

Raport 1

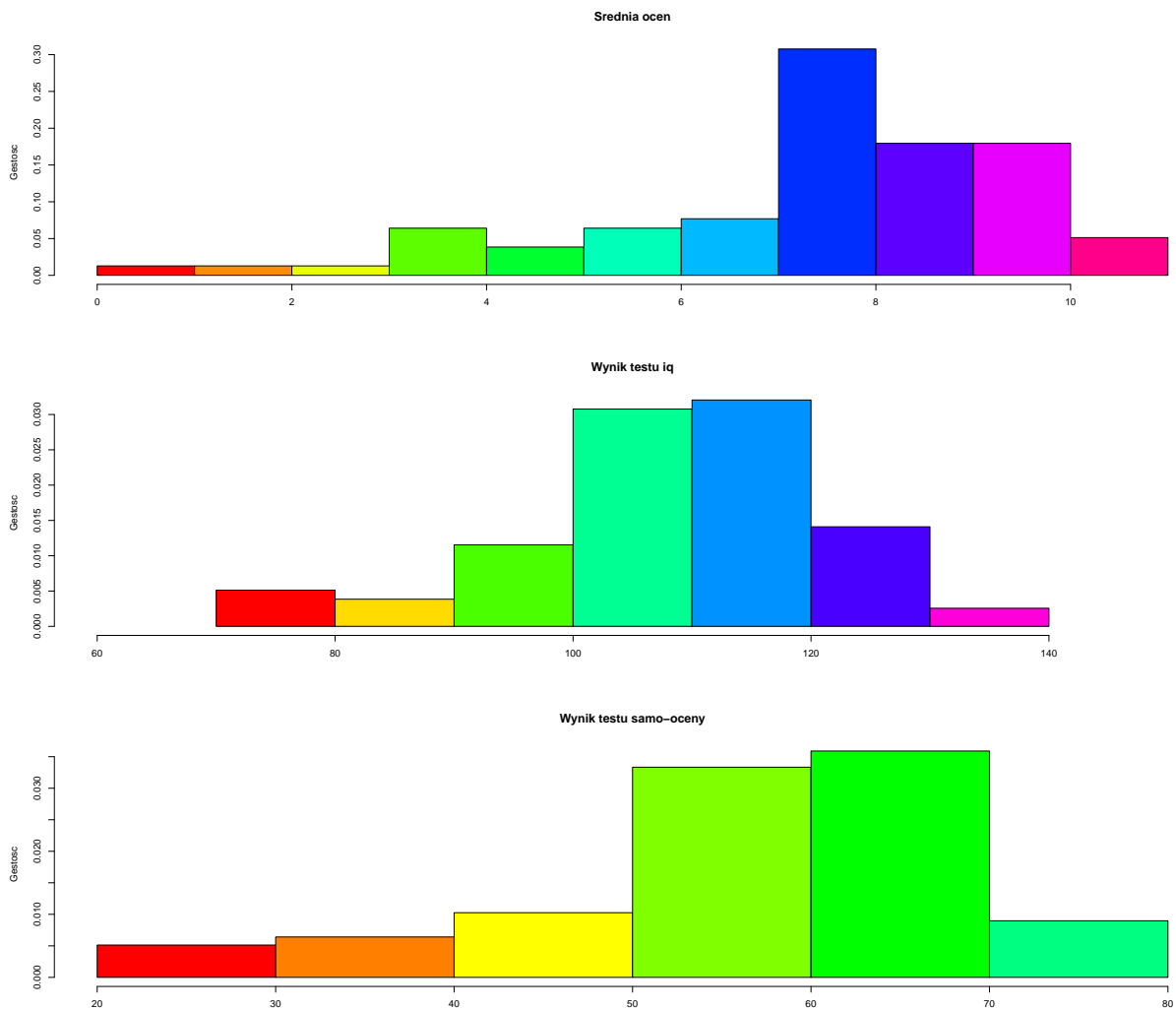
Aleksander Milach

16 marca 2018

Zadanie 1

Podpunkt a

```
par(mfrow=c(3,1))
hist(t$V2, main="Średnia ocen", col=rainbow(11), xlab="", ylab="Gęstość", freq=FALSE)
hist(t$V3, main="Wynik testu iq",xlim=c(60,150), col=rainbow(7), xlab="", ylab="Gęstość",
     freq=FALSE)
hist(t$V5, main="Wynik testu samo-oceny", breaks=8, col=rainbow(12), xlab="",
     ylab="Gęstość", freq=FALSE)
```



