

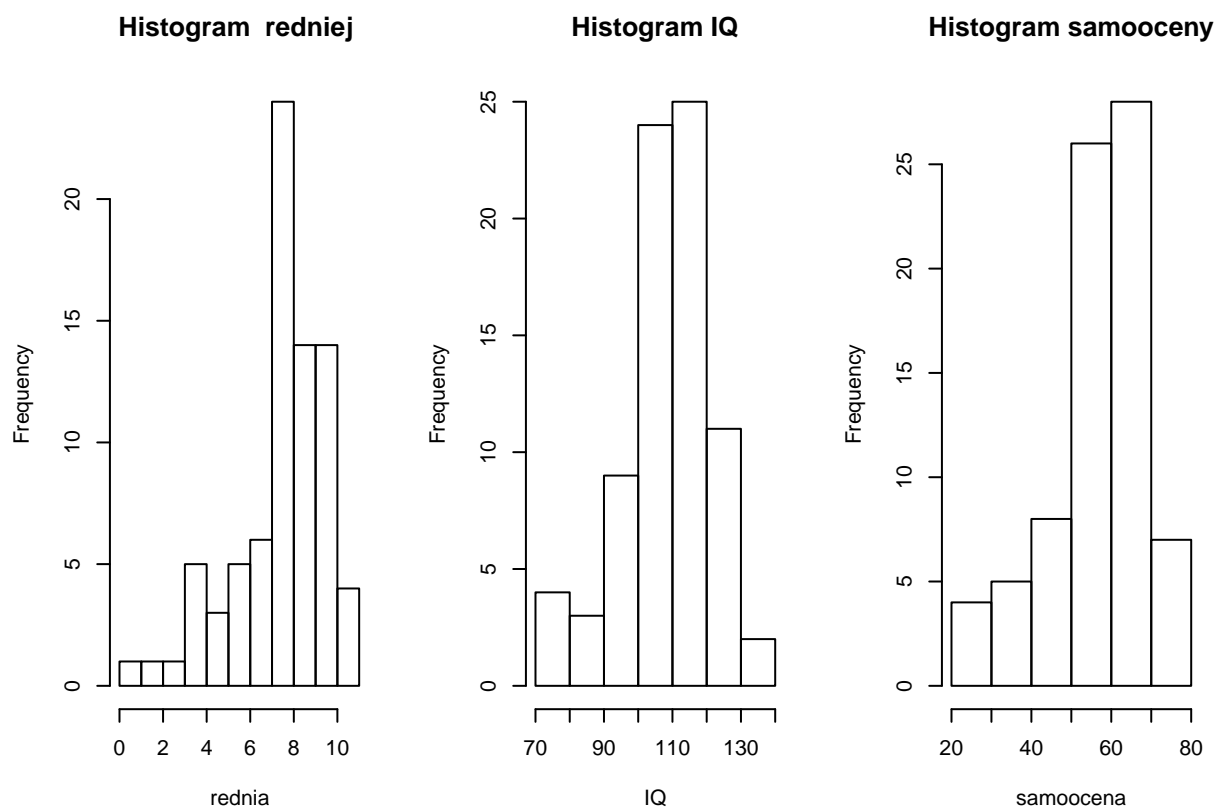
Lista1

Mikołaj Słupiński

12 marca 2018

Zadanie 1

a)



