

MVLU COLLEGE

PRACTICAL NO.3

AIM: Exploring data: View() or print() (R).

The screenshot shows the RStudio interface with the following details:

- Data View:** Displays the 'Cleaned_BMW_Sales_Data' dataset with 99 rows and 11 columns. The columns include model, year, region, color, fuel_type, transmission, engine_size_l, mileage_km, price_usd, sales_volume, and sales_classification.
- Global Environment:** Shows the 'cleaned_BMW_Sales' object with 99 observations of 11 variables.
- Console:** Displays the R version information and the command to run the script.
- Script Editor:** Contains the R code for reading the CSV file and creating the 'my_data' object.
- Output:** Shows the error message 'Error in file(ffile, "rt") : cannot open the connection' and the warning message 'In addition: warning message: In file(file, "rt") : cannot open file 'c:/users/itlab/onedrive/documents/s105/cleaned_BMW_sales_data': No such file or directory'.
- System Tray:** Shows the date and time as 24-11-2025 and 11:51.

MVLU COLLEGE

PRACTICAL NO.3

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Background Jobs
R 4.5.2 - ~
Project: (None)

Source
> my_data <- read.csv("C:/users/itlab/OneDrive/Documents/S105/cleaned_BMW_Sales_Data.csv")
> head(my_data)
#> #> model year      region color fuel_type transmission engine_size_l mileage_km price_usd sales_volume
#> 1 5 Series 2016       Asia Red Petrol   Manual        3.5    151748    98740     8300
#> 2 18 2013 North America Red Hybrid Automatic      1.6    121671    79219     3428
#> 3 5 Series 2022 North America Blue Petrol   Automatic      4.5    10991    113265     6994
#> 4 X3 2024 Middle East Blue Petrol   Automatic      1.7    27255     60971     4047
#> 5 7 Series 2020 South America Black Diesel  Manual        2.1    122131    49898     3080
#> 6 5 Series 2017 Middle East silver Diesel  Manual        1.9    171362    42926     1232
#> sales_classification
#> 1 High
#> 2 Low
#> 3 Low
#> 4 Low
#> 5 Low
#> 6 Low
#> tail(my_data)
#> #> model year      region color fuel_type transmission engine_size_l mileage_km price_usd sales_volume
#> 94 i3 2013 Africa Black Diesel  Manual        1.6    192509    72385     455
#> 95 i3 2015 Asia Silver Electric Manual        3.5    68161    63696     7218
#> 96 X6 2022 Asia White Electric  Manual        1.9    182027    35951     3389
#> 97 X5 2024 Asia Blue Hybrid   Automatic      2.6    120597    95900     8581
#> 98 18 2011 Europe Red Hybrid   Automatic      3.8    184765    44846     1431
#> 99 X5 2019 Africa Blue Electric Manual        2.1    161177    95908     2775
#> sales_classification
#> 94 Low
#> 95 High
#> 96 Low
#> 97 High
#> 98 Low
#> 99 Low
#> dim(my_data)
[1] 99 11
#> cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")
Dimensions (Rows, Columns): 99 11
#> str(my_data)
'data.frame': 99 obs. of 11 variables:
 $ model : chr "5 Series" "i8" "5 Series" "X3" ...
 $ year  : int 2016 2013 2022 2024 2020 2017 2022 2014 2016 2019 ...
 $ region: chr "Asia" "North America" "North America" "Middle East" ...
```

MVLU COLLEGE

PRACTICAL NO.3

RStudio

```

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

Source Environment History Connections Tutorial
R - R.4.5.2 ~/Project (None)

> str(my_data)
'data.frame': 99 obs. of 11 variables:
 $ model      : chr "5 series" "i8" "5 series" "x3" ...
 $ year       : int 2016 2013 2022 2024 2020 2017 2022 2014 2016 2019 ...
 $ region     : chr "Asia" "North America" "Middle East" ...
 $ color      : chr "red" "blue" "blue" ...
 $ fuel_type   : chr "Petrol" "hybrid" "petrol" "petrol" ...
 $ transmission: chr "Manual" "Automatic" "Automatic" "Automatic" ...
 $ engine_size_l: num 3.5 1.6 4.5 1.7 2.1 1.9 1.8 1.6 1.7 3 ...
 $ mileage_km : int 151748 121671 10999 27255 122131 171362 196741 121156 48073 35700 ...
 $ price_usd  : int 98740 79219 113265 60971 49898 42926 55064 102778 116482 96257 ...
 $ sales_volume: int 8300 3428 6994 4047 3080 1232 7949 632 8944 4411 ...
 $ sales_classification: chr "High" "Low" "Low" "Low" ...

