



## Module 6 - Algorithmes de base en apprentissage machine

### Exercices - Solutions

#### Exercice 1 :

1. Le téléchargement et le stockage des données IRIS est fait grâce à la ligne 3.
2. Représenter graphiquement la variation des variables `Sepal.Width` en fonction de `Sepal.Length`. Attribuez une couleur à chaque type d'espèce.

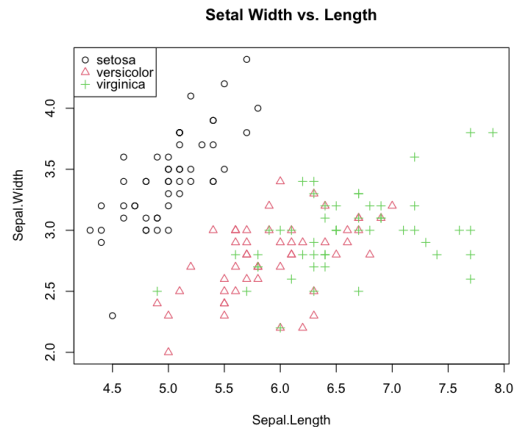


FIGURE 1 – Variation de `Sepal.Width` en fonction `Sepal.Length`

```

1 #Exercice 1
2 require("datasets")
3 data("iris")
4 str(iris)
5 head(iris)
6 Y<- iris[,"Sepal.Width"]
7 X<- iris[,"Sepal.Length"]
8 speciesID <- as.numeric(iris$Species)
9
10 plot(Y~X, col=iris[,"Species"], pch = speciesID, main = "
    Setal Width vs. Length", ylab="Sepal.Width", xlab="Sepal
    .Length")
11 legend("topleft", # specify the location of the legend
12 levels(iris$Species), # specify the levels of species
13 pch = 1:3, # specify three symbols used for the 3
    species
14 col = 1:3 # specify three colors for the 3 species
15 )
    
```

Listing 1 – Pseudo code de l'Exercice 1.2

3. La corrélation de Pearson est donnée par la commande `cor` (Ligne du Listing ). La valeur de `ycorr` est égale à `-0.1175698`. Cette valeur étant négative et faible, la relation entre `Sepal.Width` et `Sepal.Length` est donc inversement proportionnelle et faible.

```

1 #Exercice 1
2 # Correlation de Pearson
3 ycorr<- cor(Y,X, method="pearson")
4 ycorr
    
```

Listing 2 – Pseudo code de l'Exercice 1.3

4. La figure 2 illustre la pente de la droite de régression. La pente de cette droite étant négative ce qui confirme l'existence d'une corrélation négative entre `Sepal.Width` et `Sepal.Length`. Ainsi, si nous augmentons `Sepal.Width`, la valeur de `Sepal.Length` diminuera et vice-versa.

```

1 # Regression de Y sur X
2 model1<- lm(Y~X)
3 model1
4 plot(Y~X)
5 title(main = "Droite de regression de Y sur X")
6 abline(model1, col="blue", lwd=3)
    
```

Listing 3 – Pseudo code de l'Exercice 1.4

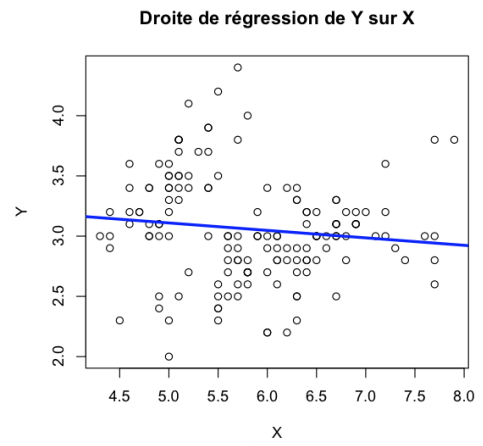


FIGURE 2 – Droite de régression de Y sur X