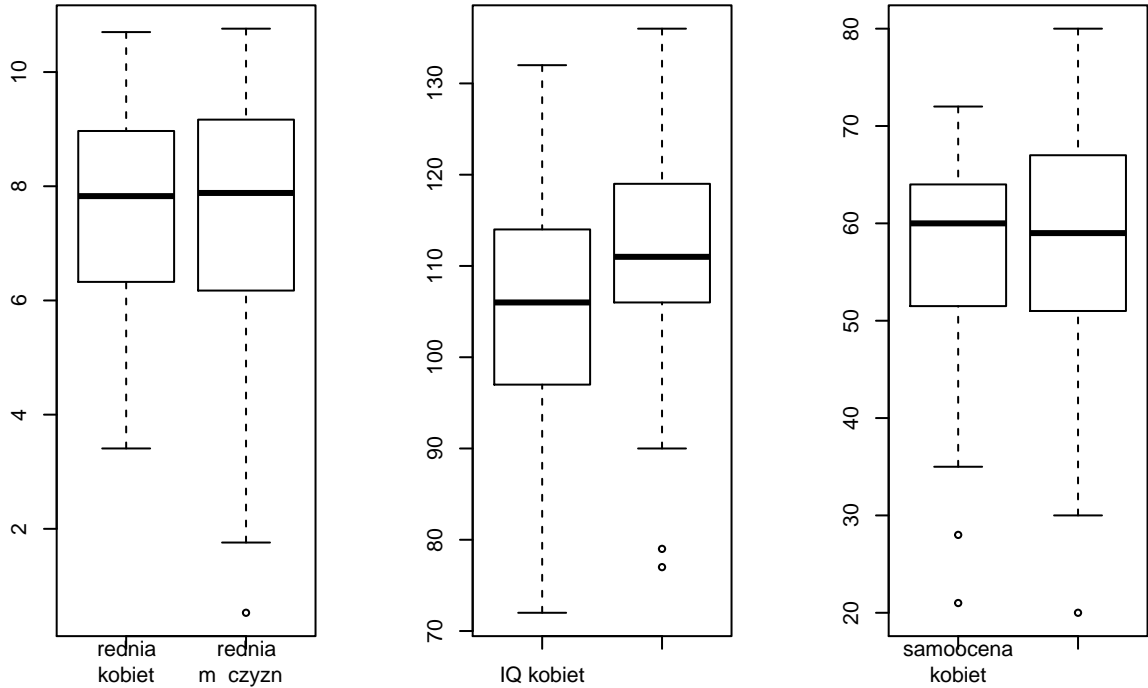
```
## [1] 7.446538 7.829000 4.408141  
## [1] 108.9231 110.0000 173.4745  
## [1] 56.96154 59.50000 154.06344
```

Otrzymaliśmy trzy wykresy o podobnym kształcie. Wszystkie są jednomodalne, ze skośnością na lewo. Najbardziej uwidacznia się to w histogramie średniej.

b)

```
male = subset(tab, tab[4] == "M")  
female = subset(tab, tab[4] == "F")  
par(mfrow = c(1,3))  
boxplot(female[[2]], male[[2]], names = c("średnia\nkobiet", "średnia\nmężczyzn"))
```

```
boxplot(female[[3]], male[[3]], names = c("IQ kobiet", "IQ mężczyzn"))
boxplot(female[[5]], male[[5]], names = c("samoocena\nkobiet", "samoocena\nmężczyzn"))
```



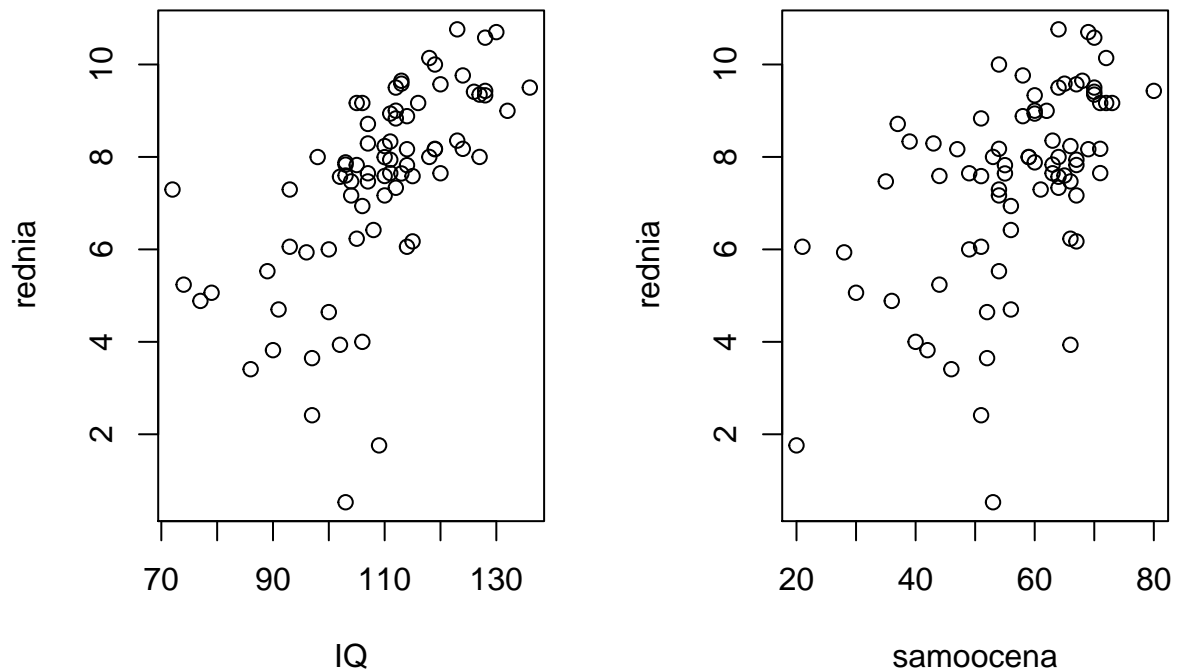
Widzimy, że mediana średniej ocen jest zbliżona zarówno w przypadku kobiet jak i mężczyzn, jednakże wśród mężczyzn występuje zdecydowanie większa wariancja i zdecydowanie większa ilość słabszych ocen.

Podobnie sprawa ma się z samooceną wśród uczniów. Kobiety mają podobną medianę, ale mniejszą wariancję. Natomiast wśród kobiet występuje więcej zjawisk odstających.

Mediana IQ jest zdecydowanie niższa niż wśród mężczyzn, występuje również większa wariancja. Co ciekawe, najwyższy wynik testu IQ wśród kobiet jest niższy niż najwyższy u mężczyzn. W szczególności najniższe wyniki są zdecydowanie niższe niż u mężczyzn.

c)

```
par(mfrow = c(1,2))
plot(tab[[3]], tab[[2]], xlab = "IQ", ylab = "średnia")
plot(tab[[5]], tab[[2]], xlab = "samoocena", ylab = "średnia")
```



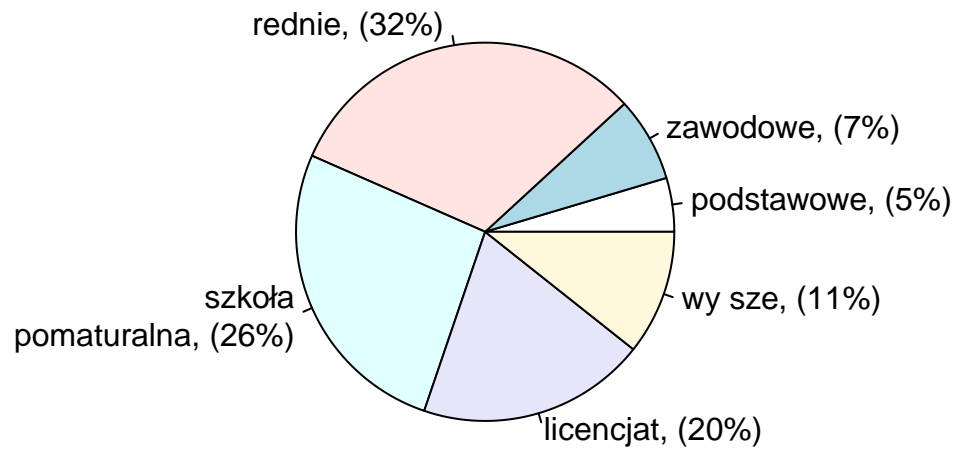
Na wykresach można zauważyć, że osoby z większym IQ mają tendencję do wyższej średniej ocen. Tak samo można zauważyć pewien związek wysokiej samooceny z wyższą średnią.

Zadanie 2

a)

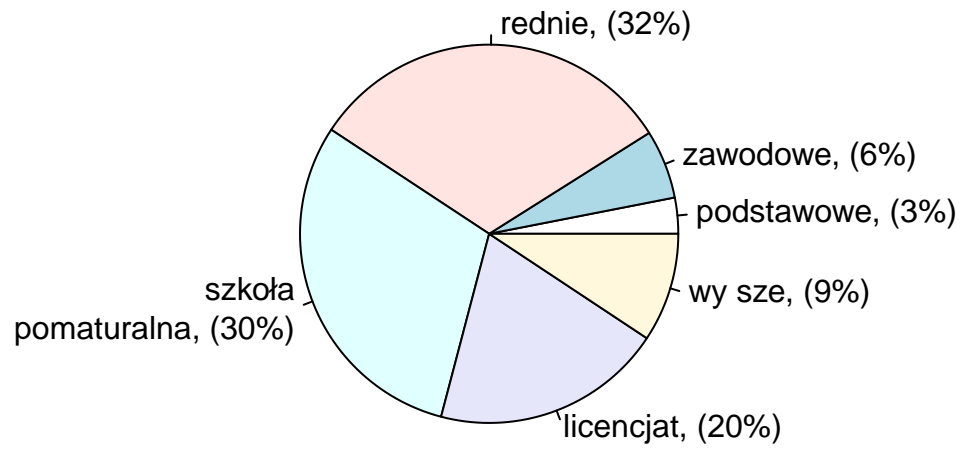
```
tab = read.table("http://www.math.uni.wroc.pl/~mbogdan/Podstawy/Dane/individuals.dat")
tab[[3]] = as.factor(tab[[3]])
levels(tab[[3]]) = c("podstawowe", "zawodowe", "średnie", "szkoła\npomaturalna", "licencjat", "wyższe")
tab[[6]] = as.factor(tab[[6]])
levels(tab[[6]]) = c("sektor prywatny", "sektor rządowy", "samo-zatrudnienie")
males = subset(tab, tab[4] == 1)
females = subset(tab, tab[4] == 2)
par(mfrow = c(1,1))
pie(table(males[[3]]), labels = paste(levels(males[[3]]), " (", round(table(males[[3]])/dim(males[3]))[1,
```

Wykształcenie u m czyzn



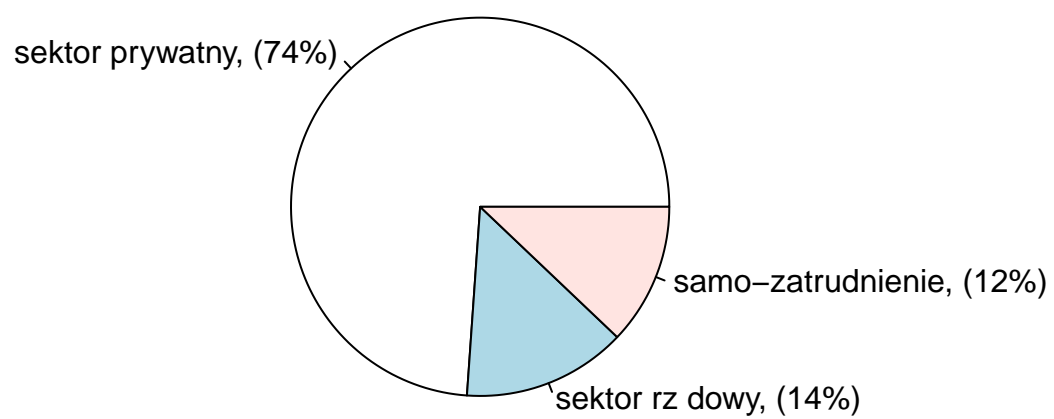
```
pie(table(females[[3]]), labels = paste(levels(females[[3]]), "%", round(table(females[[3]])/dim(females[[3]]), 1)))
```

Wykształcenie u kobiet



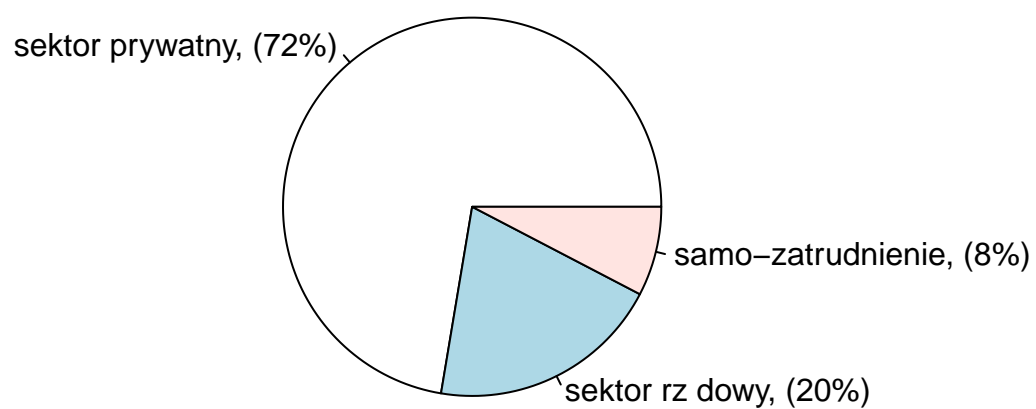
```
pie(table(males[[6]]), labels = paste(levels(males[[6]]), ", (", round(table(males[[6]])/dim(males[6]))[
```

Zatrudnienie u m czyzn



```
pie(table(females[[6]]), labels = paste(levels(females[[6]]), " (", round(table(females[[6]])/dim(fema
```

Zatrudnienie u kobiet

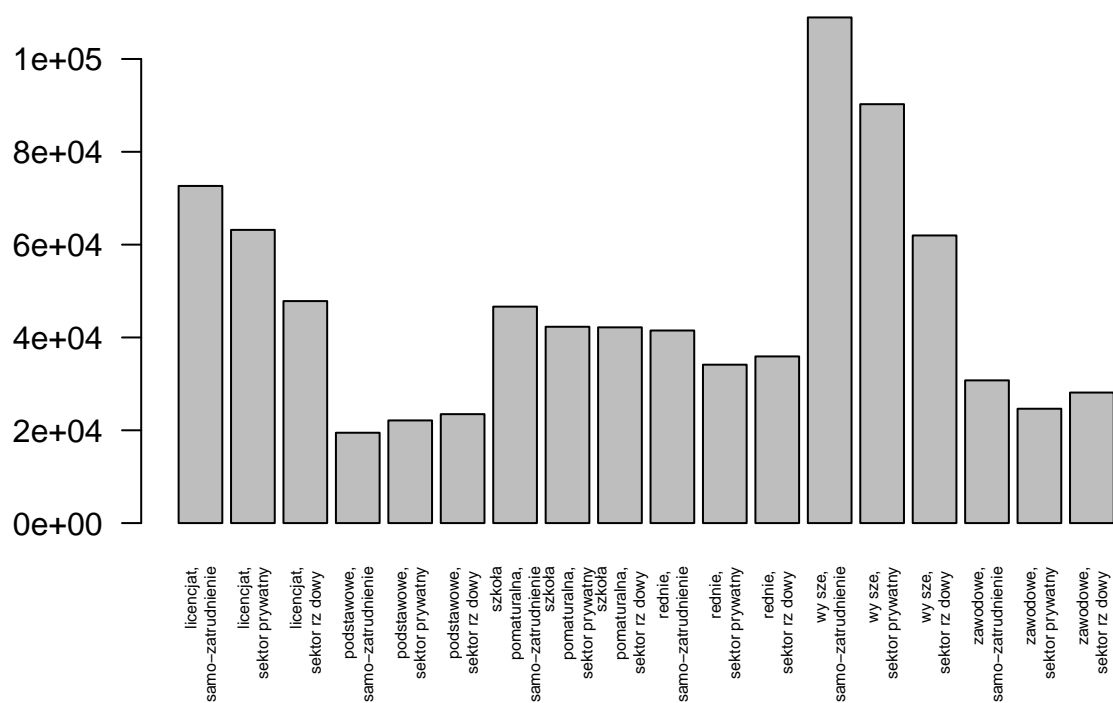


Rozkłady wykształcenia u kobiet i mężczyzn są zbliżone. Jednak wśród mężczyzn jest więcej osób z wykształceniem wyższym, natomiast wśród kobiet z wykształceniem pomaturalnym. Można zauważyć, że większość stanowisk rządowych zajmują mężczyźni.

b)

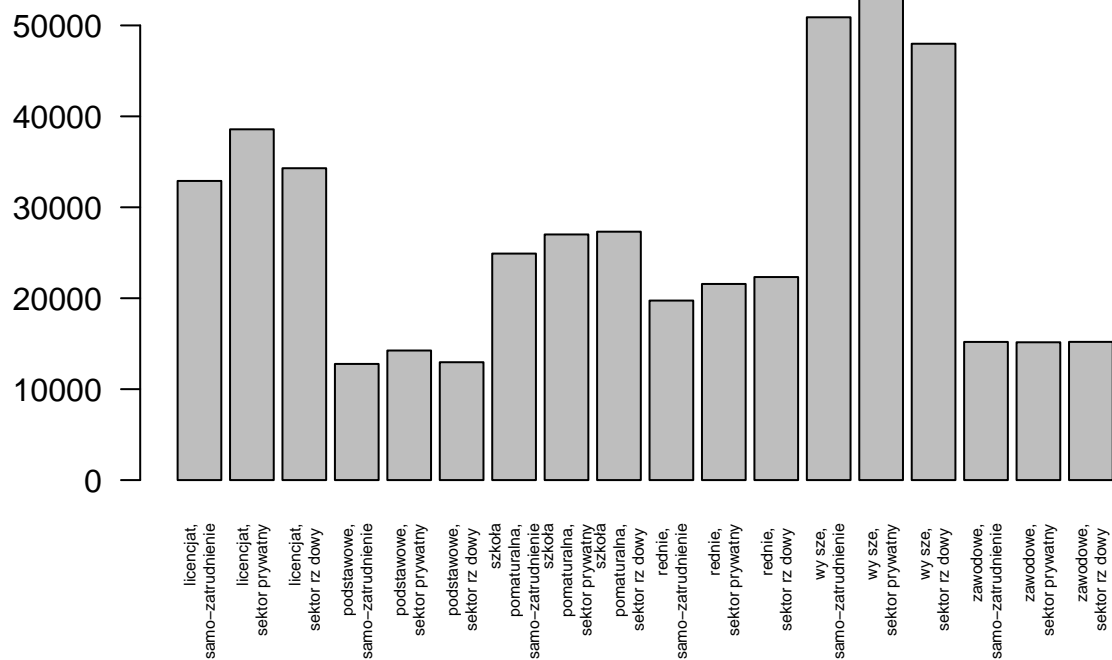
```
par(mfrow = c(1,1))  
barplot(by(males[[5]], paste(males[[3]], "\n", males[[6]], sep = ""), mean), las = 2, cex.names = 0.5, m
```

Zarobki w ród m czyzn



```
barplot(by(females[[5]], paste(females[[3]], "\n", females[[6]], sep = ""), mean), las = 2, cex.names =
```

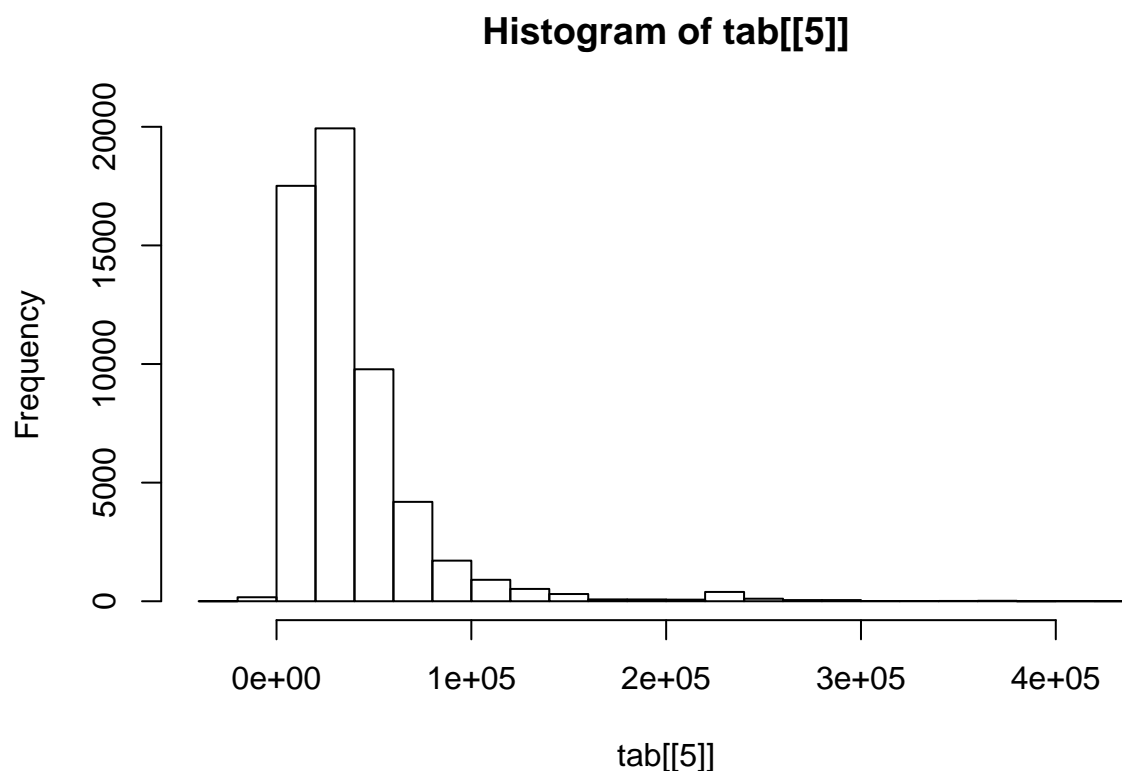

Zarobki w ród kobiet



Można zauważyć, że najwyższe zarobki są ponad pięciokrotnie wyższe od najniższych. Ponadto mężczyźni zarabiają nawet dwukrotnie więcej niż kobiety. Najmniej zarabiają osoby samo-zatrudnione z wykształceniem podstawowym oraz osoby z wykształceniem podstawowym w sektorze rządowym. Wśród mężczyzn najczęściej zarabiają osoby samo-zatrudnione z wykształceniem wyższym, natomiast wśród kobiet osoby z wykształceniem wyższym w sektorze prywatnym.

c)

```
hist(tab[[5]])
```



```
c(mean(tab[[5]]), median(tab[[5]]), mean(tab[[5]])/median(tab[[5]]), xlab = "Dochody", main = "Histogram dochodów")
```

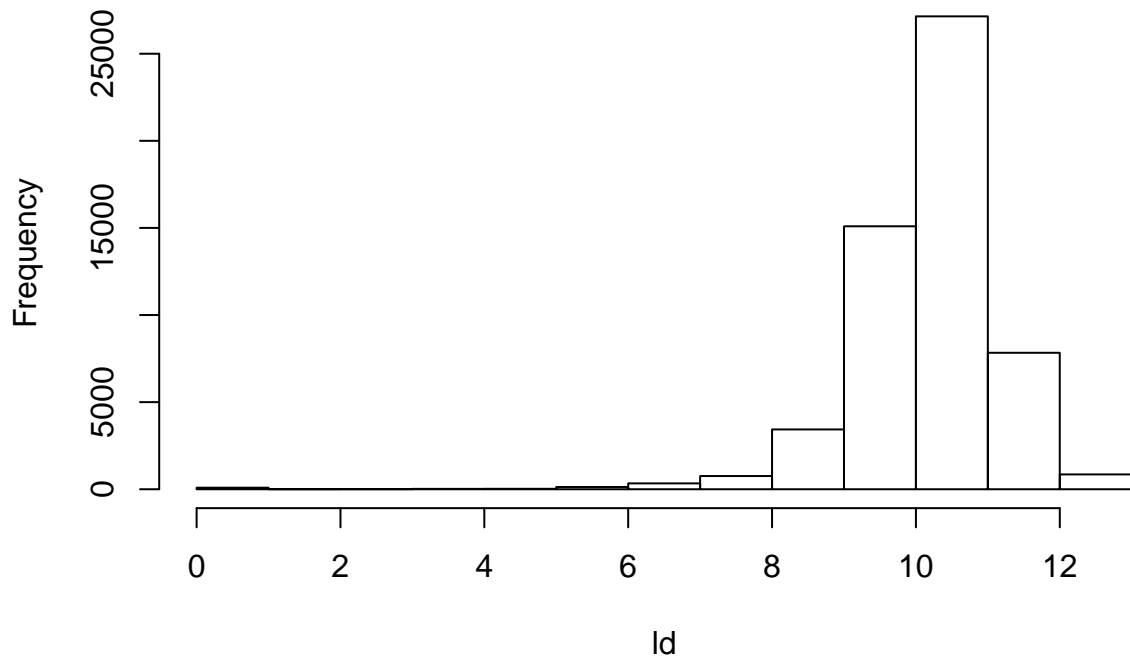
```
##
##      "37864.6094563409"      "29717"      "1.27417335048426"
##              xlab              main
##      "Dochody" "Histogram dochodów"
```

Wykres jest skośny w prawo. Oznacza to, że średnia dochodów jest zdecydowanie wyższa od mediany, co pokazuje nam również ich iloraz. Oznacza to, że jest niewiele osób bardzo bogatych, ale ich dochody są na tyle wysokie, żeby mieć znaczący wpływ na średnią.

