

Sobieski Bartłomiej PD2

21 marca 2021

1. Źródło oraz poprawiana wizualizacja.

Wizualizacja do poprawy pochodzi ze strony internetowej gazety „The Philadelphia Inquirer”, znajduje się ona w artykule, do którego link zamieszczam poniżej:

<https://www.inquirer.com/news/covid-vaccine-supply-pa-suburbs-doh-share-20210309.html>

Wspomniana wizualizacja znajduje się na następnej stronie:

First Doses Allocated to Pa., by Manufacturer

The CDC allocates doses to states and health departments nationally based on population size. The allocated totals may not match the number of doses actually delivered in any week. Totals do not include Philadelphia, which has a separate allocation from the CDC.



Chart: JOHN DUCHNESKIE / Staff Artist • Source: Centers for Disease Control and Prevention

2. Kod oraz utworzona wizualizacja.

```
## Import i obróbka danych

library(readxl)
library(tidyverse)

pfizer <- read_excel("E:/Dane o szczepionkach/Pfizer.xlsx")
moderna <- read_excel("E:/Dane o szczepionkach/Moderna.xlsx")
jj <- read_excel("E:/Dane o szczepionkach/J&J.xlsx")
pfizer

pf_pen <- pfizer %>%
  filter(Jurisdiction=="Pennsylvania") %>%
  select(Week="Week of Allocations", Pfizer="1st Dose Allocations")

mod_pen <- moderna %>%
  filter(Jurisdiction=="Pennsylvania") %>%
  select(Week="Week of Allocations", Moderna="1st Dose Allocations")

jj_pen <- jj %>%
  filter(Jurisdiction=="Pennsylvania") %>%
  select(Week="Week of Allocations", JJ="1st Dose Allocations")

tmp <- left_join(pf_pen,mod_pen)
tmp <- left_join(tmp, jj_pen)
tmp <- tmp %>% map_df(rev)
tmp

doses_alloc <- data.frame(tmp[1:12,]) %>%
  replace_na(list(Moderna = 0,JJ=0)) %>%
  mutate(Pfizer=cumsum(Pfizer), Moderna=cumsum(Moderna), JJ=cumsum(JJ)) %>%
  mutate(Total=Pfizer+Moderna+JJ)

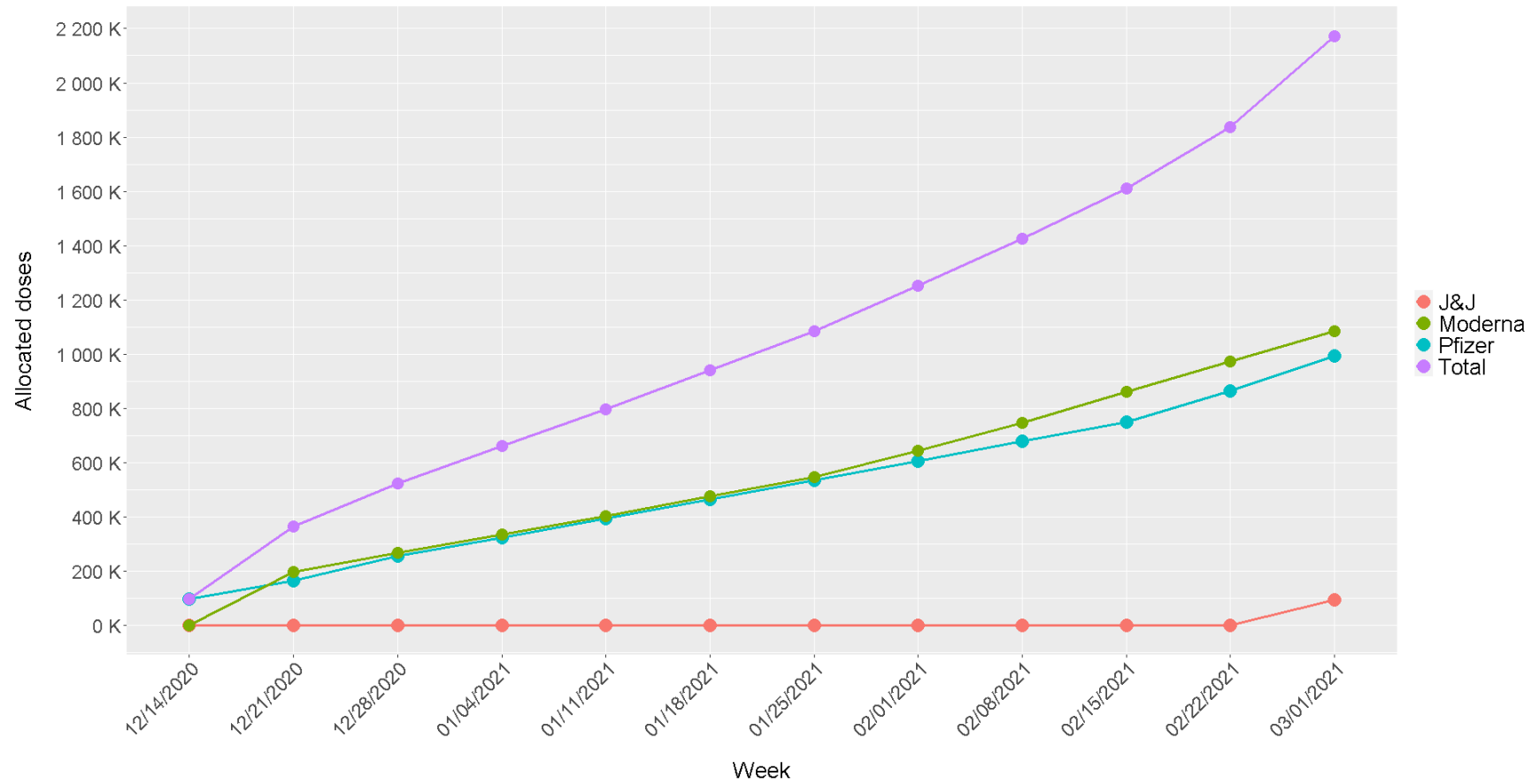
doses_alloc$Week <- factor(doses_alloc$Week, levels=unique(doses_alloc$Week))
```

