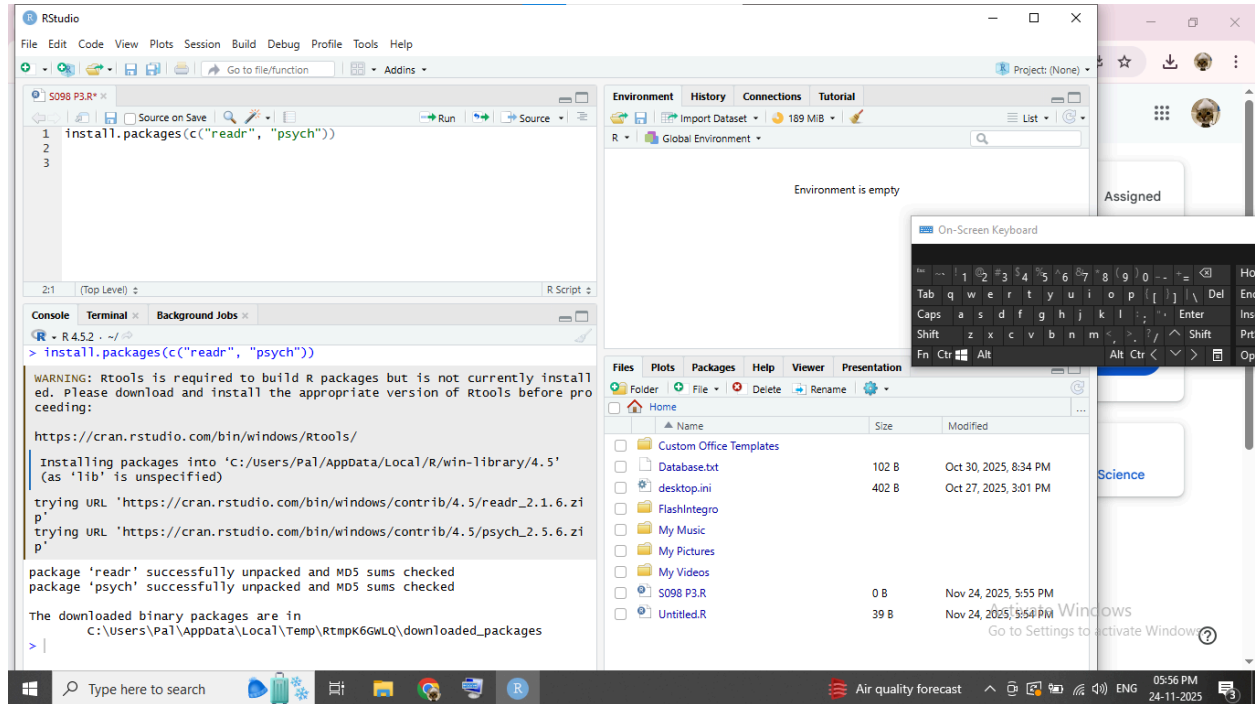


# MVLU COLLEGE

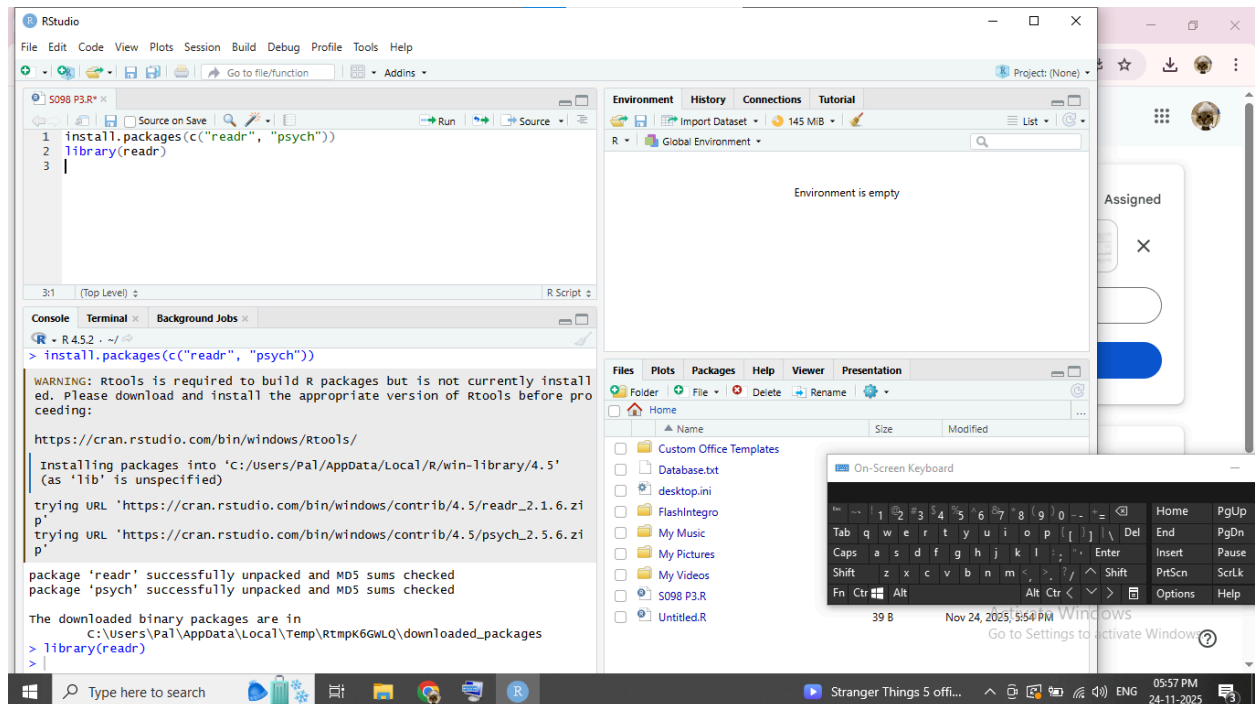
## Subject:-Data Analysis with SAS / SPSS /R

Aim:- Exploring data: View() or print() (R).

