

group1_lab1

```
rm(list=ls())
setwd("C:\\RPI\\ITWS6600 Data Analytics")
GPW3 <- read.csv("GPW3_GRUMP_SummaryInformation_2010.csv")
library(readxl)
EPIxls<- read_xls("2010EPI_data.xls",sheet="EPI2010_onlyEPIcountries")
EPI_data<- read.csv("2010EPI_data.csv", skip=1)
View(EPI_data)
attach(EPI_data)
fix(EPI_data)
EPI
```

```
##      [1] 93.5 89.1 86.4 86.0 81.1 80.6 78.2 78.1 78.1 76.8 76.3 74.7 74.5 74.2
##      [15] 73.4 73.3 73.2 73.1 73.0 72.5 72.5 71.6 71.4 71.4 70.6 69.9 69.8 69.6
##      [29] 69.4 69.3 69.3 69.2 69.1 69.1 68.7 68.4 68.3 68.2 68.2 68.0 67.8 67.4
##      [43] 67.3 67.1 67.0 66.4 66.4 65.9 65.9 65.7 65.7 65.6 65.4 65.0 65.0 64.6
##      [57] 63.8 63.7 63.6 63.5 63.5 63.4 63.1 62.9 62.5 62.4 62.2 62.0 61.2 61.0
##      [71] 60.9 60.8 60.6 60.6 60.5 60.4 60.4 60.0 59.7 59.6 59.3 59.2 59.1 59.1
##      [85] 59.0 58.8 58.2 58.1 58.0 57.9 57.3 57.3 57.1 57.0 56.4 56.3 56.1 55.9
##      [99] 55.3 54.6 54.4 54.3 54.2 54.0 54.0 51.6 51.4 51.4 51.3 51.3 51.3 51.2
##     [113] 51.1 51.1 50.8 50.3 50.1 49.9 49.8 49.2 49.0 48.9 48.3 48.3 48.0 47.9
##     [127] 47.8 47.3 47.1 47.0 45.9 44.7 44.6 44.6 44.6 44.4 44.3 44.3 44.0 43.9
##     [141] 43.1 42.8 42.3 42.3 42.0 41.9 41.8 41.7 41.3 41.0 40.8 40.7 40.2 39.6
##     [155] 39.5 39.4 38.4 37.6 36.4 36.3 33.7 33.3 32.1  NA   NA   NA   NA   NA
##     [169]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [183]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [197]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [211]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [225]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [239]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [253]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [267]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [281]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [295]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [309]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [323]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [337]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [351]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [365]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [379]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [393]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [407]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [421]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [435]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [449]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [463]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##     [477]  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
tf<- is.na(EPI)
E<- EPI[!tf]
```

```
#Exercise 1: exploring the distribution
summary(EPI)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
##    32.10  48.60   59.20   58.37  67.60   93.50   65304
```

```
fivenum(EPI, na.rm=TRUE)
```

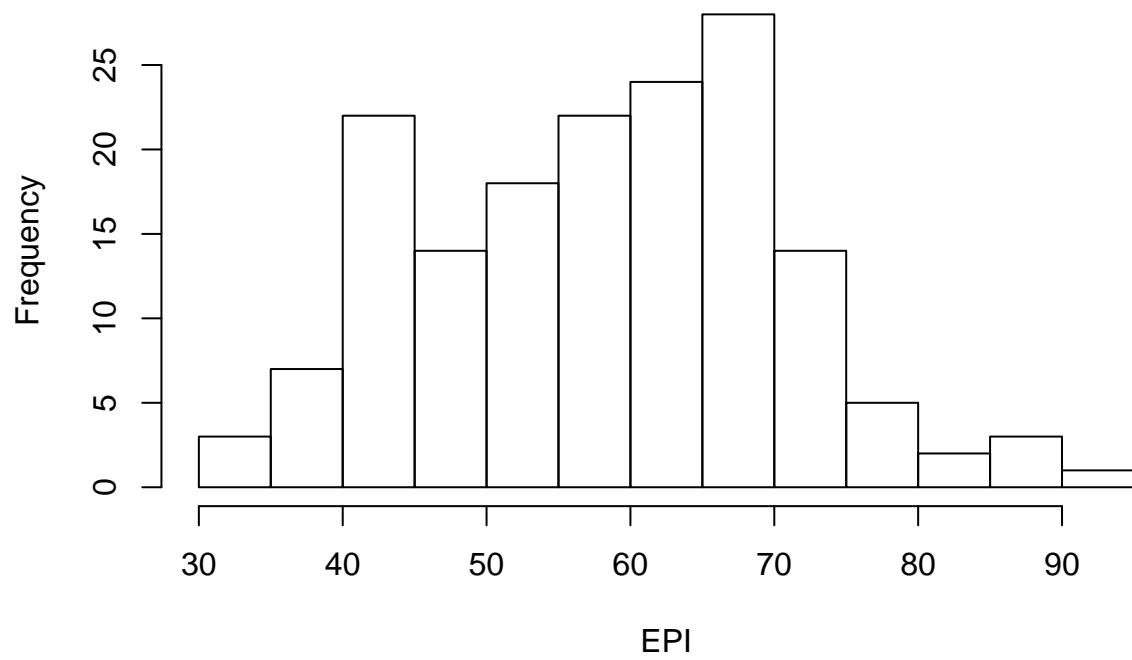
```
## [1] 32.1 48.6 59.2 67.6 93.5
```

```
stem(EPI)
```