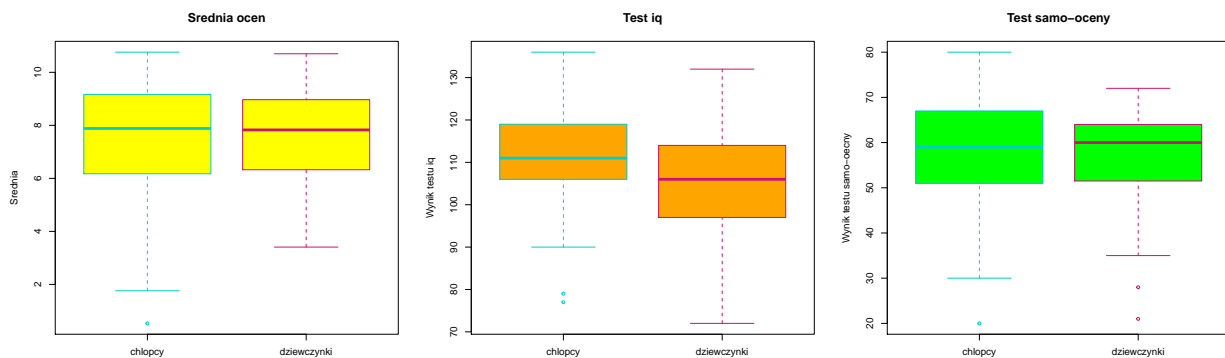
Histogram średniej ocen uczniów jest jednomodowy (skoncentrowany wokół 8), skośny w lewo, najmniejszy rozrzut, przy osi pionowej. Wykres wyników testu iq również jest jednomodowy, ale też symetryczny, ma duży rozrzut, daleko od osi pionowej. Wykres samo-oceny dzieci jednomodowy, skośny w lewo, rozrzut podobny jak przy teście iq, bliżej osi pionowej niż histogram wyników testu iq.

```
a=c(max(t$V2)-min(t$V2),max(t$V3)-min(t$V3),max(t$V5)-min(t$V5))
```

Rozrzut średniej, testu iq, testu samo-oceny wynoszą odpowiednio: 10.23, 64, 60.

Podpunkt b

```
par(mfrow=c(1,3))
boxplot(chl$V2,dz$V2,main="Średnia ocen",border=c("cyan3","deeppink3"),col=c("yellow"),
        names=c("chłopcy","dziewczynki"),ylab="Średnia")
boxplot(chl$V3,dz$V3,main="Test iq",border=c("cyan3","deeppink3"),
        col=c("orange"),names=c("chłopcy","dziewczynki"),ylab="Wynik testu iq")
boxplot(chl$V5,dz$V5,main="Test samo-oceny",border=c("cyan3","deeppink3"),
        col=c("green"),names=c("chłopcy","dziewczynki"),ylab="Wynik testu samo-oceny")
```

