Zadanie

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.2     v readr     2.1.4
v forcats   1.0.0     v stringr   1.5.0
v ggplot2   3.4.4     v tibble    3.2.1
v lubridate 1.9.2     v tidyr     1.3.0
v purrr     1.0.1
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

```
library(ggplot2)
library(dplyr)
data <- read.csv("beauty.csv", sep = ",")
```

Zadanie 1

```
any(is.na(data))
```

```
[1] FALSE
```

```
any(sapply(data, function(x) any(is.infinite(x))))
```

```
[1] FALSE
```

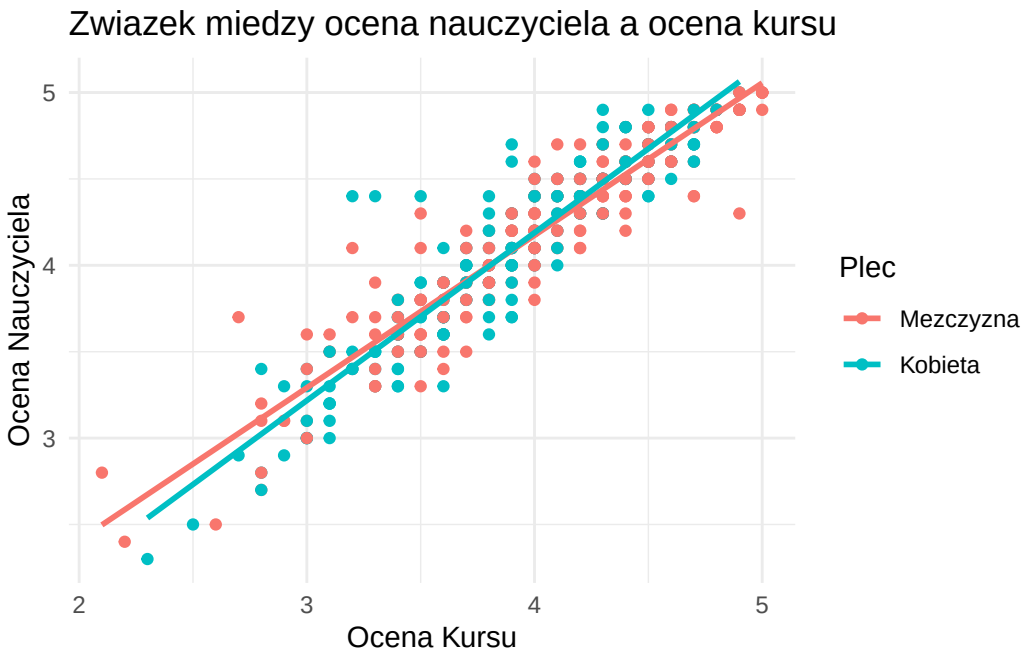
Dane wygladaja na poprawne. Nie ma wartosci NaN oraz wartosci INF

Zadanie 2

```
plot21 <- ggplot(data, aes(x = as.array(courseevaluation), y = as.array(profevaluation), c
  geom_point() +
  geom_smooth(method = 'lm', formula = 'y ~ x', se = FALSE) +
  labs(title = "Zwiazek miedzy ocena nauczyciela a ocena kursu",
       x = "Ocena Kursu",
       y = "Ocena Nauczyciela",
       color = "Plec") +
  scale_color_discrete(labels = c("1" = "Kobieta", "0" = "Mezczyzna")) +
  theme_minimal()

connection21 <- data |>
  summarise(
    correlation = cor(profevaluation, courseevaluation),
    meanprofevaluation = mean(profevaluation),
    meancourseevaluation = mean(courseevaluation)
  )

connection21
```

```
  correlation meanprofevaluation meancourseevaluation
1   0.9350966            4.17473             3.998272
```

```
plot21
```




```
connection22 <- data |>
  group_by(female, tenured, minority) |>
  summarise(
    correlation = cor(profevaluation, courseevaluation),
    meanprofevaluation = mean(profevaluation),
    meancourseevaluation = mean(courseevaluation)
  )
```

```
`summarise()` has grouped output by 'female', 'tenured'. You can override using
the `.groups` argument.
```

```
connection22
```

```
# A tibble: 8 x 6
# Groups:   female, tenured [4]
  female tenured minority correlation meanprofevaluation meancourseevaluation
   <int>   <int>    <int>       <dbl>              <dbl>                 <dbl>
1      0       0        0       0.933               4.38                  4.22
2      0       0        1       0.971               4.59                  4.43
3      0       1        0       0.947               4.16                  3.99
```

4	0	1	1	0.924	3.95	3.8
5	1	0	0	0.932	4.12	3.92
6	1	0	1	0.820	3.90	3.68
7	1	1	0	0.950	4.15	3.97
8	1	1	1	0.969	4	3.87

Maja miedzy soba jakis zwiazek i to dosc silny. Charakter tego zwiazku jest podobny dla roznych plci, doswiadczenia czy mniejszosci.

