….

Here is the code for the and the output.

# My name is

# This is my first programming assignment for AP Stats and I will copy and see if everything runs properly

# November 11, 2019

#################################### Assignment 1---students type in the data ############

# Everything here works for the latest version of R and RStudio

## The general form of a command will look like this:

## note to myself

## myGraph <- ggplot(myData, aes(variable for x axis, variable for y axis)) + geom()

## You can also use =, its the same as -<

## NOTE: DO NOT make variables names with a space, use one word or two connected with a period "."

## Here I enter the data from page 35

## The "c" function combines the data into a vector ##### Please load dplyr and ggplot2 now. ####

foreigh.born=c(2.8,7.0,15.1,3.8,27.2,10.3,12.9,8.1,18.9,9.2, 16.3,5.6,13.8,4.2,3.8,

6.3,2.7,2.9,3.2,12.2,14.1,5.9,6.6,1.8,3.3,1.9,5.6,19.1,5.4,20.1,

10.1,21.6,6.9,2.1,3.6,4.9,9.7,5.1,12.6,4.1,2.2,3.9,15.9,8.3,

3.9,10.1,12.4,1.2,4.4,2.7)

summary(foreigh.born) # Gives the five number summary.

str(foreigh.born) # the str function shows me the type of structure of the data. fivenum(foreigh.born) # gives the five number summary

mean(foreigh.born) # just shows the mean

head(foreigh.born, n=12) # shows the first 12, pick n. Used with large data files.

tail(foreigh.born) # shows the end of the data. You can pick n or leave it alone.

plot(foreigh.born) # this is R's generic scatter plot function and only shows basic information.

# we will use this more later.

hist(foreigh.born) # This is base R basic histogram function. # Below is ggplot's better graphing abilities

ggplot() + aes(foreigh.born)+ geom\_histogram(binwidth = 2.5)

# I change the variable name so I don't confuse with the prior graphs foreign.born3=c(2.8,7.0,15.1,3.8,27.2,10.3,12.9,8.1,18.9,9.

2,16.3,5.6,13.8,4.2,3.8,

6.3,2.7,2.9,3.2,12.2,14.1,5.9,6.6,1.8,3.3,1.9,5.6,19.1,5.4,20.1,

10.1,21.6,6.9,2.1,3.6,4.9,9.7,5.1,12.6,4.1,2.2,3.9,15.9,8.3,

3.9,10.1,12.4,1.2,4.4,2.7)

# This is a histogram with base R hist(foreign.born3, breaks = 10,

main = "Histogram with Base Graphics", ylim = c(0,15))

# check the structure str(foreign.born3)

# make sure it's a data frame by changing to a data.frame. fb3=as.data.frame(foreign.born3)

# I check to see the structur of fb3 str(fb3)

# I use ggplot to make a histogram similar to the book's histogram ggplot(fb3,aes(x=foreign.born3))+

geom\_histogram(color="black",fill="orange",binwidth = 3)+ labs(x="Percent of foreign born residents",y="Number of States")+ geom\_density()

# I can add a density curve to the histogtam ggplot(fb3,aes(x=foreign.born3))+

geom\_histogram(aes(y=..density..),color="black",fill="orange",binwidth = 3)+ labs(x="Percent of foreign born residents",y="Density of States")+ geom\_density(alpha=0.2,fill="#FF6666")

# Same histogram but I just change the colors a bit. ggplot(fb3, aes(x=foreign.born3)) +

geom\_histogram(aes(y=..density..),

binwidth=3,

colour="black", fill="white") + geom\_density(alpha=.2, fill="#FF6666")

# use control-l to clear the console

Some of the output:

* ##### Please load dplyr and ggplot2 now. ####

>

> foreigh.born=c(2.8,7.0,15.1,3.8,27.2,10.3,12.9,8.1,18.9,9.2, 16.3,5.6,13.8,4.2,3.8,

+ 6.3,2.7,2.9,3.2,12.2,14.1,5.9,6.6,1.8,3.3,1.9,5.6,19.1,5.4,

20.1,

+ 10.1,21.6,6.9,2.1,3.6,4.9,9.7,5.1,12.6,4.1,2.2,3.9,15.9,8.3,

+ 3.9,10.1,12.4,1.2,4.4,2.7)

>

* summary(foreigh.born) # Gives the five number summary.

Min. 1st Qu. Median Mean 3rd Qu. Max. 1.200 3.800 6.100 8.316 12.350 27.200

>

* str(foreigh.born) # the str function shows me the type of structure of the data.

num [1:50] 2.8 7 15.1 3.8 27.2 10.3 12.9 8.1 18.9 9.2 ...

>

* fivenum(foreigh.born) # gives the five number summary [1] 1.2 3.8 6.1 12.4 27.2

>

* mean(foreigh.born) # just shows the mean [1] 8.316

>

* head(foreigh.born, n=12) # shows the first 12, pick n. Used with large data files.

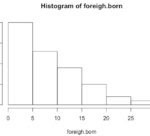
[1] 2.8 7.0 15.1 3.8 27.2 10.3 12.9 8.1 18.9 9.2 16.3 5.6

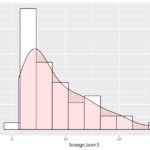
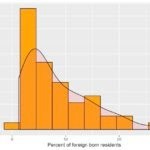
>

* tail(foreigh.born) # shows the end of the data. You can pick n or leave it alone.

[1] 3.9 10.1 12.4 1.2 4.4 2.7

Here are some of the plots using ggplot2





We have completed Unit 4 and will start Unit 5 next week. We are where we need to be at this time of the year. At this rate, we’ll finish the class on time and have a few weeks to review for the exam in May 2020.