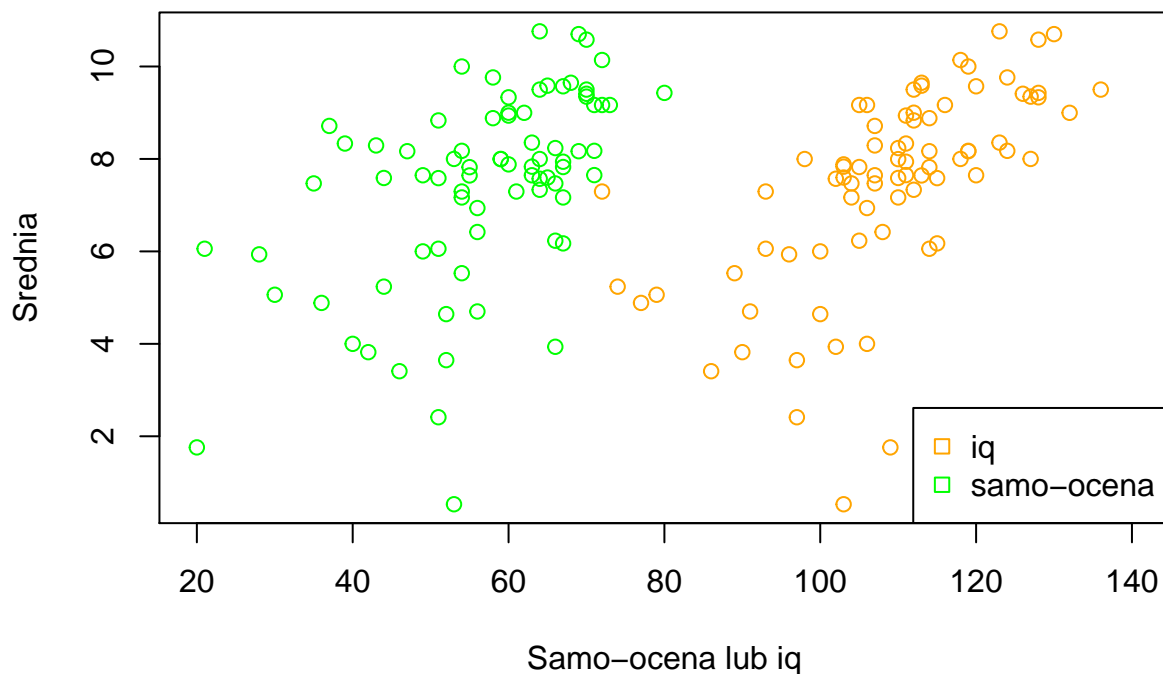


Średnia ocen chłopców ma duży rozrzut i “ogon” wskazujący na bardziej skrajne wyniki obserwacji poniżej mediany, u dziewczynek mniejszy rozrzut i wykres bardziej symetryczny, mediana niemal równa. W teście iq chłopcy wypadli zauważalnie lepiej od dziewczynek (mediana około 7 większa), duży rozrzut u obu płci, nieznacznie większy u dziewczynek. Samoocena chłopców ma znacznie większy rozrzut niż dziewczynek, zarówno u dziewczynek jak i chłopców duża różnica między pierwszym a drugim kwartylem. Średnia i mediana podobne u obu płci.

Podpunkt c

```
plot(t$V3,t$V2,xlim=c(20,140),col="orange",main="Rozrzut wyniku testu iq i testu
        samo-oceny w zależności od średniej ocen",xlab="Samo-ocena lub iq",ylab="Średnia")
points(t$V5,t$V2,col='green')
legend("bottomright",c("iq","samo-ocena"),col=c("orange","green"),pch=0)
```

Rozrzut wyniku testu iq i testu samo-oceny w zaleznosci od sredniej ocen



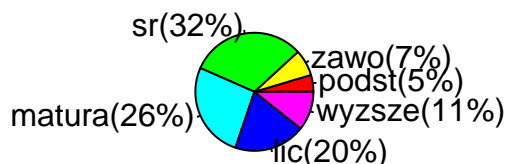
Tylko przy skrajnie niskich wartościach średniej, iq lub samooceny zdrażają się obserwacje daleko od trendu, samoocena ma większy rozrzut niż iq, niemniej w obu zależnościach wyraźny trend - proporcjonalność odpowiednich wartości.

