

Praca_domowa_II

Pawel_Malinowski_305822

22 03 2021

Wyrkes zrodlowy

Men vs. Women: Comparing Income by Industry

Median Annual Earnings For the Civilian Employed Population (16+ years old) in the U.S.

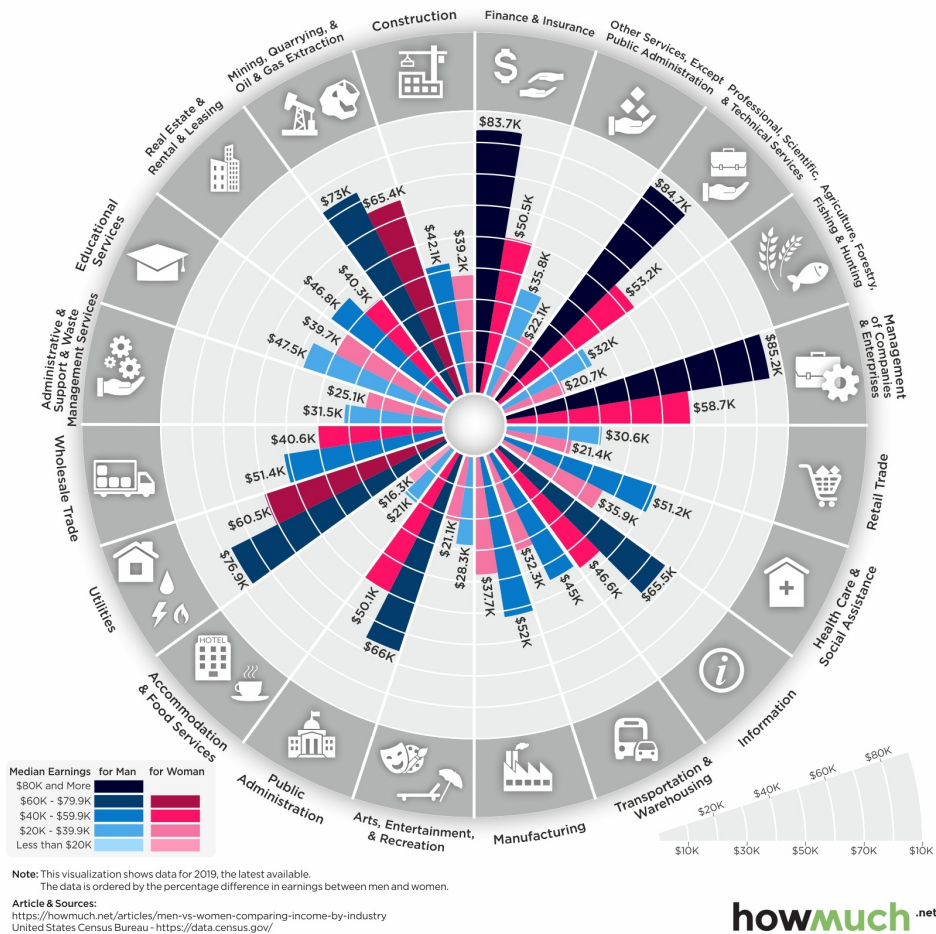


Figure 1: <https://howmuch.net/articles/men-vs-women-comparing-income-by-industry>

Wykres poprawiony, oraz kod do jego stworzenia.

