

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

Practical N0 4

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

```

#RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins

Source
Console Terminal Background Jobs
[R - R4.52 - /]

> library(dplyr)
> library(readr)
> jobs <- read.csv("ai_job_market.csv")
Rows: 2000 Columns: 12
-- Column specification --
Delimiter: ","
chr (10): company_name, industry, job_title, skills_required, experience_level, employment...
dbl (1): job_id
date (1): posted_date

# use 'spec()' to retrieve the full column specification for this data.
# Specify the column types or set 'show_col_types = FALSE' to quiet this message.
> # A tibble: 6 × 12
  job_id company_name      industry job_title skills_required experience_level employment_type
    <dbl> <chr>        <chr>     <chr>       <chr>           <chr>
1 1 Foster and Sons  Healthcare Data Ana. NumPy, Reinfor... Mid      Full-time
2 2 Boyd, Myers and R. Tech Computer Scikit-learn, ... Senior   Full-time
3 3 King Inc          Tech Quant Re. Mflow, FastAP... Entry   Full-time
4 4 Cooper, Archer an. Tech AI Produ. Scikit-learn, ... Mid      Full-time
5 5 Hall LLC          Finance Data Sci. Excel, Keras, ... Senior   Contract
6 6 Ellis PLC         E-comm... AI Produ. GCP, Excel, Sc... Senior   Remote
# i 5 more variables: location <chr>, salary_range_usd <chr>, posted_date <date>,
# company_size <chr>, tools_preferred <chr>
> high_salary <- jobs[>
+   filter(salary_range_usd > 100000,
+         experience_level == "Senior"),
+   cat("High salary + senior roles:", nrow(high_salary), "\n")
Number of high-salary jobs: 1093
> head(high_salary)
# A tibble: 6 × 12
  job_id company_name      industry job_title skills_required experience_level employment_type
    <dbl> <chr>        <chr>     <chr>       <chr>           <chr>
1 1 Foster and Sons  Healthcare Data Ana. NumPy, Reinfor... Mid      Full-time
2 2 Boyd, Myers and R. Tech Computer Scikit-learn, ... Senior   Full-time
3 4 Cooper, Archer an. Tech AI Produ. Scikit-learn, ... Mid      Full-time
4 5 Hall LLC          Finance Data Sci. Excel, Keras, ... Senior   Contract
5 6 Ellis PLC         E-comm... AI Produ. GCP, Excel, Sc... Senior   Remote
6 7 Matthews-Moses  Automot... Data Ana. Python, Flask, ... Mid      Full-time
# i 5 more variables: location <chr>, salary_range_usd <chr>, posted_date <date>,
# company_size <chr>, tools_preferred <chr>

[R - R4.52 - /]

> high_salary.senior <- jobs[>
+   filter(salary_range_usd == 100000,
+         experience_level == "Senior"),
+   cat("High salary + senior roles:", nrow(high_salary.senior), "\n")
High salary + senior roles: 336
> head(high_salary.senior)
# A tibble: 6 × 12
  job_id company_name      industry job_title skills_required experience_level employment_type
    <dbl> <chr>        <chr>     <chr>       <chr>           <chr>
1 2 Boyd, Myers and R. Tech Computer Scikit-learn, ... Senior   Full-time
2 5 Hall LLC          Finance Data Sci. Excel, Keras, ... Senior   Contract
3 6 Ellis PLC         E-comm... AI Produ. GCP, Excel, Sc... Senior   Remote
4 10 Parks LLC        Automot... Computer Excel, Reinfor... Senior   Internship
5 24 Thomas, Hurst and. Tech NLP Engi. Pytorch, CUDA, ... Senior   Internship
6 27 Le PLC            Educati... Quant Re. Azure, Flask, ... Senior   Internship
# i 5 more variables: location <chr>, salary_range_usd <chr>, posted_date <date>,
# company_size <chr>, tools_preferred <chr>
> big_companies <- jobs[>
+   filter(company_size > 500),
+   cat("Number of big companies:", nrow(big_companies), "\n")
Number of big companies: 2000
> head(big_companies)
# A tibble: 6 × 12
  job_id company_name      industry job_title skills_required experience_level employment_type
    <dbl> <chr>        <chr>     <chr>       <chr>           <chr>
1 1 Foster and Sons  Healthcare Data Ana. NumPy, Reinfor... Mid      Full-time
2 2 Boyd, Myers and R. Tech Computer Scikit-learn, ... Senior   Full-time
3 3 King Inc          Tech Quant Re. Mflow, FastAP... Entry   Full-time
4 4 Cooper, Archer an. Tech AI Produ. Scikit-learn, ... Mid      Full-time
5 5 Hall LLC          Finance Data Sci. Excel, Keras, ... Senior   Contract
6 6 Ellis PLC         E-comm... AI Produ. GCP, Excel, Sc... Senior   Remote
# i 5 more variables: location <chr>, salary_range_usd <chr>, posted_date <date>,
# company_size <chr>, tools_preferred <chr>
> us.jobs <- jobs[>
+   filter(grep("United States", location, ignore.case = TRUE)),
+   cat("Jobs located in the US:", nrow(us.jobs), "\n")
Jobs located in the US: 10
> head(us.jobs)
# A tibble: 6 × 12
  job_id company_name      industry job_title skills_required experience_level employment_type
    <dbl> <chr>        <chr>     <chr>       <chr>           <chr>
1 1 Foster and Sons  Healthcare Data Ana. NumPy, Reinfor... Mid      Full-time
2 2 Boyd, Myers and R. Tech Computer Scikit-learn, ... Senior   Full-time
3 3 King Inc          Tech Quant Re. Mflow, FastAP... Entry   Full-time
4 4 Cooper, Archer an. Tech AI Produ. Scikit-learn, ... Mid      Full-time
5 5 Hall LLC          Finance Data Sci. Excel, Keras, ... Senior   Contract
6 6 Ellis PLC         E-comm... AI Produ. GCP, Excel, Sc... Senior   Remote
# i 5 more variables: location <chr>, salary_range_usd <chr>, posted_date <date>,
# company_size <chr>, tools_preferred <chr>

```

Name :- Vinod Mali

Roll No :- S092

Subject :- Data Analysis with SAS / SPSS /R

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

Practical N0 4

The screenshot shows the RStudio IDE interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help, and Addins. The title bar says "RStudio". The left sidebar has tabs for Source, Console, Terminal, and Background Jobs. The main console window displays R code and its output. The code involves filtering job data based on company size, industry, and location. It also creates variables like 'big_companies', 'high_salary', and 'industry_Filter'. The right sidebar shows the Global Environment pane with various objects listed, such as 'high_salary' (1093 obs. of 12 variables), 'high_salary_senior' (336 obs. of 12 variables), and 'jobs' (2000 obs. of 12 variables). Below the environment is a file browser showing local files like 'prac_1.txt', 'Practical_No_3.R', and 'Processed_Student_Mental_Health.csv'. The bottom status bar shows system information including battery level, signal strength, and date/time.

Name :- Vinod Mali
Roll No :- S092
Subject :- Data Analysis with SAS / SPSS /R