Zadanie 2

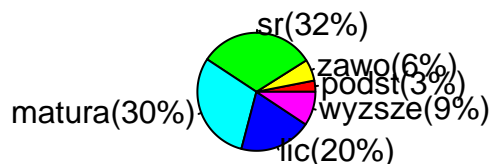
Podpunkt a

```
par(mfrow=c(1,2))
pie(table(panowie[[3]]),labels=paste0(levels(panowie[[3]]),
                                     "(",round(table(panowie[[3]])/dim(panowie)[1]*100,"%)" ),
     main="Wykształcenie mężczyzn",col=rainbow(6),radius=.4)
pie(table(panie[[3]]),labels=paste0(levels(panie[[3]]),
                                     "(",round(table(panie[[3]])/dim(panie)[1]*100,"%)" ),
     main="Wykształcenie kobiet",col=rainbow(6),radius=.4)
```

Wykształcenie mężczyzn

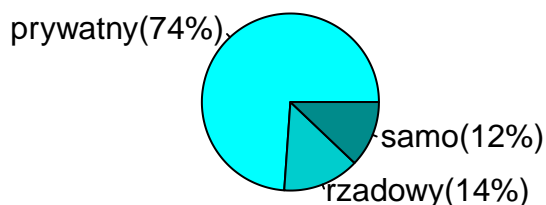


Wykształcenie kobiet

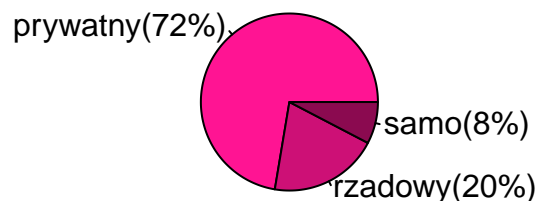


```
pie(table(panowie[[6]]),labels=paste0(levels(panowie[[6]]),
                                     "(",round(table(panowie[[6]])/dim(panowie)[1]*100),"%)" ),
     main="Zatrudnienie mezczyzn",col=c("cyan1","cyan3","cyan4"),radius=.6)
pie(table(panie[[6]]),labels=paste0(levels(panie[[6]]),
                                     "(",round(table(panie[[6]])/dim(panie)[1]*100),"%)" ),
     main="Zatrudnienie kobiet",col=c("deeppink1","deeppink3","deeppink4"),radius=.6)
```

