

Question 1 .....	1
Question 2 .....	2
Question 3 .....	2
Question 4 .....	3
Question 5 .....	3
Question 6 .....	4
Question 7 .....	4
Question 9 .....	5

```
%*****
%***** TP 2 *****
%*****
```

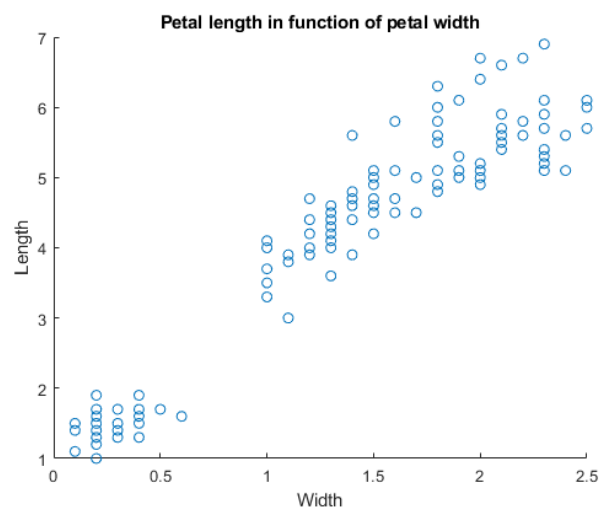
```
load ("fisheriris.mat")
```

```
sepalLength = meas(:,1:1);
sepalWidth = meas(:,2:2);
petalLength = meas(:,3:3);
petalWidth = meas(:,4:4);
```

## Question 1

```
figure(1)
scatter(petalWidth, petalLength);
title('Petal length in function of petal width');
xlabel('width');
ylabel('Length');
```

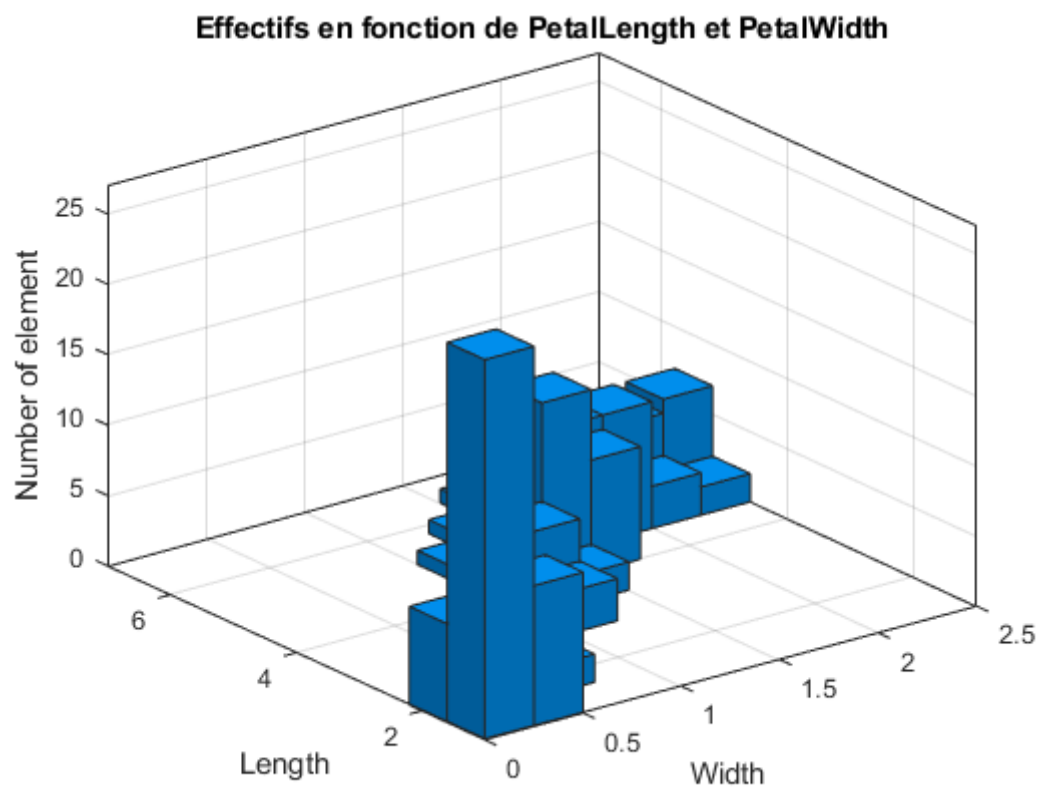
```
%Les points semblent former un droite. A première vu la longueur du petal
%semble être proportionnelle a sa largeur. En revanche on voit deux
%groupes de valeurs distinctes. Ces valeurs appartiennent a deux population
%differentes, la corrélation reste donc a démontrer.
```



