Task 3 for R language and data analysis

Qiang Shen

Dec.14.2017

Use the knowledge learned from course 3 to solve the following questions.

1.In build-in dataset named **iris** in R:

(1): Firstly, please generate a new variable called *PW.level* using the following criteria. For observations who belong to *setosa*, please classify those who have *Petal.Width* equal or large than 0.5 as *c*, equal or large than 0.3 but smaller than 0.5 as *b* and the rest observations as *a* .

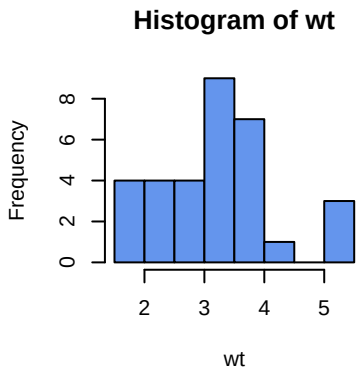
For observations who belong to *versicolor*, please classify those who have *Petal.Width* equal or large than 1.8 as *c*, equal or large than 1.4 but smaller than 1.8 as *b* and the rest as *a* .

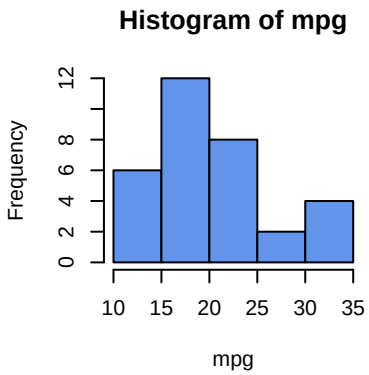
For observations who belong to *virginica*, please classify those who have *Petal.Width* equal or large than 2.2 as *c*, equal or large than 1.8 but smaller than 2.2 as *b* and the rest as *a* .

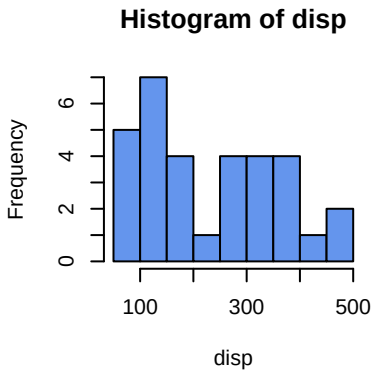
(2): Secondly, please calculate the mean and standard deviation of the four variables

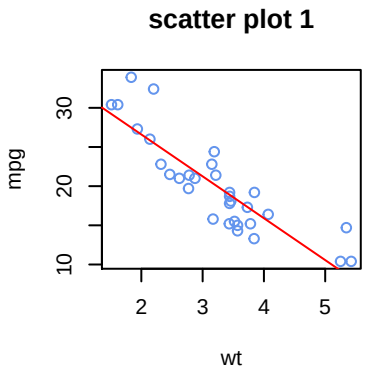
Sepal.Length, *Sepal.Width*, *Petal.Width* and *Petal.Length* in the groups collapsed by the variables **Species** and **PW.level**. (Note: don't forget to add labels for the output results)

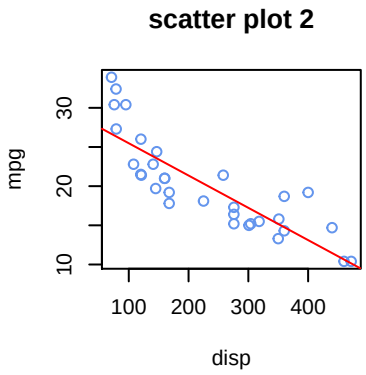
2. Please generate 5 plots based on your own preference and then combine them as the way demonstrated in the figure below. (Note: each plot has the same height and width.)


Histogram of wt
Frequency
0
2
4
6
8
2
3
4
5
wt


Histogram of mpg
Frequency
0
2
4
6
8
12
10
15
20
25
30
35
mpg


Histogram of disp
Frequency
0
2
4
6
100
300
500
disp


scatter plot 1
mpg
10
20
30
2
3
4
5
wt


scatter plot 2
mpg
10
20
30
100
200
300
400
disp