**Zach Adair**

**MSDS 660**

**Regis University**

**3/12/2019**

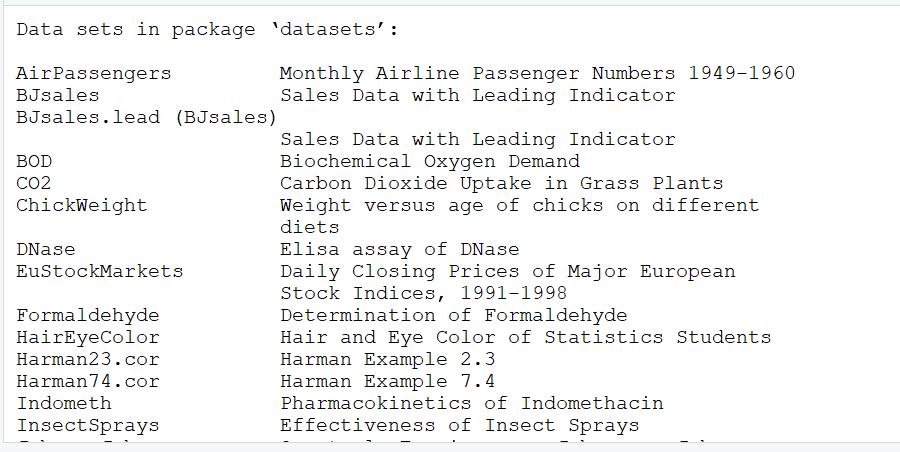
## Week 2 Assignment – 15 Commands on the Iris Dataset within R



1. The library loads built in commands and datasets into R.

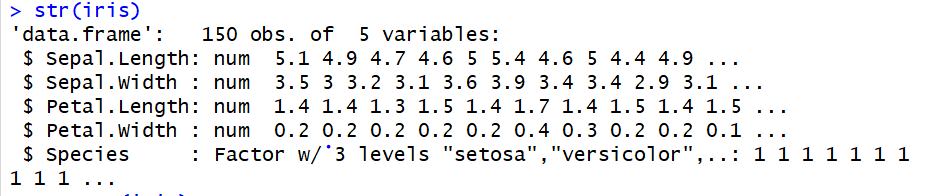


1. data() gives a list of all available datasets.

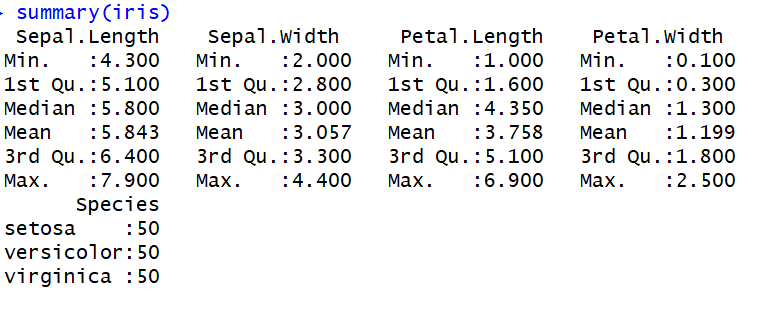




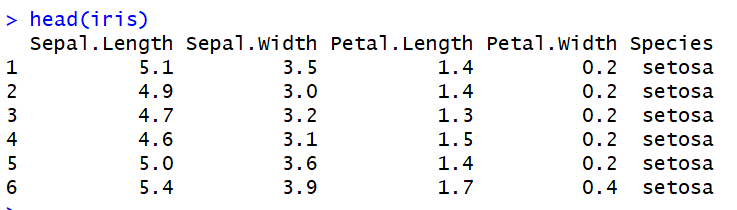
1. data(iris) ensures we have the iris dataset loaded in R.



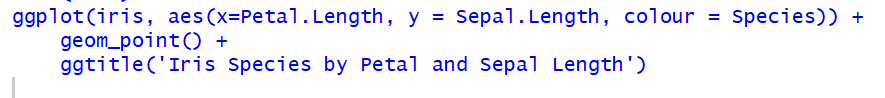
1. str(iris) displays the abbreviated contents of lists.