`install.packages(c("readr", "psych"))`



`library(readr)`

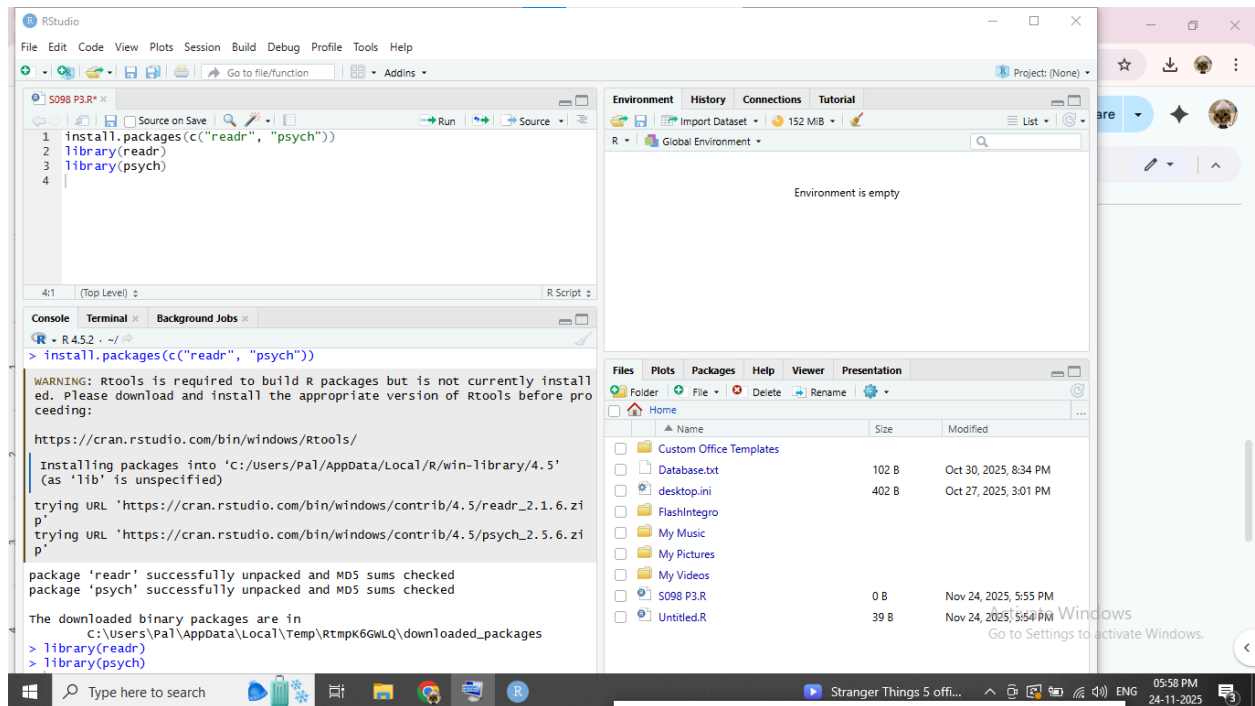


Vikas Pal  
S098

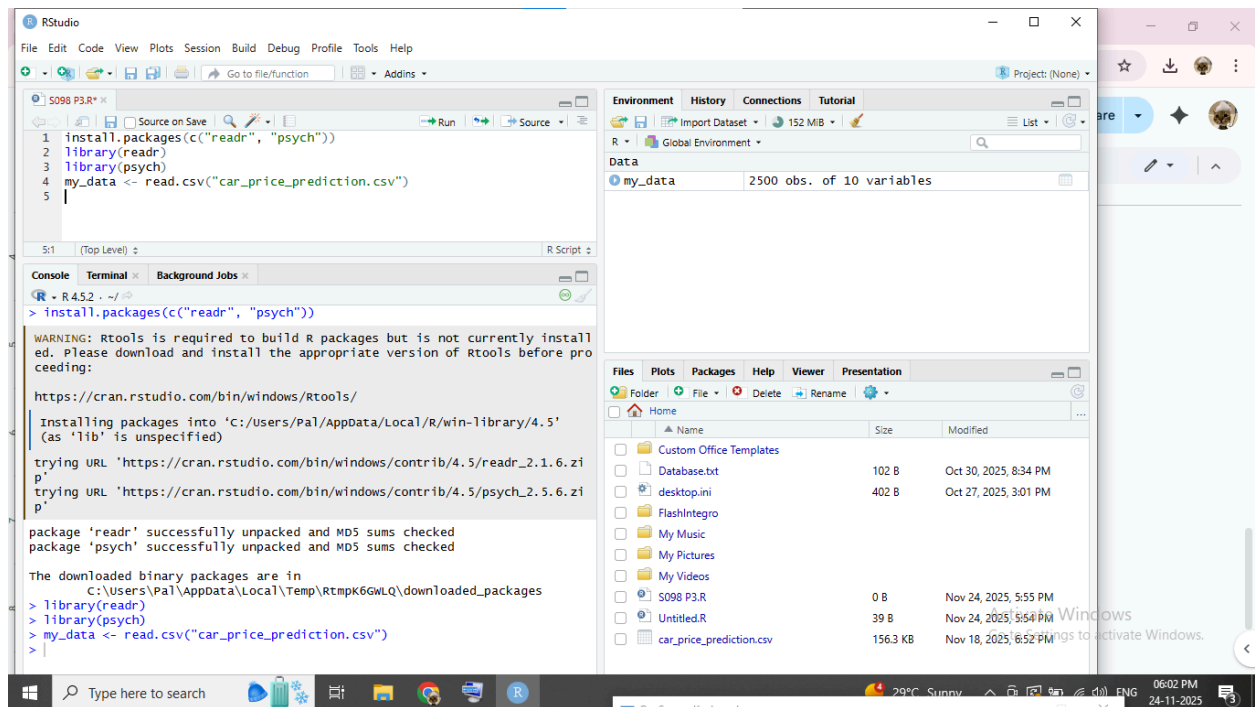
# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

library(psych)



my\_data <- read.csv("car\_price\_prediction.csv")



