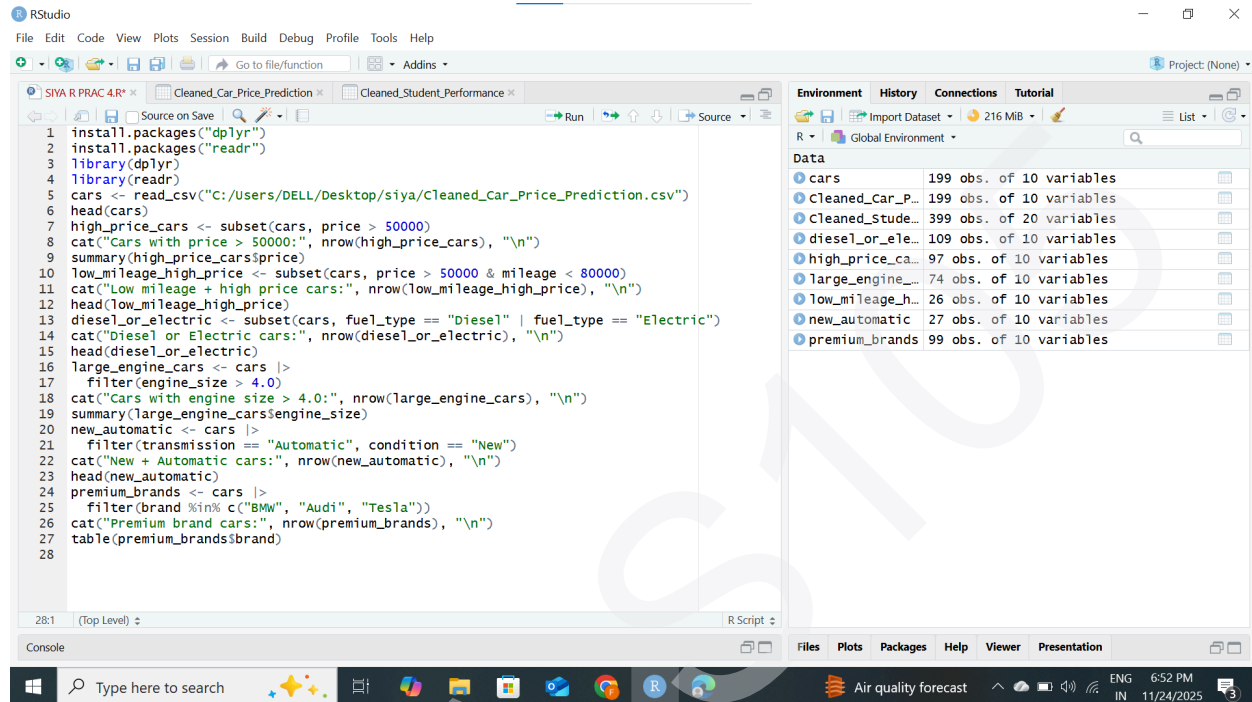


MVLU COLLEGE

PRACTICAL NO. 4

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

#CODE:



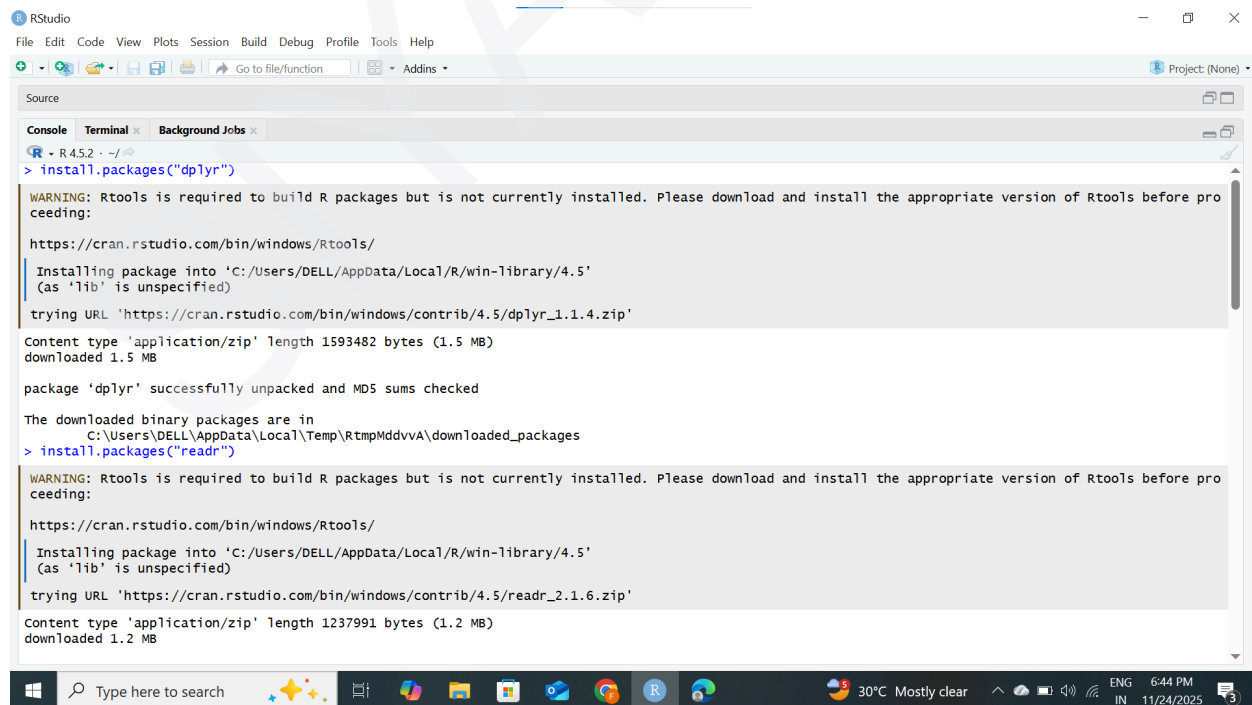
The screenshot shows the RStudio interface with a script editor on the left and the Environment pane on the right. The script editor contains R code for loading data, filtering cars by price, mileage, fuel type, engine size, and transmission, and identifying premium brands. The Environment pane shows the objects created in the global environment.

```
1 install.packages("dplyr")
2 install.packages("readr")
3 library(dplyr)
4 library(readr)
5 cars <- read_csv("C:/Users/DELL/desktop/siya/Cleaned_Car_Price_Prediction.csv")
6 head(cars)
7 high_price_cars <- subset(cars, price > 50000)
8 cat("Cars with price > 50000:", nrow(high_price_cars), "\n")
9 summary(high_price_cars$price)
10 low_mileage_high_price <- subset(cars, price > 50000 & mileage < 80000)
11 cat("Low mileage + high price cars:", nrow(low_mileage_high_price), "\n")
12 head(low_mileage_high_price)
13 diesel_or_electric <- subset(cars, fuel_type == "Diesel" | fuel_type == "Electric")
14 cat("Diesel or Electric cars:", nrow(diesel_or_electric), "\n")
15 head(diesel_or_electric)
16 large_engine_cars <- cars |>
17   filter(engine_size > 4.0)
18 cat("Cars with engine size > 4.0:", nrow(large_engine_cars), "\n")
19 summary(large_engine_cars$engine_size)
20 new_automatic <- cars |>
21   filter(transmission == "Automatic", condition == "New")
22 cat("New + Automatic cars:", nrow(new_automatic), "\n")
23 head(new_automatic)
24 premium_brands <- cars |>
25   filter(brand %in% c("BMW", "Audi", "Tesla"))
26 cat("Premium brand cars:", nrow(premium_brands), "\n")
27 table(premium_brands$brand)
28
```

Environment pane data:

Object	Variables	Observations
cars	10	199
Cleaned_Car_P...	10	199
Cleaned_Stude...	20	399
diesel_or_ele...	10	109
high_price_ca...	10	97
large_engine_...	10	74
low_mileage_h...	10	26
new_automatic	10	27
premium_brands	10	99

#OUTPUT:

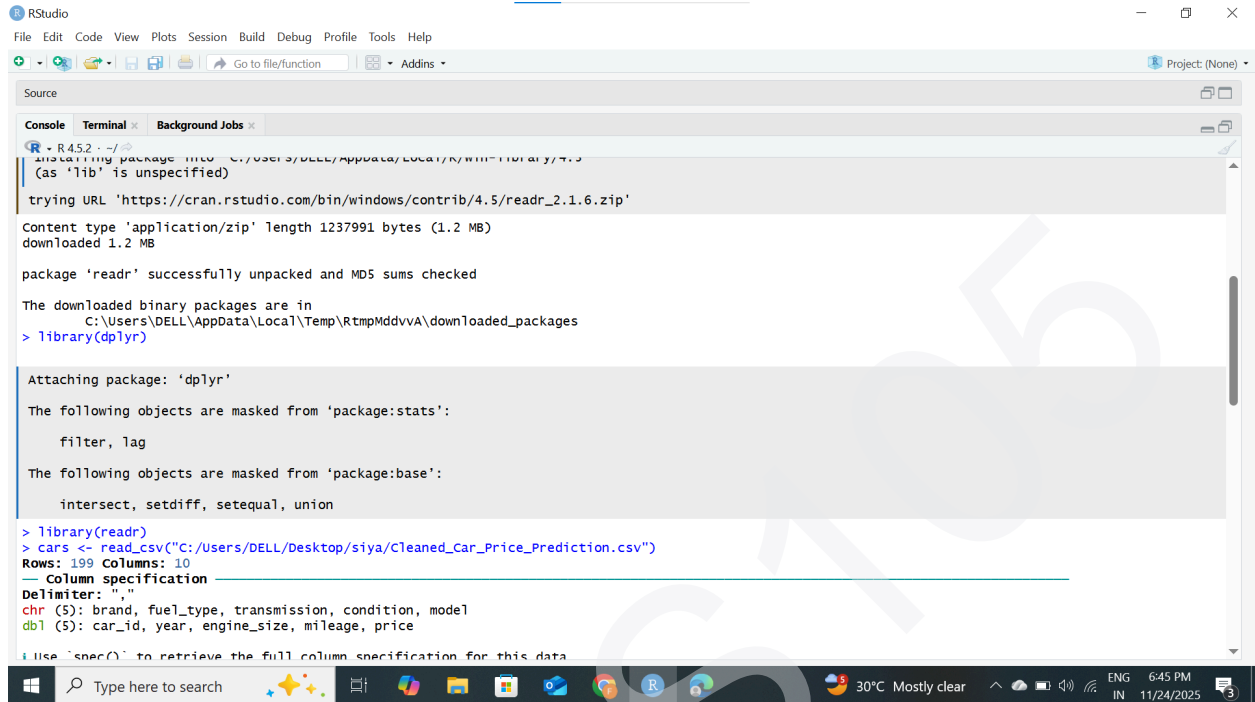


The screenshot shows the RStudio console output for the R code. It displays the installation of the 'dplyr' and 'readr' packages, followed by the execution of the code to filter cars by price, mileage, fuel type, engine size, and transmission, and to identify premium brands.

```
R - R 4.5.2 - ~/bin
> 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/DELL/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/dplyr_1.1.4.zip'
Content type 'application/zip' length 1593482 bytes (1.5 MB)
downloaded 1.5 MB
package 'dplyr' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:/Users/DELL/AppData/Local/Temp/RtmpMddvva/downloaded_packages
> install.packages("readr")
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/DELL/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/readr_2.1.6.zip'
Content type 'application/zip' length 1237991 bytes (1.2 MB)
downloaded 1.2 MB
```

MVLU COLLEGE

PRACTICAL NO. 4



```
R - R4.5.2 - ~/R
Installing package into 'C:/Users/DELL/AppData/Local/R/win-library/7.3'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/readr_2.1.6.zip'
Content type 'application/zip' length 1237991 bytes (1.2 MB)
downloaded 1.2 MB
package 'readr' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:\Users\DELL\AppData\Local\Temp\RtmpMddvVA\downloaded_packages
> 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

> library(readr)
> cars <- read_csv("C:/Users/DELL/Desktop/siya/Cleaned_Car_Price_Prediction.csv")
Rows: 199 Columns: 10
  Column specification
Delimiter: ","
chr (5): brand, fuel_type, transmission, condition, model
dbl (5): car_id, year, engine_size, mileage, price
Use `spec()` to retrieve the full column specification for this data
```

