Lab 1

Kent Codding

6/Sep/2023

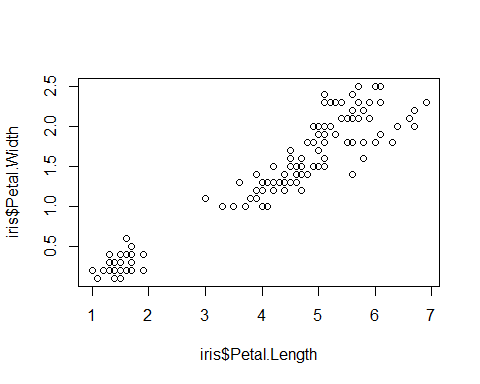
Here I will create a simple summary table of the data in the cars data set.

summary(iris)

## Sepal.Length Sepal.Width Petal.Length Petal.Width   
## Min. :4.300 Min. :2.000 Min. :1.000 Min. :0.100   
## 1st Qu.:5.100 1st Qu.:2.800 1st Qu.:1.600 1st Qu.:0.300   
## Median :5.800 Median :3.000 Median :4.350 Median :1.300   
## Mean :5.843 Mean :3.057 Mean :3.758 Mean :1.199   
## 3rd Qu.:6.400 3rd Qu.:3.300 3rd Qu.:5.100 3rd Qu.:1.800   
## Max. :7.900 Max. :4.400 Max. :6.900 Max. :2.500   
## Species   
## setosa :50   
## versicolor:50   
## virginica :50   
##   
##   
##

Now lets make a plot of petal width (y-axis) versus petal length (x-axis).

plot(x = iris$Petal.Length, y = iris$Petal.Width)



#in the plot command above, I used $ to call each variable in the data frame by the column name

# Here I will make a bullet list of the 3 topics on the syllabus that I think may be most useful to my research:

• Plotting in R

• Model Selection in R

• Zero truncated and Zero inflated models

# Level of experience with R, programming

As a Biology Major and Data Science Minor, I this is my 3rd class in R. I will take my 2nd python-based class in the Spring Semester: Applied Machine Learning. However, I gained valuable experience and R skills last summer working as a part-time intern for ERG, where I analyzed, manipulated, and transformed large datasets to extract important information. I actually published an R package on my github - <https://github.com/kentcod> - that has some cool functions. I am currently attempting to use python to predict a target variable (levels of plant invasion) based on a variety of features for my research, but the accuracy at the moment is low and inconsistent… so I have much more work to do with that.

# Stats experience

I have taken two other classes with Professor Lamar: BIOL 325 and BIOL 327. Thus, I have experience with frequentism and some bayesian (although bayesian honestly was quite difficult to understand). I am actually TAing BIOL 327 - Intro to Biostatistics this fall.