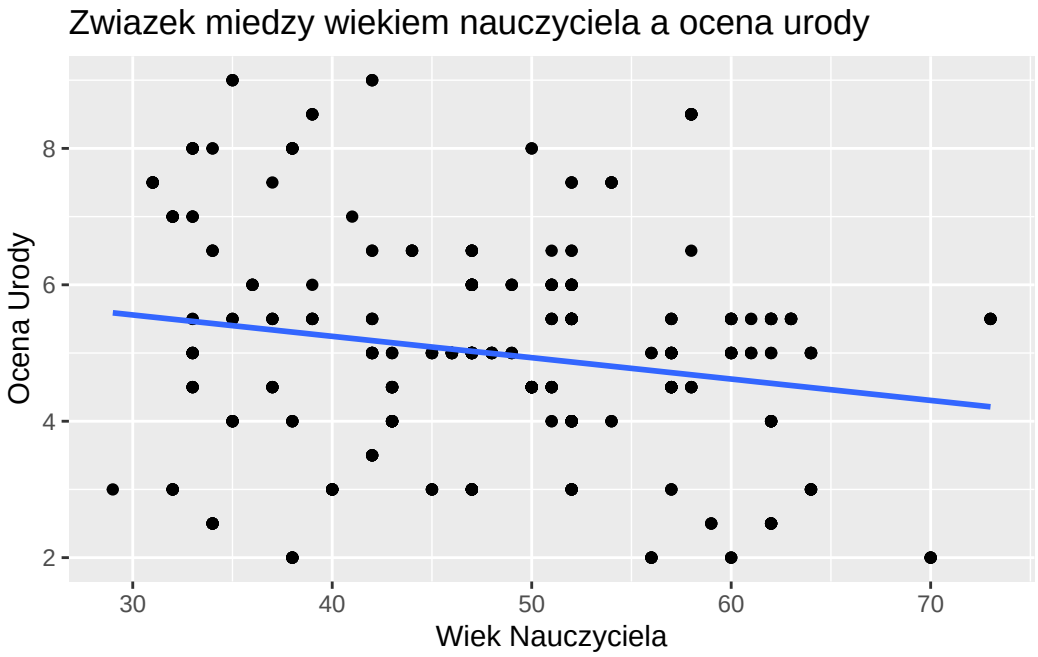
Zadanie 3

Zadanie 4

```
newdata = data |>
  mutate(meanbeauty2upper = (beautyf2upper + beautym2upper) / 2)

connection41 <- lm(meanbeauty2upper ~ age, data = newdata)
summary(connection41)
```

```
Call:
lm(formula = meanbeauty2upper ~ age, data = newdata)

Residuals:
    Min      1Q  Median      3Q     Max
-3.3074 -0.8689 -0.0255  0.9444  3.8191

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.497643   0.360172  18.040  < 2e-16 ***
age         -0.031323   0.007299  -4.291 2.16e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.538 on 461 degrees of freedom
Multiple R-squared:  0.03841,   Adjusted R-squared:  0.03633
F-statistic: 18.42 on 1 and 461 DF,  p-value: 2.165e-05
```

```r
plot41 <- ggplot(newdata, aes(x = as.array(age), y = as.array(meanbeauty2upper))) +
  geom_point() +
  geom_smooth(method = "lm", formula = 'y ~ x', se = FALSE) +
  labs(title = "Zwiazek miedzy wiekiem nauczyciela a ocena urody",
       x = "Wiek Nauczyciela",
       y = "Ocena Urody"
  )

plot41
```

