

Kamil Jankowski, Artur Basiński

1. Liczba kobiet oraz mężczyzn w badanych zestawie danych:

- a. Kobiet – 256
- b. Mężczyzn – 256

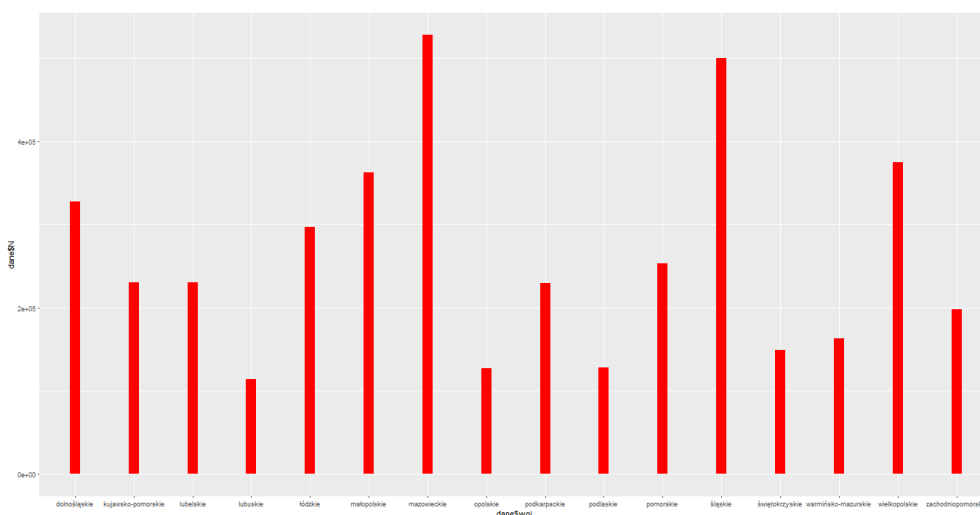
```
0 1
256 256
> |
```

2. Liczba danych dla każdego z województw wynosi 32, czyli dla każdego roku po 8 obserwacji:

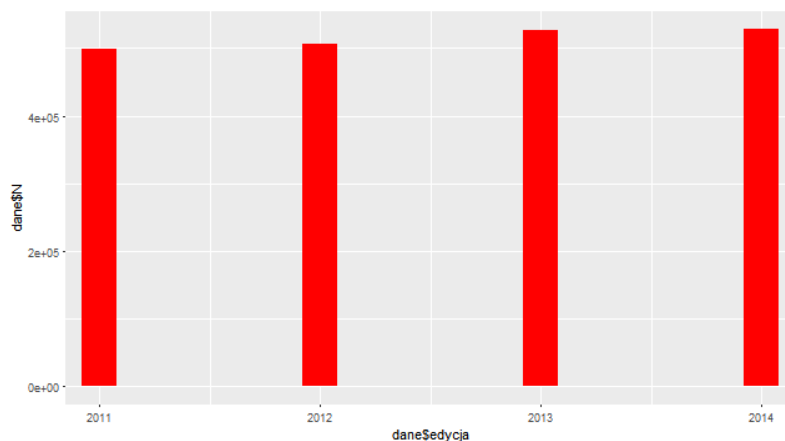
```
> tabela_woj
```

dolnośląskie	kujawsko-pomorskie	lubelskie	lubuskie	łódzkie	małopolskie
32	32	32	32	32	32
mazowieckie	opolskie	podkarpackie	podlaskie	pomorskie	śląskie
32	32	32	32	32	32
świętokrzyskie	warmińsko-mazurskie	wielkopolskie	zachodniopomorskie		
32	32	32	32		

3. Dla lat 2011-2014 poniższy wykres przedstawia wielkość N dla poszczególnych województw. Jak można zaobserwować najmniejsze wartości należą do województw lubuskiego, opolskiego i podlaskiego. Najwyższe wartości reprezentują natomiast województwa mazowieckie i śląskie.



