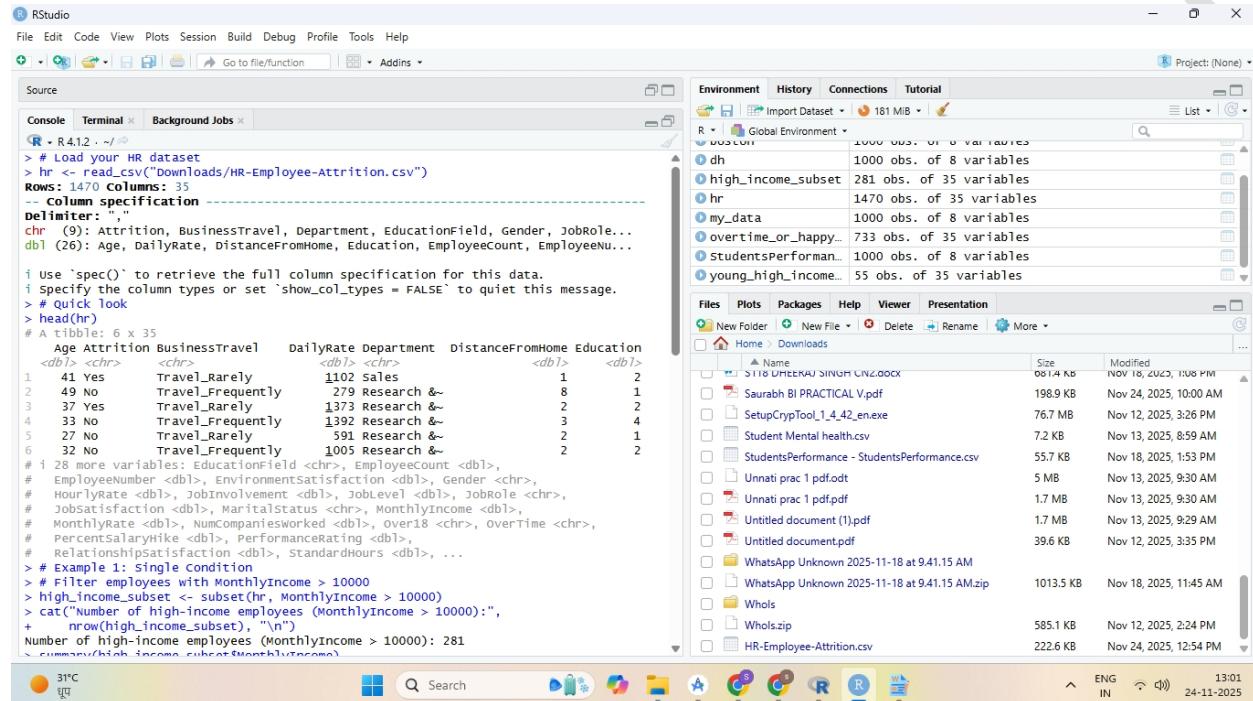


PRACTICAL NO 4

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



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Terminal Background Jobs
# Load your HR dataset
hr <- read_csv("Downloads/HR-Employee-Attrition.csv")
Rows: 1470 Columns: 35
-- Column specification --
Delimiter: ","
chr (9): Attrition, BusinessTravel, Department, EducationField, Gender, JobRole...
dbl (26): Age, DailyRate, DistanceFromHome, Education, EmployeeCount, EmployeeNu...
# A tibble: 6 x 35
# ... with 28 more variables: EducationField <chr>, Employeecount <dbl>,
#   EmployeeNumber <dbl>, Environmentsatisfaction <dbl>, Gender <chr>,
#   HourlyRate <dbl>, JobInvolvement <dbl>, JobLevel <dbl>, Jobrole <chr>,
#   Jobsatisfaction <dbl>, Maritalstatus <chr>, Monthlyincome <dbl>,
#   Monthlyrate <dbl>, NumCompaniesworked <dbl>, over18 <chr>, Overtime <chr>,
#   PercentSalaryhike <dbl>, PerformanceRating <dbl>,
#   Relationshipsatisfaction <dbl>, Standardhours <dbl>, ...
# Example 1: Single condition
# Filter employees with MonthlyIncome > 10000
high_income_subset <- subset(hr, MonthlyIncome > 10000)
cat("Number of high-income employees (MonthlyIncome > 10000):",
+ nrow(high_income_subset), "\n")
Number of high-income employees (MonthlyIncome > 10000): 281
summary(high_income_subset$MonthlyIncome)

```

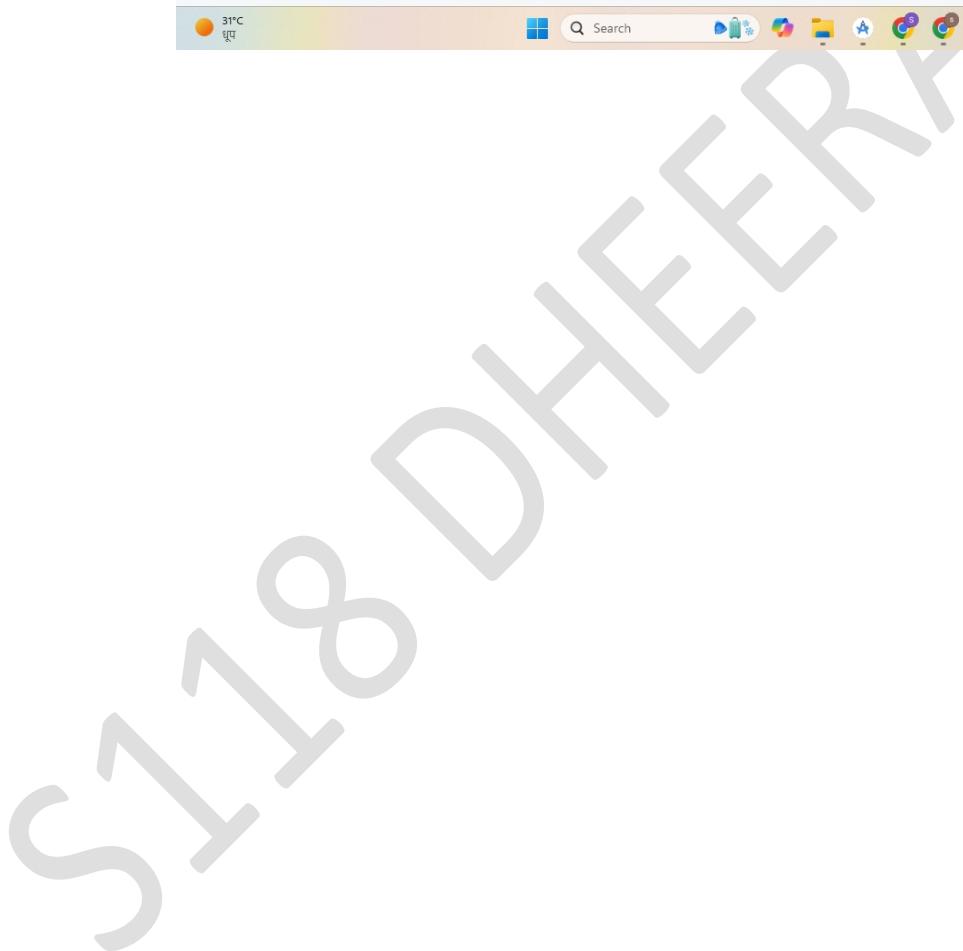
```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Terminal Background Jobs
# Filter employees with MonthlyIncome > 10000
high_income_subset <- subset(hr, MonthlyIncome > 10000)
cat("Number of high-income employees (MonthlyIncome > 10000):",
+ nrow(high_income_subset), "\n")
Number of high-income employees (MonthlyIncome > 10000): 281
summary(high_income_subset$MonthlyIncome)
Min. 1st Qu. Median Mean 3rd Qu. Max.
10000 11245 14118 14826 17875 19999
# Example 2: Multiple Conditions (AND)
# Employees with MonthlyIncome > 10000 AND YearsAtCompany < 5
young_high_income_subset <- subset(hr,
+                         MonthlyIncome > 10000 & YearsAtCompany < 5)
cat("Employees earning >10000 AND less than 5 years in company:",
+ nrow(young_high_income_subset), "\n")
Employees earning >10000 AND less than 5 years in company: 55
head(young_high_income_subset)
