**Practical no 6**

**#Practical of simple/multiple Linear Regression**

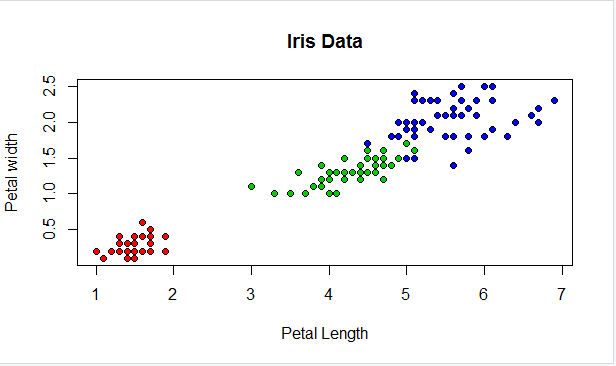
**#”least squares fit” function**

> lsfit(iris$Petal.Length,iris$Petal.Width)$coefficients

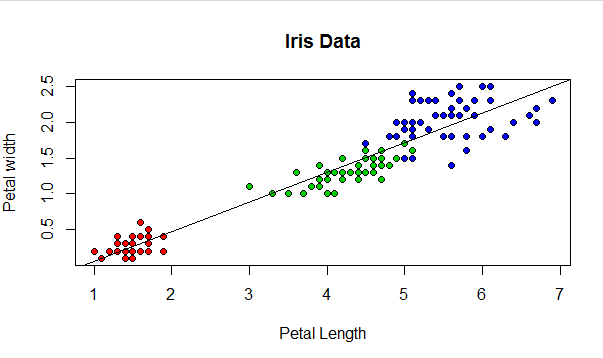
Intercept X

-0.3630755 0.4157554

>plot(iris$Petal.Length,iris$Petal.Width,pch=21,bg=c("red","green3","blue")[unclass(iris$Species)],main="Iris Data",xlab = "Petal Length",ylab = "Petal width")



> abline(lsfit(iris$Petal.Length,iris$Petal.Width)$coefficient,col="black")



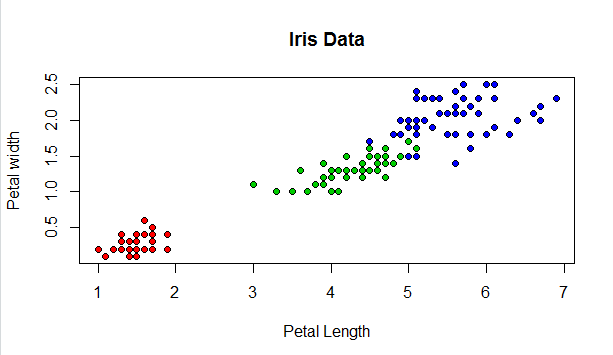
**#Create simple regression model**

> lm(Petal.Width~Petal.Length,data=iris)$coefficient

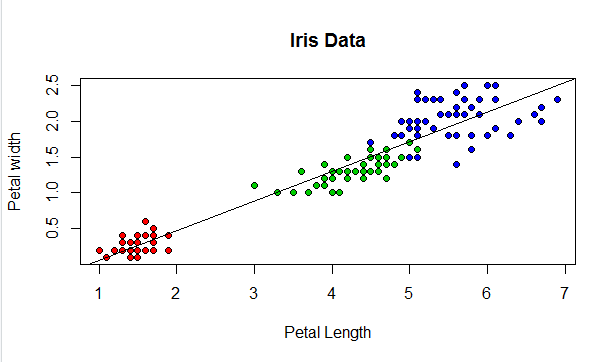
(Intercept) Petal.Length

-0.3630755 0.4157554

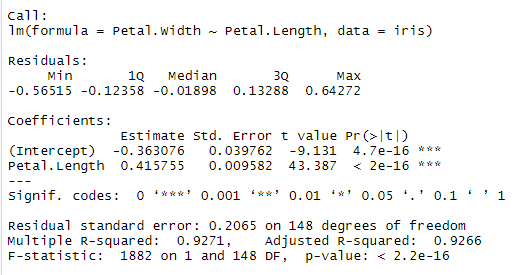
>plot(iris$Petal.Length,iris$Petal.Width,pch=21,bg=c("red","green3","blue")[unclass(iris$Species)],main="Iris Data",xlab = "Petal Length",ylab="Petal width")