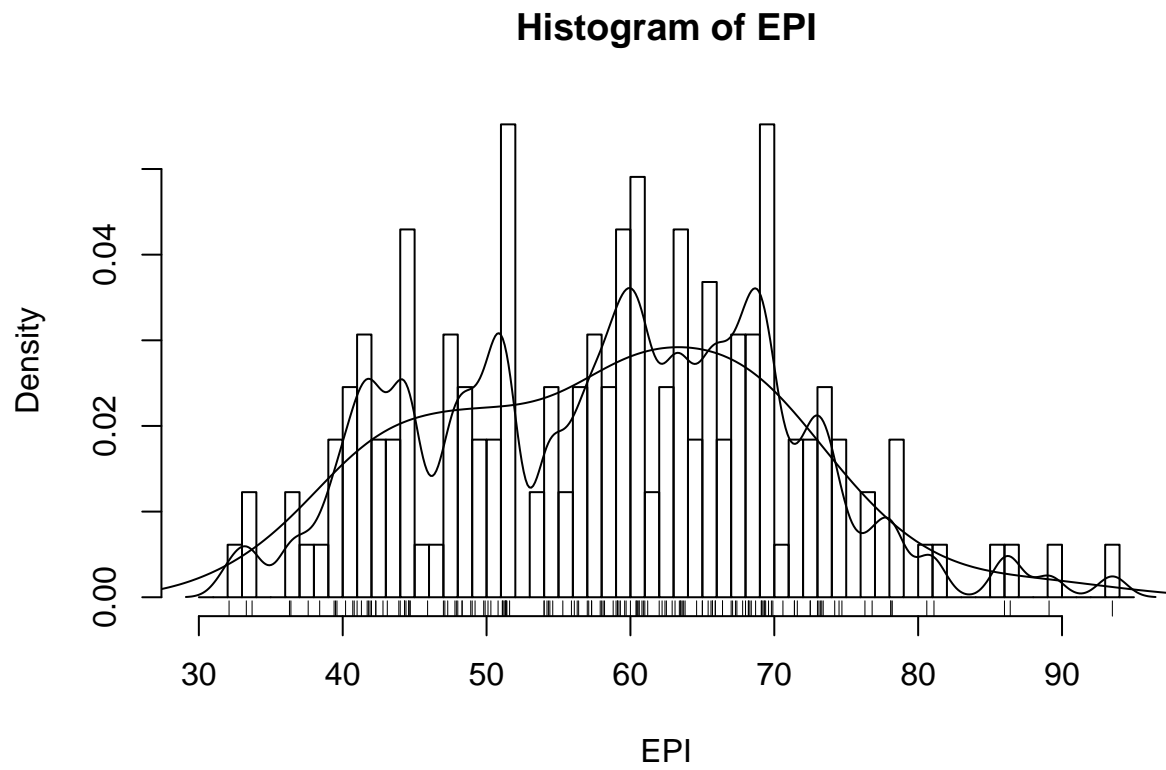
```
##
## The decimal point is 1 digit(s) to the right of the |
##
## 3 | 234
## 3 | 66889
## 4 | 00011112222223344444
## 4 | 5555677788888999
## 5 | 0000111111111244444
## 5 | 55666677778888999999
## 6 | 0000011111111222333344444
## 6 | 5555666666677778888889999999
## 7 | 000111233333334
## 7 | 5567888
## 8 | 11
## 8 | 669
## 9 | 4
```

```
hist(EPI)
```