## Question 2

```
figure(2);  
histogram2(petalwidth, petalLength, 'NumBins', 10);  
title('Effectifs en fonction de PetalLength et PetalWidth');  
xlabel('width');  
ylabel('Length');  
zlabel('Number of element');
```

%La position de chaque colonnes represente sa largeurs et sa longueur, sa  
%hauteur represente le nombre d'effectifs dans la chacune des catégories.



## Question 3

cf question 1

#### Question 4

```
%Recuperation de la matricede variance/covariance  
[corr] = corrcoef(petalwidth, petalLength);  
covXY = corr(2)
```

```
% Calcul du coefficient de correlation  
r = covXY / (std(petalLength) .* std(petalwidth))
```

```
covXY =
```

```
0.9629
```

```
r =
```

```
0.7156
```

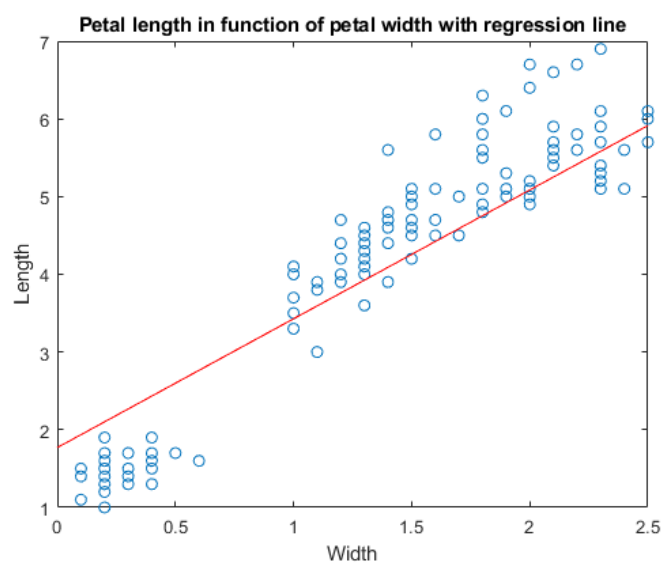
#### Question 5

```
% Calcul des paramètres de l'equation de la droite de regression  
a = r*std(petalLength)/std(petalwidth);  
b = mean(petalLength) - r*std(petalLength)/std(petalwidth)*mean(petalwidth);
```

```
% Calcul de deux point pour tracer la droite
```

```
p1 = [0, 2.5];  
p2 = [b, 2.5*a + b];
```

```
plot(petalwidth, petalLength, 'o', p1, p2, 'r-');  
title('Petal length in function of petal width with regression line');  
xlabel('width');  
ylabel('Length');
```



## Question 6

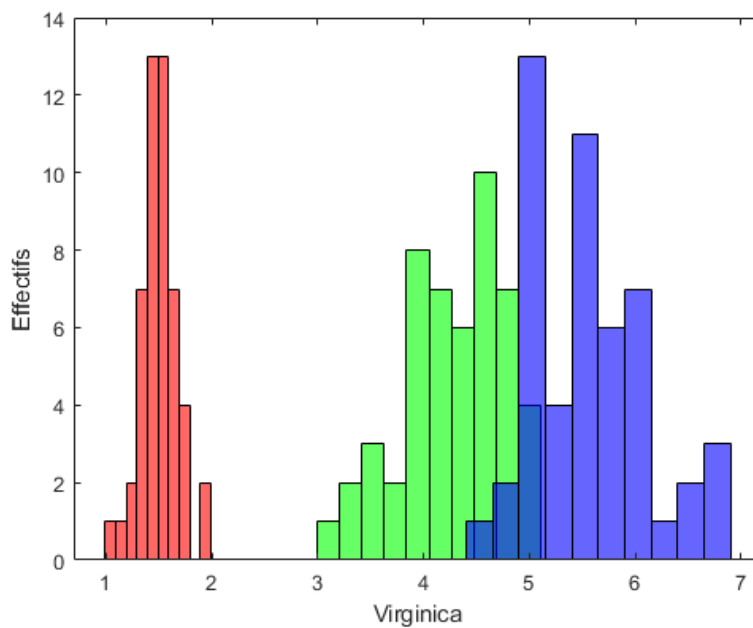
%Le coefficient de corrélation vaut  $r = 0.7156$  ce qui est relativement élevé. Les deux variables sont donc corrélées positivement (si l'un augmente l'autre augmente également).

