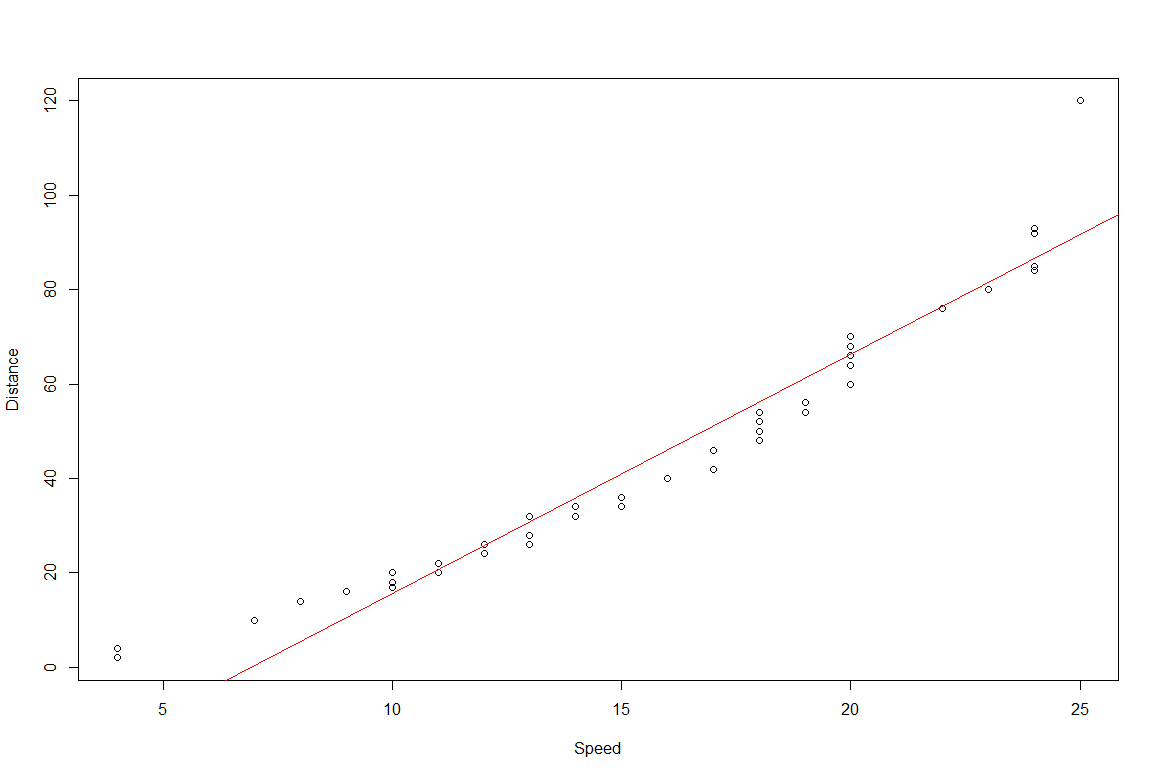
Course 3 Task 1 - Keegan Gelvoria Informal Report