Zatrudnienie mężczyzn



Zatrudnienie kobiet

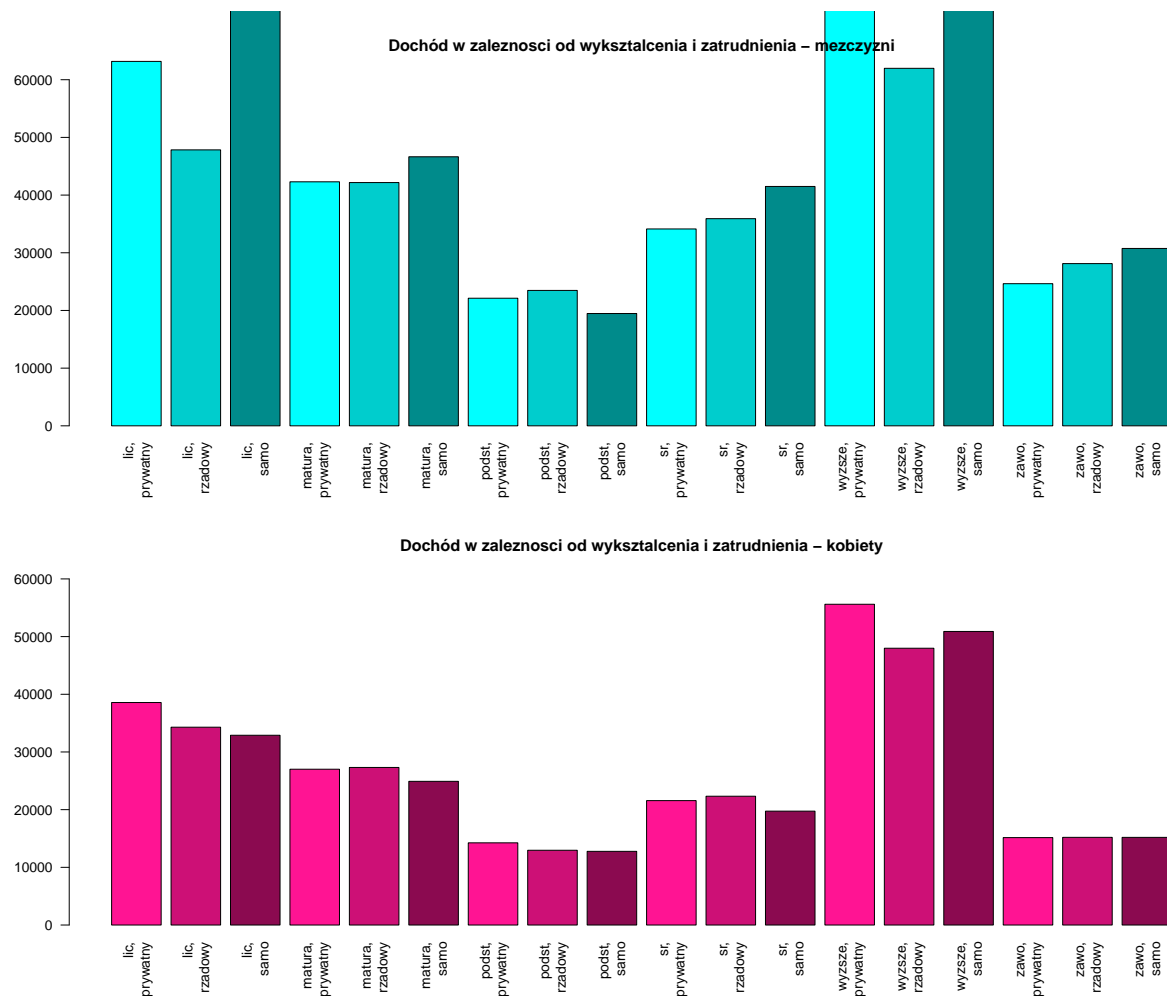


Prawie 2 razy więcej mężczyzn niż kobiet ma tylko wykształcenie podstawowe, więcej kobiet zakończyło edukację na egzaminie maturalnym. Większy procent osób z wyższym wykształceniem jest wśród mężczyzn. Podobny odsetek zatrudnionych w prywatnym u obu płci (znaczną większość). Różnica występuje w pozostałych sektorach zatrudnienia: 1,5 raza większy procent mężczyzn niż kobiet jest samozatrudnionych oraz

1,5 raza większy procent kobiet niż mężczyzn w sektorze rządowym.

Podpunkt b

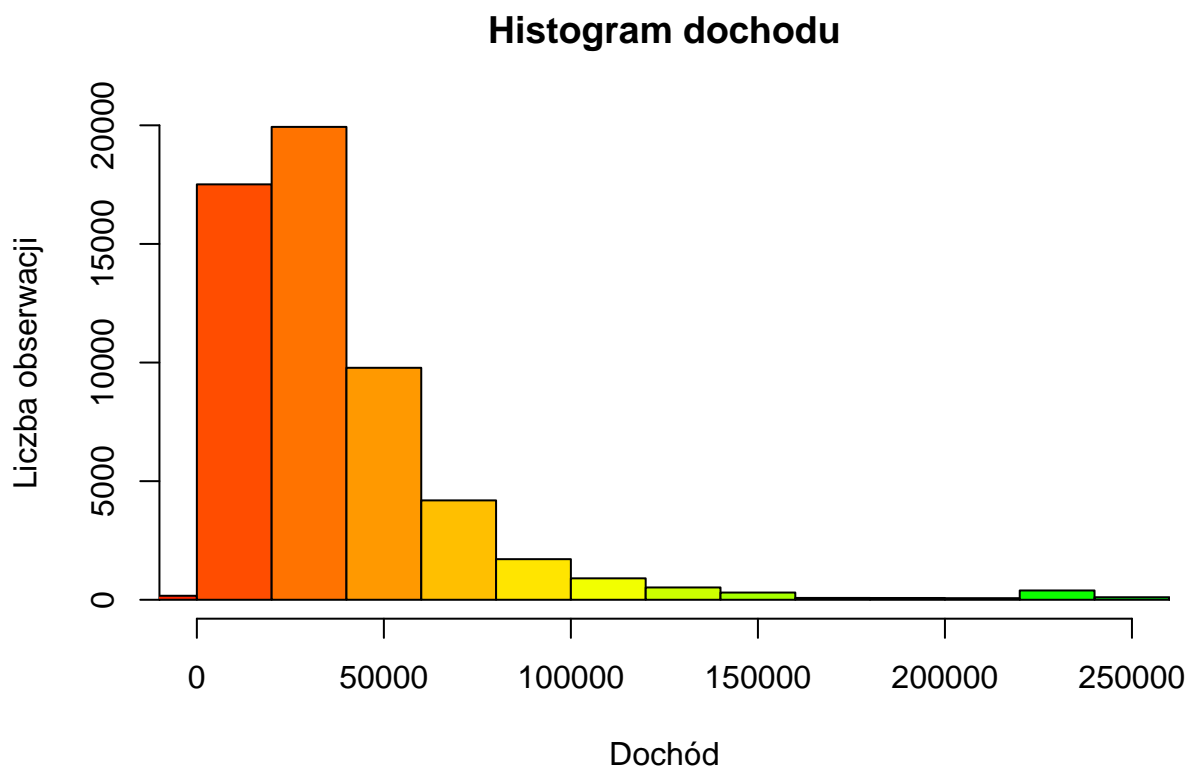
```
par(mfrow=c(2,1))
barplot(by(panowie[[5]],paste0(panowie[[3]],"\n",panowie[[6]]),mean),
        las=2,cex.names=1,ylim=c(0,60000),main=
        "Dochód w zależności od wykształcenia i zatrudnienia - mężczyźni",col=
        c("cyan1","cyan3","cyan4"))
barplot(by(panie[[5]],paste0(panie[[3]],"\n",panie[[6]]),mean),
        las=2,cex.names=1,ylim=c(0,60000),main=
        "Dochód w zależności od wykształcenia i zatrudnienia - kobiety",col=
        c("deeppink1","deeppink3","deeppink4"))
```