wykres punktowy

```
doses_alloc %>%
  ggplot(aes(x=Week,y=value))+
  geom_point(aes(y=JJ, color="J&J"),size=5)+
  geom_line(aes(y=JJ,group = 1,color="J&J"),size=1.1)+
  geom_point(aes(y=Pfizer, color="Pfizer"),size=5)+
  geom_line(aes(y=Pfizer,group = 1,color="Pfizer"),size=1.1)+
  geom_point(aes(y=Moderna, color="Moderna"),size=5)+
  geom_line(aes(y=Moderna,group = 1,color="Moderna"),size=1.1)+
  geom_point(aes(y=Total,color="Total"),size=5)+
  geom_line(aes(y=Total,group = 1,color="Total"),size=1.1)+
  labs(title="First doses of COVID-19 vaccine allocated to Pennsylvania, by manufacturer",
        subtitle="Data from Philadelphia not included",
        x="Week",y="Allocated doses")+
  theme(plot.title = element_text(size = 25, face = "bold",hjust=1,
                                   margin = margin(t = 0, r = 20, b = 0, l = 0)),
        plot.subtitle = element_text(size = 16,hjust=0.5,
                                       margin = margin(t = 0, r = 20, b = 30, l = 0)),
        axis.title.y = element_text(margin = margin(t = 0, r = 20, b = 0, l = 0), size=21),
        axis.title.x = element_text(margin = margin(t = 20, r = 0, b = 0, l = 0), size=21),
        axis.text.y = element_text(size=18),
        axis.text.x = element_text(size=18,angle=45,hjust=1),
        legend.text = element_text(size = 21),
        legend.title = element_blank())+
  scale_y_continuous(breaks = round(seq(0, 2200,by = 200),1),label = unit_format(unit = "K"))
```

First doses of COVID-19 vaccine allocated to Pennsylvania, by manufacturer

Data from Philadelphia not included



3. Podsumowanie

Powyższa propozycja poprawy wizualizacji pozwala względem pierwotnego wykresu:

- odczytać dokładne (co do 200K) wartości na osi OY, które różnią się między sobą o stałą wartość (pierwotny wykres ma niedokładne wartości i nieregularne różnice)
- uświadomić odbiorcę o dyskretnym charakterze danych, których wartości zostały zaznaczone punktami, a następnie połączone ciągłą linią, by w łatwy sposób ukazać tendencje wzrostowe (pierwotna wizualizacja prezentowała w pewnym sensie wykres gęstości, jednak wartości na osi OY były dopasowane do wykresu punktowego)
- odczytać, kiedy dany producent nie dostarczał żadnych szczepionek i w którym momencie rozpoczął ich dystrybucję, a także dokładnie określić jego udział w całkowitej liczbie wydanych próbek w każdym tygodniu (pierwotny wykres sumy wszystkich próbek miał wydzielone kolorami obszary, które tylko w bardzo poglądowy sposób pozwalały określić, jaka część szczepionek pochodzi od konkretnego producenta)
- pominąć przetwarzanie przez odbiorcę nadmiernych informacji zawartych w podtytule, które utrudniają jego wstępny odbiór i pogarszają czytelność całego wykresu.