Vikas Pal  
S098

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

head(my\_data)

The screenshot shows the RStudio interface. The console window displays the following output:

```
> install.packages(c("readr", "psych"))
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 packages into 'C:/Users/Pal/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/readr_2.1.6.6.zip'
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/psych_2.5.6.6.zip'
package 'readr' successfully unpacked and MD5 sums checked
package 'psych' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:/Users/Pal/AppData/Local/Temp/RtmpK6GWLQ/downloaded_packages
> library(readr)
> library(psych)
> my_data <- read.csv("car_price_prediction.csv")
> head(my_data)
```

The Environment pane shows the 'my\_data' object with 2500 observations and 10 variables.

tail(my\_data)

The screenshot shows the RStudio interface. The console window displays the following output:

```
> install.packages(c("readr", "psych"))
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 packages into 'C:/Users/Pal/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/readr_2.1.6.6.zip'
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/psych_2.5.6.6.zip'
package 'readr' successfully unpacked and MD5 sums checked
package 'psych' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:/Users/Pal/AppData/Local/Temp/RtmpK6GWLQ/downloaded_packages
> library(readr)
> library(psych)
> my_data <- read.csv("car_price_prediction.csv")
> tail(my_data)
```

The Environment pane shows the 'my\_data' object with 2500 observations and 10 variables.

Car.ID	Brand	Year	Engine.Size	Fuel.Type	Transmission	Mileage
2495	Mercedes	2001	2.3	Petrol	Manual	162586
2496	Audi	2020	2.4	Petrol	Automatic	22650
2497	Audi	2001	5.7	Hybrid	Manual	77701
2498	Ford	2021	1.1	Hybrid	Manual	272827
2499	Audi	2002	4.5	Diesel	Manual	229164
2500	Toyota	2005	4.6	Diesel	Automatic	80978

Vikas Pal  
S098

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

dim(my\_data)

The screenshot shows the RStudio interface. The console window displays the following output:

```
R - R 4.5.2 ~ / ~  
Installing packages into 'C:/Users/Pal/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'  
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/psych_2.5.6.zip'  
package 'readr' successfully unpacked and MD5 sums checked  
package 'psych' successfully unpacked and MD5 sums checked  
The downloaded binary packages are in  
C:/Users/Pal/AppData/Local/Temp/RtmpK6WLQ/downloaded_packages  
> library(readr)  
> library(psych)  
> my_data <- read_csv("car_price_prediction.csv")  
>  
> my_data <- read_csv("car_price_prediction.csv")  
> tail(my_data)  
  Car.ID  Brand Year Engine.Size Fuel.Type Transmission Mileage  
2495 2495 Mercedes 2001      2.3   Petrol      Manual    162586  
2496 2496   Audi  2020      2.4   Petrol    Automatic    22650  
2497 2497   Audi  2001      5.7   Hybrid      Manual    77701  
2498 2498   Ford  2021      1.1   Hybrid      Manual    272827  
2499 2499   Audi  2002      4.5   Diesel      Manual    229164  
2500 2500  Toyota 2005      4.6   Diesel    Automatic     80978  
  Condition Price Model  
2495   Used 90378.98 E=Class  
2496 Like New 61384.10 Q5  
2497 Like New 24710.35 A3  
2498 Like New 29902.45 Fiesta  
2499 Like New 46085.67 Q5  
2500   Used 16594.14 RAV4  
>  
> dim(my_data)  
[1] 2500 10
```

The Environment pane on the right shows the variable `my_data` with 2500 observations and 10 variables.

cat("Dimensions (Rows, Columns): ", dim(my\_data), "\n")

The screenshot shows the RStudio interface with a script editor containing the following code:

```
1 install.packages(c("readr", "psych"))  
2 library(readr)  
3 library(psych)  
4 my_data <- read_csv("car_price_prediction.csv")  
5 head(my_data)  
6 tail(my_data)  
7 dim(my_data)  
8 cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")  
9
```

