Week 2 tasks

teodor dyakov

10/15/2020

```
library(MASS)

#namirame parvo vsichki stoynosti v kolonata Smoke
unique(survey$Smoke)
```

```
## [1] Never Regul Occas Heavy <NA>
## Levels: Heavy Never Occas Regul
```

```
#Zadacha 1
#veroyatnosta sluchayno izbran chovek da e redoven pushach
sum(survey$Smoke == 'Regul', na.rm = TRUE)/nrow(survey)
```

```
## [1] 0.07172996
```

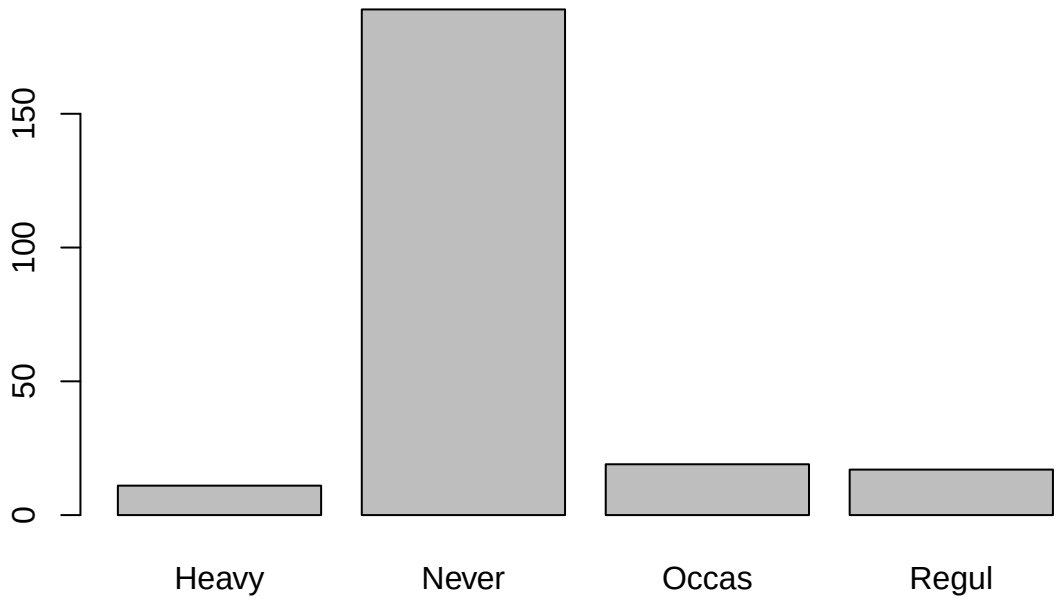
```
maleSmokers = sum(survey$Smoke %in% c("Regul", "Heavy") & survey$Sex == 'Male', na.rm = TRUE)

#veroyatnostta sluchayno ibzbran chovek da e mazh i redoven pushach
maleSmokers/nrow(survey)
```

```
## [1] 0.07594937
```

