

# Praca Domowa 2 - Raport

Kinga Ułasik IiAD

## Błędna wizualizacja

Wybrany przez mnie wykres przedstawia ilość wykonanych testów na covid19 na milion mieszkańców dla różnych państw (Brazylia, Argentyna, USA, Włochy, Niemcy, Norwegia). Obrazek został umieszczony na Reddit-cie, oryginalnie został pokazany na argentyńskim kanale telewizyjnym C5N.

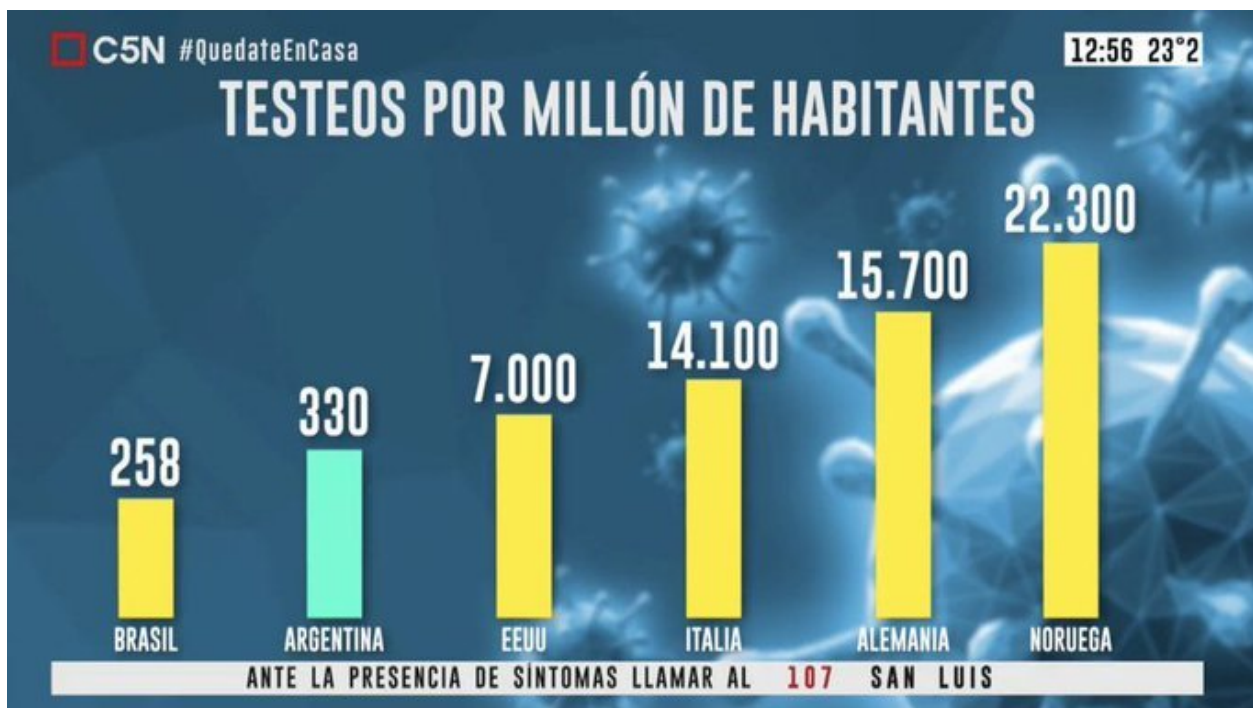


Figure 1: Nieprawidłowy wykres

[Link do Reddita z wykresem](#)

[Link do strony C5N](#)

## Poprawiony wykres i kod w R

Mój kod do poprawionego wykresu

```
Pais <- c( "Brasil", "Argentina", "EEUU", "Italia", "Alemania", "Noruega")
Testeos <- c(258, 330, 7000, 14100, 15700, 22300)
colores <- c("gold1", "lightblue", "gold1", "gold1", "gold1", "gold1" )
CovidPruebas <- data.frame(Pais, Testeos, colores)
CovidPruebas$Pais <-
  factor(CovidPruebas$Pais, levels = c("Brasil","Argentina","EEUU","Italia","Alemania","Noruega"))

ggplot(CovidPruebas, aes(x = Pais, y = Testeos)) +
  geom_bar(stat = "identity", fill = colores) +
  geom_text(aes(label = Testeos), vjust=-0.3, size=3.5) +
  theme_minimal()
```