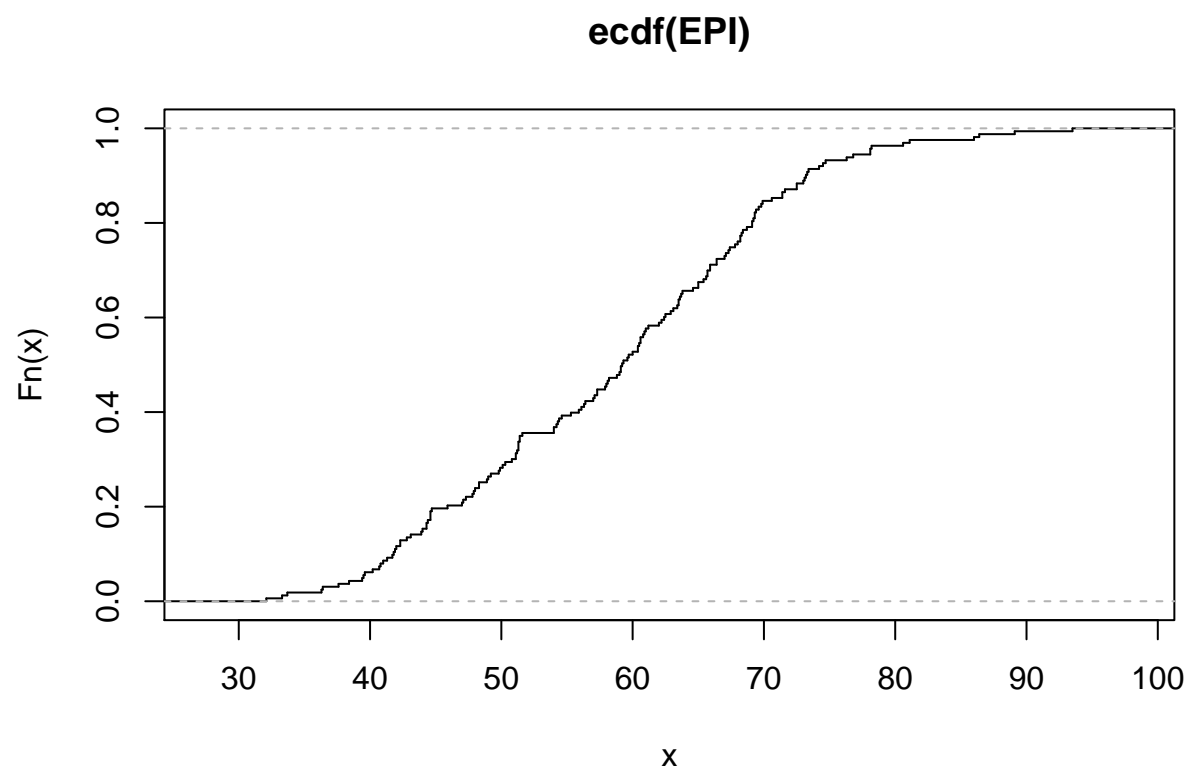

Histogram of EPI



```
hist(EPI,seq(30.,95.,1.0), prob=TRUE)
lines(density(EPI,na.rm = T,bw=1.))
lines(density(EPI,na.rm = T,bw="SJ"))
rug(EPI)
```