# A tibble: 6 x 35
# ... with 28 more variables: EducationField <chr>, Employeecount <dbl>,
#   EmployeeNumber <dbl>, Environmentsatisfaction <dbl>, Gender <chr>,
#   HourlyRate <dbl>, JobInvolvement <dbl>, JobLevel <dbl>, Jobrole <chr>,
#   Jobsatisfaction <dbl>, Maritalstatus <chr>, Monthlyincome <dbl>,
#   Monthlyrate <dbl>, NumCompaniesworked <dbl>, over18 <chr>, Overtime <chr>,
#   PercentSalaryhike <dbl>, PerformanceRating <dbl>,
#   Relationshipsatisfaction <dbl>, StandardHours <dbl>, ...
# Example 3: Multiple Conditions (OR)
# Employees with Overtime == "Yes" OR JobSatisfaction > 3

```

Sheth L.U.J. College of Arts & Sir M.V. College Of Science & Commerce

SUBJECT NAME: Data Analysis with SAS / SPSS / R



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
R > R 4.1.2 : ~
5 59 No Non_Travel 1420 Human Reso~ 2 4
6 51 Yes Travel_Frequently 1150 Research & 8 4
# 28 more variables: EducationField <chr>, EmployeeCount <dbl>,
# EmployeeNumber <dbl>, EnvironmentSatisfaction <dbl>, Gender <chr>,
# HourlyRate <dbl>, JobInvolvement <dbl>, JobLevel <dbl>, JobRole <chr>,
# Jobsatisfaction <dbl>, MaritalStatus <chr>, MonthlyIncome <dbl>,
# MonthlyRate <dbl>, NumCompaniesWorked <dbl>, Over18 <chr>, Overtime <chr>,
# PercentSalaryHike <dbl>, PerformanceRating <dbl>,
# RelationshipSatisfaction <dbl>, StandardHours <dbl>, ...
> # Example 3: Multiple Conditions (OR)
> # Employees with overtime == "Yes" OR jobsatisfaction > 3
> overtime_or_happy_subset <- subset(hr,
+                                     Overtime == "yes" | jobsatisfaction > 3)
> cat("Employees working overtime or highly satisfied:")
+     nrow(overtime_or_happy_subset), "\n")
Employees working overtime or highly satisfied: 733
> head(overtime_or_happy_subset)
# A tibble: 6 x 35
   Age Attrition BusinessTravel DailyRate Department DistanceFromHome Education
   <dbl> <chr> <chr> <dbl> <chr> <dbl> <dbl>
1 41 Yes Travel_Rarely 1102 Sales 1 2
2 37 Yes Travel_Rarely 1373 Research & 2 2
3 33 No Travel_Frequently 1392 Research & 3 4
4 32 No Travel_Frequently 1005 Research & 2 2
5 59 No Travel_Rarely 1324 Research & 3 3
6 29 No Travel_Rarely 153 Research & 15 2
# 18 more variables: EducationField <chr>, EmployeeCount <dbl>,
# EmployeeNumber <dbl>, EnvironmentSatisfaction <dbl>, Gender <chr>,
# HourlyRate <dbl>, JobInvolvement <dbl>, JobLevel <dbl>, JobRole <chr>,
# Jobsatisfaction <dbl>, MaritalStatus <chr>, MonthlyIncome <dbl>,
# MonthlyRate <dbl>, NumCompaniesWorked <dbl>, Over18 <chr>, Overtime <chr>,
# PercentSalaryHike <dbl>, PerformanceRating <dbl>,
# RelationshipSatisfaction <dbl>, StandardHours <dbl>, ...
> |
```

