

Raport nr 1

Kinga Heda, Karol Cieřlik

Niniejszy raport przedstawia analizę różnic w wynagrodzeniach ze względu na lata doświadczenia zawodowego oraz wykonywany zawód, uwzględniając dane z Głównego Urzędu Statystycznego (GUS). W kontekście równości płci w wynagrodzeniach, szczególną uwagę poświęcono wskaźnikowi luki płacowej (GPG – Gender Pay Gap), który stanowi istotny wskaźnik monitorujący nierówności w wynagrodzeniach między kobietami a mężczyznami.

Luka płacowa w swojej niekorygowanej formie jest definiowana jako różnica między średnimi godzinowymi zarobkami brutto mężczyzn i kobiet, wyrażona jako procent średnich godzinowych zarobków brutto mężczyzn. Wyliczana jest dla przedsiębiorstw zatrudniających co najmniej 10 pracowników. W raporcie wykorzystano ten wskaźnik jako punkt odniesienia do oceny różnic płacowych w wybranych grupach zawodowych oraz w zależności od stażu pracy.

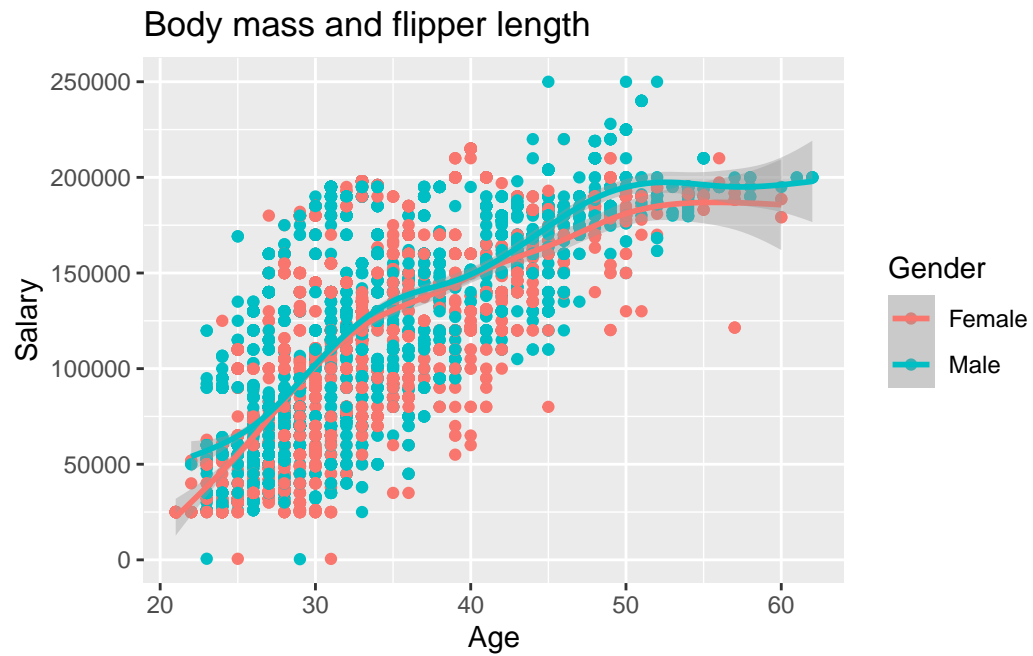
Dane te mają kluczowe znaczenie w analizie strukturalnych uwarunkowań wynagrodzeń, pozwalając zidentyfikować obszary wymagające działań w zakresie polityki równościowej i zarządzania zasobami ludzkimi.

```
dane_cleaned <- dane %>%  
  mutate(Gender = factor(Gender),  
         Education.Level = factor(Education.Level)) %>%  
  filter(!is.na(Gender),  
         !is.na(Education.Level),  
         !is.na(Years.of.Experience),  
         !is.na(Salary))
```

