

# TWD PD1

*Hubert Baniecki*

*14 października 2018*

## The 2018 Forbes 400

W tej pracy domowej zwizualizuję dane z artykułu:

[www.forbes.pl/rankingi/lista-forbes-400-jeff-bezos-najbogatszym-czlowiekiem-usa-2018/few6e1f](http://www.forbes.pl/rankingi/lista-forbes-400-jeff-bezos-najbogatszym-czlowiekiem-usa-2018/few6e1f)

([www.forbes.com/forbes-400/](http://www.forbes.com/forbes-400/))

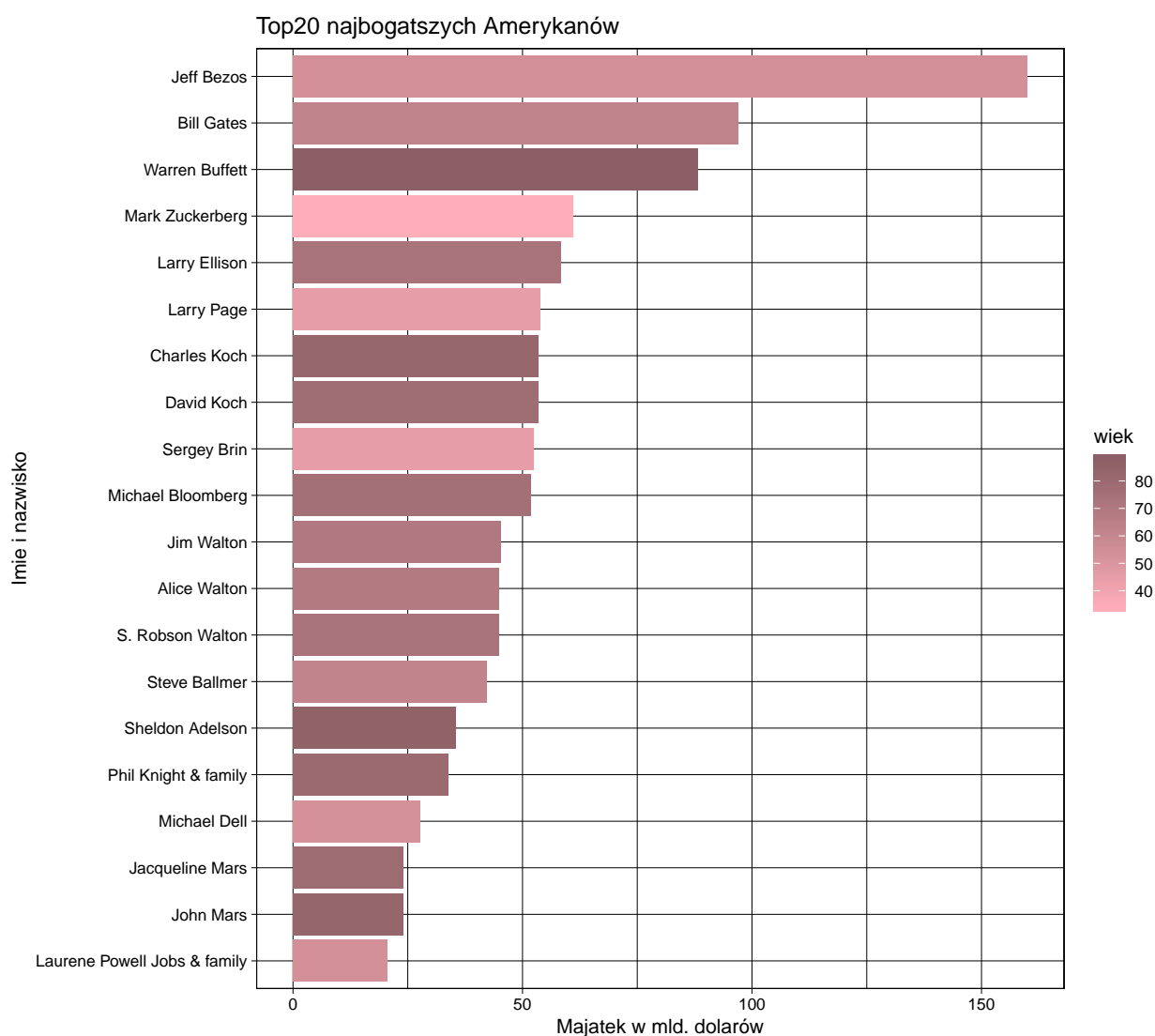
Zawiera on listę 400 najbogatszych ludzi Ameryki posortowaną względem wielkości ich majątku. Znajdziemy w nim również dane dotyczące wieku, źródła majątku czy stanu w którym mieszkają. Interesującą statystyką jest wskaźnik filantropii oceniany w skali 1-5, na podstawie którego możemy dowiedzieć się, jak bardzo dana osoba angażuje się w pomoc innym. Na potrzeby tego zadania wybrałem dane pierwszych 100 osób i zapisałem je w pliku *“top100.csv”*.

```
input <- read.csv("top100.csv")
colnames(input) <- c("pozycja", "imie.nazwisko", "majatek",
                    "wiek", "zrodlo.majatku", "filantropia")
input$imie.nazwisko <- factor(input$imie.nazwisko, levels=input$imie.nazwisko)
```

## Ranking top20

Poniższy wykres przedstawia porównanie majątków 20 najbogatszych ludzi w Ameryce. Zmieniłem motyw na linedraw, aby ułatwić oszacowanie wartości. Kolor słupka ilustruje wiek danej osoby.

```
temp <- input[1:20,]  
ggplot(temp, aes(x=imie.nazwisko, y=majatek, fill = wiek)) +  
  geom_col() +  
  coord_flip() +  
  scale_x_discrete(limits = rev(levels(temp$imie.nazwisko))[81:100]) +  
  scale_fill_gradient(low="lightpink1", high="lightpink4") +  
  xlab("Imię i nazwisko") +  
  ylab("Majątek w mld. dolarów") +  
  ggtitle("Top20 najbogatszych Amerykanów") +  
  theme_linedraw()
```



## Jak korelują wiek, majątek i filantropia?

Poniższy wykres przedstawia zależność majątku od wieku 100 najbogatszych ludzi w Ameryce. Kolor punktu ilustruje skalę filantropii. Możemy zauważyć, że większość osób na górze listy ma ponad 60 lat. O ile wiek zdaje się nie mieć wpływu na skalę filantropii, to ludzie z większym majątkiem chętniej dzielą się nim z innymi. Białe linie na wykresie wyznaczają medianę majątku oraz medianę wieku.

```
temp <- na.omit(input[,c(2,3,4,6)])
ggplot(temp, aes(x=wiek, y=majatek, color=filantropia, label=imie.nazwisko)) +
  geom_point(size=3, shape=15) +
  geom_hline(yintercept = median(temp$majatek), color="white", size=1) +
  geom_vline(xintercept = median(temp$wiek), color="white", size=1) +
  theme_dark() +
  scale_color_gradient(low="yellow", high="red") +
  ylab("Majątek w mld. dolarów") + ylim(min(temp$majatek),max(temp$majatek)) +
  xlab("Wiek") + xlim(min(temp$wiek),max(temp$wiek)) +
  theme(legend.position = "bottom") +
  ggtitle("Top100 najbogatszych Amerykanów") +
  geom_text_repel(data=temp[1:10,], color="white") + labs(color = "Skala filantropii")
```