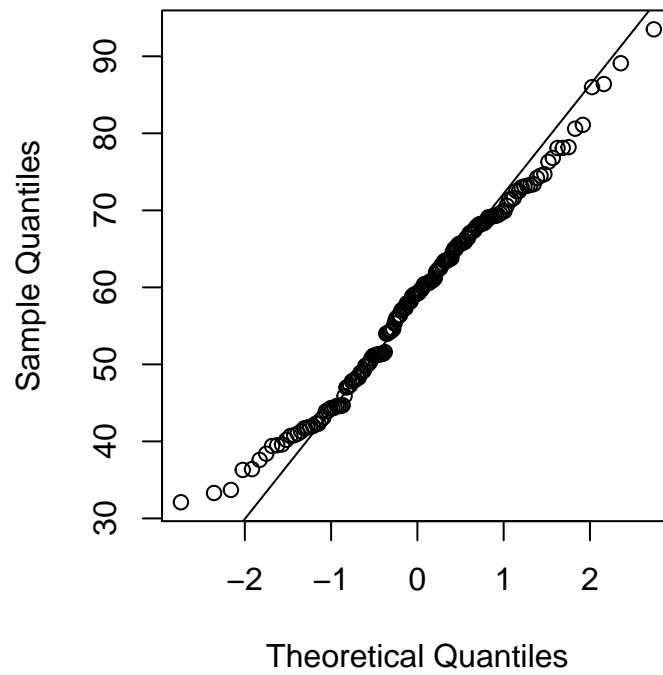


```
#Exercise 1: fitting a distribution beyond histograms  
plot(ecdf(EPI),do.points=F, verticals = T)
```

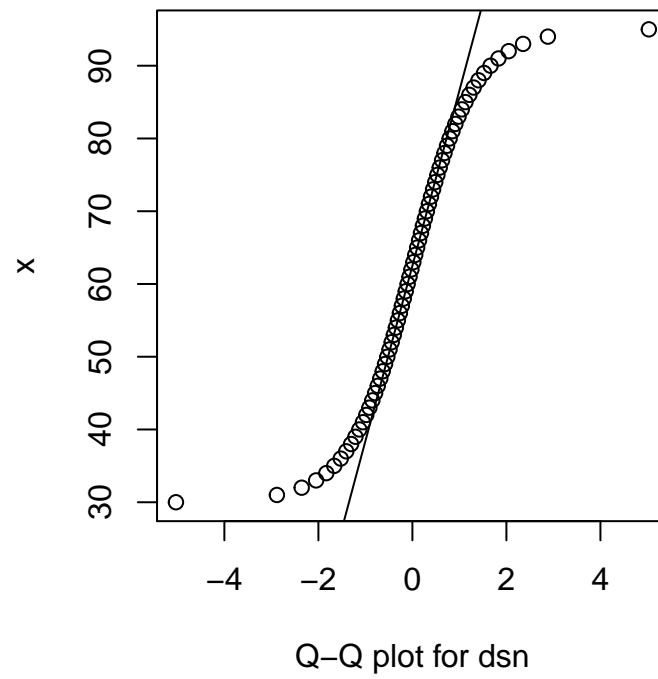


```
par(pty="s")  
qqnorm(EPI)  
qqline(EPI)
```

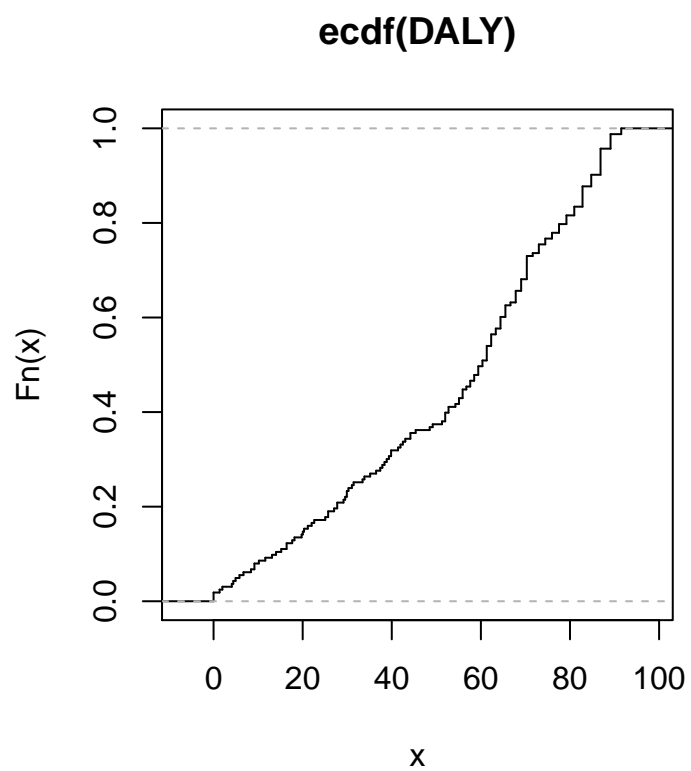
Normal Q-Q Plot



```
x<- seq(30,95,1)
qqplot(qt(ppoints(250),df=5),x,xlab= "Q-Q plot for dsn")
qqline(x)
```

