Data Viz Notes: September 21, 2016

Rmarkdown: has R chunks been names cars

* Control alt I = this will create a R chunk for you
* Echo = true means it will show coding
* Echo = false means it will not show the coding
* He prefers to use R markdown version instead of regular version
* Interactive documents will always have HTML format that will be created on Rmarkdown

Usually import data from excel or CSV

?iris = will give you information on the data that’s present

Plot(iris$sepal.width)

Plot(iris$sepal.length)

Library.(ggplot2)

ggplot(iris,aes(Sepal.Length, y=1))+geom\_point()

> ggplot(iris,aes(Sepal.Length, Sepal.Width))+geom\_point()

We get our packages from Gethub

Dim()= control and enter at the same time. This gives you the dimension

Nrow()= how many rows there are

Ncol()= how many columns there are

Name() = name of all the variables

Head(iris,6) give top 6 of dataset. Can change the number to get as many top #’s as you want

Tail (iris,1) bottom 1 of dataset.

Iris[1;2,]

Want 4th row and 3rd column🡪 iris[4,3]

Str() =is for structure of the dataset. Know why type of variables, numerical or categorical

Factor in this feature is categorical

Summary() = gives count for non-numerical numbers.

He wants us to explore the cars dataset in a Rmarkdown files, and use the functions we did in class. Turn it in next Wednesday. Use book by yechng zhao to help? – the 7th link on the textbook list on the homepage.

Data Viz Notes: 9/28/2016

Difference between bar graph and histogram🡪 histograms have two quantitative axis and continuous

1. make sure to save the file

2. determine working directory- know where files are

Getwd()

Session:set working directory – changed working directory to docuemente- class folder- data viz

Control -alt I : gives the set up for entering data

Control enter” do this in RMarkdown top file and will copy and do the command in the bottom console

**Out of diamond tables we want:** size, structure of dataset, names of values, distribution of variables, correlations (numeric in nature), summary

= is the same function as <-