d)

```
ld = log(tab[5][tab[5] > 0])
hist(ld)
```

Histogram of Id



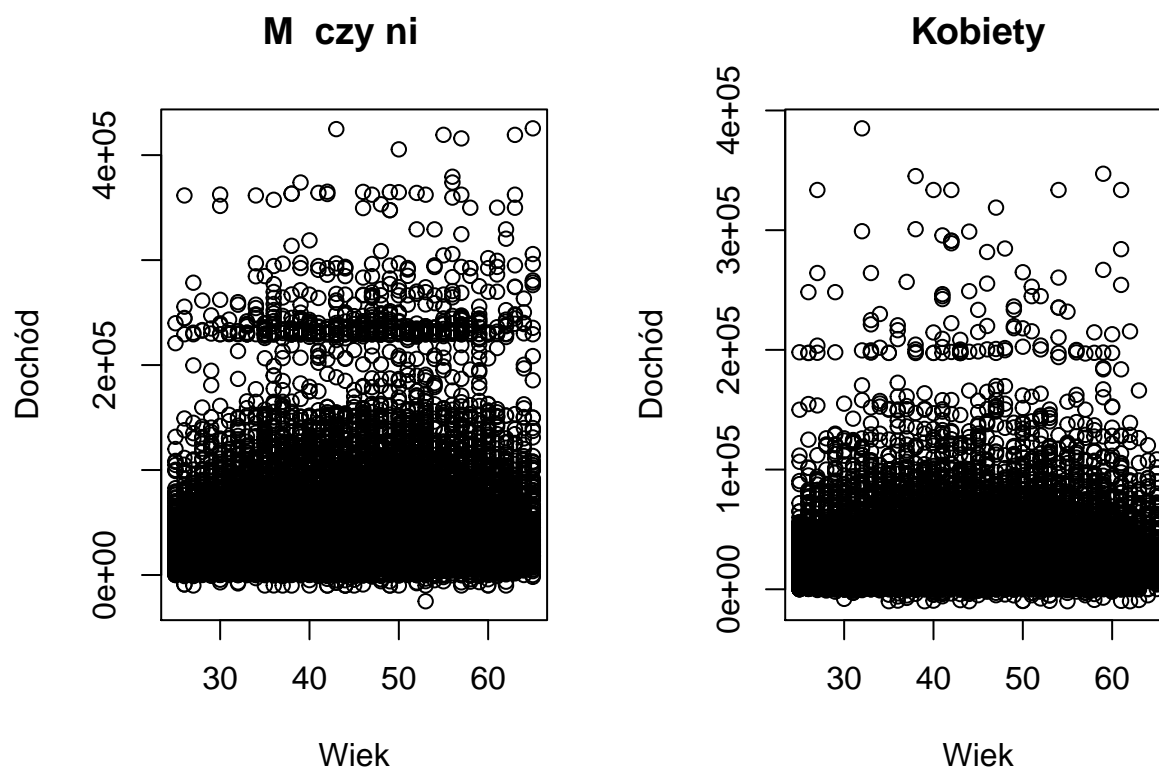
```
c(mean(Id), median(Id), mean(Id)/median(Id))
```

```
## [1] 10.1777570 10.3060819 0.9875486
```

Uzyskaliśmy wykres jednomodalny ze skośnością w lewo. Średnia zbliżyła nam się do mediany. Można to wytłumaczyć tym, że logarytm jest funkcją wolnorosnącą.

e)

```
par(mfrow = c(1,2))
plot(ages[2], incomes[5], xlab = "Wiek", ylab = "Dochód", main = "Mężczyźni")
plot(ages[2], incomes[5], xlab = "Wiek", ylab = "Dochód", main = "Kobiety")
```



Jest zdecydowanie większe zagęszczenie górnych zarabków dla mężczyzn. Widać również, że wśród górnych zarbków może istnieć zależność między wysokością zarobku, a wiekiem mężczyzny, gdyż najwyższe zarobki osiągają jedynie mężczyźni po 40 roku życia. Również wśród niższych dochodów widać tendencję do zarabiania więcej przez panów po 40 roku życia. Widzimy, że wśród kobiet jest o wiele niższa tendencja niż u mężczyzn do zarabiania w okolicach mediany.