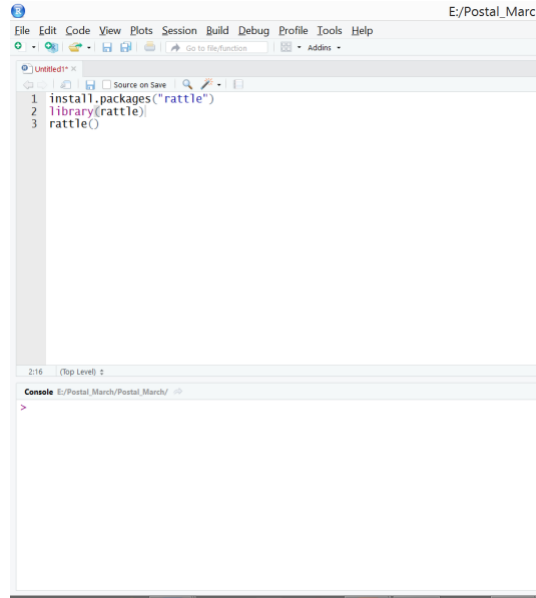


IMPORTANT: We trust that you will not read through this answer key until you have completed your own assignment.

Step 1. Install Rattle

Install Rattle through RStudio by running the following commands:

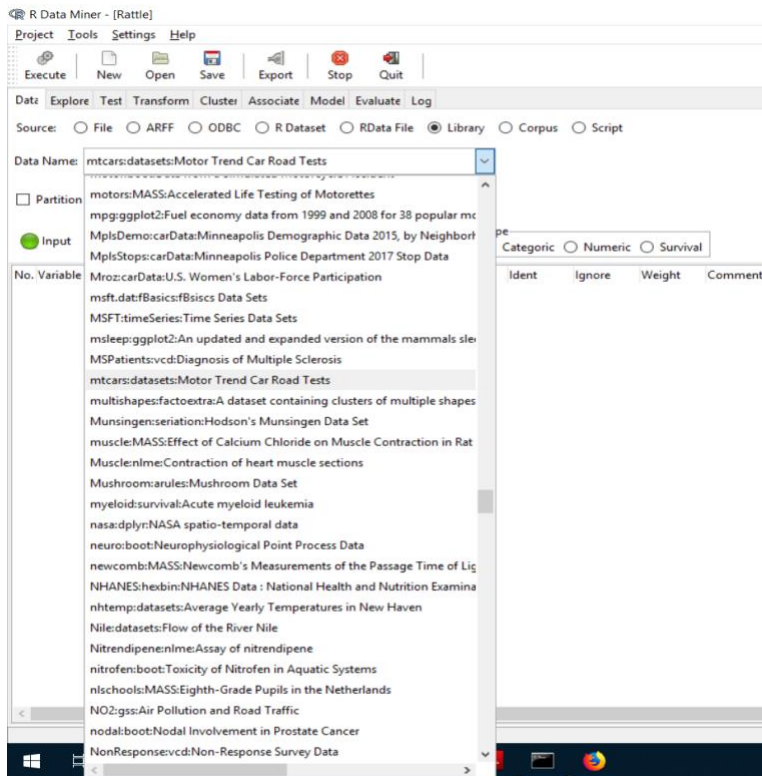


The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for running, saving, and other functions. The main editor window displays three lines of R code: `1 install.packages("rattle")`, `2 library(rattle)`, and `3 rattle()`. The bottom console window shows the prompt `>` and the file path `E:/Postal_Marc/Postal_Marc/`.