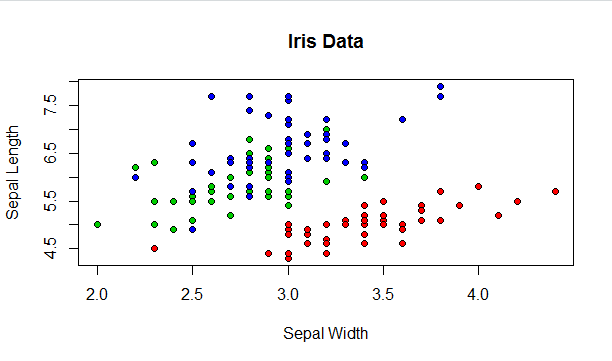
> abline(lm(Petal.Width~Petal.Length,data=iris)$coefficient,col="black")



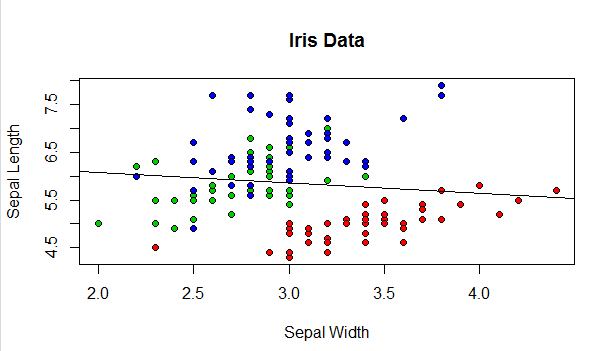
> summary(lm(Petal.Width~Petal.Length,data=iris))



> plot(iris$Sepal.Width,iris$Sepal.Length,pch=21,bg=c("red","green3","blue")[unclass(iris$Species)],main="Iris Data",xlab="Sepal Width",ylab="Sepal Length")

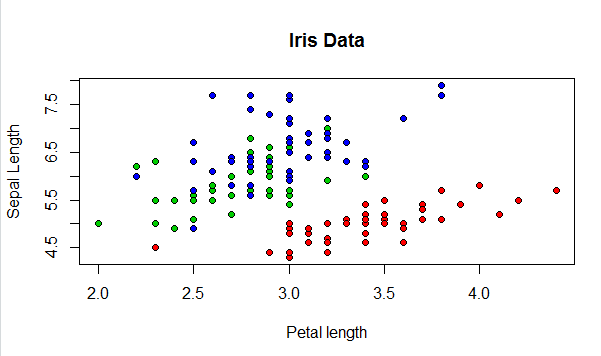


> abline(lm(Sepal.Length~Sepal.Width,data=iris)$coefficients,col="black")

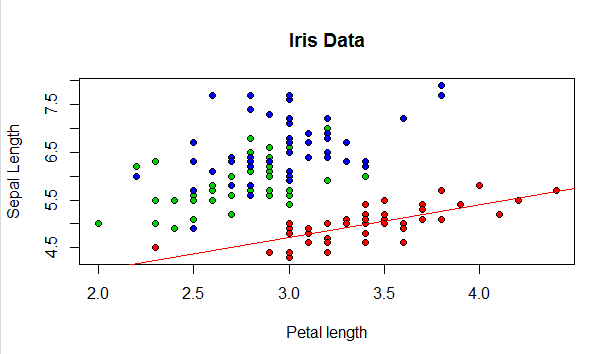


**#What happens if we divide the data up by species and run three separate linear regression?**

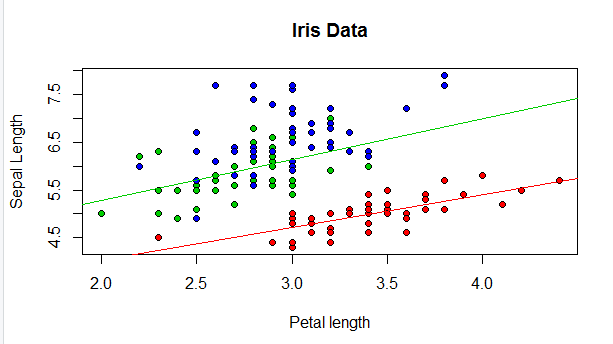
> plot(iris$Sepal.Width,iris$Sepal.Length,pch=21,bg=c("red","green3","blue")[unclass(iris$Species)],main="Iris Data",xlab="Petal length",ylab="Sepal Length")