Po znormalizowaniu osi Y słupki mężczyzn są ogólnie wyższe. Mężczyźni po studiach (o ile nie licencjaci w sektorze rządowym oraz kobiety z wykształceniem wyższym stanowią elitę - zarabiają co najmniej 10000 więcej od reszty społeczeństwa. Zarobki osób z wykształceniem podstawowym lub zawodowym oraz maturalnym i średnim są porównywalne.

Podpunkt c

```
hist(f[[5]],main="Histogram dochodu",col=rainbow(40),xlab="Dochód",  
     ylab="Liczba obserwacji",xlim=c(0,250000))
```



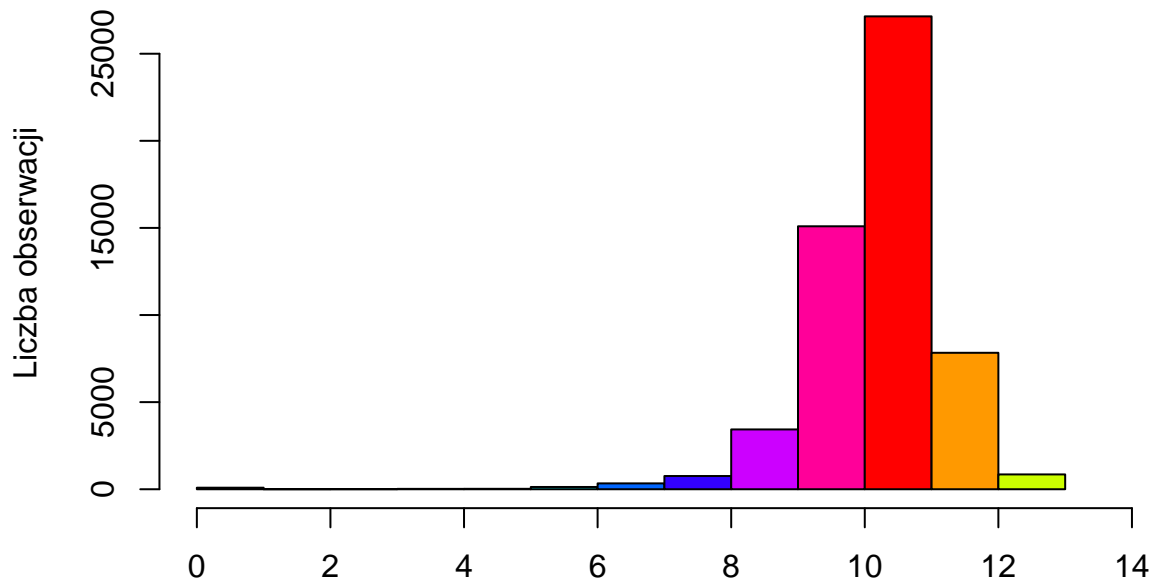
```
a=c(mean(f[[5]]),median(f[[5]]),mean(f[[5]])/median(f[[5]]),max(f[[5]])-min(f[[5]]))
```

