

MVLU COLLEGE.

PRACTICAL NO :- 04

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

```
CODE —install.packages("dplyr")
```

```
library(dplyr)
```

```
library(readr) # For efficient reading
```

```
student <- read_csv("C:/Users/itlab/Downloads/S100/StudentPerformanceFactors.csv")
```

```
head(student)
```

```
# Example 1: Single Condition (Pipe Operator |>)
```

```
# Filter students with exam score above 80
```

```
high_exam_subset <- subset(student, Exam_Score > 80)
```

```
cat("Number of students with exam score > 80:", nrow(high_exam_subset), "\n")
```

```
head(high_exam_subset)
```

```
# Example 2: Multiple Conditions (AND)
```

```
# Students who studied > 20 hours AND had attendance > 90%
```

```
high_study_high_attendance <- subset(student,
```

```
Hours_Studied > 20 & Attendance > 90)
```

```
cat("High study + high attendance:",
```

```
    nrow(high_study_high_attendance), "\n")
```

```
head(high_study_high_attendance)
```

```
# Example 3: Multiple Conditions (OR)
```

```
# Students who sleep more than 8 hours OR do extracurricular activities
```

```
sleep_or_extracurricular <- subset(student,
```

```
    Sleep_Hours > 8 | Extracurricular_Activities == "Yes")
```

```
cat("Sleep > 8 hours OR extracurricular:",
```

```
    nrow(sleep_or_extracurricular), "\n")
```

```
head(sleep_or_extracurricular)
```

```
# METHOD 2: Using dplyr::filter()
```

```
# -----
```

NANDINI PANDIT S100

Data Analysis with SAS / SPSS /R Practical

MVLU COLLEGE.

```
# Example 1: Single Condition (Pipe Operator)
# Students scoring above 85 marks in exams
high_exam_score <- student |>
  filter(Exam_Score > 80)
cat("Students with exam score > 80:",
  nrow(high_exam_score), "\n")
head(high_exam_score)

# Example 2: Multiple Conditions (AND using commas)
# Students with low sleep (<6 hours) AND low motivation
low_sleep_low_motivation <- student |>
  filter(Sleep_Hours < 8, Motivation_Level == "Low")

cat("Low sleep + low motivation:",
  nrow(low_sleep_low_motivation), "\n")
head(low_sleep_low_motivation)

# Example 3: Using %in% operator
# Filter by school type: Public or Private
school_type_filter <- student |>
  filter(School_Type %in% c("Public_School", "Private_School"))

cat("School type Public or Private:",
  nrow(school_type_filter), "\n")
table(school_type_filter$School_Type)
```

MVLU COLLEGE.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Console Terminal Background Jobs

```

R > R 4.5.2 - /~/.R/4.5.2/library/tidyverse/v3.3.0/Rcpp_1.0.8.2/headers.h:1:10: fatal error: 'Rcpp/headers.h' file not found
 #include "headers.h"
          ^~~~~~
compilation terminated.
Error: object 'high_performance_filter' not found

```

Source Environment History Connections Tutorial

Data

- high_exam_score 43 obs. of 20 variables
- high_study_high_att... 727 obs. of 20 variables
- high_study_subset 3063 obs. of 20 variables
- low_sleep_low_motiv... 1226 obs. of 20 variables
- Mental_health 101 obs. of 11 variables
- my_data 1000 obs. of 14 variables
- sales_data 1000 obs. of 14 variables
- school_type_filter 0 obs. of 20 variables
- sleep_or_extracurri... 4363 obs. of 20 variables
- student 6607 obs. of 20 variables
- student.Mental.heal... 101 obs. of 11 variables

Files Plots Packages Help Viewer Presentation

NIITY +0.04%

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Console Terminal Background Jobs

```

R > R 4.5.2 - /~/.R/4.5.2/library/tidyverse/v3.3.0/Rcpp_1.0.8.2/headers.h:1:10: fatal error: 'Rcpp/headers.h' file not found
 #include "headers.h"
          ^~~~~~
compilation terminated.
Error: object 'high_study_high_attendance' not found