The console window displays the output of the script:

```
R - R 4.5.2 ~ / ~  
my_data <- read_csv("car_price_prediction.csv")  
> tail(my_data)  
  Car.ID  Brand Year Engine.Size Fuel.Type Transmission Mileage  
2495 2495 Mercedes 2001      2.3   Petrol      Manual    162586  
2496 2496   Audi  2020      2.4   Petrol    Automatic    22650  
2497 2497   Audi  2001      5.7   Hybrid      Manual    77701  
2498 2498   Ford  2021      1.1   Hybrid      Manual    272827  
2499 2499   Audi  2002      4.5   Diesel      Manual    229164  
2500 2500  Toyota 2005      4.6   Diesel    Automatic     80978  
  Condition Price Model  
2495   Used 90378.98 E=Class  
2496 Like New 61384.10 Q5  
2497 Like New 24710.35 A3  
2498 Like New 29902.45 Fiesta  
2499 Like New 46085.67 Q5  
2500   Used 16594.14 RAV4  
>  
> dim(my_data)  
[1] 2500 10  
> cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")  
Dimensions (Rows, Columns): 2500 10  
>
```

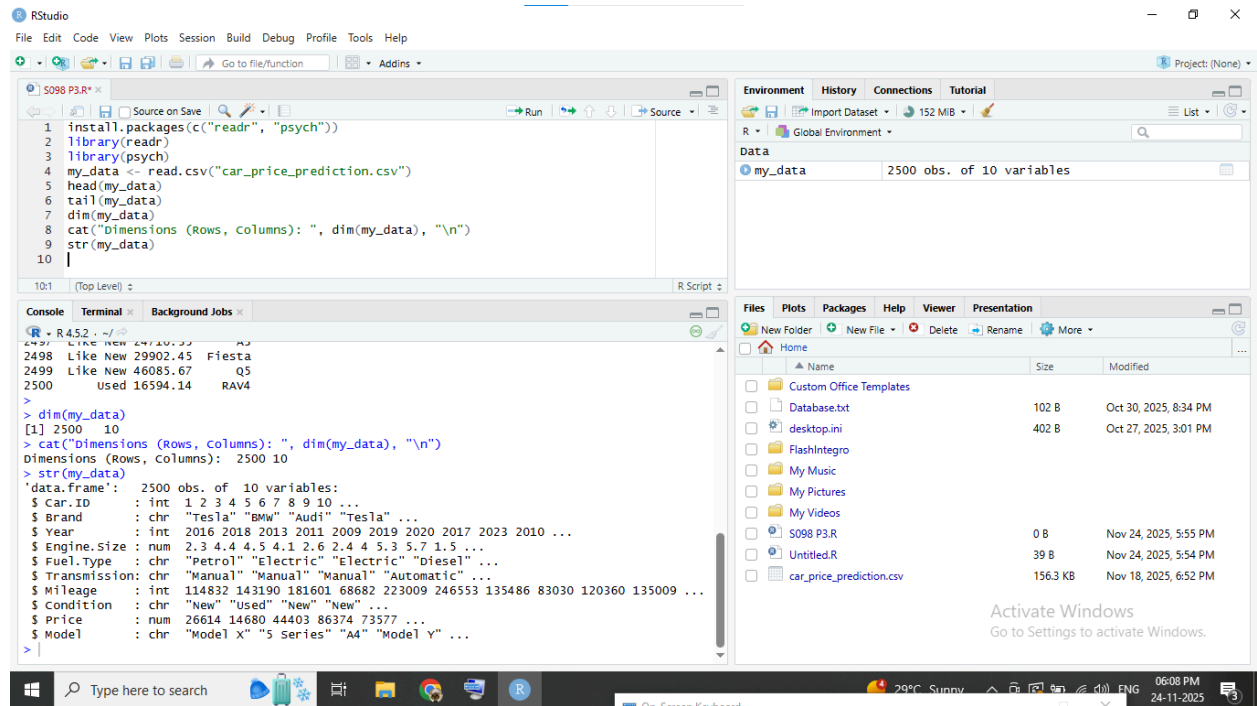
The Environment pane on the right shows the variable `my_data` with 2500 observations and 10 variables.