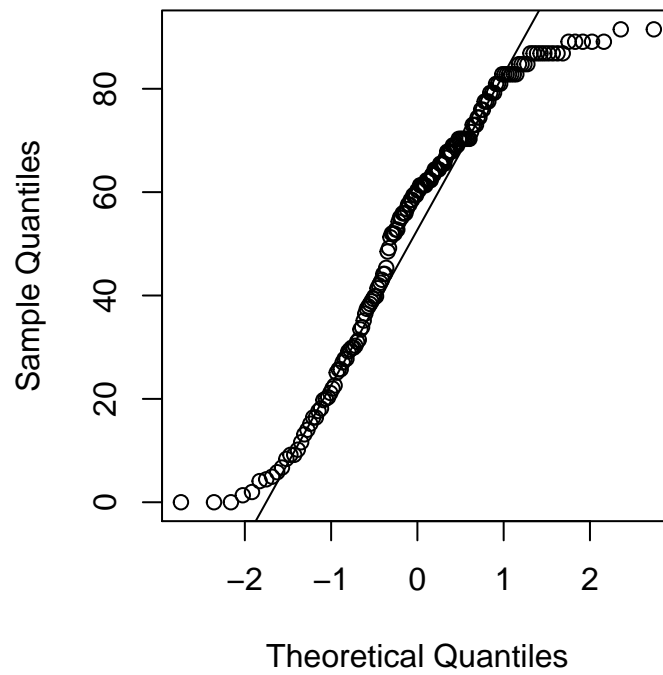


```
plot(ecdf(DALY), do.points=F, verticals = T)
```