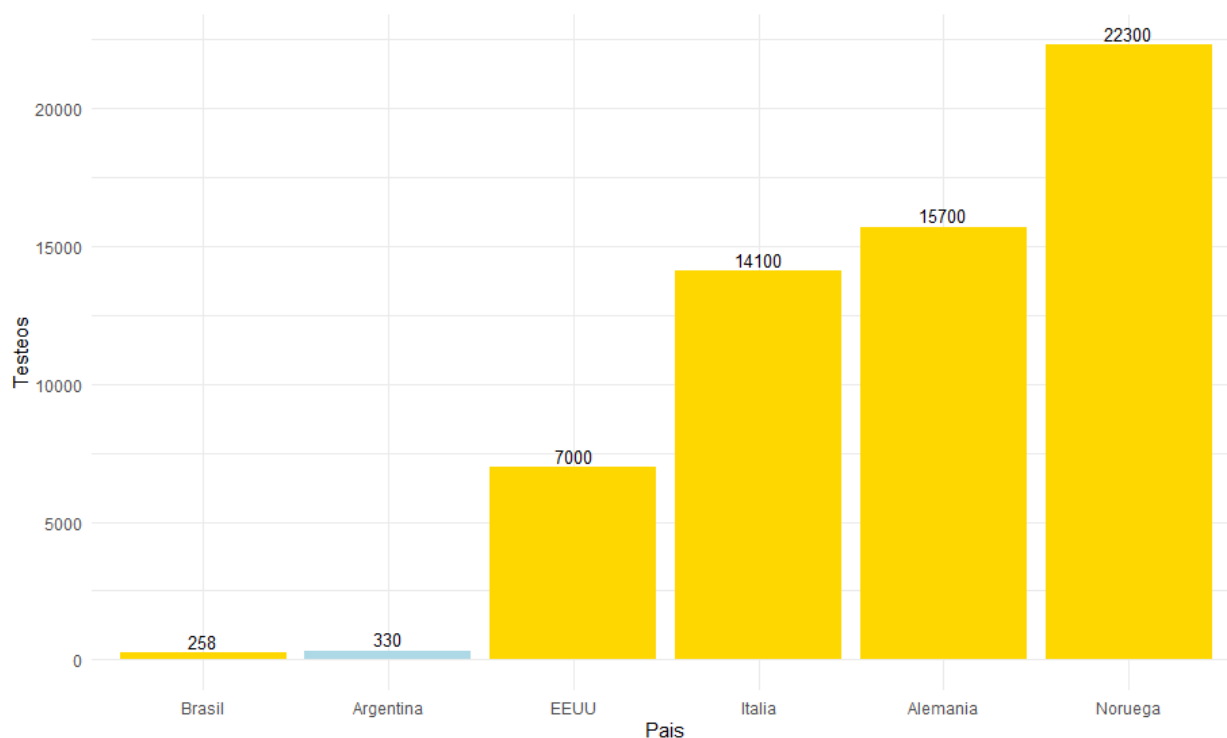


Figure 2: Poprawiony wykres

## Dlaczego mój wykres jest bardziej poprawny?

Na wykresie umieszczonym przez C5N wysokości słupków nie są adekwatne do wartości na tych słupkach. Słupki dla Argentyny i USA są bardzo podobnej wielkości, a przecież 7000 jest około 21 razy większe niż 330. Dalej, słupki dla USA i Włoch różnią się niewiele jeśli chodzi o wysokość, jednak wykonanych testów we Włoszech było ponad dwa razy więcej. Jako, że C5N jest argentyńskim kanałem, prawdopodobnie chodziło o przedstawienie tego właśnie kraju w jak najlepszym świetle, stąd te „dziwne” proporcje (a raczej ich brak). Poprawiony przeze mnie wykres ma zachowane prawidłowe proporcje wysokości słupków, zachowując styl poprzedniego wykresu (podpisane kraje, Argentyna podkreślona innym kolorem)