Male Athlete ARG Argentina
4 1540537 ALVAREZ Santiago ALVAREZ S Santia... Male Athlete ARG Argentina
5 1540538 SCHULZ German SCHULZ G German... Male Athlete ARG Argentina
6 1540603 ELIZALDE Tomas ELIZALDE T Tomas ... Male Athlete ARG Argentina
# i 9 more variables: country_full <chr>, nationality <chr>,
# nationality_full <chr>, nationality_code <chr>, height <dbl>, weight <dbl>,
# disciplines <chr>, events <chr>, birth_date <date>
> selected_countries <- Olympics |>
+ filter(country_code %in% c("MEX", "USA", "IND"))
> cat("Athletes from selected countries:", nrow(selected_countries), "\n")
Athletes from selected countries: 840
> table(selected_countries$country_code)

IND MEX USA
112 108 620
> archery_athletes <- Olympics |>
+ filter(grepl("Archery", disciplines))
> cat("Archery Athletes:", nrow(archery_athletes), "\n")
Archery Athletes: 128
> head(archery_athletes)
# A tibble: 6 × 17
  code name      short_name name_tv gender `function` country_code country
  <dbl> <chr>        <chr>    <chr>   <chr>    <chr>    <chr>
1 1533420 VALENCIA Alejan. VALENCIA A Alejan. Female Athlete MEX Mexico
2 1533429 RUIZ Angela RUIZ A Angela. Female Athlete MEX Mexico
3 1533430 GRANDE Matias GRANDE M Matias. Male Athlete MEX Mexico
4 1536460 ROJAS Carlos ROJAS C Carlos. Male Athlete MEX Mexico
5 1536467 MARTINEZ WING B. MARTINEZ ... Bruno ... Male Athlete MEX Mexico
6 1536479 VAZQUEZ Ana VAZQUEZ A Ana VA. Female Athlete MEX Mexico
# i 9 more variables: country_full <chr>, nationality <chr>,
# nationality_full <chr>, nationality_code <chr>, height <dbl>, weight <dbl>,
# disciplines <chr>, events <chr>, birth_date <date>

28°C Clear
Search
21:56
ENG IN 24-11-2025
```