Histogram dochodu skupiony blisko osi Y, skośny w prawo, ma bardzo duży rozrzut, średnia dużo wyższa od mediany. Średnia, mediana, współczynnik zmienności i rozrzut wynoszą kolejno 37864.61, 29170, 1.274173, 450508.

Podpunkt d

```
ld=log(f[[5]][f[[5]]>0])  
hist(ld,main="Logarytm dochodu",col=rainbow(10),xlab="",ylab="Liczba obserwacji",xlim=c(0,14))
```

Logarytm dochodu

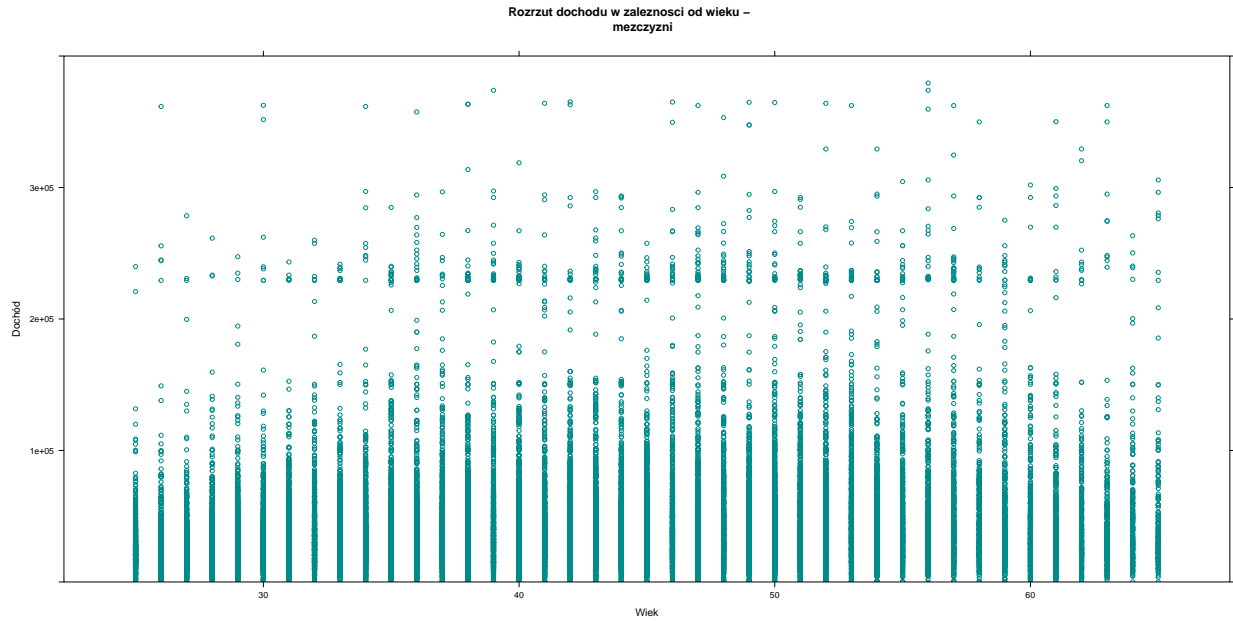


```
a=c(mean(ld),median(ld),mean(ld)/median(ld),max(ld)-min(ld))
```