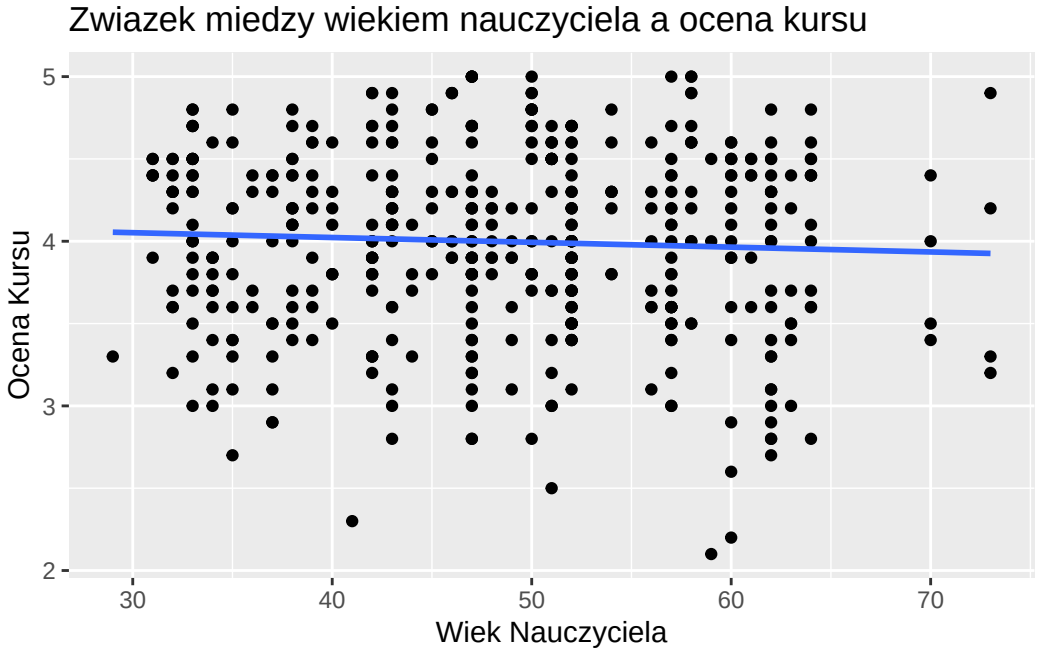


Tak, im jest starszy tym jest gorzej oceniany, sa od tego wyjatki jak np ocena 8.5 w wieku 58 lat

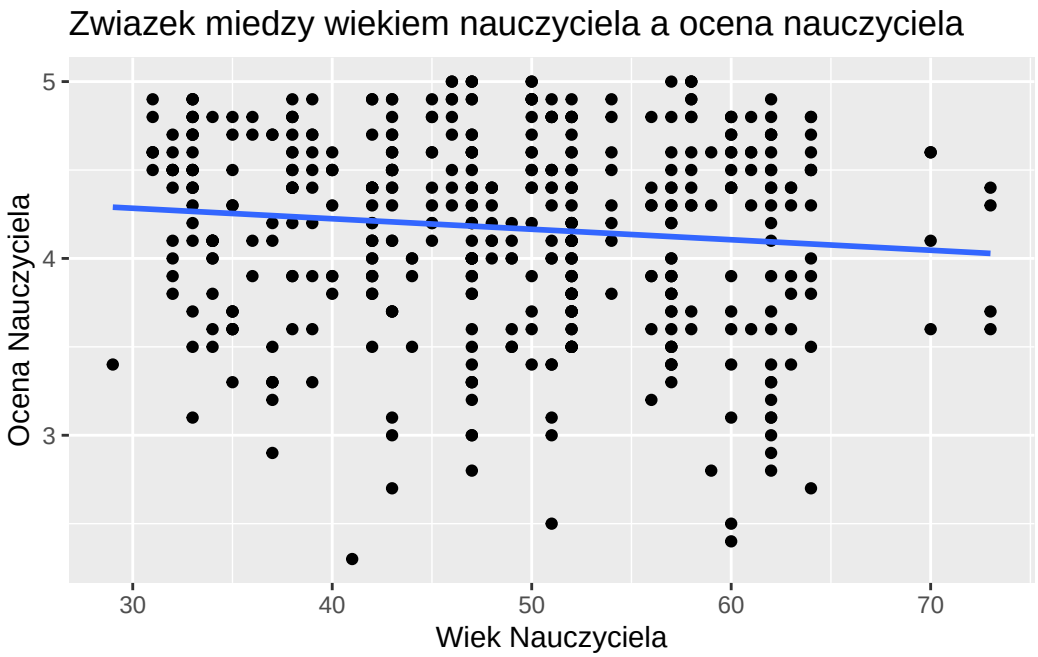
Zadanie 5

```r
plot51 <- ggplot(data, aes(x = as.array(age), y = as.array(courseevaluation))) +
  geom_point() +
  geom_smooth(method = "lm", formula = 'y ~ x', se = FALSE) +
  labs(title = "Zwiazek miedzy wiekiem nauczyciela a ocena kursu",
```

```
       x = "Wiek Nauczyciela",
       y = "Ocena Kursu"
  )

plot52 <- ggplot(data, aes(x = as.array(age), y = as.array(profevaluation))) +
  geom_point() +
  geom_smooth(method = "lm", formula = 'y ~ x', se = FALSE) +
  labs(title = "Zwiazek miedzy wiekiem nauczyciela a ocena nauczyciela",
       x = "Wiek Nauczyciela",
       y = "Ocena Nauczyciela"
  )

plot51
```




```
plot52
```



Tak, im nauczyciel jest starszy tym gorsza jego ocena i ocena kursu

Zadanie 6

```
connection61 <- newdata |>
  summarise(
    correlation = cor(meanbeauty2upper, courseevaluation),
    meanbeauty = mean(meanbeauty2upper),
    meancourseevaluation = mean(courseevaluation)
  )
connection61
```

```
  correlation meanbeauty meancourseevaluation
1   0.1849776   4.982721             3.998272
```

Jest slaby zwiazek pomiedzy uroda nauczyciela a ocena kursu