Environment History Connections Tutorial

R Global Environment 1000 obs. of 35 variables

- dh 1000 obs. of 8 variables
- high_income_subset 281 obs. of 35 variables
- hr 1470 obs. of 35 variables
- my_data 1000 obs. of 8 variables
- overtime_or_happy... 733 obs. of 35 variables
- Studentsperforman... 1000 obs. of 8 variables
- young_high_income... 55 obs. of 35 variables

Files Plots Packages Help Viewer Presentation

Home Downloads

Name	Size	Modified
S118 DHEERAJ SINGH LIN2.QQX	9814 KB	NOV 18, 2025, 10:08 PM
Saurabh BI PRACTICAL V.pdf	198.9 KB	Nov 24, 2025, 10:00 AM
SetupCrypTool_1_4_42_en.exe	76.7 MB	Nov 12, 2025, 3:26 PM
Student Mental health.csv	7.2 KB	Nov 13, 2025, 8:59 AM
StudentsPerformance - StudentsPerformance.csv	55.7 KB	Nov 18, 2025, 1:53 PM
Unnati prac 1.pdf.odt	5 MB	Nov 13, 2025, 9:30 AM
Unnati prac 1.pdf.pdf	1.7 MB	Nov 13, 2025, 9:30 AM
Untitled document (1).pdf	1.7 MB	Nov 13, 2025, 9:29 AM
Untitled document.pdf	39.6 KB	Nov 12, 2025, 3:35 PM
WhatsApp Unknown 2025-11-18 at 9:41:15 AM	1013.5 KB	Nov 18, 2025, 11:45 AM
WhatsApp Unknown 2025-11-18 at 9:41:15 AM.zip		
Whols	585.1 KB	Nov 12, 2025, 2:24 PM
Whols.zip	222.6 KB	Nov 24, 2025, 12:54 PM

31°C ENG IN 13:02 24-11-2025