4. Sprawdzając wartości N dla poszczególnych roczników dochodzimy do wniosku, że rok 2013 i 2014 dysponują wyższymi wartościami niż lata poprzedzające:



5. W związku z tym porównując dane dla roczników 2013 i 2014, weryfikujemy, iż to rok 2013 posiadał wyższe wartości:

```

· wiek3 <- subset(dane, edycja == "2013")
· summary(wiek3)
      edycja      woj.      wiek      plec      N
Min.   :2013   Length:128   Min.   :1.00   Min.   :0.0   Min.   : 51948
1st Qu.:2013   Class :character 1st Qu.:1.75   1st Qu.:0.0   1st Qu.:110123
Median :2013   Mode  :character  Median :2.50   Median :0.5   Median :167537
Mean   :2013                      Mean  :2.50   Mean   :0.5   Mean   :192752
3rd Qu.:2013                      3rd Qu.:3.25   3rd Qu.:1.0   3rd Qu.:242418
Max.   :2013                      Max.   :4.00   Max.   :1.0   Max.   :526761
· wiek4 <- subset(dane, edycja == "2014")
·
·
· wiek4 <- subset(dane, edycja == "2014")
· summary(wiek4)
      edycja      woj.      wiek      plec      N
Min.   :2014   Length:128   Min.   :1.00   Min.   :0.0   Min.   : 51097
1st Qu.:2014   Class :character 1st Qu.:1.75   1st Qu.:0.0   1st Qu.:111029
Median :2014   Mode  :character  Median :2.50   Median :0.5   Median :163418
Mean   :2014                      Mean  :2.50   Mean   :0.5   Mean   :191326
3rd Qu.:2014                      3rd Qu.:3.25   3rd Qu.:1.0   3rd Qu.:235801
Max.   :2014                      Max.   :4.00   Max.   :1.0   Max.   :527725

```

6. Następnie wykonaliśmy analizy poszczególnych lat wg. podziału na kobiety (0) i mężczyzn (1) (porównujemy z wartością średnią dla lat 2011-2014 dla poszczególnych płci):

```

#2011
sr2011 <- subset(dane, edycja == "2011" & plec == 0)
summary(sr2011)
#srednia - 198688
198688/200030
100-99.3291
#roznica = 0,67%

#2012
sr2012 <- subset(dane, edycja == "2012" & plec == 0)
summary(sr2012)
#srednia - 199689
199689/200030
100-99.82953
#roznica = 0,17%

#2013
sr2013 <- subset(dane, edycja == "2013" & plec == 0)
summary(sr2013)
#srednia - 200920
200920/200030
100-100.4449
#roznica = 0,44%

#2014
sr2014 <- subset(dane, edycja == "2014" & plec == 0)
summary(sr2014)
#srednia - 200825
200825/200030
100-100.3974
#roznica = 0,40%

#2011
sr2011M <- subset(dane, edycja == "2011" & plec == 1)
summary(sr2011M)
#srednia - 186292
186292/184498
100-100.9724
#roznica = 0,97%

#2012
sr2012M <- subset(dane, edycja == "2012" & plec == 1)
summary(sr2012M)
#srednia - 185290
185290/184498
100-100.4293
#roznica = 0,43%

#2013
sr2013M <- subset(dane, edycja == "2013" & plec == 1)
summary(sr2013M)
#srednia - 184584
200920/184498
100-100.89009
#roznica = 0,89%

#2014
sr2014M <- subset(dane, edycja == "2014" & plec == 1)
summary(sr2014M)
#srednia - 181827
181827/184498
100-98.55229
#roznica = 1,45%

```

7. Następnie porównujemy wartości ze względu na zmienną wiek i dochodzimy do wniosku iż rok 2014 cechował się najwyższą różnicą:

```
#2011
wiek1 <- subset(dane, edycja == "2011")
summary(wiek1)
192490/192264
100-100.12
#roznica - 0,12%

#2012
wiek2 <- subset(dane, edycja == "2012")
summary(wiek2)
192490/192264
100-100.12
#roznica - 0,12%

#2013
wiek3 <- subset(dane, edycja == "2013")
summary(wiek3)
192752/192264
100-100.25
#roznica - 0,25%

#2014
wiek4 <- subset(dane, edycja == "2014")
summary(wiek4)
191326/192264
100-99.51
#roznica - 0,49%
```

8. Wykres wartości N ze względu na grupę wiekową:

