Στατιστική Υπολογιστική
Εργασία 1

Αθανάσιος Θεοχάρης
AEM:2961

ENTER DAAAAAAAAAAAAAAAAAAAATE

Καταχώρηση και προετοιμασία Δεδομένων

Καταχώρηση δεδομένων

Η διεύθυνση του αρχείου EXERCISE_1_WAGES.csv πρέπει να είναι σωστή.

```
EXERCISE_1_WAGES <- read.csv("~/R-course-projects/EXERCISE_1_WAGES.csv")
```

Αλλαγή δεδομένων στις κατηγορηματικές στήλες.

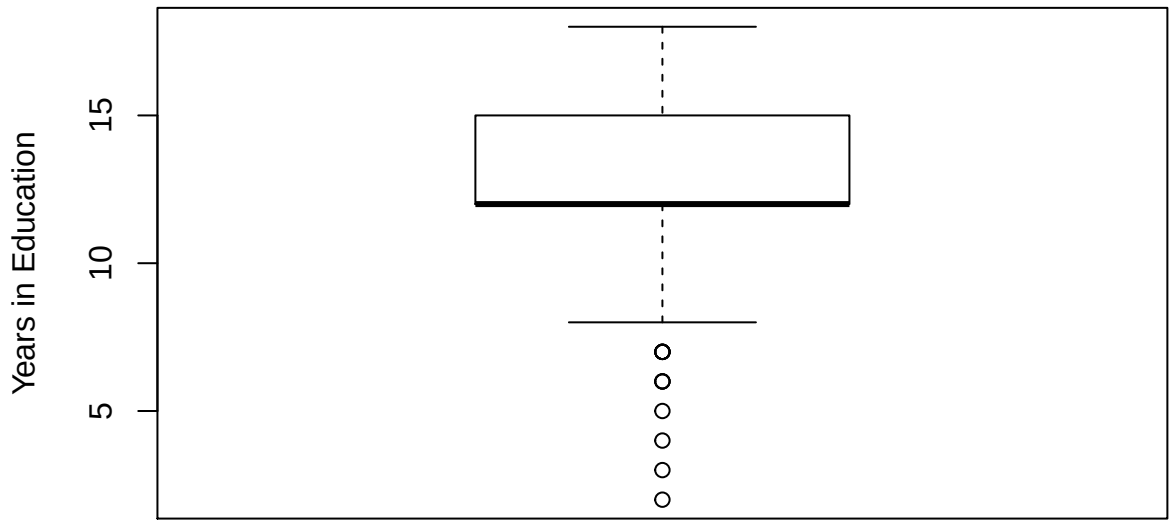
```
EXERCISE_1_WAGES$SEX = factor(EXERCISE_1_WAGES$SEX,  levels=c(0,1), labels =c("Male","Female"))
EXERCISE_1_WAGES$RACE = factor(EXERCISE_1_WAGES$RACE,  levels=c(1,2,3), labels =c("Other","Hispanic", "
EXERCISE_1_WAGES$OCCUPATION = factor(EXERCISE_1_WAGES$OCCUPATION,  levels=c(1,2,3,4,5,6), labels =c("Ma
EXERCISE_1_WAGES$SECTOR = factor(EXERCISE_1_WAGES$SECTOR,  levels=c(0,1,2), labels =c("Other","Manufactu
EXERCISE_1_WAGES$MARR = factor(EXERCISE_1_WAGES$MARR,  levels=c(0,1), labels =c("Unmarried","Married"))
```

Προβολή δεδομένων

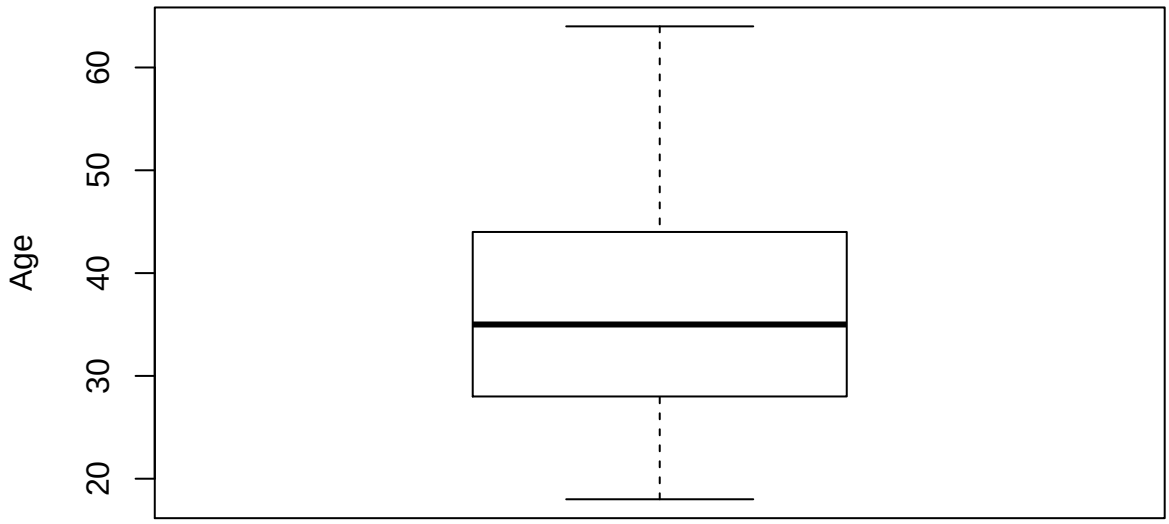
```
View(EXERCISE_1_WAGES)
```

Γραφικές παραστάσεις μεταβλητών

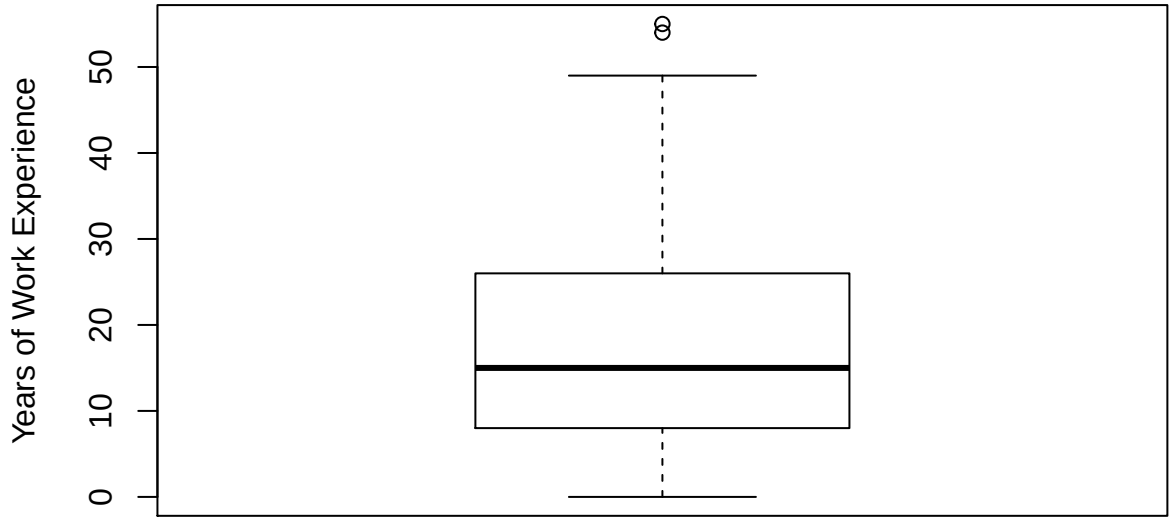
```
boxplot(EXERCISE_1_WAGES$EDUCATION, ylab="Years in Education")
```



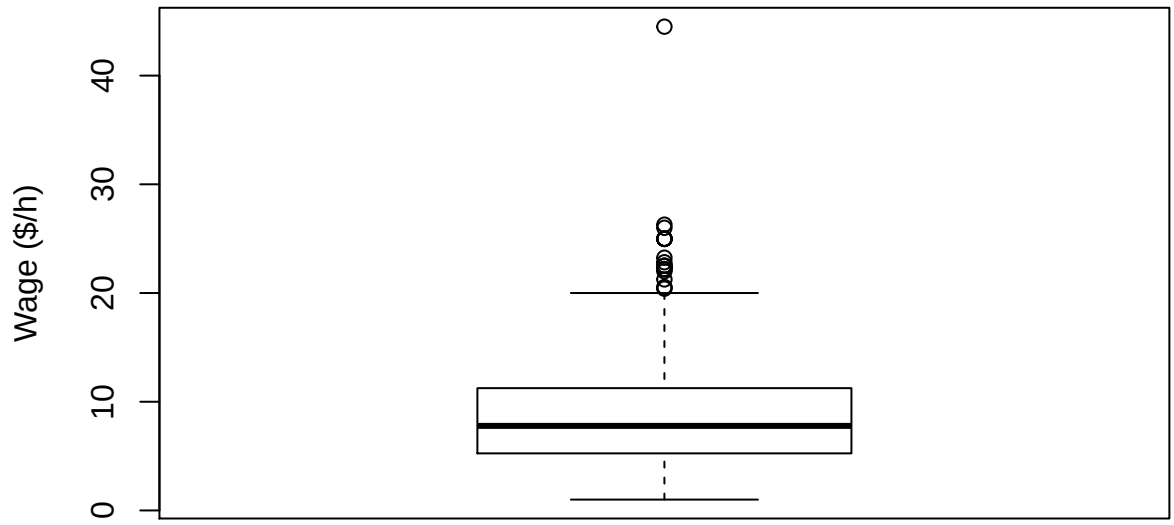
```r
boxplot(EXERCISE_1_WAGES$AGE, ylab="Age")
```



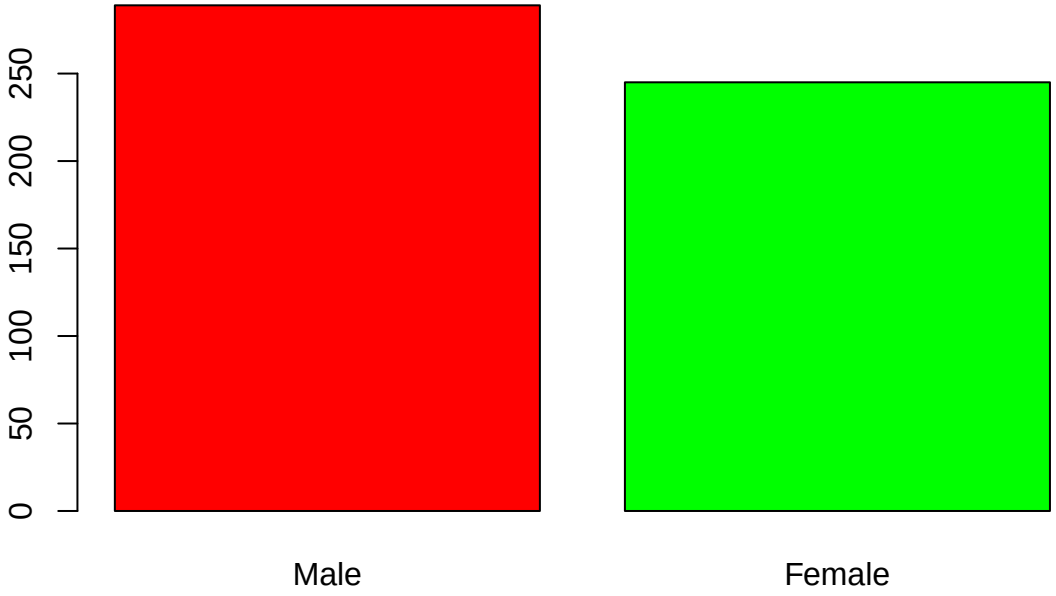

```r
boxplot(EXERCISE_1_WAGES$EXPERIENCE, ylab="Years of Work Experience")
```



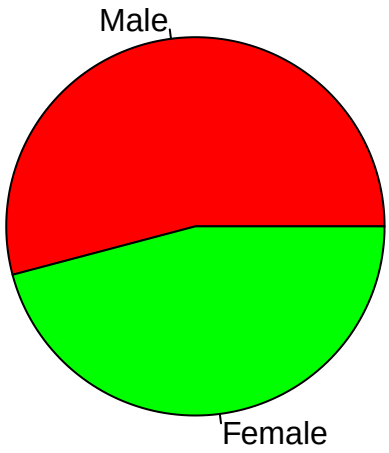

```r
boxplot(EXERCISE_1_WAGES$WAGE, ylab="Wage ($/h)")
```



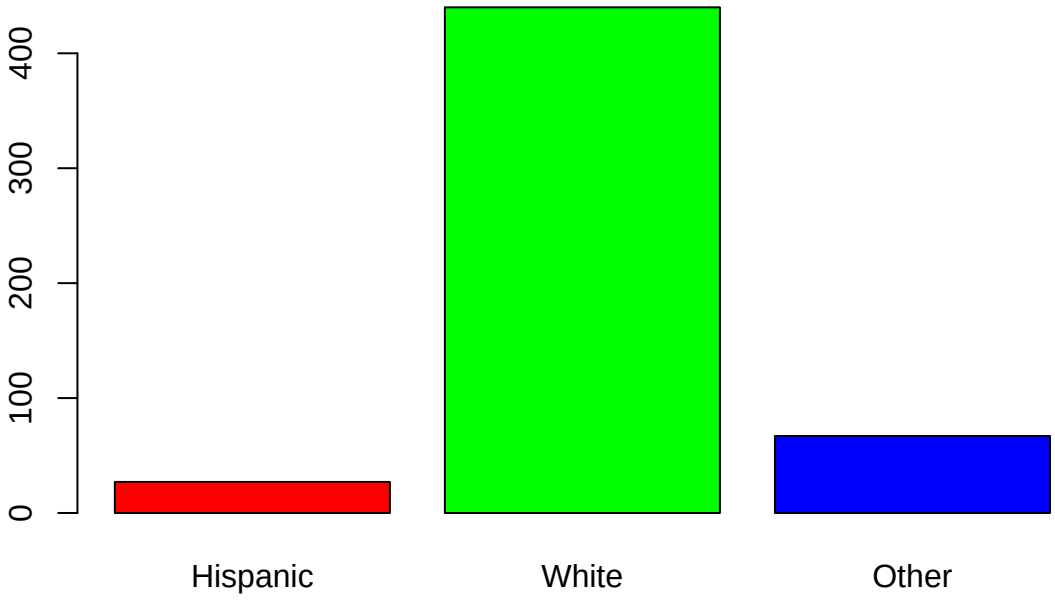

```
table = table(EXERCISE_1_WAGES$SEX)
barplot(c(table["Male"], table["Female"]), col = c("red","green"))
```



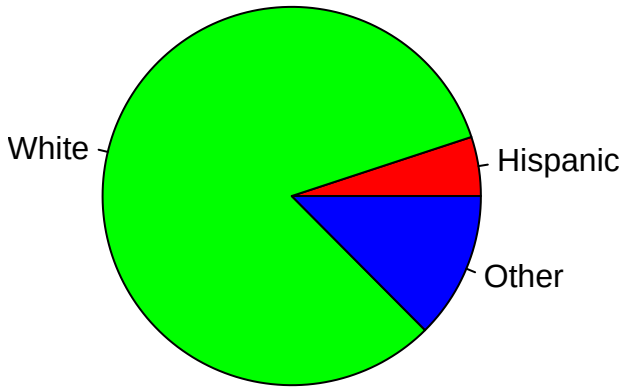

```
pie(c(table["Male"], table["Female"]), col = c("red","green"))
```



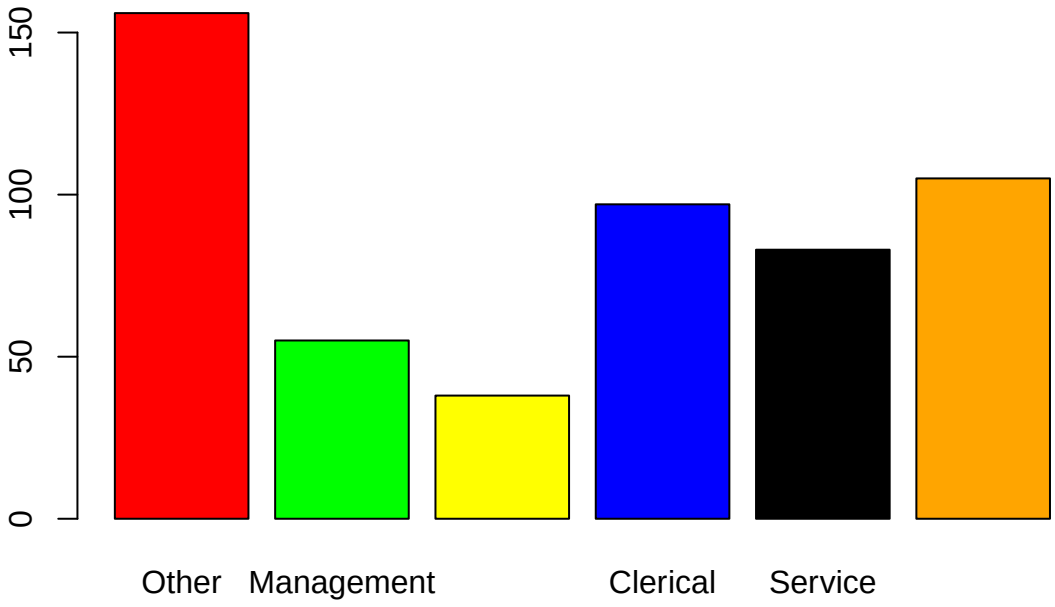
```
table = table(EXERCISE_1_WAGES$RACE)
barplot(c(table["Hispanic"], table["White"], table["Other"]), col = c("red","green","blue"))
```



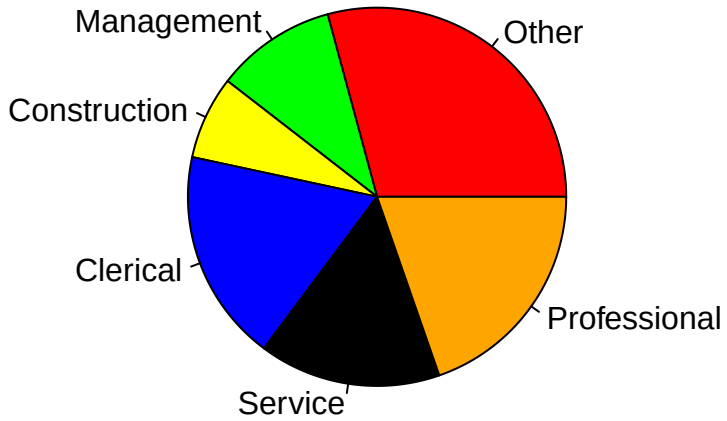

```
pie(c(table["Hispanic"], table["White"], table["Other"]), col = c("red","green","blue"))
```



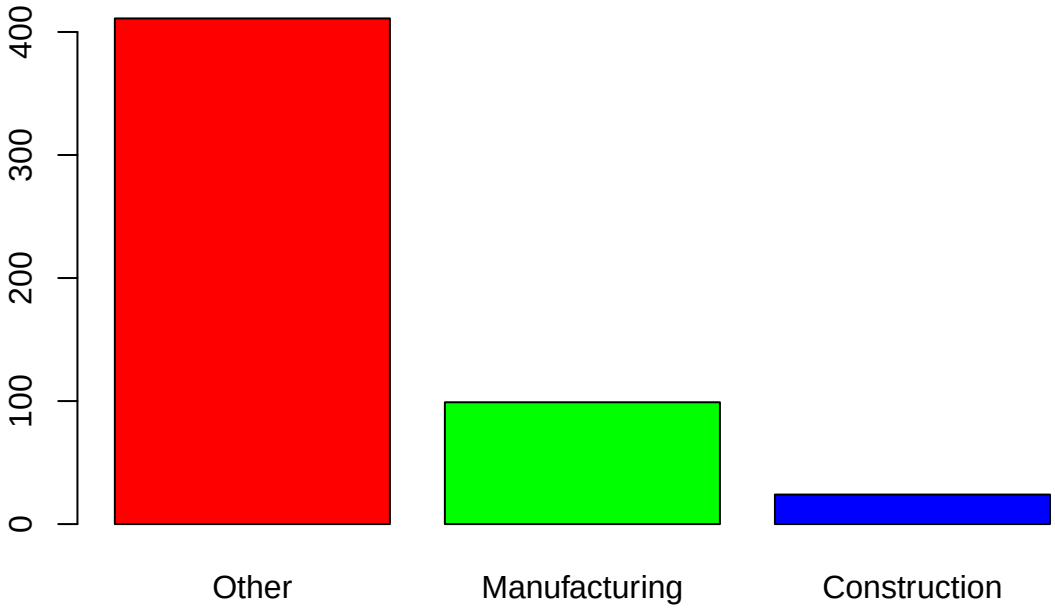

```
table = table(EXERCISE_1_WAGES$OCCUPATION)
barplot(c(table["Other"], table["Management"], table["Construction"], table["Clerical"], table["Service
```



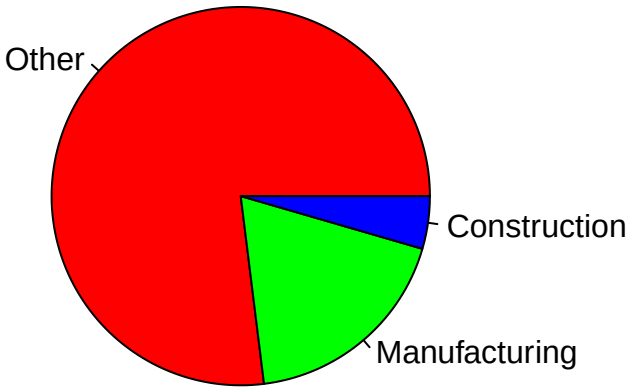

```
pie(c(table["Other"], table["Management"], table["Construction"], table["Clerical"], table["Service"],
```