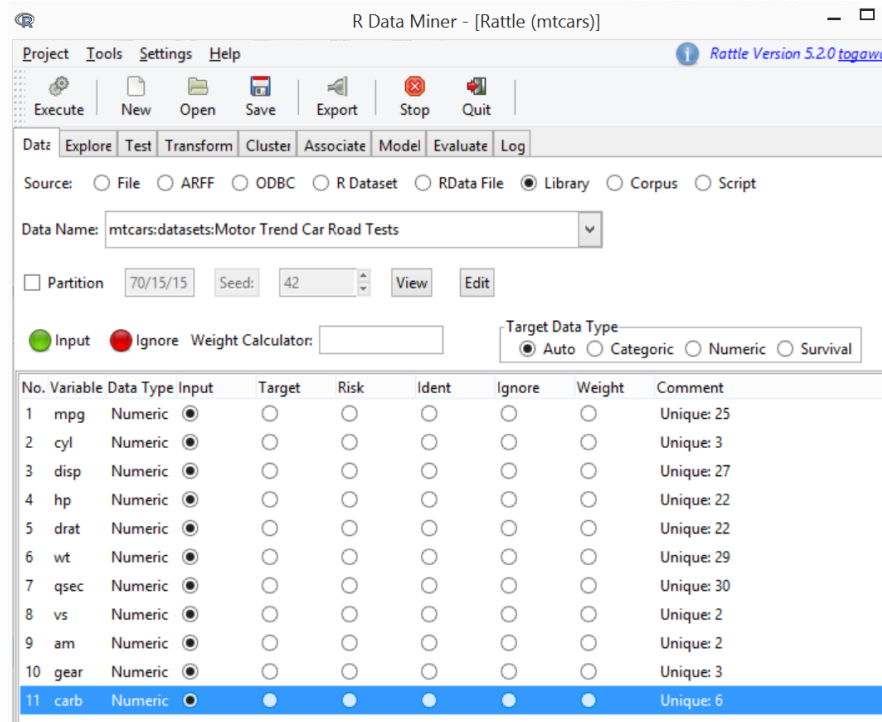
```
1 install.packages("rattle")
2 library(rattle)
3 rattle()
```

Step 2. Import the Motor Trend Car Road Test (mtcars) dataset from R library

Action 1. Select the data (mtcars:datasets:Motor Trend Car Road Test) from the dropdown menu called Data Name.



Action 2. Click “Execute”. This will show you the variables in the dataset.

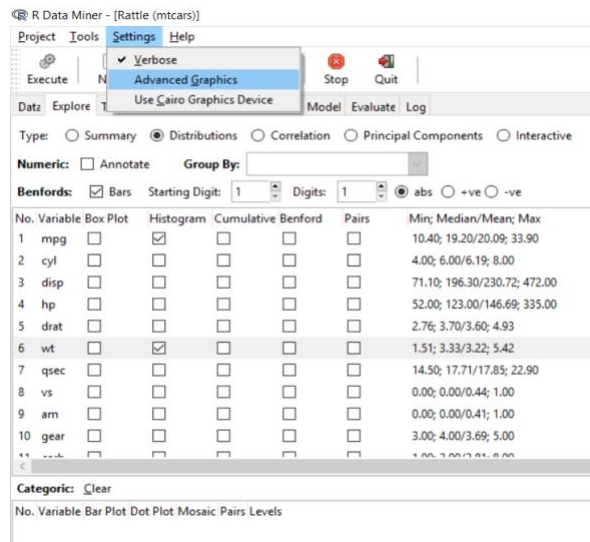


Step 3. Plot histograms and boxplots for mpg (miles per gallon) and wt (weight)

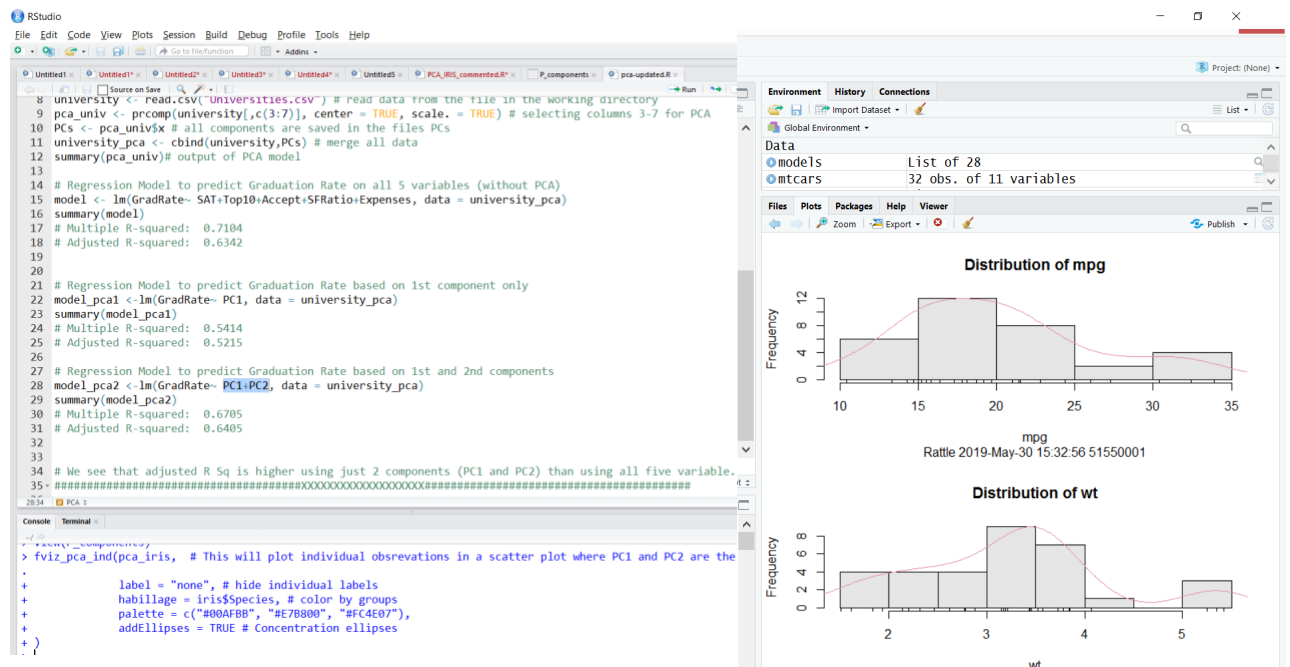
Action 1: Uncheck the Advanced Graphics option from Settings.