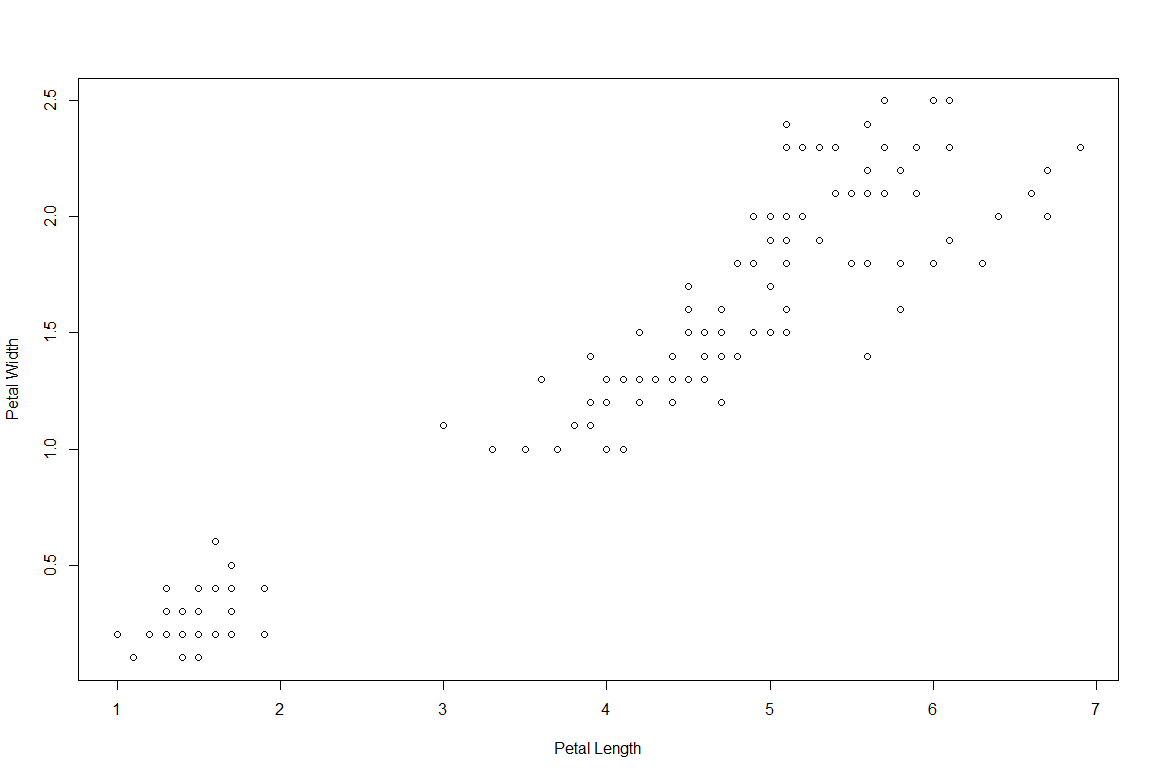
Predictions - R Tutorial

In the R Tutorial, I was able to create a model for predictions concerning how far a car can travel based on speed. This model was extremely accurate, with a multiple R-Squared value of 0.92. The model also showed that there is a statistically significant effect of Speed on Distance with a p-value of <2.2e-16.



Predictions - Find the Error

In the Find the Errors Task, I was able to create a model for predictions of petal width based on petal length. The model was not very accurate with a multiple r-squared of 0.0895. The model also showed that there is no statistically significant effect of Petal Length on Petal Width with a p-value of 0.08083.



Errors/Warnings and How They Were Overcome

* Error importing CSV file
  + Required name to be in quotes
* Error due to syntax/spelling
  + summary(risDataset) instead of summary (IrisDataset
  + str(IrisDatasets) instead of str(IrisDataset)
  + trainSizes instead of trainSize
  + predictions instead of prediction
  + plot(IrisDataset$Sepal.Length
    - Missing end parentheses
* Error in creation of a histogram of Species because column had data as characters while histograms require numerical data. I was able to change the data into numerical values by defining numerical values for each of the 3 species and map them in a new SpeciesNum column. A histogram was able to be created using this new SpeciesNum column.
* Error trying to create a qqnorm for the whole Iris dataset. This was fixed by creating normal quantile plots for each numerical column.
* Error trying to change the species values from characters to numerical values because it didn’t know what to change the characters to in numerical value. I had to re-import the whole dataset.
* Error in creating the first train set which was set at 20% when train set should be more than the test set size
  + trainSize <- round(nrow(IrisDataset) \* 0.2)
  + This was fixed in the seed 405 by using a 70% train size
* Error in eval structure
  + LinearModel<- lm(trainSet$Petal.Width ~ testingSet$Petal.Length)
  + Fixed
    - LinearModel<-lm(Petal.Width~ Petal.Length, trainSet)
* prediction<-predict(LinearModeltestSet)
  + Fixed
    - prediction<-predict(LinearModel,testSet)
* Predictions was supposed to be used but there was not Predictions object
  + Had to restructure to object Prediction

Takeaways

* R and Rstudio installation was straightforward based off of the documentation
* The tutorial was very helpful in walking through the necessary steps and building off of our python knowledge to this new language. I would recommend this initial tutorial.
* Main lessons from this experience
  + Walk through tutorial slowly to fully digest the differences between R and Python
  + Try to learn R studio functionalities
    - Code into the script so that can be saved
      * Coded initially into the console and lost progress when it wasn't saved