

## #ZADANIE 1

```
?geom_boxplot()
```

```
cholesterol_age$AgeGroup
```

```
ploot <- ggplot( data = cholesterol_age , mapping = aes(y = Before,  
                                                       x = AgeGroup))
```

```
ploot <- ploot + geom_boxplot() + theme_classic()
```

```
ploot <- ploot + ylab("Cholesterol before the experiment")
```

```
plot(ploot)
```

```
# granice pudełka: lewy bok: pierwszy kwartył, prawy bok: trzeci kwartył
```

```
# wąsy: zmienność poza górnym i dolnym kwartylem
```

```
# pozioma linia: mediana
```

## #ZADANIE 2

```
diet      <- read.table(file = "Diet_R.csv",  
                        sep = ",",  
                        header = T,  
                        stringsAsFactors = T)
```

```
diet$gender<- as.character(diet$gender)
```

```
diet$Diet<- as.character(diet$Diet)
```

```
na.omit(diet$Diet)
```

```
wkrs <- ggplot(na.omit(diet), aes(x=Diet, y=weight6weeks, fill=gender))
```

```
wkrs <- wkrs + geom_boxplot()+geom_jitter( size=1, width=0.3) + theme_classic()
```

```
wkrs <- wkrs + xlab("Diet type")
wkrs <- wkrs + ylab("Weight after 6 weeks")
plot(wkrs)
```

#ZADANIE 3

```
dane <- data.frame(kat = as.factor(c(4,2,2,6,2,2,4,4,6)),
                  x = c(3,5,4,5,7,8,6,4,1))
sapply (dane, class)
dane$kat <- as.numeric(as.character(dane$kat))
sapply (dane, class)
dane
```

#ZADANIE 4

```
#for (INDEKS in LISTA_WARTOSCI) {
# INSTRUKCJA1
# INSTRUKCJA2
#}
cholesterol_age$Before
msr <- mean(cholesterol_age$Before)
wektorBoot <- vector()
for (i in 1:1000) {
  s <- sample((cholesterol_age$Before),18, replace = TRUE)
  wektorBoot[i] <- mean(s)
}
wektorBoot
?abline
msr
kres <- hist( wektorBoot, col = 'gray', breaks = 35 )
abline( v = msr, col="red")
```

## #ZADANIE 4.2

```
median50 <- vector()
median1000 <- vector()

for (i in 1:50) {
  s <- sample((cholesterol_age$Before),18, replace = TRUE)
  median50[i] <- median(s)
}
for (i in 1:1000) {
  s <- sample((cholesterol_age$Before),18, replace = TRUE)
  median1000[i] <- median(s)
}
mesr <- median(cholesterol_age$Before)
median50
median1000
df <- data.frame(median50, median1000)
class(df)

wykres <- ggplot() + theme_classic() + xlab("median")
wykres <- wykres + geom_density(data = df, aes(median50,colour="median50"),bw = 0.15)
wykres <- wykres + geom_density(data = df, aes(median1000,colour="median1000"),bw = 0.15)
wykres <- wykres + geom_vline(xintercept=mesr, size=1.5, color="red")
wykres <- wykres + xlim(2,10)
wykres

# wierzcholek krzywej - wartosc pojawiajaca sie najczesciej jest bardziej zblizona
```

# prawdziwej mediany dla 1000 estymatorów

# krzywa dla 1000 estymatorów przyjmuje też szerszy zakres możliwych wartości