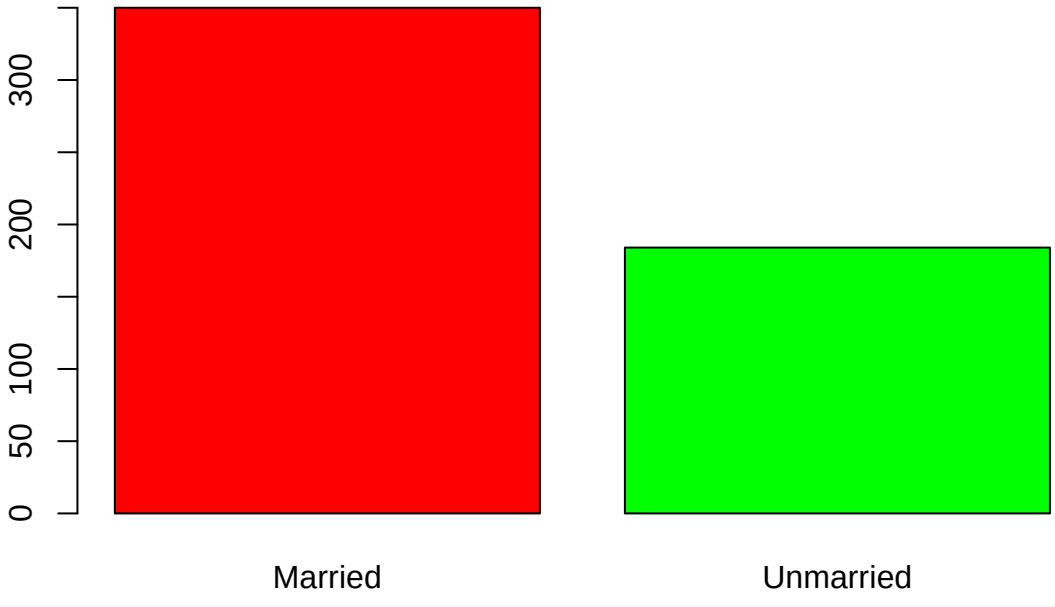

```
table = table(EXERCISE_1_WAGES$SECTOR)
barplot(c(table["Other"], table["Manufacturing"], table["Construction"]), col = c("red","green", "blue"
```



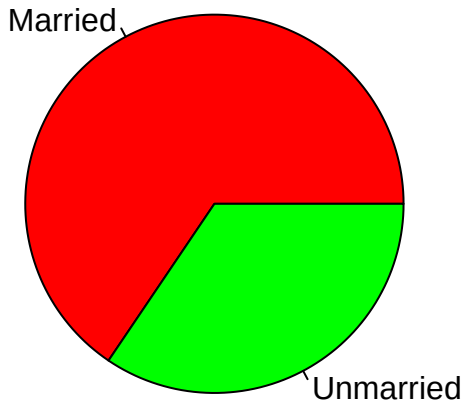

```
pie(c(table["Other"], table["Manufacturing"], table["Construction"]), col = c("red","green", "blue"))
```



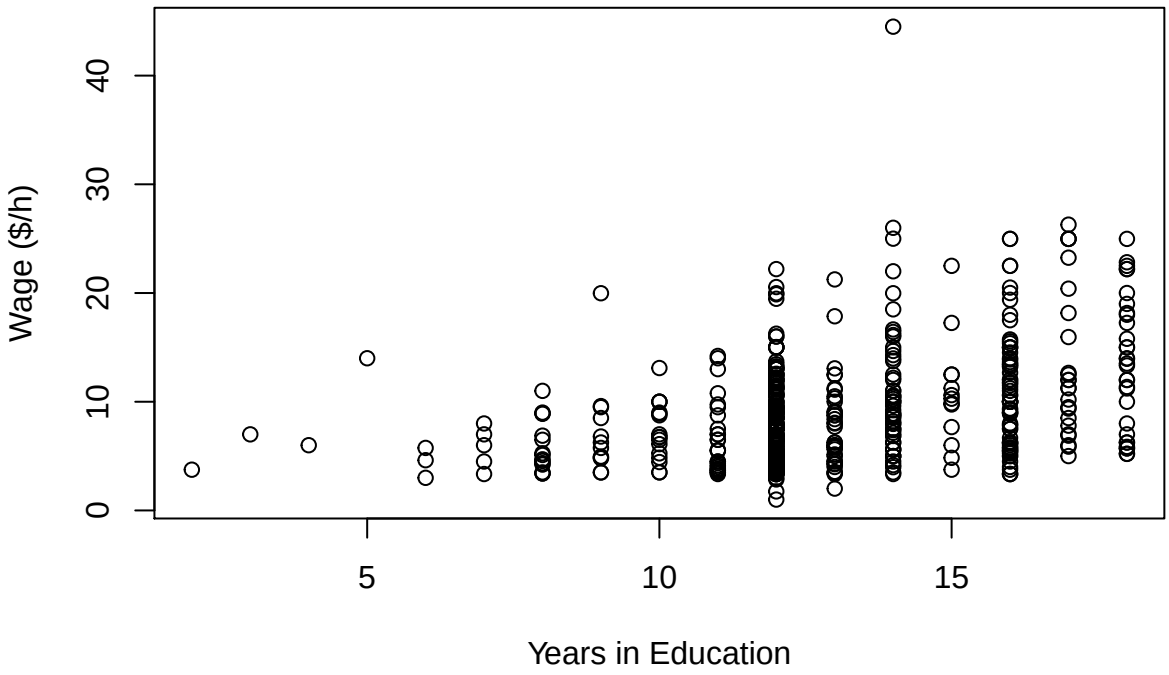

```
table = table(EXERCISE_1_WAGES$MARR)
barplot(c(table["Married"], table["Unmarried"]), col = c("red","green"))
```