Histogram:

Action 2: Explore → Distributions → Histogram (Check the box for mpg and wt) → Execute



Output: Available in the plot/graph section in RStudio.

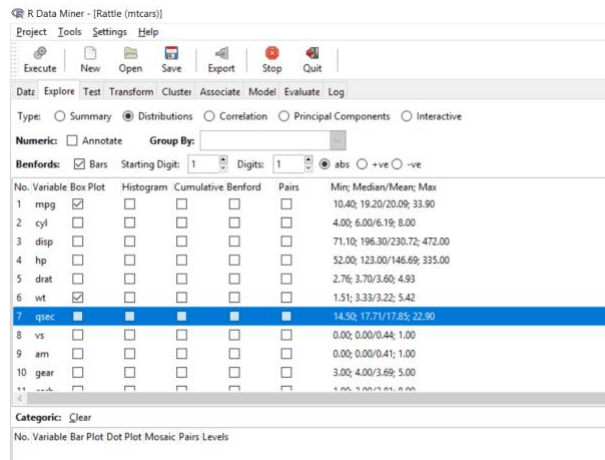


Interpretation (mpg): mpg(miles per gallon) doesn't follow normal distribution and it is somewhat right skewed.

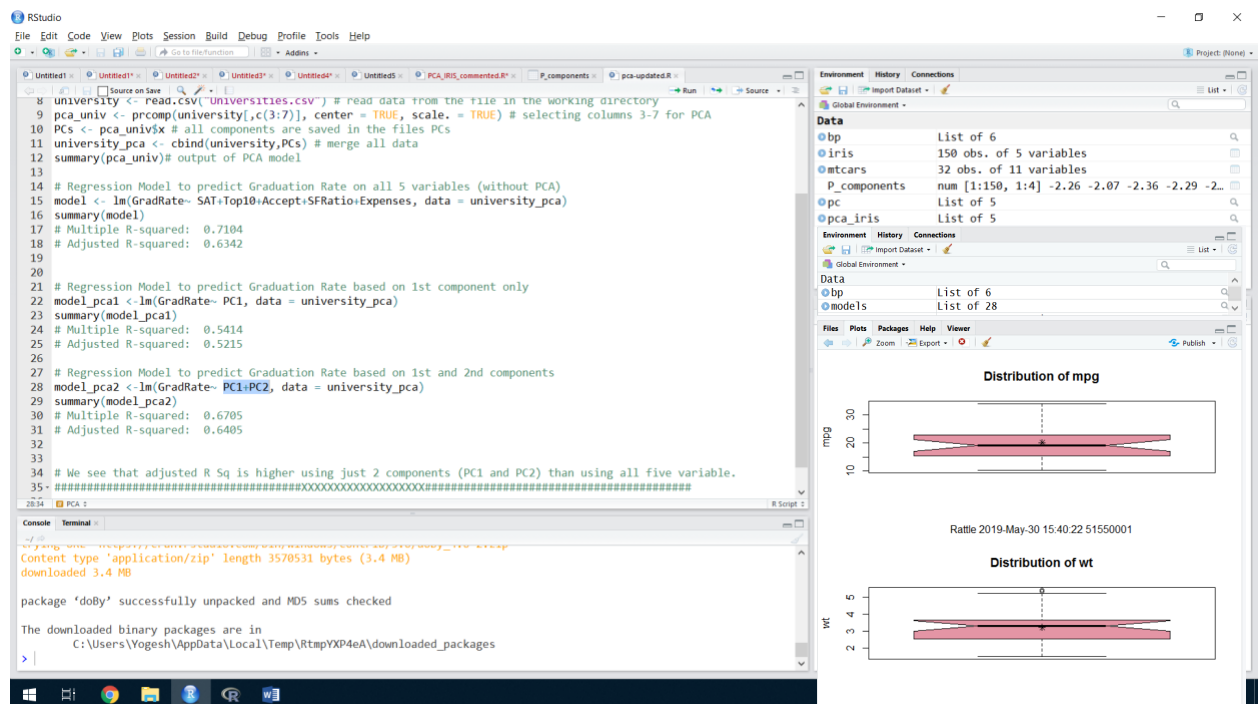
Interpretation (wt): There are more cars with the weight between 3 and 4 units than others. Mostly, a uniform distribution.