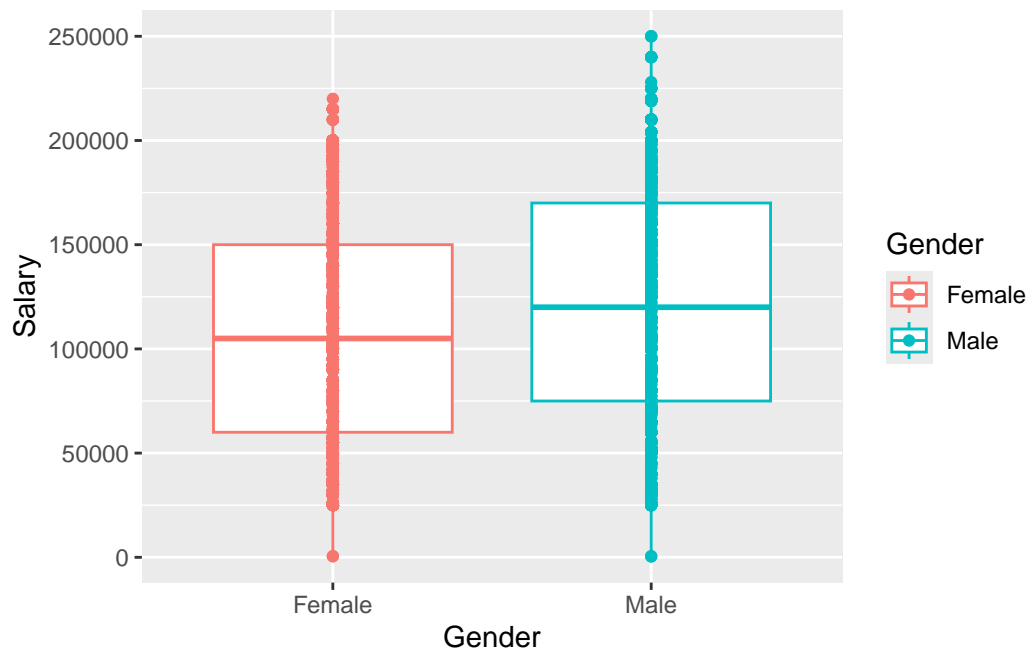
ssas

```
dane_cleaned %>%  
  ggplot(mapping=aes(x = Age, y = Salary, colour = Gender)) +  
  geom_point()+  
  labs(  
    title = "Body mass and flipper length")+  
  geom_smooth(se = TRUE)
```

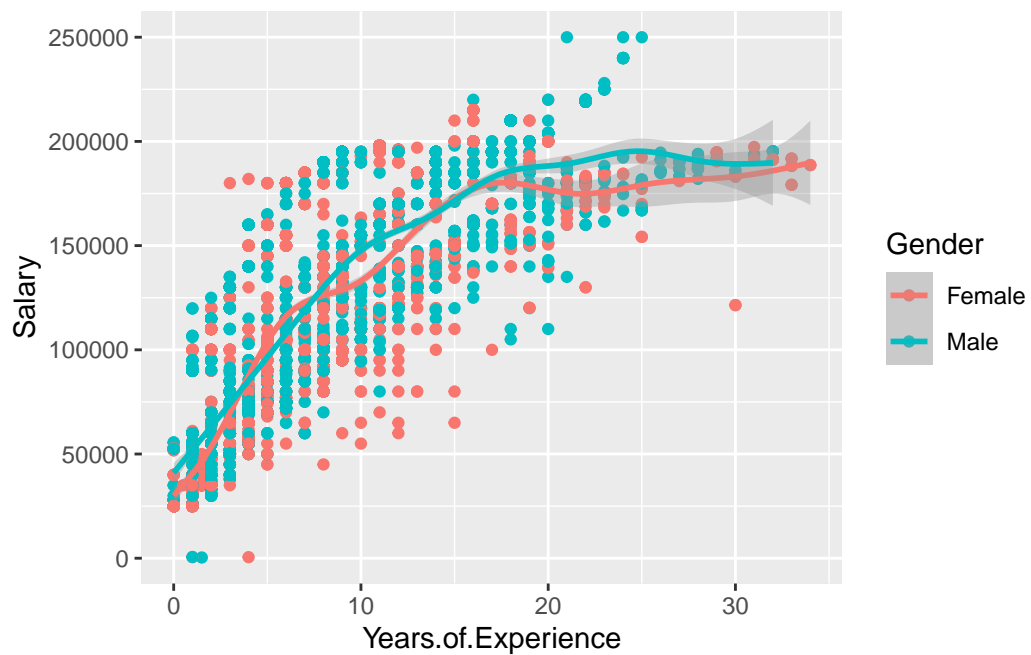
```
`geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



```
dane_cleaned %>%  
  ggplot(aes(x = Gender, y = Salary, colour = Gender)) +  
  geom_boxplot()+  
  geom_point()
```



```
dane_cleaned %>%
  ggplot(mapping=aes(y = Salary, x = Years.of.Experience, colour = Gender)) +
  geom_point() +
  geom_smooth(se = TRUE)
```



```
seto <- dane_cleaned$Salary[dane_cleaned$Gender == "Female"]
versi <- dane_cleaned$Salary[dane_cleaned$Gender == "Male"]
t_test_result <- t.test(seto, versi)

print(t_test_result)
```

Welch Two Sample t-test

```
data: seto and versi
t = -10.477, df = 6400.5, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -16033.92 -10979.48
sample estimates:
mean of x mean of y
 107889.0  121395.7
```

```
dane_cleaned %>%
  ggplot(mapping = aes(x = Salary, fill = Gender)) +
  geom_density(alpha = 0.5)
```