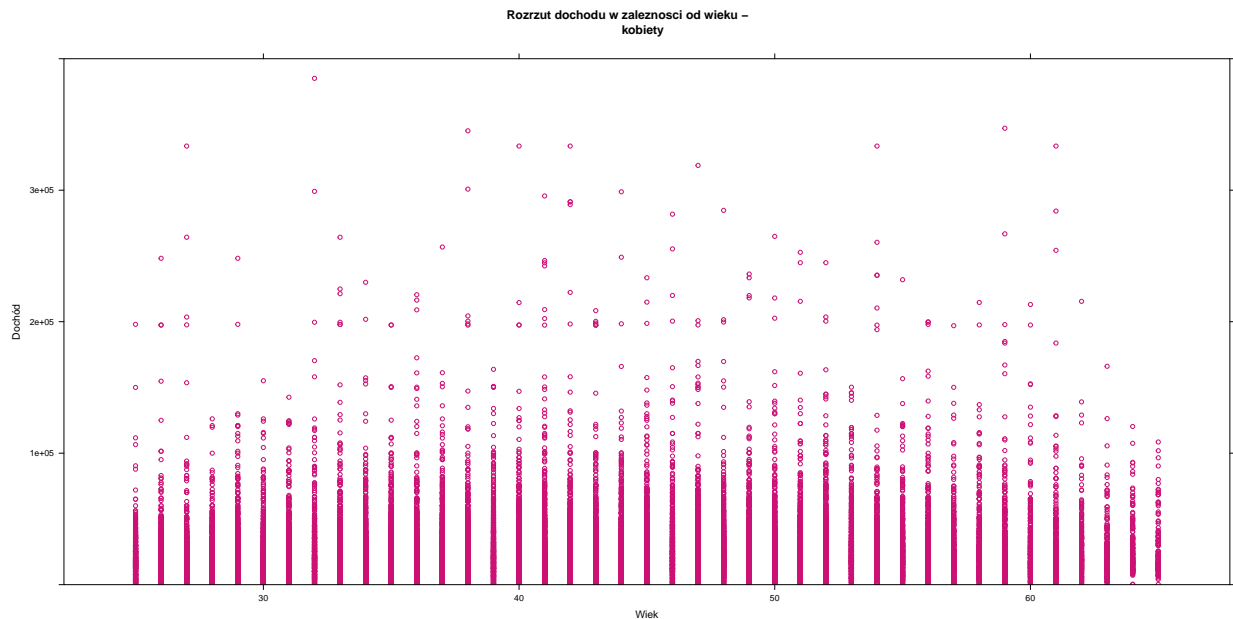
Histogram oddalony od osi y, jednomodowy symetryczny z środkiem w 10; znaczna większość ma logarytm dochodu między 9 a 12. Mediana bardzo nieznacznie wyższa od średniej, wciąż histogram ma bardzo duży rozrzut. Kolejno średnia, mediana, współczynnik zmienności i rozrzut: 10.177757, 10.3060819, 0.9875486, 12.9610437 .

Podpunkt e

```
library(lattice)
par(mfrow=c(1,2))
xyplot(panowie[[5]]~panowie[[2]],main="Rozrzut dochodu w zależności od wieku -
mężczyźni",xlab="Wiek",ylab="Dochód",ylim=c(0,400000),col="cyan4",auto.key=TRUE)
```



```
xyplot(panie[[5]]~panie[[2]],main="Rozrzut dochodu w zależności od wieku -  
kobiety",xlab="Wiek",ylab="Dochód",ylim=c(0,400000),col="deeppink3",auto.key=TRUE)
```



U mężczyzn do 33 roku życia dochód zauważalnie rośnie, pojawia się też elitarna grupa zarabiających około 23000, po 60 roku życia dochody maleją. Na wykresie dochodu kobiet zagęszczenie obserwacji jest niższe (u mężczyzn na około 10000, u kobiet około 8000), podobny rozrzut obserwacji, ale u kobiet wraz ze wzrostem dochodu ilość obserwacji maleje szybciej niż u mężczyzn. U obu grup można wyróżnić wyraźny trend wzrostu płacy do 30 roku życia i jej spadek po 60 roku życia.