Box Plot:

Action 1: Explore → Distributions → Box Plot (Check the box for mpg and wt) → Execute



Output: Available in the plot/graph section in RStudio.

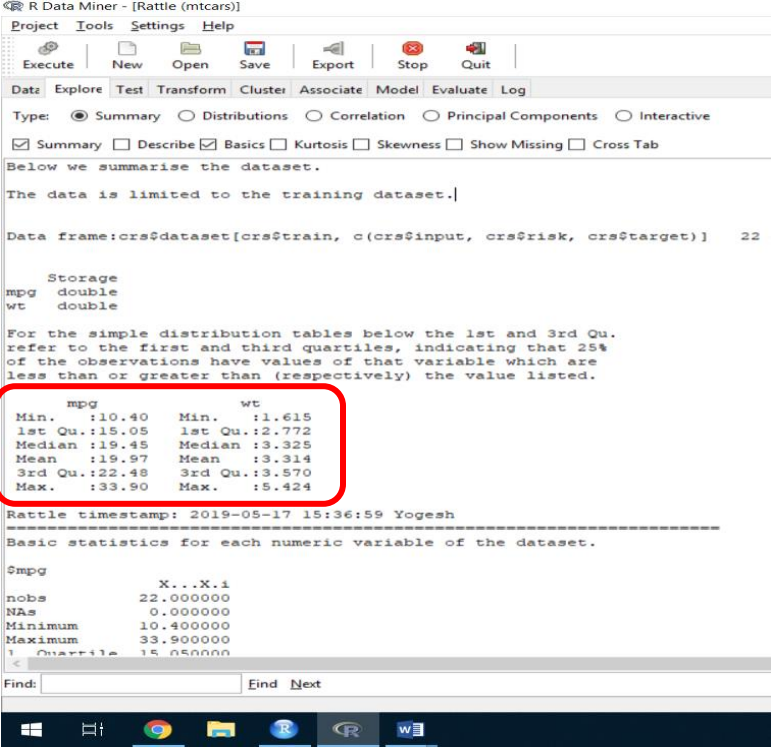


Interpretation(mpg): There are no outliers.

Interpretation(wt): There is an outlier in data, an observation with very high weight in comparison with others.

Step 4. Calculate summary statistics such as mean, median, 1st quartile, 3rd quartile, etc. for the following variables: mpg and wt

Action 1: Explore → Summary → Basics



```
R Data Miner - [Rattle (mtcars)]
Project Tools Settings Help
Execute New Open Save Export Stop Quit
Data Explore Test Transform Cluster Associate Model Evaluate Log
Type: Summary Distributions Correlation Principal Components Interactive
Summary Describe Basics Kurtosis Skewness Show Missing Cross Tab
Below we summarise the dataset.
The data is limited to the training dataset.
Data frame:crs$dataset[crs$train, c(crs$input, crs$risk, crs$target)] 22
Storage
mpg double
wt double
For the simple distribution tables below the 1st and 3rd Qu.
refer to the first and third quartiles, indicating that 25%
of the observations have values of that variable which are
less than or greater than (respectively) the value listed.
mpg wt
Min. :10.40 Min. :1.615
1st Qu.:15.05 1st Qu.:2.772
Median :19.45 Median :3.325
Mean :19.97 Mean :3.314
3rd Qu.:22.48 3rd Qu.:3.570
Max. :33.90 Max. :5.424
Rattle timestamp: 2019-05-17 15:36:59 Yogesh
Basic statistics for each numeric variable of the dataset.
$mpg
X...X.i
nobs 22.000000
NAs 0.000000
Minimum 10.400000
Maximum 33.900000
1st Quartile 15.050000
```

Interpretation:

Mean and median are almost similar for mpg

Mean and median are almost similar for wt

The range for mpg: 33.9 – 10.4

The range for wt: 5.4 – 1.6