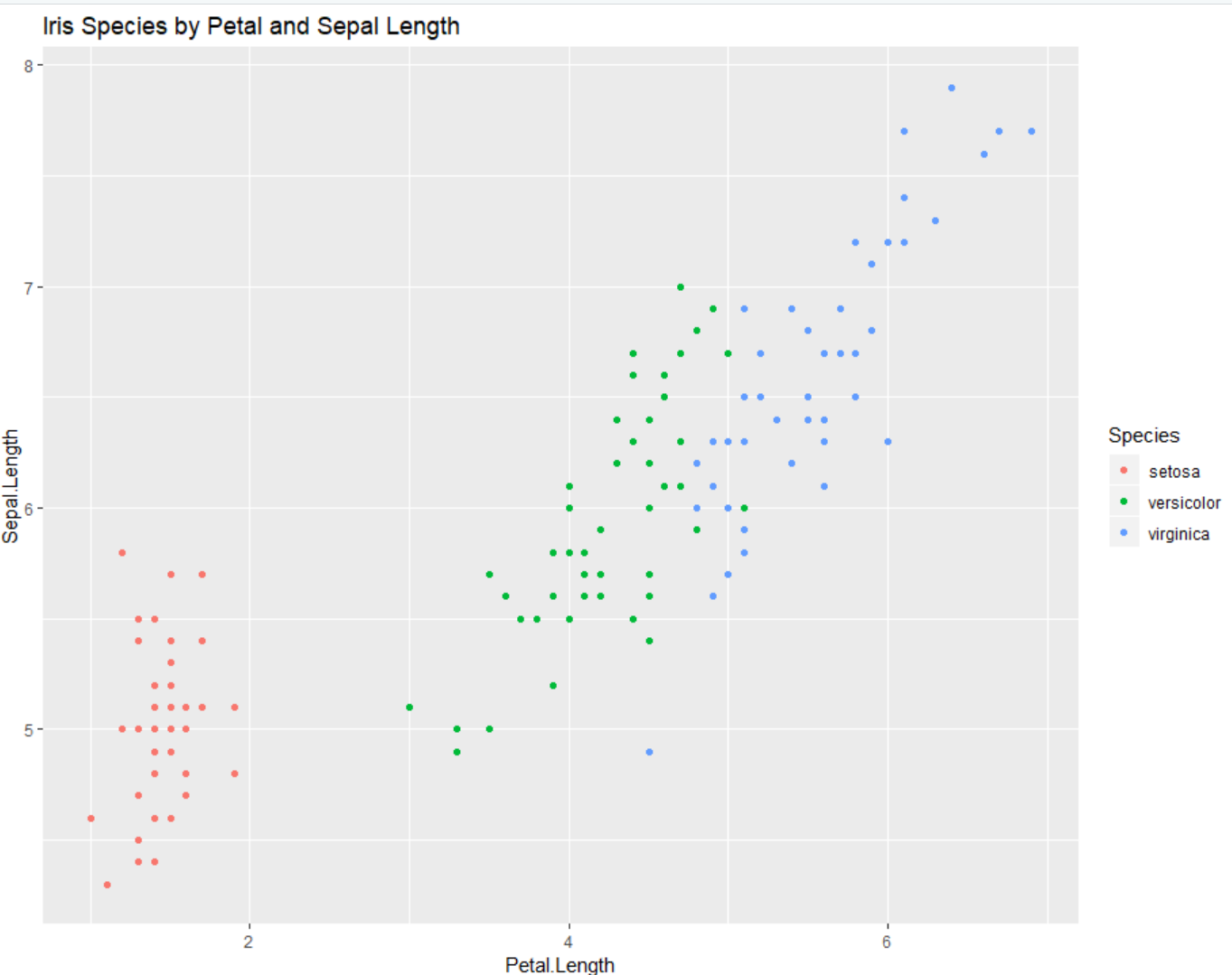
1. summary(iris) summarizes the data and shows the minimum, 1st quartile, median, mean, 3rd quartile and maximum of the different fields as well as a disparity of the different species types.

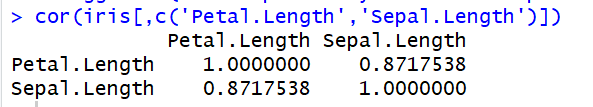


1. head(iris) shows a brief look at the data and its first 6 rows.



1. The ggplot will form a scatterplot.





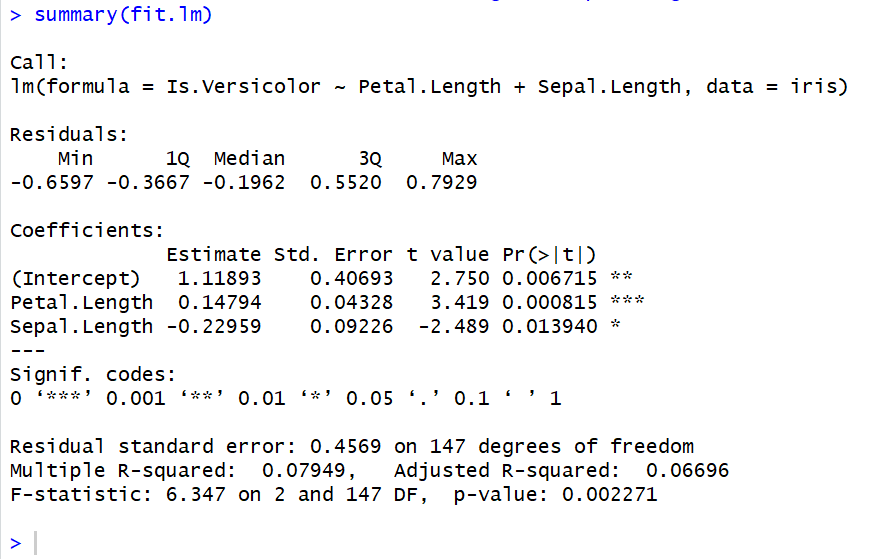
1. Correlation of the Iris dataset Petal Length vs Sepal Length.