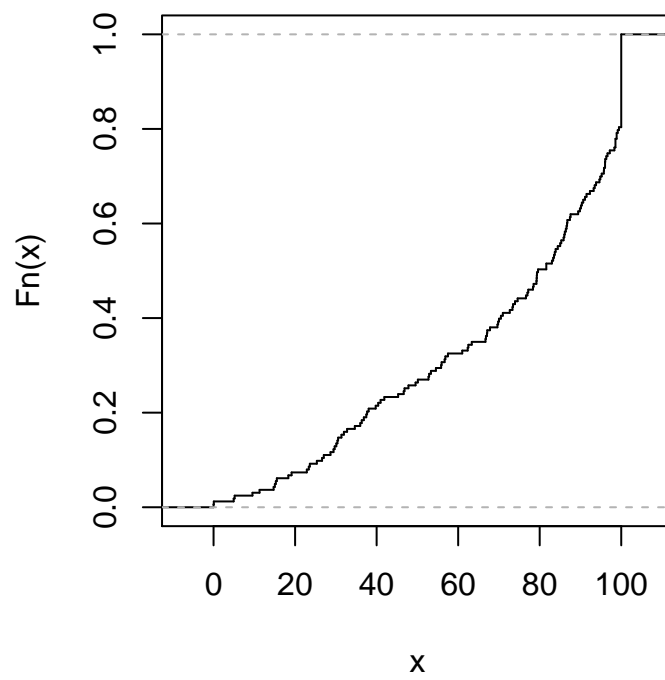
```
qqnorm(DALY)  
qqline(DALY)
```

Normal Q-Q Plot



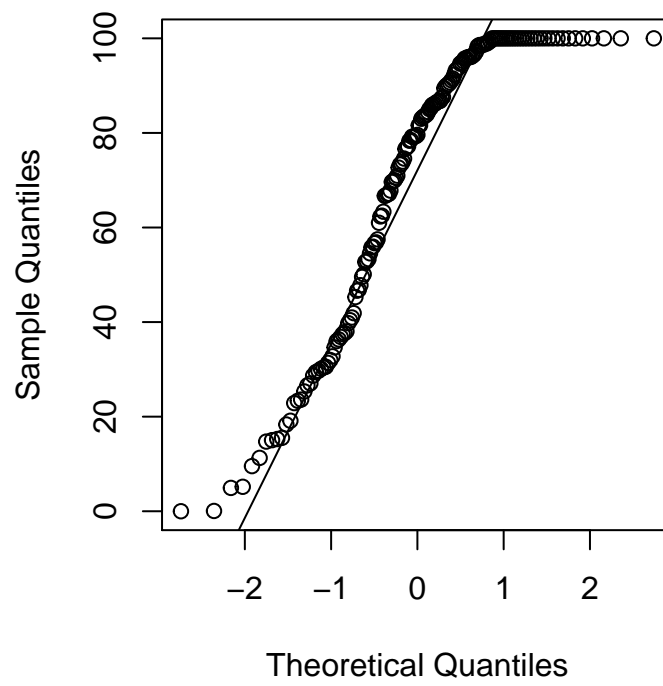
```
plot(ecdf(WATER_H),do.points=F, verticals = T)
```

ecdf(WATER_H)

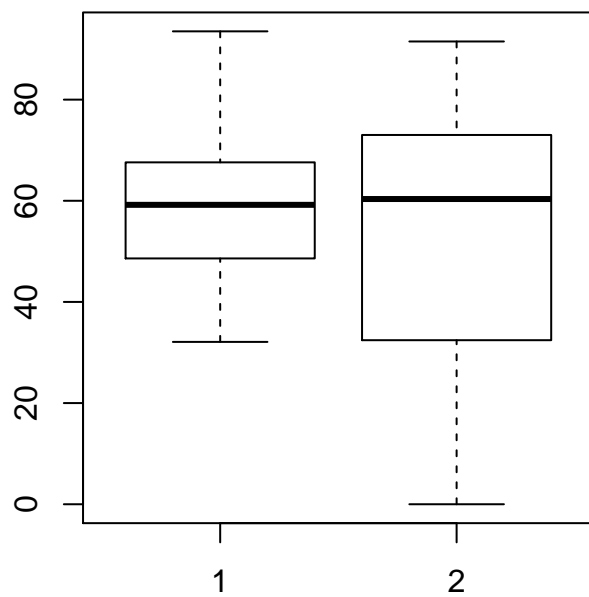


```
qqnorm(WATER_H)  
qqline(WATER_H)
```

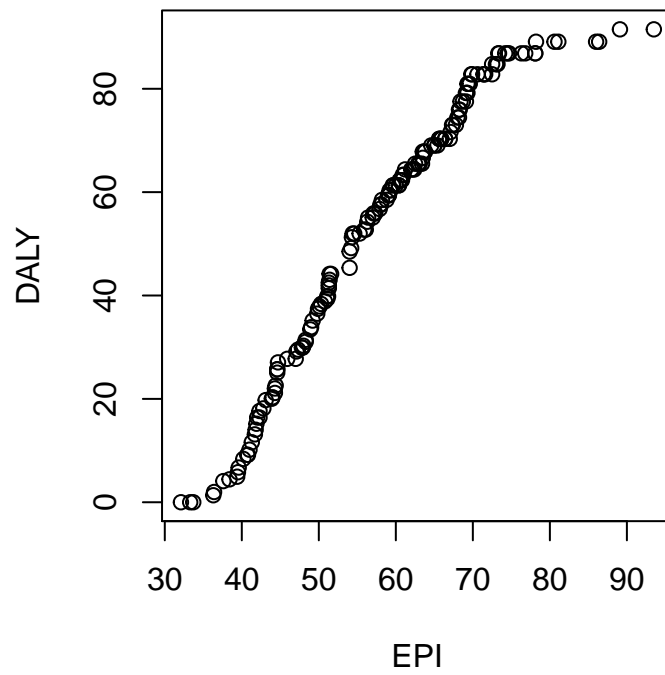

Normal Q-Q Plot



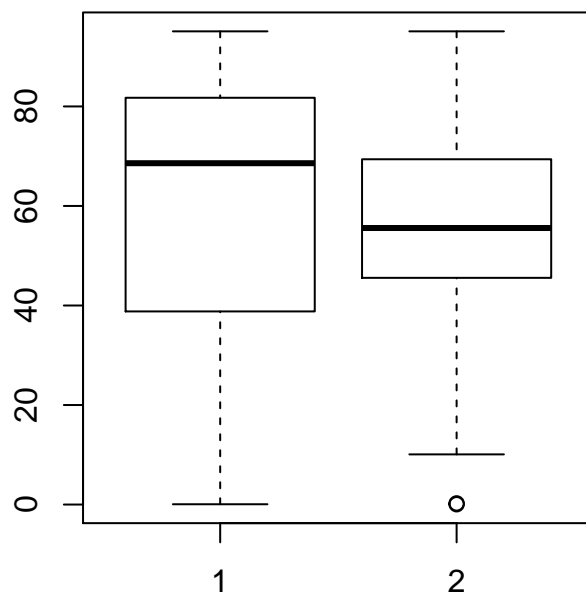
```
boxplot(EPI,DALY)
```



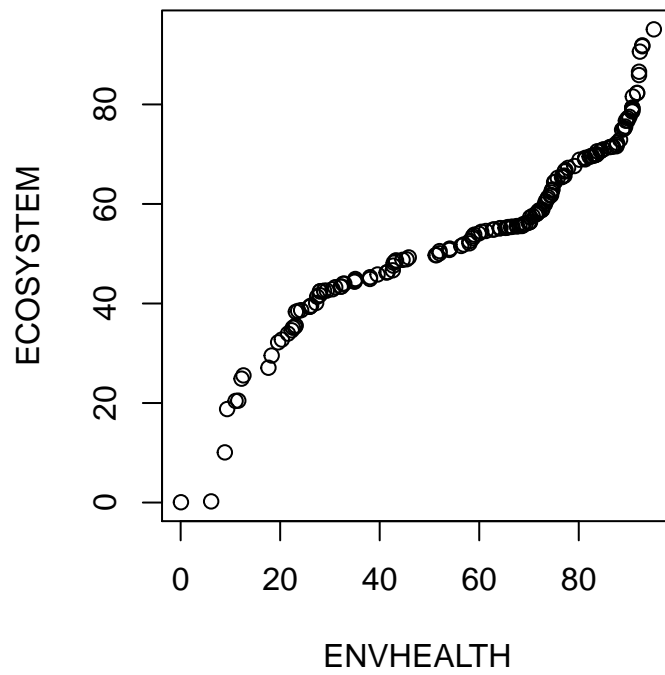
```
qqplot(EPI,DALY)
```



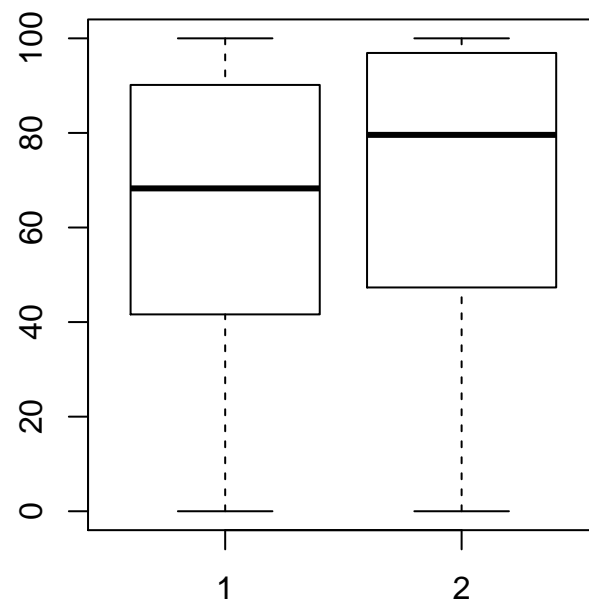
```
boxplot(ENVHEALTH,ECOSYSTEM)
```



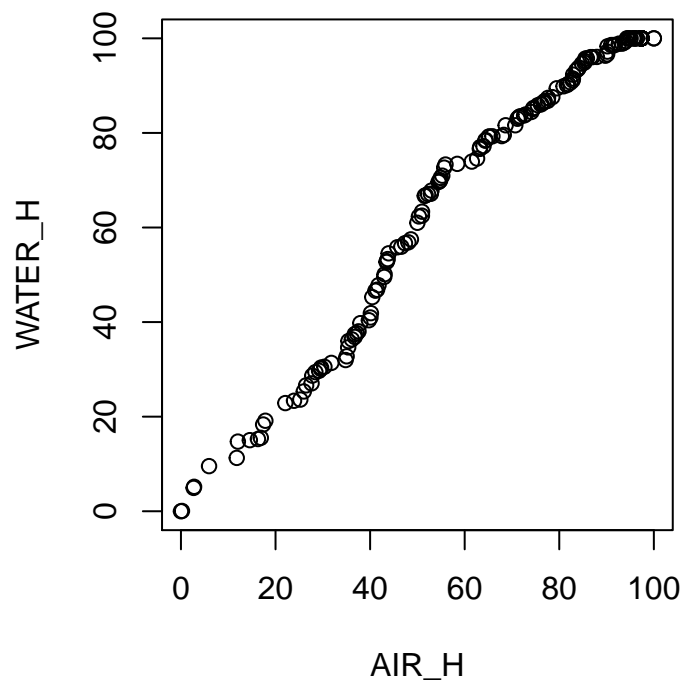
```
qqplot(ENVHEALTH, ECOSYSTEM)
```



```
boxplot(AIR_H,WATER_H)
```



```