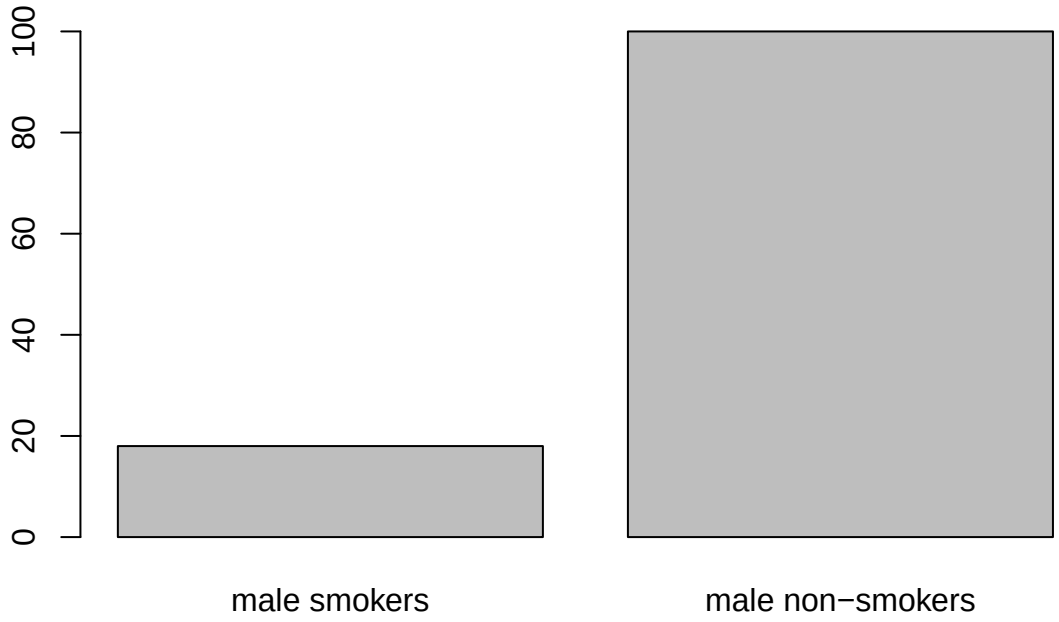
```
#broyat na vsichki mazhe
maleNumber = sum(survey$Sex == 'Male', na.rm = TRUE)

#veroyatnostta sluchayno izbran mazh da e redoven pushach
prob = maleSmokers/maleNumber

#broyat na vsichki redovni pushachi
regulOrHeavyCount = sum(survey$Smoke %in% c('Regul', 'Heavy'), na.rm = TRUE)

#veroyatnosta sluchayno izbran redoven pushach da e mazh
maleSmokers/regulOrHeavyCount
```

```
## [1] 0.6428571
```

```
#Zad 2
#barplot na broya na tipovete pushachi
barplot(table(survey$Smoke))
```




```
#pushene v zavisimost ot pola
barplot(c(maleSmokers, maleNumber - maleSmokers), names.arg = c('male smokers', 'male non-smokers'))
```




```
#Zad 3
#nekvi statistiki
summary(survey$Height)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   150.0   165.0   171.0   172.4   180.0   200.0      28
```

```
heightMean = mean(survey$Height, na.rm = TRUE)
heightSd = sd(survey$Height, na.rm = TRUE)


maleHeights = survey[survey$Sex == 'Male', 'Height']
summary(maleHeights)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   154.9   172.8   180.0   178.8   185.0   200.0      13
```

```
sd(maleHeights, na.rm = TRUE)
```

```
## [1] 8.380252
```

```
femaleHeights = survey[survey$Sex == 'Female', 'Height']
summary(femaleHeights)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   150.0   162.6   166.8   165.7   170.0   180.3      17
```