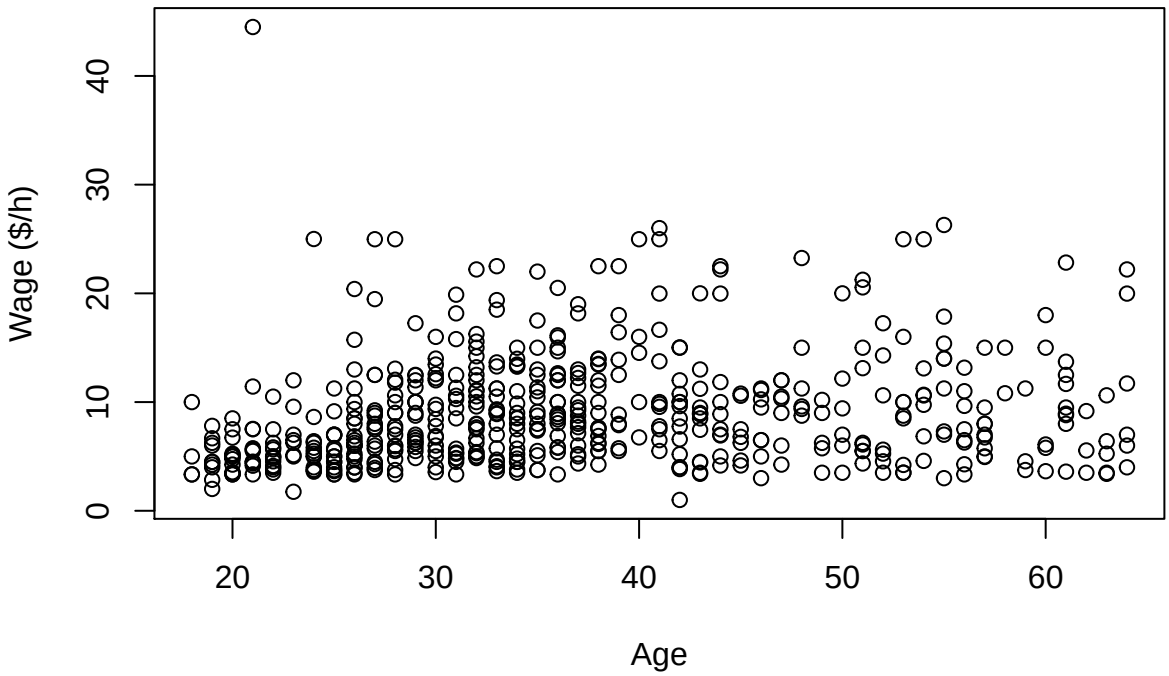

```
pie(c(table["Married"], table["Unmarried"]), col = c("red","green"))
```




```
plot(EXERCISE_1_WAGES$EDUCATION, EXERCISE_1_WAGES$WAGE, xlab = "Years in Education", ylab = "Wage ($/h)
```

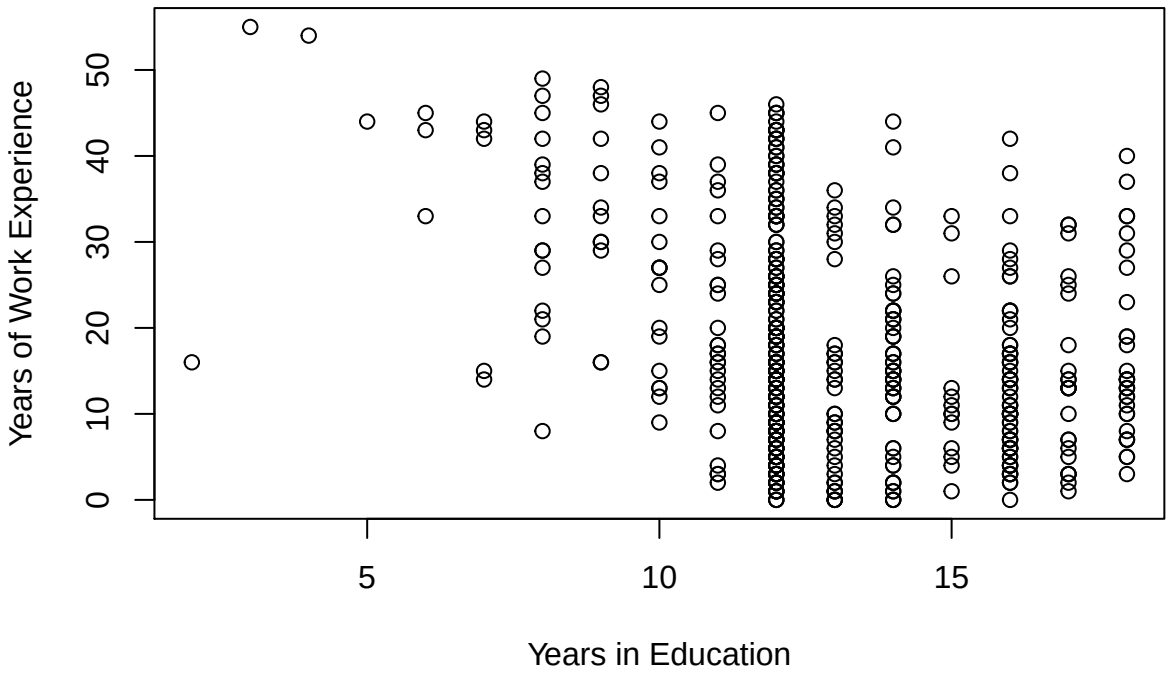



```
plot(EXERCISE_1_WAGES$AGE, EXERCISE_1_WAGES$WAGE, xlab = "Age", ylab = "Wage ($/h)")
```

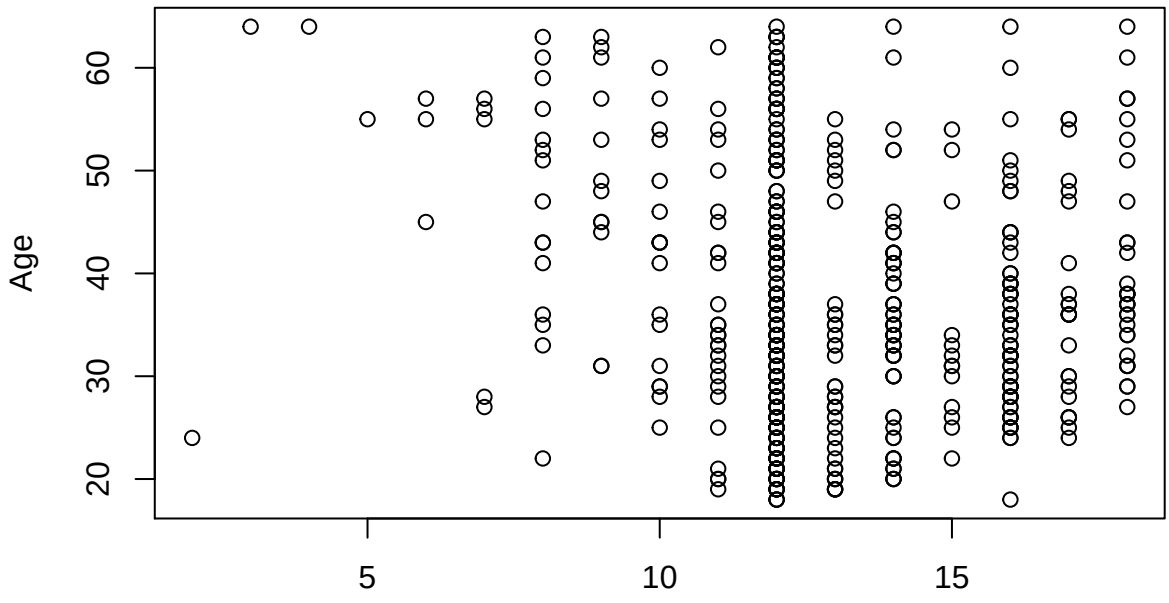



```
plot(EXERCISE_1_WAGES$EXPERIENCE, EXERCISE_1_WAGES$WAGE, xlab = "Years of Work Experience", ylab = "Wag
```