```

Source Environment History Connections Tutorial

Data

- high_exam_score 43 obs. of 20 variables
- high_study_high_att... 727 obs. of 20 variables
- high_study_subset 3063 obs. of 20 variables
- low_sleep_low_motiv... 1226 obs. of 20 variables
- Mental_health 101 obs. of 11 variables
- my_data 1000 obs. of 14 variables
- sales_data 1000 obs. of 14 variables
- school_type_filter 0 obs. of 20 variables
- sleep_or_extracurri... 4363 obs. of 20 variables
- student 6607 obs. of 20 variables
- student.Mental.heal... 101 obs. of 11 variables

Files Plots Packages Help Viewer Presentation

SENSEX +0.06%

MVLU COLLEGE.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Console Terminal Background Jobs

```

> # i abbreviated name: 'Extracurricular_Activities'
> # i 14 more variables: Previous_Scores <dbl>, Motivation_Level <chr>, Internet_Access <chr>,
> # Tutoring_sessions <dbl>, Family_Income <chr>, Teacher_Quality <chr>, School_Type <chr>,
> # Peer_Influence <chr>, Physical_Activity <dbl>, Learning_Disabilities <chr>,
> # Parental_Education_Level <chr>, Distance_From_Home <chr>, Gender <chr>, Exam_Score <dbl>
>
> # Example 1: single condition (Pipe operator)
> # Students scoring above 85 marks in exams
> high_exam_score <- student |>
+ filter(Exam_Score > 80)
> view(high_exam_score)
>
> cat("Students with exam score > 80:", nrow(high_exam_score), "\n")
Students with exam score > 80: 43
>
> # Example 2: Multiple conditions (AND using commas)
> # Students with low sleep (<6 hours) AND low motivation
> low_sleep_low_motivation <- student |>
+ filter(Sleep_Hours < 6, Motivation_Level == "Low")
> view(low_sleep_low_motivation)
> head(low_sleep_low_motivation)
# A tibble: 6 × 20
  Hours_Studied Attendance Parental_Involvement Access_to_Resources Extracurricular_Activities Sleep_Hours
  <dbl> <dbl> <chr> <chr> <chr> <chr> <dbl>
1 23     84 Low   High  NO    7
2 29     84 Medium Low   Yes   7
3 17     97 Medium High  Yes   6
4 18     66 High  High  NO    4
5 13     72 Low   Medium NO   6
6 21     65 Medium Low   Yes   7
# i abbreviated name: 'Extracurricular_Activities'
# i 14 more variables: Previous_Scores <dbl>, Motivation_Level <chr>, Internet_Access <chr>,
# Tutoring_sessions <dbl>, Family_Income <chr>, Teacher_Quality <chr>, School_Type <chr>,
# Peer_Influence <chr>, Physical_Activity <dbl>, Learning_Disabilities <chr>,
# Parental_Education_Level <chr>, Distance_From_Home <chr>, Gender <chr>, Exam_Score <dbl>
>
```

Files Plots Packages Help Viewer Presentation

SENSEX 13:06 24-11-2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Console Terminal Background Jobs

```

> # i abbreviated name: 'Extracurricular_Activities'
> # i 14 more variables: Previous_Scores <dbl>, Motivation_Level <chr>, Internet_Access <chr>,
> # Tutoring_sessions <dbl>, Family_Income <chr>, Teacher_Quality <chr>, School_Type <chr>,
> # Peer_Influence <chr>, Physical_Activity <dbl>, Learning_Disabilities <chr>,
> # Parental_Education_Level <chr>, Distance_From_Home <chr>, Gender <chr>, Exam_Score <dbl>
>
> # Example 2: Multiple conditions (AND using commas)
> # Students with low sleep (<6 hours) AND low motivation
> low_sleep_low_motivation <- student |>
+ filter(Sleep_Hours < 6, Motivation_Level == "Low")
> head(low_sleep_low_motivation)
# A tibble: 6 × 20
  Hours_Studied Attendance Parental_Involvement Access_to_Resources Extracurricular_Activities Sleep_Hours
  <dbl> <dbl> <chr> <chr> <chr> <chr> <dbl>
1 23     84 Low   High  NO    7
2 29     84 Medium Low   Yes   7
3 17     97 Medium High  Yes   6
4 18     66 High  High  NO    4
5 13     72 Low   Medium NO   6
6 21     65 Medium Low   Yes   7
# i abbreviated name: 'Extracurricular_Activities'
# i 14 more variables: Previous_Scores <dbl>, Motivation_Level <chr>, Internet_Access <chr>,
# Tutoring_sessions <dbl>, Family_Income <chr>, Teacher_Quality <chr>, School_Type <chr>,
# Peer_Influence <chr>, Physical_Activity <dbl>, Learning_Disabilities <chr>,
# Parental_Education_Level <chr>, Distance_From_Home <chr>, Gender <chr>, Exam_Score <dbl>
>
> Private
Error: object 'Private' not found
```