## Question 7

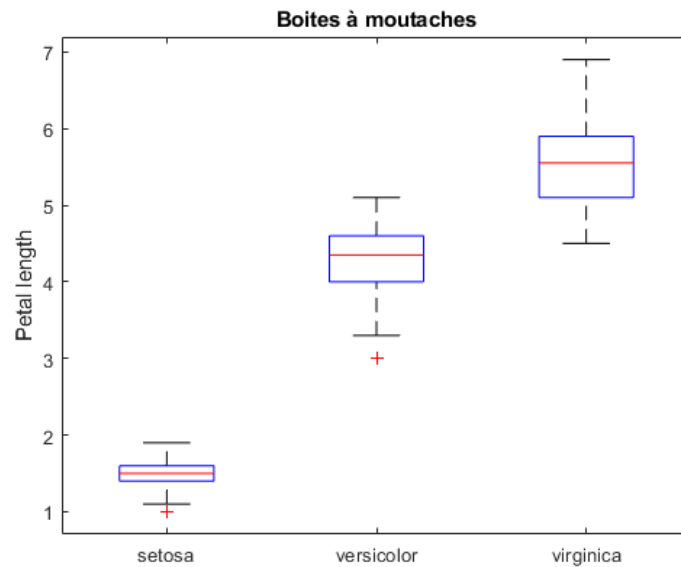
```
setosaPL = petalLength(1:50);
versicolorPL = petalLength(51:100);
virginicaPL = petalLength(101:150);
petal_length = [setosaPL; versicolorPL; virginicaPL];
figure(4)
subplot(1,1,1)

histogram(setosaPL, 'FaceColor', 'r', 'NumBins', 10)
xlabel('setosa')
ylabel("Effectifs")
hold on
histogram(versicolorPL, 'FaceColor', 'g', 'NumBins', 10)
xlabel('versicolor')
ylabel("Effectifs")
hold on
histogram(virginicaPL, 'FaceColor', 'b', 'NumBins', 10)
xlabel('virginica')
ylabel("Effectifs")

%Question 8
figure(5)
boxplot(y, species)
ylabel("Petal length")
title('Boîtes à moutaches')
```



## Question 8



## Question 9

Calcul de la variance intr-classe

```
% Calcul des moyennes
mean_petal = mean(petal_length);
mean_setosa = mean(setosaPL);
mean_versicolor = mean(versicolorPL);
mean_virginica = mean(virginicaPL);

% Calcul des variances
var_inter = 50.*(mean_setosa - mean_petal).^2 + 50.*(mean_versicolor - mean_petal).^2 +
50.*(mean_virginica - mean_petal).^2;
var_intra = (50.*var(setosaPL) + 50.*var(versicolorPL) + 50.*var(virginicaPL))./150;
var_inter = var_inter ./ 150;
my_variance = var_intra + var_inter;

% Difference entre mon calcul et le calcul de variance par matlab
var(petal_length) - my_variance

% Calcul du coefficient de corrélation entre nos variables qualitatives et
% quantitative
S_yx = sqrt(var_inter/my_variance)
```

```
Difference entre les deux méthodes de calcul =
0.0171
```

```
S_yx =
0.9697
```

La valeur étant proche de 1, les variables qualitatives sont bien distinctes. En conséquence l'espèce d'iris influe sur la longueur du pétal.