Vikas Pal  
S098

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

str(my\_data)



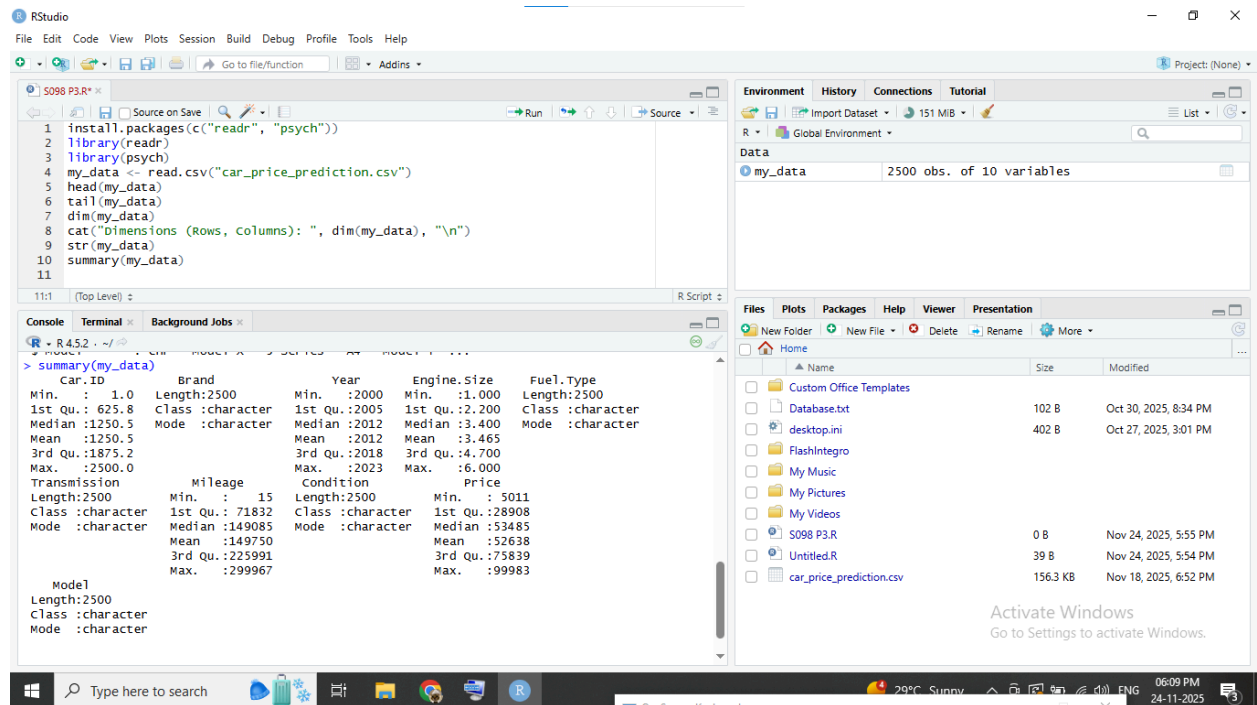
The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for loading packages, reading a CSV file, and using the `str()` function on `my_data`.
- Environment:** Shows `my_data` as a data frame with 2500 observations and 10 variables.
- Console:** Displays the output of the `str(my_data)` command, showing the structure of the data frame with 2500 observations and 10 variables.

```
1 install.packages(c("readr", "psych"))
2 library(readr)
3 library(psych)
4 my_data <- read.csv("car_price_prediction.csv")
5 head(my_data)
6 tail(my_data)
7 dim(my_data)
8 cat("Dimensions (Rows, columns): ", dim(my_data), "\n")
9 str(my_data)
10
```

```
> dim(my_data)
[1] 2500 10
> cat("Dimensions (Rows, columns): ", dim(my_data), "\n")
Dimensions (Rows, columns): 2500 10
> str(my_data)
'data.frame':   2500 obs. of  10 variables:
 $ car.ID       : int  1 2 3 4 5 6 7 8 9 10 ...
 $ Brand        : chr  "Tesla" "BMW" "Audi" "Tesla" ...
 $ Year         : int  2016 2018 2013 2011 2009 2019 2020 2017 2023 2010 ...
 $ Engine.Size  : num  2.3 4.4 4.5 4.1 2.6 2.4 4 5.3 5.7 1.5 ...
 $ Fuel.Type    : chr  "Petrol" "Electric" "Electric" "Diesel" ...
 $ Transmission: chr  "Manual" "Manual" "Manual" "Automatic" ...
 $ Mileage      : int  114832 143190 181601 68682 223009 246553 135486 83030 120360 135009 ...
 $ Condition    : chr  "New" "Used" "New" "New" ...
 $ Price        : num  26614 14680 44403 86374 73577 ...
 $ Model        : chr  "Model X" "5 Series" "A4" "Model Y" ...
```