MVLU COLLEGE

PRACTICAL NO. 4

```
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 - R4.5.2 - /...
Column specification
Delimiter: ,
chr (5): brand, fuel_type, transmission, condition, model
dbl (5): car_id, year, engine_size, mileage, price

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.
> head(cars)
# A tibble: 6 x 10
  car_id brand    year engine_size fuel_type transmission mileage condition price model
  <dbl> <chr>    <dbl>    <dbl> <chr>    <chr>    <dbl> <chr>    <dbl> <chr>
1     1 Tesla    2016      2.3 Petrol    Manual    114832 New    26614. Model X
2     2 BMW      2018      4.4 Electric Manual    143190 Used    14680. 5 Series
3     3 Audi      2013      4.5 Electric Manual    181601 New    44403. A4
4     4 Tesla    2011      4.1 Diesel    Automatic 68682 New    86374. Model Y
5     5 Ford      2009      2.6 Diesel    Manual    223009 Like New 73577. Mustang
6     6 Audi      2019      2.4 Diesel    Automatic 246553 Like New 88970. Q7

> high_price_cars <- subset(cars, price > 50000)
> cat("Cars with price > 50000:", nrow(high_price_cars), "\n")
Cars with price > 50000: 97
> summary(high_price_cars$price)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
50288  66045  73779   75105  85267  99372
> low_mileage_high_price <- subset(cars, price > 50000 & mileage < 80000)
> cat("Low mileage + high price cars:", nrow(low_mileage_high_price), "\n")
Low mileage + high price cars: 26
> head(low_mileage_high_price)
# A tibble: 6 x 10
  car_id brand    year engine_size fuel_type transmission mileage condition price model
  <dbl> <chr>    <dbl>    <dbl> <chr>    <chr>    <dbl> <chr>    <dbl> <chr>
1     4 Tesla    2011      4.1 Diesel    Automatic 68682 New    86374. Model Y
2    22 Mercedes 2022      2.3 Electric Manual    12150 Used    61393. E-Class
```

```
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 - R4.5.2 - /...
> head(low_mileage_high_price)
# A tibble: 6 x 10
  car_id brand    year engine_size fuel_type transmission mileage condition price model
  <dbl> <chr>    <dbl>    <dbl> <chr>    <chr>    <dbl> <chr>    <dbl> <chr>
1     4 Tesla    2011      4.1 Diesel    Automatic 68682 New    86374. Model Y
2    22 Mercedes 2022      2.3 Electric Manual    12150 Used    61393. E-Class
3    26 Toyota    2017      5.2 Electric Automatic 18325 Used    70177. Camry
4    29 Mercedes 2007      5.9 Diesel    Manual    17669 Used    78308. GLC
5    33 Toyota    2022      1.5 Hybrid    Automatic 50812 Used    92010. Corolla
6    37 Honda     2014      2.3 Diesel    Manual    25395 New    76380. Fit

> diesel_or_electric <- subset(cars, fuel_type == "Diesel" | fuel_type == "Electric")
> cat("Diesel or Electric cars:", nrow(diesel_or_electric), "\n")
Diesel or Electric cars: 109
> head(diesel_or_electric)
# A tibble: 6 x 10
  car_id brand    year engine_size fuel_type transmission mileage condition price model
  <dbl> <chr>    <dbl>    <dbl> <chr>    <chr>    <dbl> <chr>    <dbl> <chr>
1     2 BMW      2018      4.4 Electric Manual    143190 Used    14680. 5 Series
2     3 Audi      2013      4.5 Electric Manual    181601 New    44403. A4
3     4 Tesla    2011      4.1 Diesel    Automatic 68682 New    86374. Model Y
4     5 Ford      2009      2.6 Diesel    Manual    223009 Like New 73577. Mustang
5     6 Audi      2019      2.4 Diesel    Automatic 246553 Like New 88970. Q7
6     7 Audi      2020      4.0 Electric Automatic 135486 Used    63499. Q5

Error in large_engine_cars :
  The pipe operator requires a function call as RHS (<input>:2:1)

> cat("Cars with engine size > 4.0:", nrow(large_engine_cars), "\n")
```

MVLU COLLEGE

PRACTICAL NO. 4

```
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 4.5.2 - ~/

> large_engine_cars <- cars |>
+ filter(engine_size > 4.0)
> cat("Cars with engine size > 4.0:", nrow(large_engine_cars), "\n")
Cars with engine size > 4.0: 74
> summary(large_engine_cars$engine_size)
      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 4.100   4.425   4.750   4.918   5.400   5.900
> new_automatic <- cars |>
+ filter(transmission == "Automatic", condition == "New")
> cat("New + Automatic cars:", nrow(new_automatic), "\n")
New + Automatic cars: 27
> head(new_automatic)
# A tibble: 6 x 10
  car_id brand  year engine_size fuel_type transmission mileage condition price model
  <dbl> <chr>  <dbl>   <dbl> <chr>   <chr>      <dbl> <chr>   <dbl> <chr>
1     4 Tesla  2011     4.1 Diesel Automatic    68682 New    86374. Model Y
2     8 Tesla  2017     5.3 Hybrid Automatic    83030 New   17381. Model Y
3    13 Ford   2006     4.7 Petrol Automatic   114360 New    74766. Fiesta
4    15 BMW    2014     2 Electric Automatic    65018 New    35221. X3
5    17 Mercedes 2017     4.5 Electric Automatic   136817 New   14728. GLA
6    19 Honda  2011     3 Electric Automatic    86984 New    47792. Civic
> premium_brands <- cars |>
+ filter(brand %in% c("BMW", "Audi", "Tesla"))
> cat("Premium brand cars:", nrow(premium_brands), "\n")
Premium brand cars: 99
> table(premium_brands$brand)

Audi  BMW  Tesla
 28    34    37
> |
```