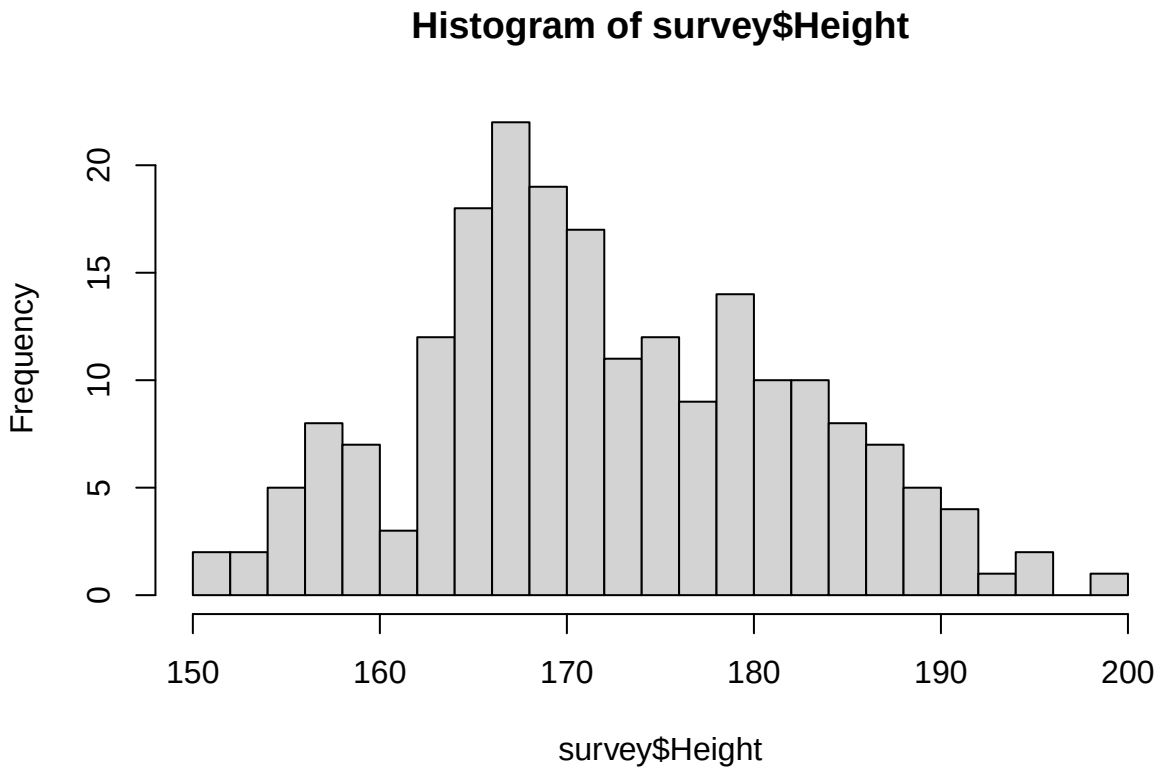
```r
sd(femaleHeights, na.rm = TRUE)
```

```
## [1] 6.151777
```

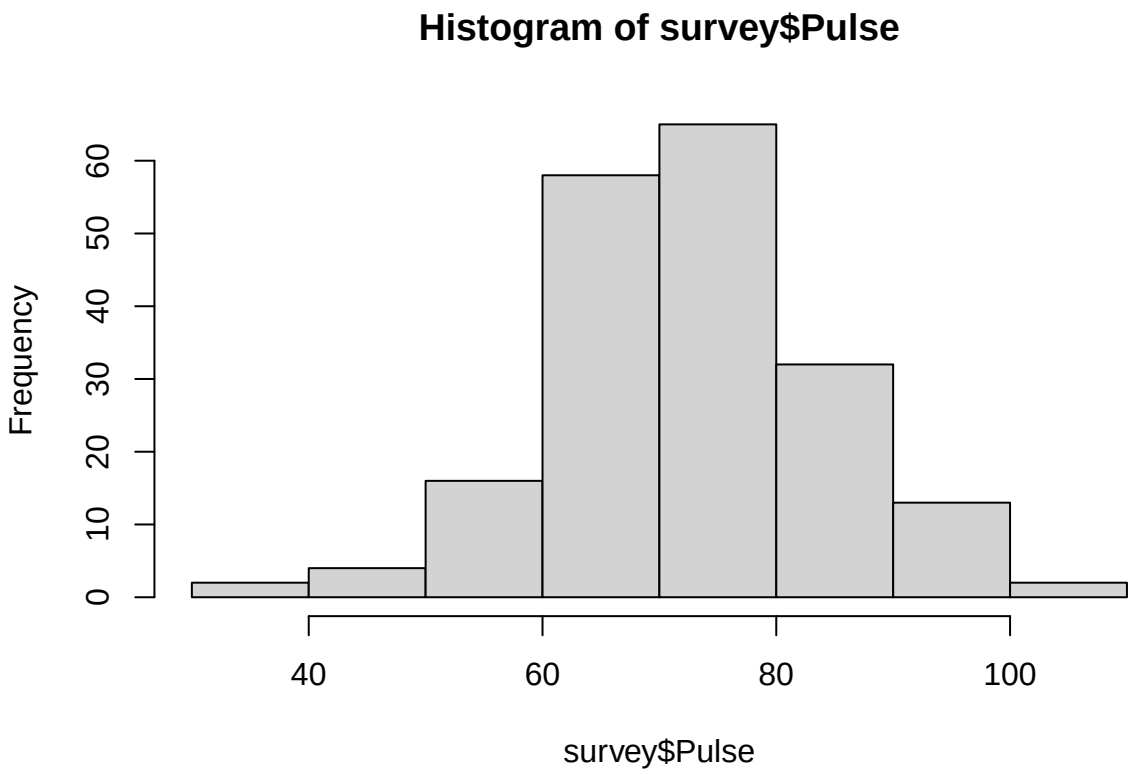
```r
#kakva chast ot studentite se razlichavat ot sr. visochina s <= 1 standartno otklonenie?
sum(abs(survey$Height - heightMean) <= heightSd, na.rm = TRUE)/sum(!is.na(survey$Height))
```

```
## [1] 0.6842105
```