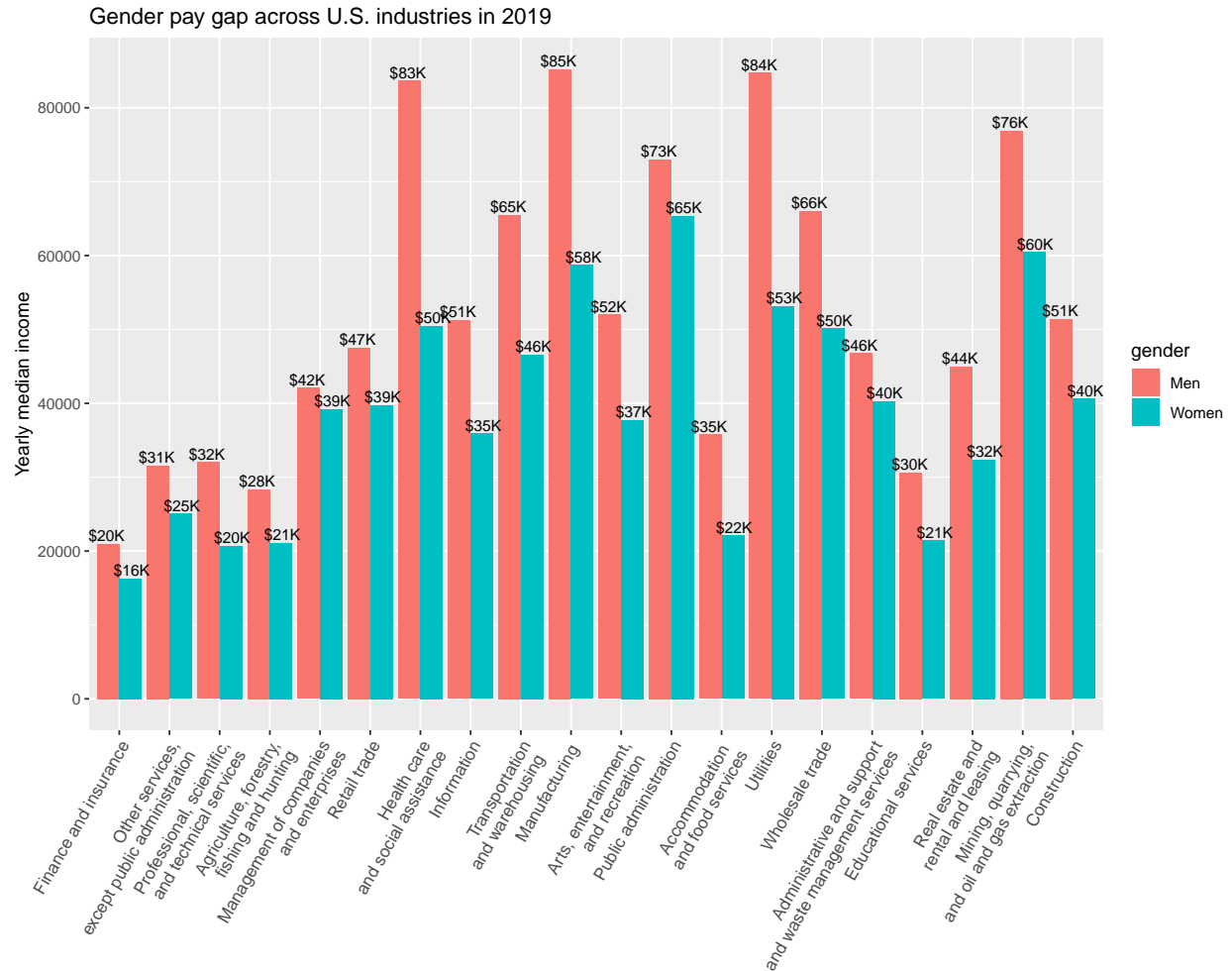
```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.3      v purrr 0.3.4
## v tibble 3.1.0      v dplyr 1.0.5
## v tidyr 1.1.3       v stringr 1.4.0
## v readr 1.4.0       v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()      masks stats::lag()

library(readxl)
library(stringr)
dane<-read_excel("C:\\Users\\pawel\\OneDrive\\Pulpit\\dane.xlsx")
dane_wydluzone <- dane %>%
  pivot_longer(Men:Women, names_to = "gender", values_to = "median_income") %>%
  mutate(zapis = str_c("$",median_income %/% 1000, "K"))
ggplot(data=dane_wydluzone, aes(x=Industry, y=median_income, fill=gender))+
  geom_bar( stat = 'identity', position ="dodge")+
  geom_text(aes(label=zapis), position=position_dodge(width=1), vjust=-0.25,size=(3))+
  theme(axis.text.x = element_text(angle=60, hjust=1,size = (10)),axis.title.x=element_blank()+
  labs(y="Yearly median income",title = "Gender pay gap across U.S. industries in 2019")+
  scale_x_discrete(labels=c("Finance and insurance", "Other services,\n except public administration"
    , "Professional, scientific,\n and technical services",
    "Agriculture, forestry,\n fishing and hunting",
    "Management of companies\n and enterprises","Retail trade",
    "Health care\n and social assistance", "Information",
    "Transportation\n and warehousing ", "Manufacturing ",
    "Arts, entertainment,\n and recreation ", "Public administration",
    "Accommodation \nand food services","Utilities", "Wholesale trade",
    "Administrative and support\n and waste management services",
    "Educational services", "Real estate and\n rental and leasing",
    "Mining, quarrying,\n and oil and gas extraction", "Construction"))
```



Dokonane zmiany i ich uzasadnienie.

Co nie podoba mi się w pierwotnym wykresie:

*Jest to wykres słupkowy “ukolowiony” czyli koniec i początek osi X został ze sobą połączony. Powoduje to, że taka sama różnica płac przy płacach niższych wydaje się mniejsza niż przy płacach wyższych. Co więcej procentowo różnica ta będzie większa właśnie dla płac niższych.

*Zastosowano zmiany koloru, która powoduje zwrócenie uwagi na branże, w których obydwie płcie zarabiają dużo co nie jest tematem tego wykresu.

*Rok, z którego pochodzą dane jest zapisany w adnotacji a nie w tytule, mimo tego, że wykres powstał w roku 2021 a dane zostały zebrane w 2019

Jakie zmiany wprowadziłem:

*Stworzyłem wykres słupkowy z prostą osią x

*Nie dodawałem zmiany nasycenia koloru

*Umieściłem datę pobrania danych w tytule.