Show Traceback Rerun with Debug

```

>
>
> cat("Schools that are Public or Private:", nrow(school_type_filter), "\n")
Error: object 'school_type_filter' not found
```

Show Traceback Rerun with Debug

```

> # Example 3: Using %in% operator
> # Filter by school type: Public or Private
> school_type_filter <- student |>
+ filter(School_Type %in% c("Public_School", "Private_School"))
> view(school_type_filter)
> View(school_type_filter)
> View(school_type_filter)
> table(school_type_filter$School_Type)
< table of extent 0 >
>
```

Files Plots Packages Help Viewer Presentation

SENSEX 13:06 24-11-2025

MVLU COLLEGE.

RStudio Session 1 (Top):

```
R > View(school_type_filter)
> View(school_type_filter)
> View(school_type_filter)
> table(school_type_filter$school_Type)
<table of extent 0 x 1>
> View(high_exam_score)
> View(high_exam_score)
```

RStudio Session 2 (Bottom):

```
R > View(school_type_filter)
> View(school_type_filter)
> View(school_type_filter)
> table(school_type_filter$school_Type)
<table of extent 0 x 1>
> View(high_exam_score)
> View(high_exam_score)
```

NANDINI PANDIT S100
Data Analysis with SAS / SPSS /R Practical

MVLU COLLEGE.

The screenshot shows an RStudio interface with the following components:

- Top Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Project Explorer:** Shows multiple files: R_Prac4.R, school_type_filter, low_sleep_low_motivation, high_exam_score, sleep_or_extracurricular, high_stuc, and high_exam_score.
- Data View:** A data frame titled "Data" containing 4,363 entries and 20 columns. The columns include Peer_Influence, Physical_Activity, Learning_Disabilities, Parental_Education_Level, Distance_from_Home, Gender, and Exam_Score. The data shows various combinations of these variables, such as Neutral/No/Postgraduate/Near/Male/74, Negative/Yes/High School/Moderate/Male/71, etc.
- Environment View:** Shows the global environment with objects like high_exam_score, high_study_subset, low_sleep_low_motivation, Mental_health, my_data, sales_data, school_type_filter, sleep_or_extracurricular, student, and student.Mental.health.
- Console View:** Displays R code and its output. The session starts with "R 4.5.2 - ~/". It then lists several commands: View(school_type_filter), View(school_type_filter), View(school_type_filter), title(school_type_filter\$school_Type), table of extent 0, View(high_exam_score), View(high_exam_score), and View(high_exam_score). The console ends with a blank line " > |".
- Bottom Status Bar:** Shows system information including battery level (SENSEX +0.06%), network connection, and date/time (13:06, 24-11-2025).

MVLU COLLEGE.

RStudio Environment View (Top Window):

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores
1	24	98	Medium	Medium	Yes	7	91
2	23	98	Medium	Medium	Yes	8	71
3	26	91	Medium	Low	Yes	8	71
4	25	98	Low	High	Yes	5	56
5	22	98	Medium	High	No	5	69
6	22	98	Medium	Low	Yes	7	52
7	29	96	Medium	Low	Yes	7	72
8	29	92	Low	Medium	No	4	91
9	23	98	Medium	Low	Yes	7	54
10	22	99	Medium	Low	Yes	6	92
11	31	100	Medium	Medium	No	7	59
12	35	99	High	High	Yes	7	85
13	27	97	Low	High	Yes	8	81
14	21	96	Medium	Medium	No	5	64
15	21	91	High	Medium	No	6	100
16	22	95	Medium	High	Yes	6	69
17	31	92	Medium	Medium	Yes	4	76
18	22	95	Medium	High	No	9	88
19	29	96	Medium	Medium	Yes	9	55

RStudio Console View (Bottom Window):

```
R > view(school_type_filter)
> View(school_type_filter)
> View(school_type_filter)
> table(school_type_filter$school_Type)
> table(extent_0)
> View(high_exam_score)
> View(high_exam_score)
> |
```

NANDINI PANDIT S100
Data Analysis with SAS / SPSS /R Practical