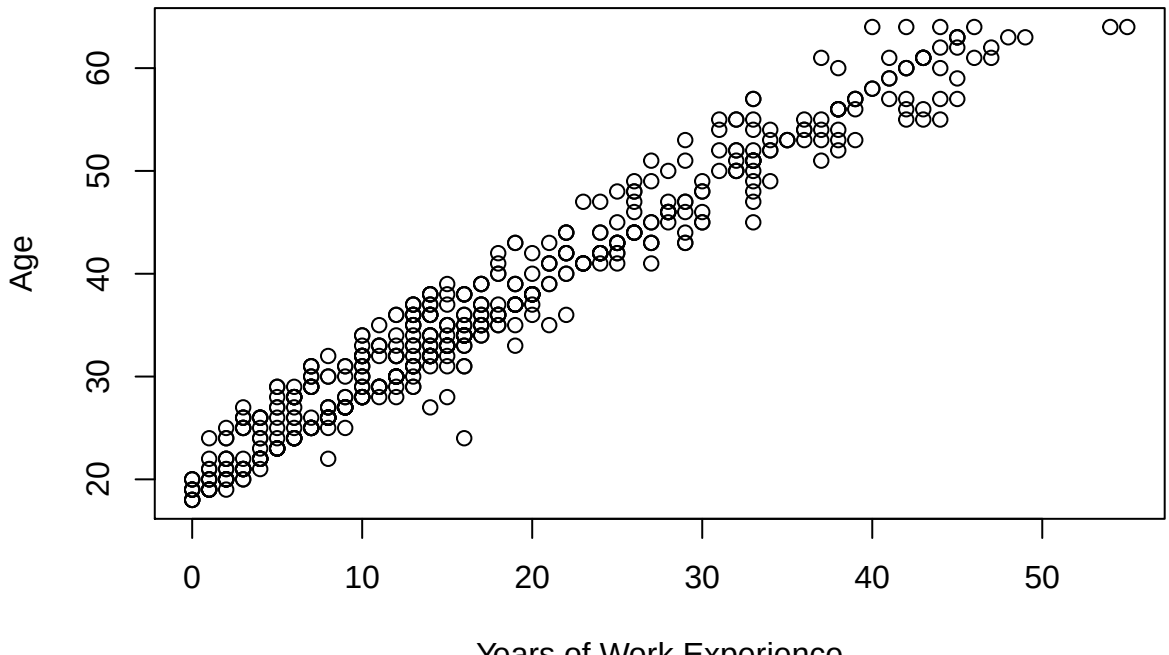

```
plot(EXERCISE_1_WAGES$EDUCATION, EXERCISE_1_WAGES$EXPERIENCE, xlab = "Years in Education", ylab = "Years
```




```
plot(EXERCISE_1_WAGES$EDUCATION, EXERCISE_1_WAGES$AGE, xlab = "Years in Education", ylab = "Age")
```



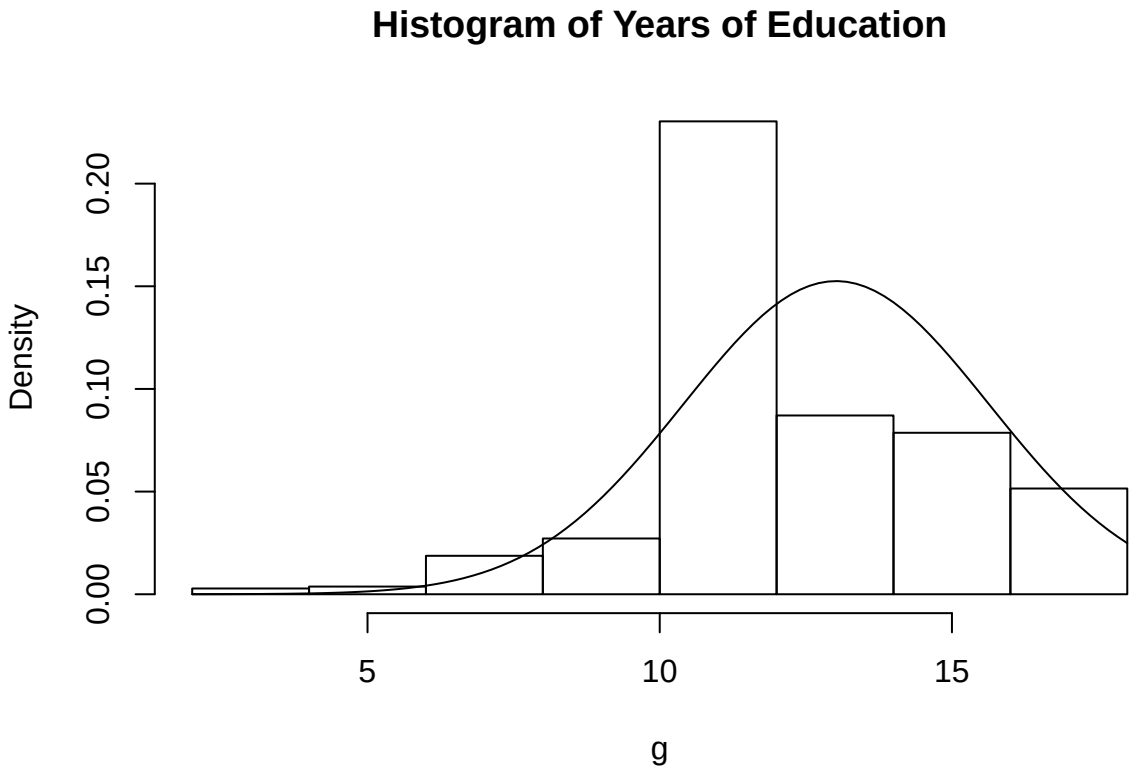

```
plot(EXERCISE_1_WAGES$EXPERIENCE, EXERCISE_1_WAGES$AGE, xlab = "Years of Work Experience", ylab = "Age")
```