qqplot(AIR_H, WATER_H)
```

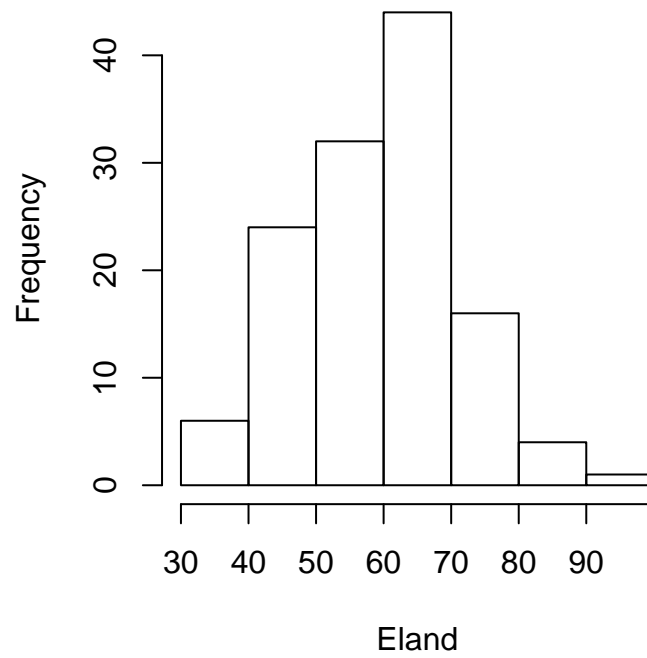


```
help("distributions")
```

```
## starting httpd help server ... done
```

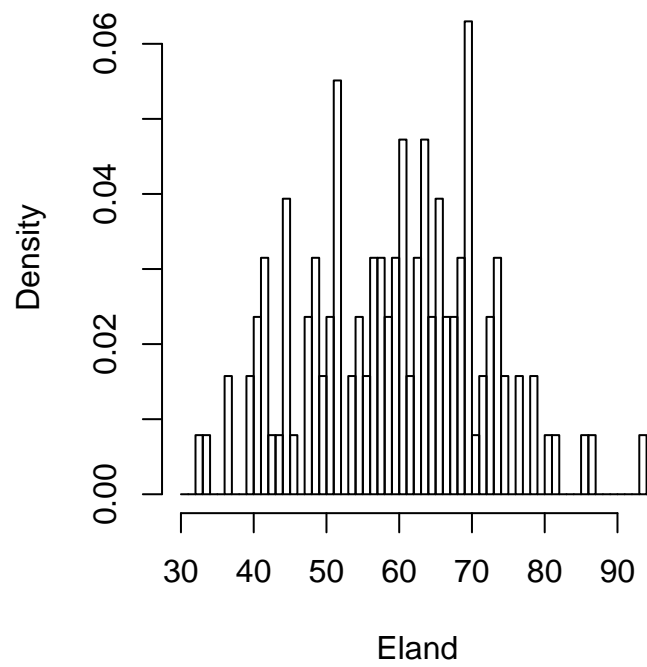
```
EPILand<-EPI[!Landlock]  
Eland<- EPILand[!is.na(EPILand)]  
hist(Eland)
```

Histogram of Eland



```
hist(Eland, seq(30.,95.,1.0), prob=T)
```


Histogram of Eland



```
EPI_South_Asia <- EPI[EPI_regions=="South Asia"]
```