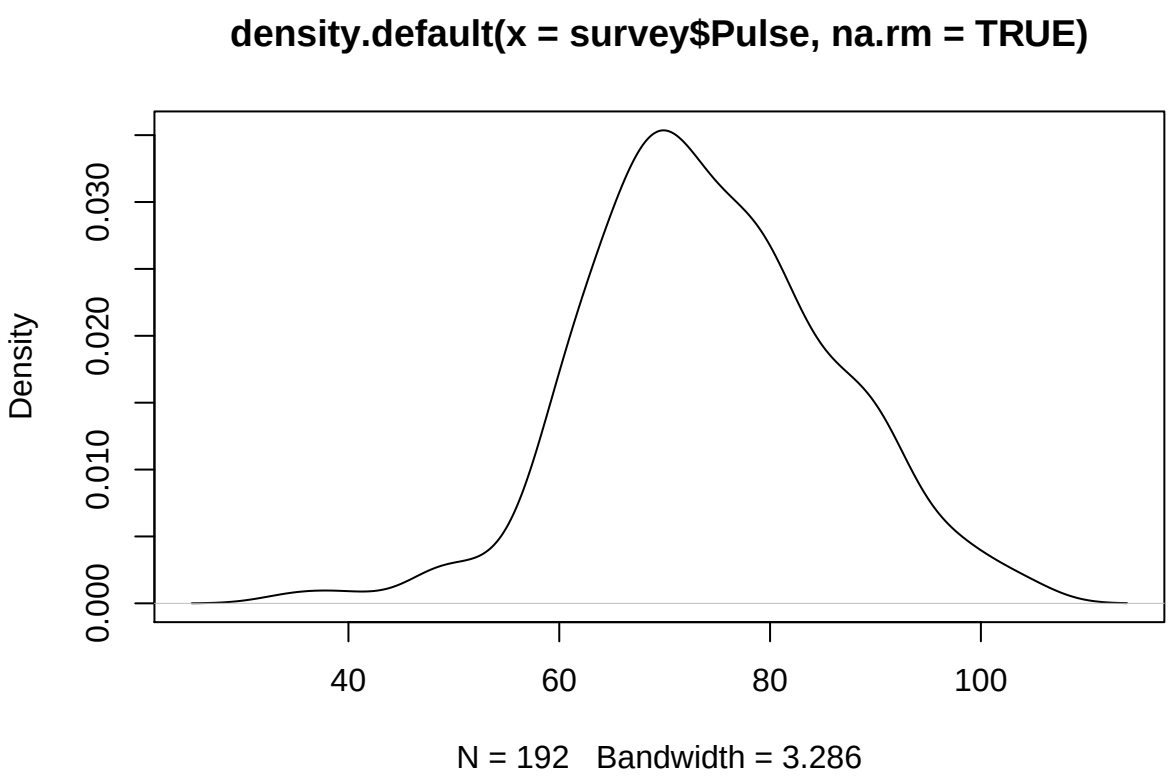
```r
#Zad 4
hist(survey$Height, breaks = 20)
```




```r
#Zad 5
#histograma na pulsa na studentite
hist(survey$Pulse)
```




```
#grafika platnostta na razpredelenieto na pulsa
plot(density(survey$Pulse, na.rm = TRUE))
```


density.default(x = survey$Pulse, na.rm = TRUE)
Density
0.000
0.010
0.020
0.030
40
60
80
100
N = 192 Bandwidth = 3.286