summary(my\_data)



The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for loading packages, reading a CSV file, and using the `summary()` function on `my_data`.
- Environment:** Shows `my_data` as a data frame with 2500 observations and 10 variables.
- Console:** Displays the output of the `summary(my_data)` command, showing summary statistics for each variable.

```
1 install.packages(c("readr", "psych"))
2 library(readr)
3 library(psych)
4 my_data <- read.csv("car_price_prediction.csv")
5 head(my_data)
6 tail(my_data)
7 dim(my_data)
8 cat("Dimensions (Rows, columns): ", dim(my_data), "\n")
9 str(my_data)
10 summary(my_data)
11
```

```
> summary(my_data)
   car.ID      Brand      Year      Engine.Size      Fuel.Type      Length:2500
Min.   : 1.0    Min.   :2000   Min.   :1.000   Min.   :28908
1st Qu.: 625.8  1st Qu.:2005   1st Qu.:2.200   1st Qu.:53485
Median :1250.5 Median :2012   Median :3.400   Median :52638
Mean   :1250.5 Mean   :2012   Mean   :3.465   Mean   :52638
3rd Qu.:1875.2 3rd Qu.:2018   3rd Qu.:4.700   3rd Qu.:75839
Max.   :2500.0 Max.   :2023   Max.   :6.000   Max.   :99983

Transmission      Mileage      Condition      Price
Length:2500      Min.   : 15      Length:2500   Min.   :5011
Class :character  1st Qu.: 71832   Class :character  1st Qu.:28908
Mode  :character  Median :149085   Mode  :character  Median :53485
                  Mean   :149750   Mean   :52638
                  3rd Qu.:225991   3rd Qu.:75839
                  Max.   :299967   Max.   :99983

Model
Length:2500
Class :character
Mode  :character
```

Vikas Pal  
S098

# MVLU COLLEGE

## Subject:-Data Analysis with SAS / SPSS /R

```
names(my_data)
```

```
cat("Column Names: ", names(my_data), "\n")
```

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code to load the 'psych' package, read 'car\_price\_prediction.csv', and inspect the data using `dim()`, `str()`, `summary()`, and `names()`.
- Console:** Displays the output of the code, showing the dimensions (2500 rows, 10 columns) and a summary of the data, including the column names: Car.ID, Brand, Year, Engine.Size, Fuel.Type, Transmission, Mileage, Condition, Price, and Model.
- Environment:** Shows the 'my\_data' object with 2500 observations and 10 variables.
- Files:** Lists files in the current directory, including 'car\_price\_prediction.csv'.

```
describe(my_data)
```

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code to install 'readr', load 'psych' and 'readr', read 'car\_price\_prediction.csv', and use `describe()` to get summary statistics.
- Console:** Displays the output of `describe(my_data)`, providing a detailed summary of the data, including mean, standard deviation, median, and range for each variable.
- Environment:** Shows the 'my\_data' object with 2500 observations and 10 variables.
- Files:** Lists files in the current directory, including 'car\_price\_prediction.csv'.

Vikas Pal  
S098