> summary(my_data)
model      year       region     color      fuel_type    transmission
Length:99  Min.   :2010   Class :character  Length:99  Length:99  Length:99
Class :character 1st Qu.:2014   Class :character  Class :character  Class :character  Class :character
Mode  :character Median:2017   Mode :character  Mode :character  Mode :character  Mode :character
Mean   :2017
3rd Qu.:2021
Max.   :2024
engine_size_l mileage_km   price_usd  sales_volume sales_classification
Min. :1.500  Min. : 1428  Min. : 53893.14 105162.00 103746.31 57410.72 1428.0 196741.0 195313.0 -0.03
1st Qu.:2.550 1st Qu.: 61221 1st Qu.: 54020 1st Qu.:2526 1st Qu.:2526 1st Qu.:2526 1st Qu.: 10.0 11.0 10.0 -0.06
Median :3.200  Median :105162  Median : 82675  Median :4823  Median :4823  Median :4823  Median : 14.0 14.0 14.0 -0.01
Mean   :3.195  Mean   :103617  Mean   : 79109  Mean   :5029  Mean   :5029  Mean   :5029  Mean   : 14.0 14.0 14.0 -0.01
3rd Qu.:3.800 3rd Qu.:142628 3rd Qu.:101498 3rd Qu.:7793 3rd Qu.:7793 3rd Qu.:7793 3rd Qu.: 19.0 20.0 19.0 -0.05
Max.   :4.900  Max.   :196741  Max.   :1119927  Max.   :9797  Max.   :9797  Max.   :9797  Max.   : 20.0 20.0 20.0 -0.13
> cat("Column Names: ", names(my_data), "\n")
Column Names: model year region color fuel_type transmission engine_size_l mileage_km price_usd sales_volume sales_classification
> describe(my_data)
vars n mean sd median trimmed mad min max range skew
model* 1 99 6.34 2.94 6.0 6.38 2.97 1.0 11.0 10.0 -0.06
year 2 99 2017.23 4.26 2017.0 2017.23 5.93 2010.0 2024.0 14.0 -0.01
region* 3 99 3.75 1.64 4.0 3.80 1.48 1.0 6.0 5.0 -0.16
color* 4 99 3.55 1.76 4.0 3.56 1.48 1.0 6.0 5.0 -0.11
fuel_type* 5 99 2.51 1.08 3.0 2.51 1.48 1.0 4.0 3.0 -0.04
transmission* 6 99 1.55 0.50 2.0 1.56 0.00 1.0 2.0 1.0 -0.18
engine_size_l 7 99 3.19 0.92 3.2 3.19 0.89 1.5 4.9 3.4 -0.05
mileage_km 8 99 103617.16 53893.14 105162.00 103746.31 57410.72 1428.0 196741.0 195313.0 -0.03
price_usd 9 99 79109.01 27034.89 82675.0 79609.12 29804.71 31815.0 119927.0 88112.0 -0.13
sales_volume 10 99 5028.74 2918.70 4823.0 5028.46 3955.58 173.0 9797.0 9624.0 0.00
sales_classification* 11 99 1.67 0.47 2.0 1.70 0.00 1.0 2.0 1.0 -0.70
kurtosis se
model* -1.10 0.49
year -1.18 0.43
region* -1.23 0.36
color* -1.28 0.18
fuel_type* -1.30 0.11
transmission* -1.99 0.05
engine_size_l -0.91 0.09
mileage_km -0.89 5416.46
price_usd -1.27 2717.11
sales_volume -1.30 293.34
sales_classification* -1.53 0.05
>
```

RStudio

```

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

Source Environment History Connections Tutorial
R - R.4.5.2 ~/Project (None)

