

# Abou Jaoudé Yann TDA ESILV A3

## Statistique de la donnée TP1

```
library(MASS)
library(factoextra)
library(FactoMineR)
library(plot3D)
library(car)
library(rgl)
library(ggplot2)
library(gridExtra)
library(cowplot)
```

### #question 1

```
data(survey)
summary(survey)
```

# Il y a un melange un variable quantitative et qualitative

#l'ACP sur l'ensemble des données n'est donc pas possible

Sex	Qualitative
Wr.Hnd	Quantitative
NW.Hnd	Quantitative
W.Hnd	Qualitative
Fold	Qualitative
Pulse	Quantitative
Clap	Qualitative
Exer	Qualitative
Smoke	Qualitative
Height	Quantitative
M.I.	Qualitative
Age	Quantitative

### #question 2

```
base.0 = survey[c(1,2,3,10)]
```

### #question 3

```
base.1 = na.exclude(base.0)
print(nrow(base.1))
# Il y a bien 207 individus et 4 variables
```

### #question 4

```
vectorSex= base.1[c(1)]
couleur = ifelse(vectorSex=="Female","red","blue")
```

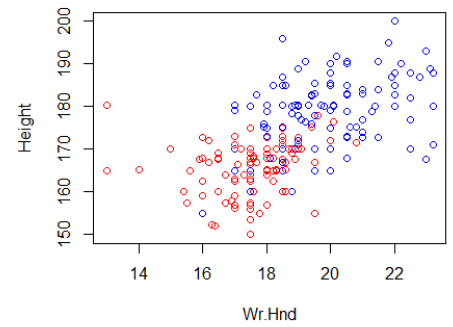
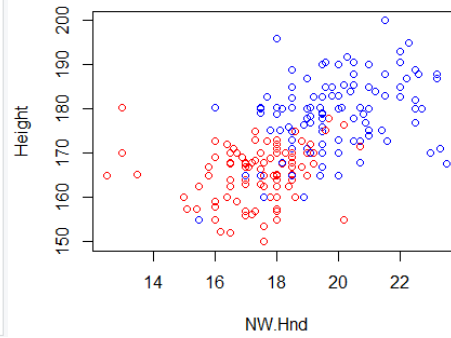
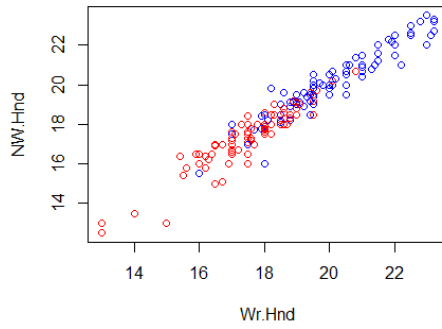
### #question 5

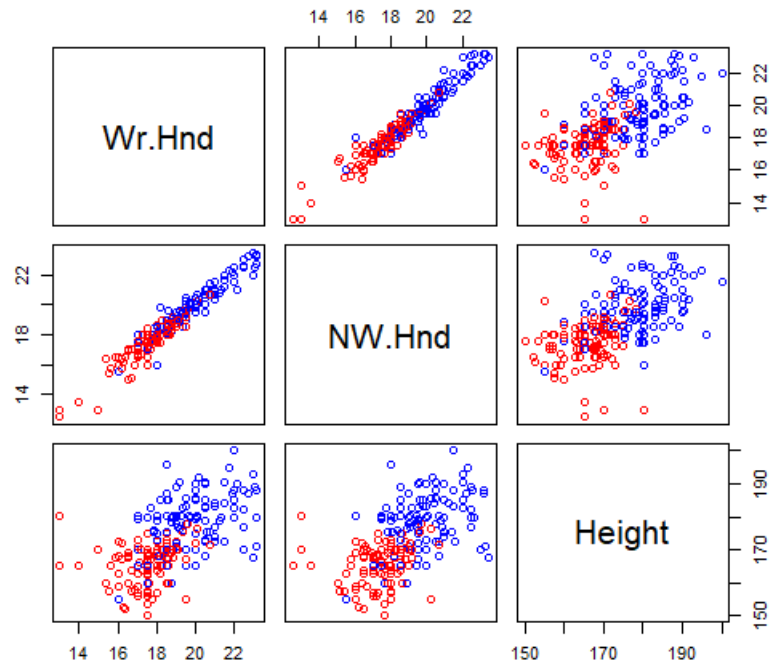
```
x = base.1[c(2,3,4)]
```

```
plot(x[c(1,2)],col=couleur)
```

```
plot(x[c(2,3)],col=couleur)
```

```
plot(x[c(1,3)],col=couleur)
```





#question 6

plot3d(x[c(1,2,3)],col=couleur)

# Les données sont rognées

# Non, cette représentation n'est pas correcte pour visualiser correctement les données

# il faudrait mettre les données à l'échelle.

plot3d(x[c(1,2,3)],xlim=range(x),ylim=range(x),zlim=range(x),col=couleur)

