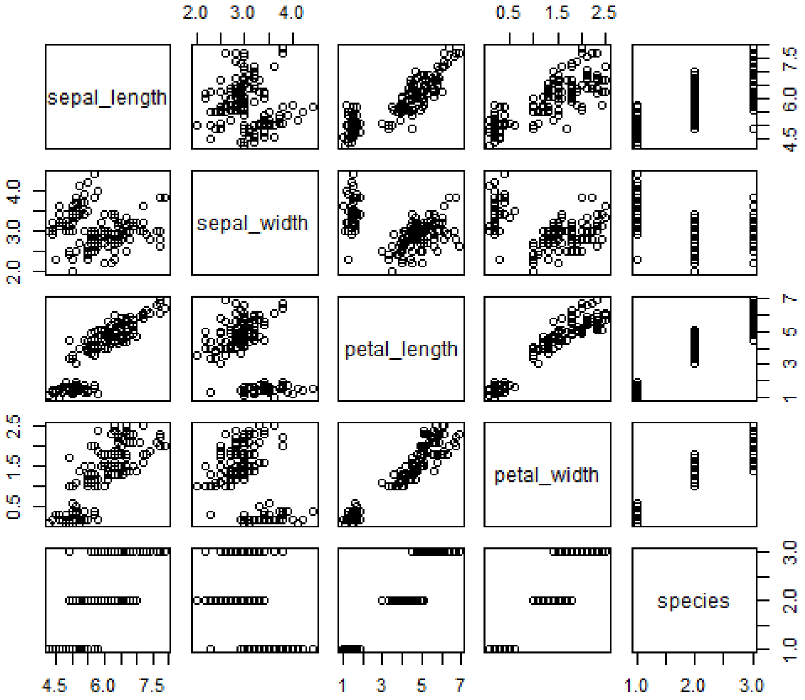
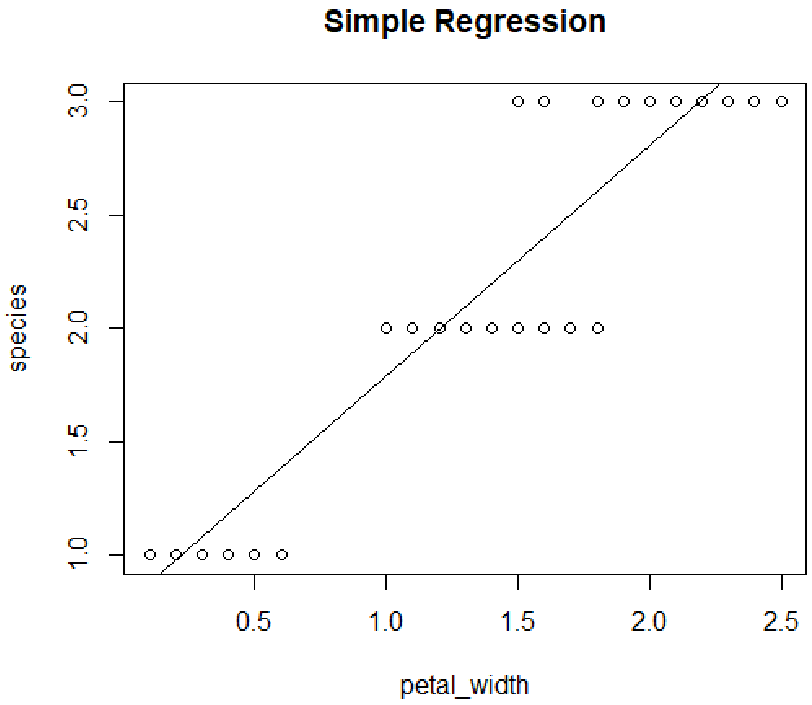
Data Analytics Assignment 4:

Simple and Multiple Linear Regression

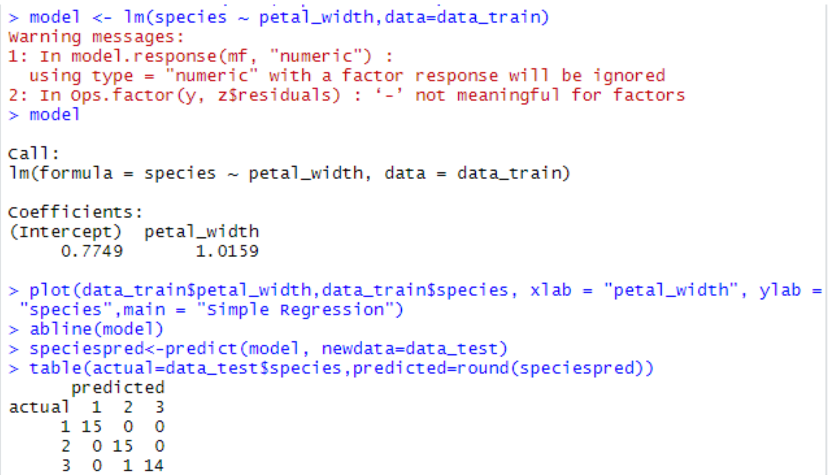
Initial plot of data prior to train-test split



Visualisation of simple linear regression fit line:



Code and confusion matrix for the classification using petal_width:



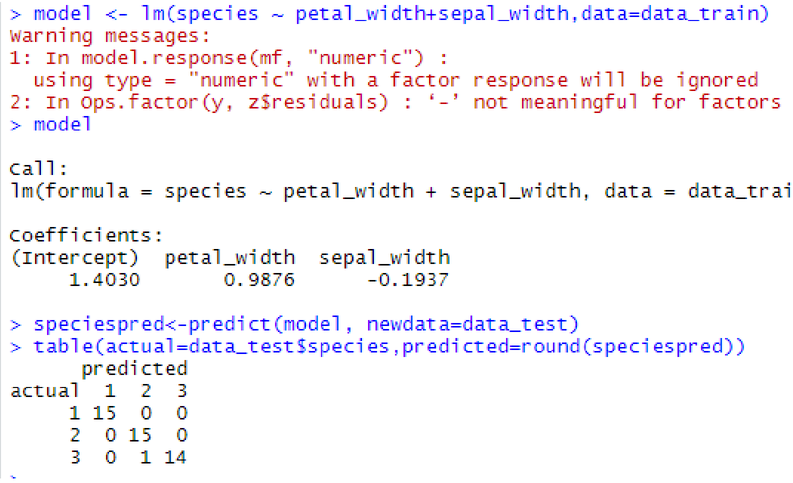
```
> model <- lm(species ~ petal_width,data=data_train)
Warning messages:
1: In model.response(mf, "numeric") :
  using type = "numeric" with a factor response will be ignored
2: In Ops.factor(y, z$residuals) : ‘-’ not meaningful for factors
> model

Call:
lm(formula = species ~ petal_width, data = data_train)

Coefficients:
(Intercept)  petal_width  
     0.7749       1.0159  

> plot(data_train$petal_width,data_train$species, xlab = "petal_width", ylab =
 "species",main = "Simple Regression")
> abline(model)
> speciespred<-predict(model, newdata=data_test)
> table(actual=data_test$species,predicted=round(speciespred))
      predicted
actual  1  2  3
     1 15  0  0
     2  0 15  0
     3  0  1 14
```

Code and confusion matrix for multiple linear regression using petal_width and sepal_width:



```
> model <- lm(species ~ petal_width+sepal_width,data=data_train)
Warning messages:
1: In model.response(mf, "numeric") :
  using type = "numeric" with a factor response will be ignored
2: In Ops.factor(y, z$residuals) : ‘-’ not meaningful for factors
> model

Call:
lm(formula = species ~ petal_width + sepal_width, data = data_train)

Coefficients:
(Intercept)  petal_width  sepal_width  
     1.4030       0.9876      -0.1937  

> speciespred<-predict(model, newdata=data_test)
> table(actual=data_test$species,predicted=round(speciespred))
      predicted
actual  1  2  3
     1 15  0  0
     2  0 15  0
     3  0  1 14
```