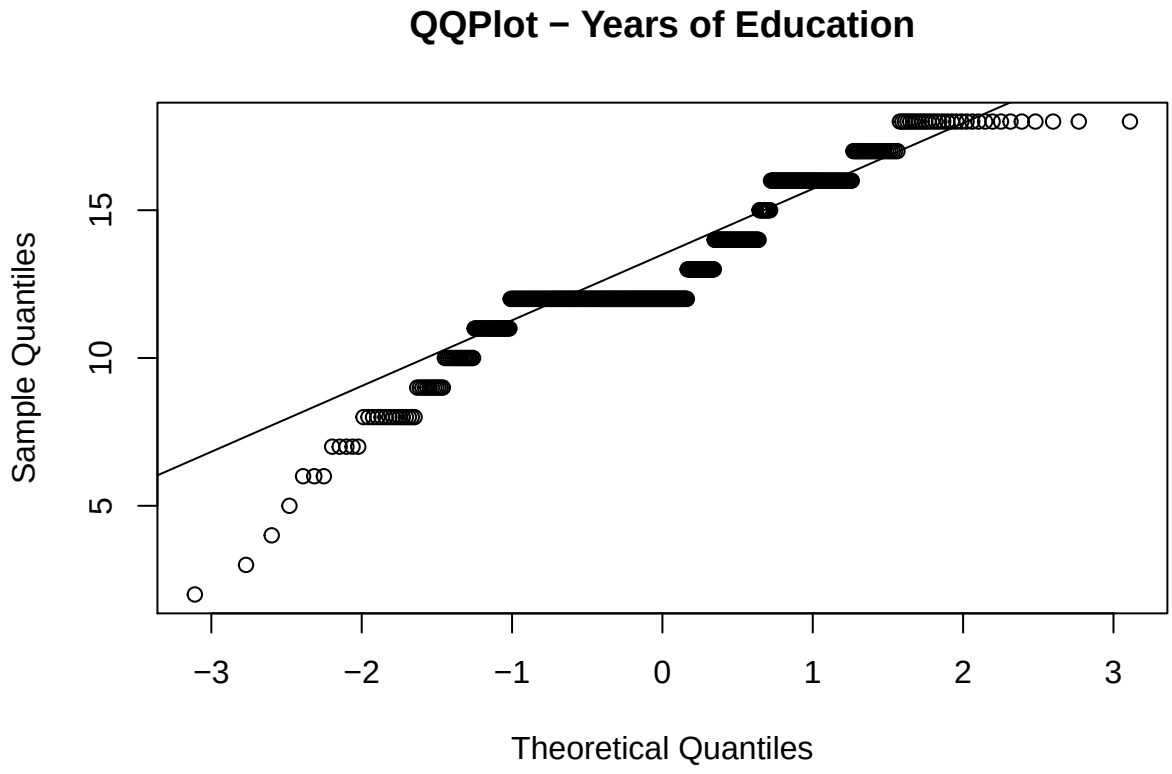
##

Έλεγχος κανονικής κατανομής

```
g<-EXERCISE_1_WAGES$EDUCATION
h<-hist(g, freq = FALSE, main = "Histogram of Years of Education")
x<-seq(min(g), max(g), by=0.02)
curve(dnorm(x, mean = mean(g), sd = sd(g)), add=TRUE)
```



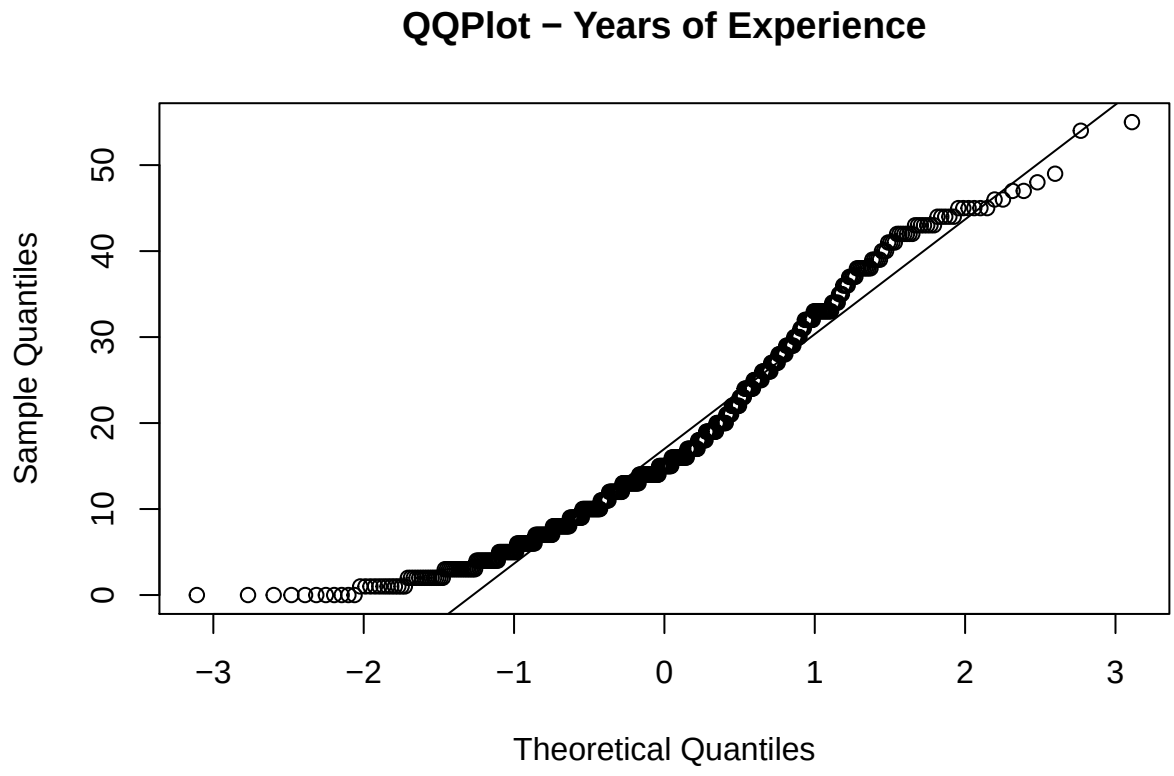

```
qqnorm(g, main = "QQPlot - Years of Education")
qqline(g)
```




```
g<-EXERCISE_1_WAGES$EXPERIENCE
h<-hist(g, freq = FALSE, main = "Histogram of Years of Experience")
```

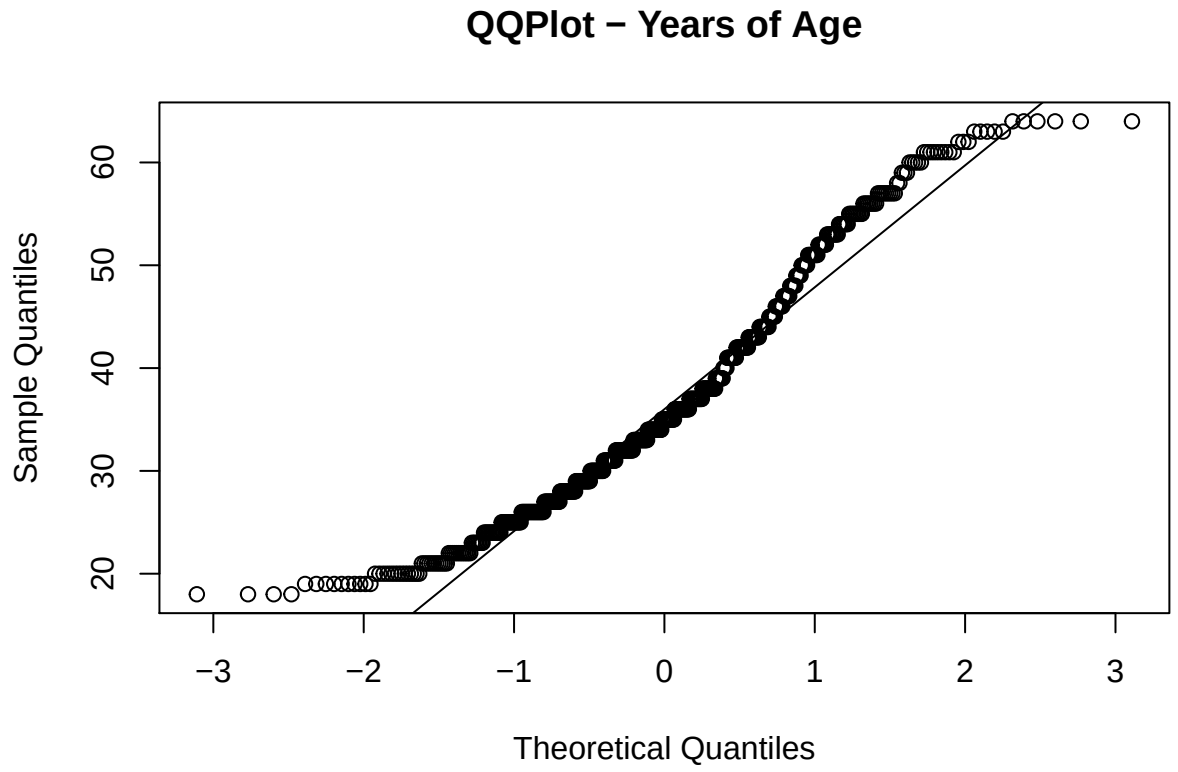
```
x<-seq(min(g), max(g), by=0.02)
curve(dnorm(x, mean = mean(g), sd = sd(g)), add=TRUE)
```


```
qqnorm(g, main = "QQPlot - Years of Experience")
qqline(g)
```