> str(my_data)
'data.frame': 99 obs. of 11 variables:
 $ model      : chr "5 series" "i8" "5 series" "x3" ...
 $ year       : int 2016 2013 2022 2024 2020 2017 2022 2014 2016 2019 ...
 $ region     : chr "Asia" "North America" "Middle East" ...
 $ color      : chr "red" "blue" "blue" ...
 $ fuel_type   : chr "Petrol" "hybrid" "petrol" "petrol" ...
 $ transmission: chr "Manual" "Automatic" "Automatic" "Automatic" ...
 $ engine_size_l: num 3.5 1.6 4.5 1.7 2.1 1.9 1.8 1.6 1.7 3 ...
 $ mileage_km : int 151748 121671 10999 27255 122131 171362 196741 121156 48073 35700 ...
 $ price_usd  : int 98740 79219 113265 60971 49898 42926 55064 102778 116482 96257 ...
 $ sales_volume: int 8300 3428 6994 4047 3080 1232 7949 632 8944 4411 ...
 $ sales_classification: chr "High" "Low" "Low" "Low" ...

> summary(my_data)
model      year       region     color      fuel_type    transmission
Length:99  Min.   :2010   Class :character  Length:99  Length:99  Length:99
Class :character 1st Qu.:2014   Class :character  Class :character  Class :character  Class :character
Mode  :character Median:2017   Mode :character  Mode :character  Mode :character  Mode :character
Mean   :2017
3rd Qu.:2021
Max.   :2024
engine_size_l mileage_km   price_usd  sales_volume sales_classification
Min. :1.500  Min. : 1428  Min. : 53893.14 105162.00 103746.31 57410.72 1428.0 196741.0 195313.0 -0.03
1st Qu.:2.550 1st Qu.: 61221 1st Qu.: 54020 1st Qu.:2526 1st Qu.:2526 1st Qu.:2526 1st Qu.: 10.0 11.0 10.0 -0.06
Median :3.200  Median :105162  Median : 82675  Median :4823  Median :4823  Median :4823  Median : 14.0 14.0 14.0 -0.01
Mean   :3.195  Mean   :103617  Mean   : 79109  Mean   :5029  Mean   :5029  Mean   :5029  Mean   : 14.0 14.0 14.0 -0.01
3rd Qu.:3.800 3rd Qu.:142628 3rd Qu.:101498 3rd Qu.:7793 3rd Qu.:7793 3rd Qu.:7793 3rd Qu.: 19.0 20.0 19.0 -0.05
Max.   :4.900  Max.   :196741  Max.   :1119927  Max.   :9797  Max.   :9797  Max.   :9797  Max.   : 20.0 20.0 20.0 -0.13
> cat("Column Names: ", names(my_data), "\n")
Column Names: model year region color fuel_type transmission engine_size_l mileage_km price_usd sales_volume sales_classification
> describe(my_data)
vars n mean sd median trimmed mad min max range skew
model* 1 99 6.34 2.94 6.0 6.38 2.97 1.0 11.0 10.0 -0.06
year 2 99 2017.23 4.26 2017.0 2017.23 5.93 2010.0 2024.0 14.0 -0.01
region* 3 99 3.75 1.64 4.0 3.80 1.48 1.0 6.0 5.0 -0.16
color* 4 99 3.55 1.76 4.0 3.56 1.48 1.0 6.0 5.0 -0.11
fuel_type* 5 99 2.51 1.08 3.0 2.51 1.48 1.0 4.0 3.0 -0.04
transmission* 6 99 1.55 0.50 2.0 1.56 0.00 1.0 2.0 1.0 -0.18
engine_size_l 7 99 3.19 0.92 3.2 3.19 0.89 1.5 4.9 3.4 -0.05
mileage_km 8 99 103617.16 53893.14 105162.00 103746.31 57410.72 1428.0 196741.0 195313.0 -0.03
price_usd 9 99 79109.01 27034.89 82675.0 79609.12 29804.71 31815.0 119927.0 88112.0 -0.13
sales_volume 10 99 5028.74 2918.70 4823.0 5028.46 3955.58 173.0 9797.0 9624.0 0.00
sales_classification* 11 99 1.67 0.47 2.0 1.70 0.00 1.0 2.0 1.0 -0.70
kurtosis se
model* -1.10 0.49
year -1.18 0.43
region* -1.23 0.36
color* -1.28 0.18
fuel_type* -1.30 0.11
transmission* -1.99 0.05
engine_size_l -0.91 0.09
mileage_km -0.89 5416.46
price_usd -1.27 2717.11
sales_volume -1.30 293.34
sales_classification* -1.53 0.05
>
```