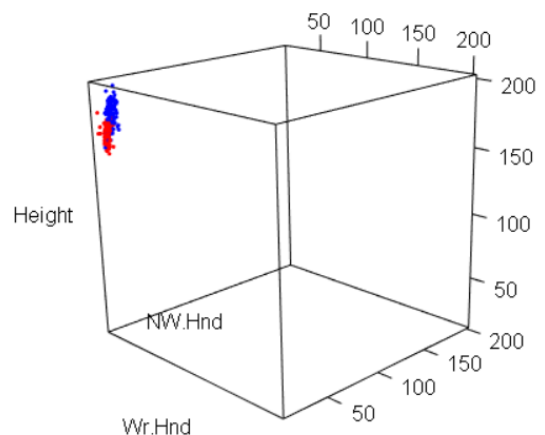
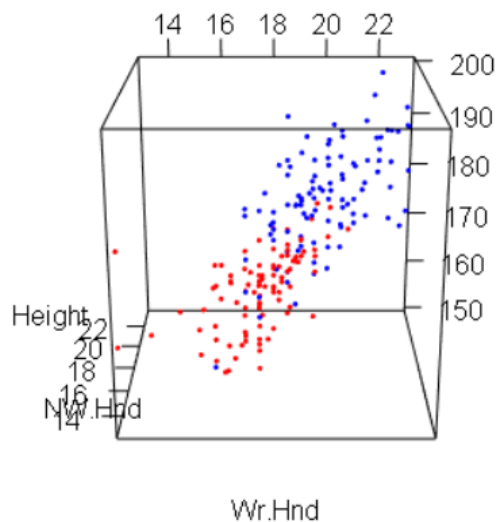
#Toute les données se retrouve dans un coin du graphique

moywr=mean(x[,1])

moynw=mean(x[,2])

moyheight=mean(x[,3])

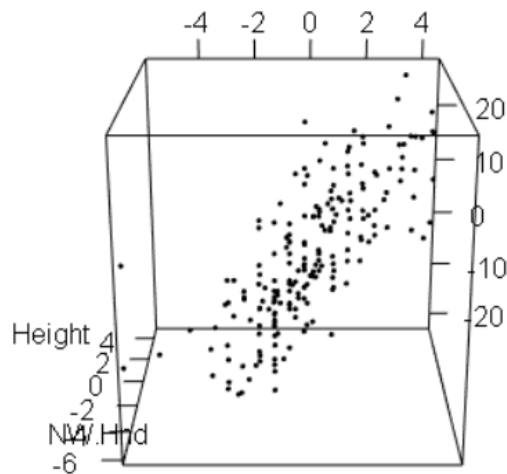
#La moyenne height est 10 fois plus élevée que les autres



### #Question 7

```
x.c = scale(x,center=TRUE,scale=FALSE)
print(x.c)
plot3d(x.c)
varwr = var(x[,1])
varnw = var(x[,2])
varheight = var(x[,3])
```

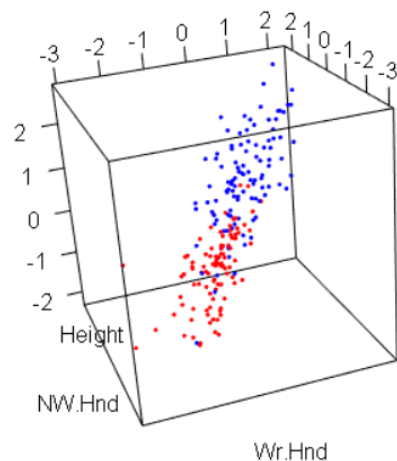
# ça ne va pas parceque l'écart type  
n'est pas bon  
# Je propose de réduire avant de  
centré



### # Question 8

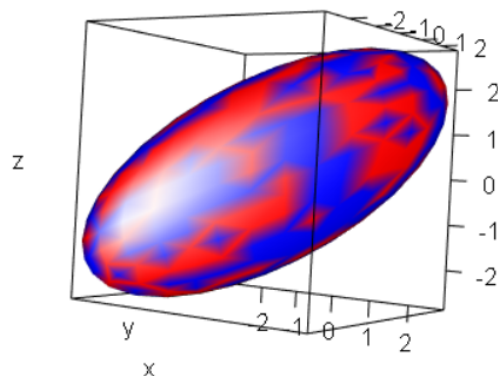
```
x.cr = scale(x)
plot(x.cr, col=couleur)
```

# Le graphique est plus représentatif des  
données désormais

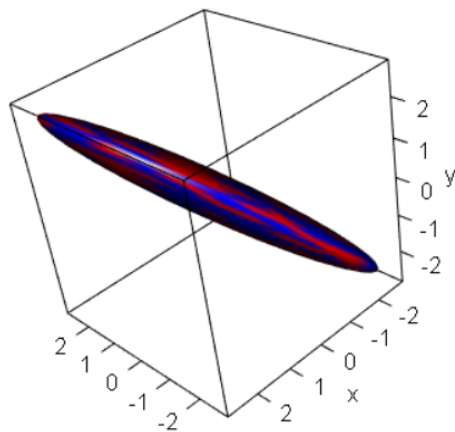
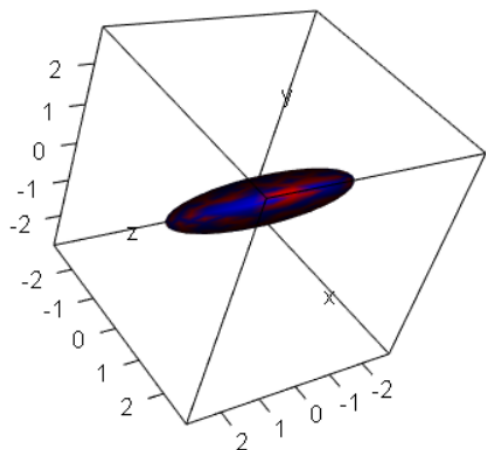


### # Question 9

# Le nuage de point précédent semble éllipsoïdale  
plot3d(ellipse3d(cor(x.cr),level = 0.95, col=couleur))



### # Question 10



Je choisis le plan (1,2)  
Il représente l'ellipse en mettant en valeur  
sa longueur et sa largeur.  
Ainsi, L'amplitude de la plage de données  
est plus élevée.