1. Sets the iris data set by the number of species it has in the dataset.



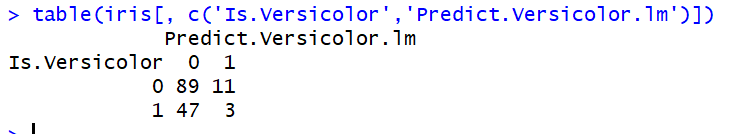
1. Fits soon to be model by the Petal length and Sepal length.
2. Summary data based on the fit of the new linear model



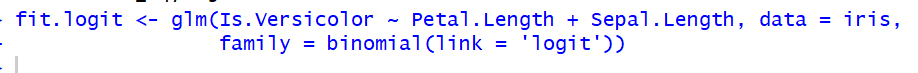
1. Set a new model based on the old one that now has a predictor variable.



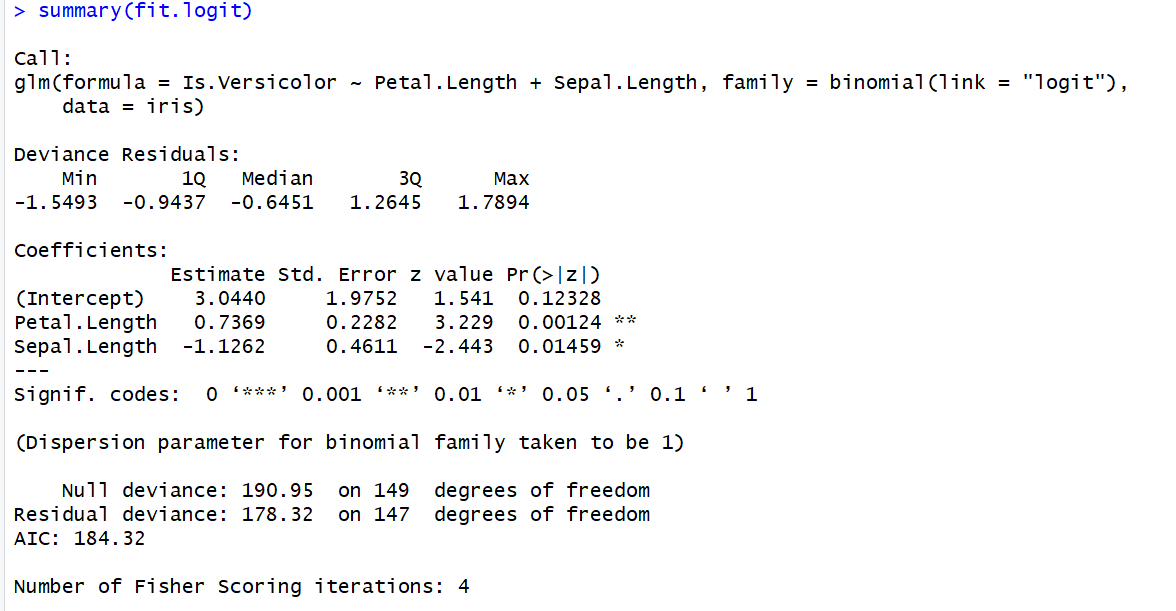
1. Creating a table of the new variable trying to predict versicolor iris species. The predictive model correctly predicted 92 out of the 150 total iris flowers which isn’t very good.



1. Create a new model this time attempting logistical regression, assigning the variable fit.logit



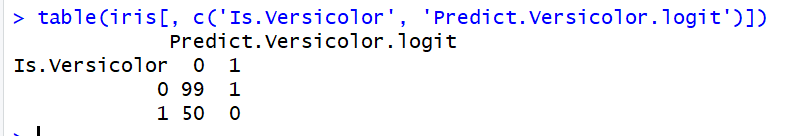
1. Now were going to look at a summary of our new variable.



1. Set predictor model to compare with fit.logit



1. Create a table to judge the predictor model



The table didn’t really do a good job predicting because it only predicted 1 to actually be a Versicolor iris.