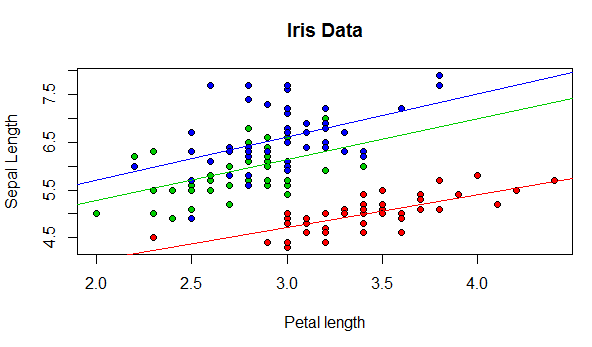
> abline(lm(Sepal.Length~Sepal.Width,data=iris[which(iris$Species=="setosa"),])$coefficients,col="red")



> abline(lm(Sepal.Length~Sepal.Width,data=iris[which(iris$Species=="versicolor"),])$coefficients,col="green3")



> abline(lm(Sepal.Length~Sepal.Width,data=iris[which(iris$Species=="virginica"),])$coefficients,col="blue")



**#The coefficients doing separate per species regression of Sepal.Length and Sepal.Width are:**

> lm(Sepal.Length~Sepal.Width,data=iris[which(iris$Species=="setosa"),])$coefficients

(Intercept) Sepal.Width

2.6390012 0.6904897

> lm(Sepal.Length~Sepal.Width,data=iris[which(iris$Species=="versicolor"),])$coefficients

(Intercept) Sepal.Width

3.5397347 0.8650777

> lm(Sepal.Length~Sepal.Width,data=iris[which(iris$Species=="virginica"),])$coefficients

(Intercept) Sepal.Width

3.9068365 0.9015345

> lm(Sepal.Length~Sepal.Width:Species+Species-1,data=iris)$coefficients

Speciessetosa Speciesversicolor

2.6390012 3.5397347

Speciesvirginica Sepal.Width:Speciessetosa

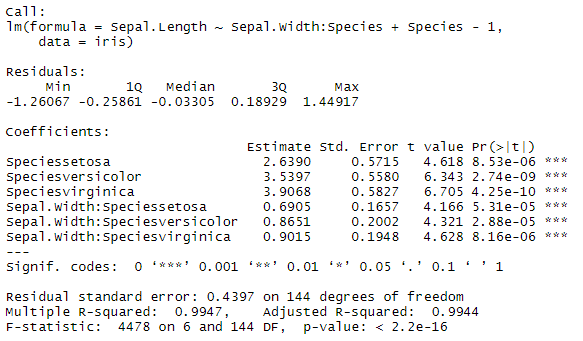
3.9068365 0.6904897

Sepal.Width:Speciesversicolor Sepal.Width:Speciesvirginica

0.8650777 0.9015345

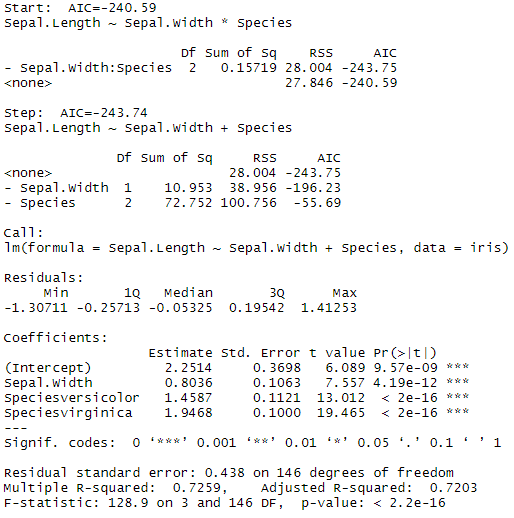
**#Using the summary command on the linear model object gives:**

> summary(lm(Sepal.Length~Sepal.Width:Species+Species-1,data=iris))



**#Simplify with AIC(Akaike Information Criterion)**

> summary(step(lm(Sepal.Length~Sepal.Width\*Species,data=iris)))



**#I just introduced a model of the form Sepal.Length~Sepal.Width:Species+Species-1**

**#which gave identical coefficients to those found doing species specific regressions:**

|  |
| --- |
| > lm(Sepal.Length~Sepal.Width:Species+Species-1,data=iris)$coefficients  Speciessetosa Speciesversicolor  2.6390012 3.5397347  Speciesvirginica Sepal.Width:Speciessetosa  3.9068365 0.6904897  Sepal.Width:Speciesversicolor Sepal.Width:Speciesvirginica  0.8650777 0.9015345  > lm(Sepal.Length~Sepal.Width:Species+Species,data=iris)$coefficients  (Intercept) Speciesversicolor  2.6390012 0.9007335  Speciesvirginica Sepal.Width:Speciessetosa  1.2678352 0.6904897  Sepal.Width:Speciesversicolor Sepal.Width:Speciesvirginica  0.8650777 0.9015345 |
|  |
| |  | | --- | |  | |