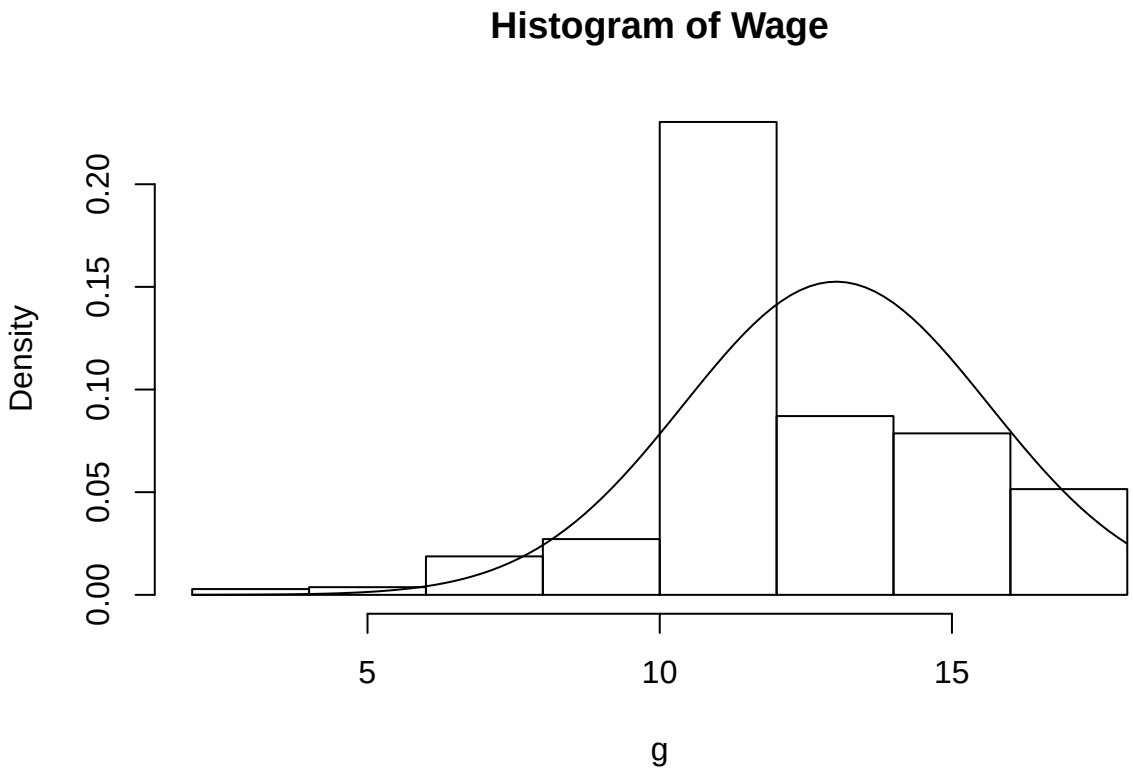


```r
g<-EXERCISE_1_WAGES$AGE
h<-hist(g, freq = FALSE, main = "Histogram of Age")
x<-seq(min(g), max(g), by=0.02)
curve(dnorm(x, mean = mean(g), sd = sd(g)), add=TRUE)
```

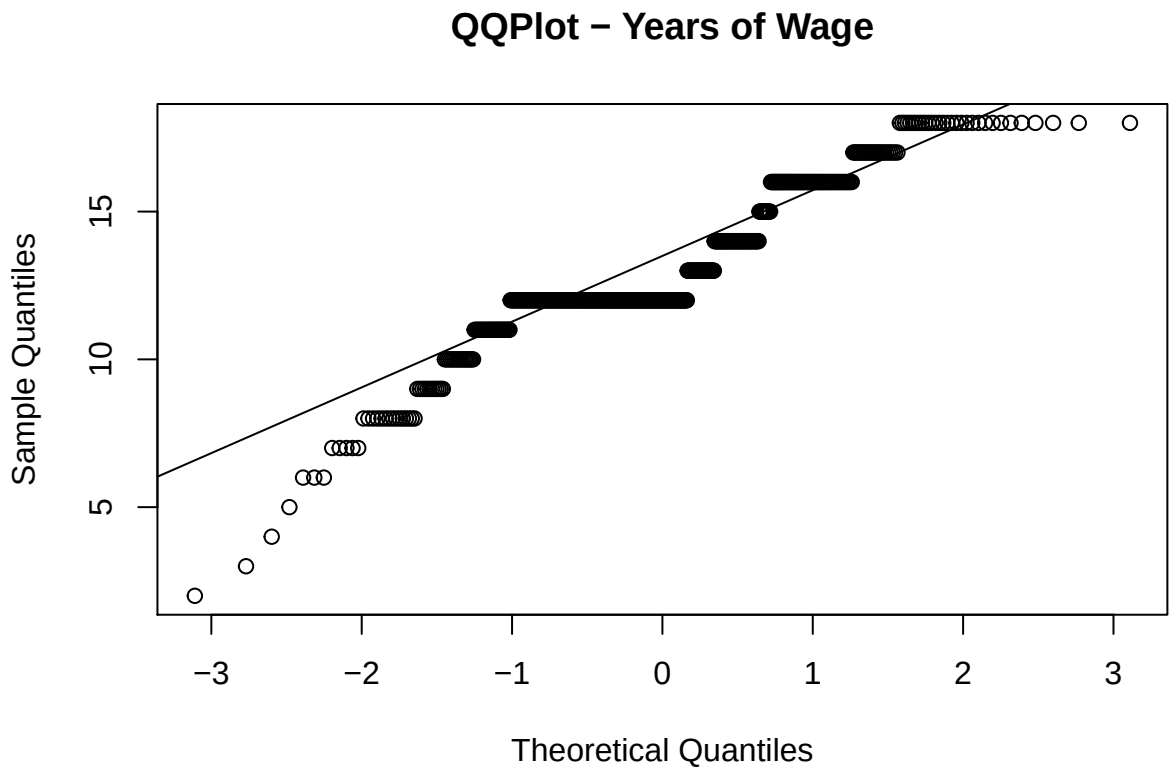

```r
qqnorm(g, main = "QQPlot - Years of Age")
qqline(g)
```




```
g<-EXERCISE_1_WAGES$EDUCATION
h<-hist(g, freq = FALSE, main = "Histogram of Wage")
x<-seq(min(g), max(g), by=0.02)
curve(dnorm(x, mean = mean(g), sd = sd(g)), add=TRUE)
```



```
qqnorm(g, main = "QQPlot - Years of Wage")
qqline(g)
```



Εύρεση δ.ε. a<-numeric(10000) > for(i in 1:10000) a[i]<-mean(sample(speed,replace=T)) > hist(a) > max(a)