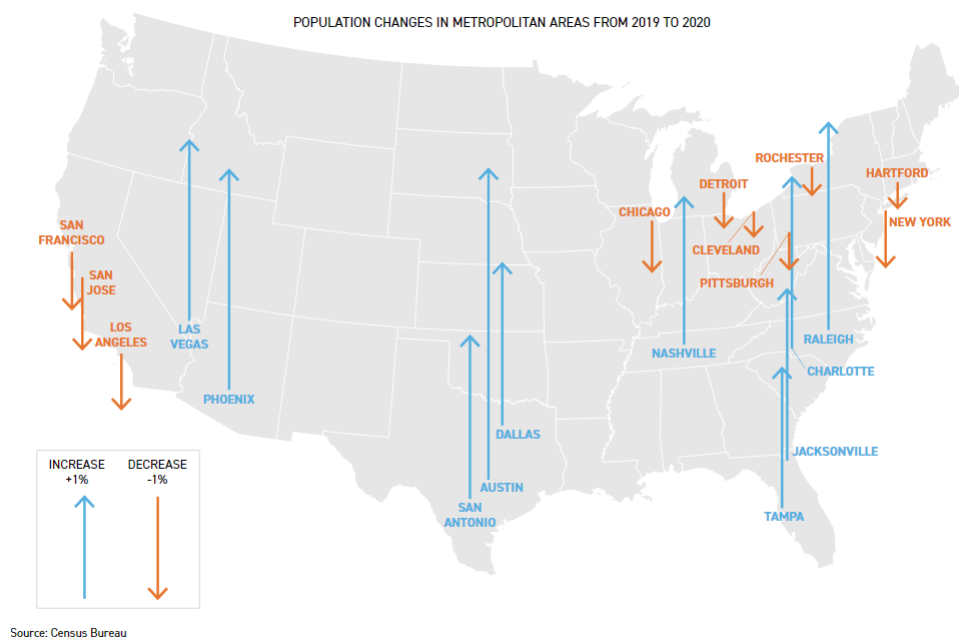


TWD - HW2

Michał Gromadzki

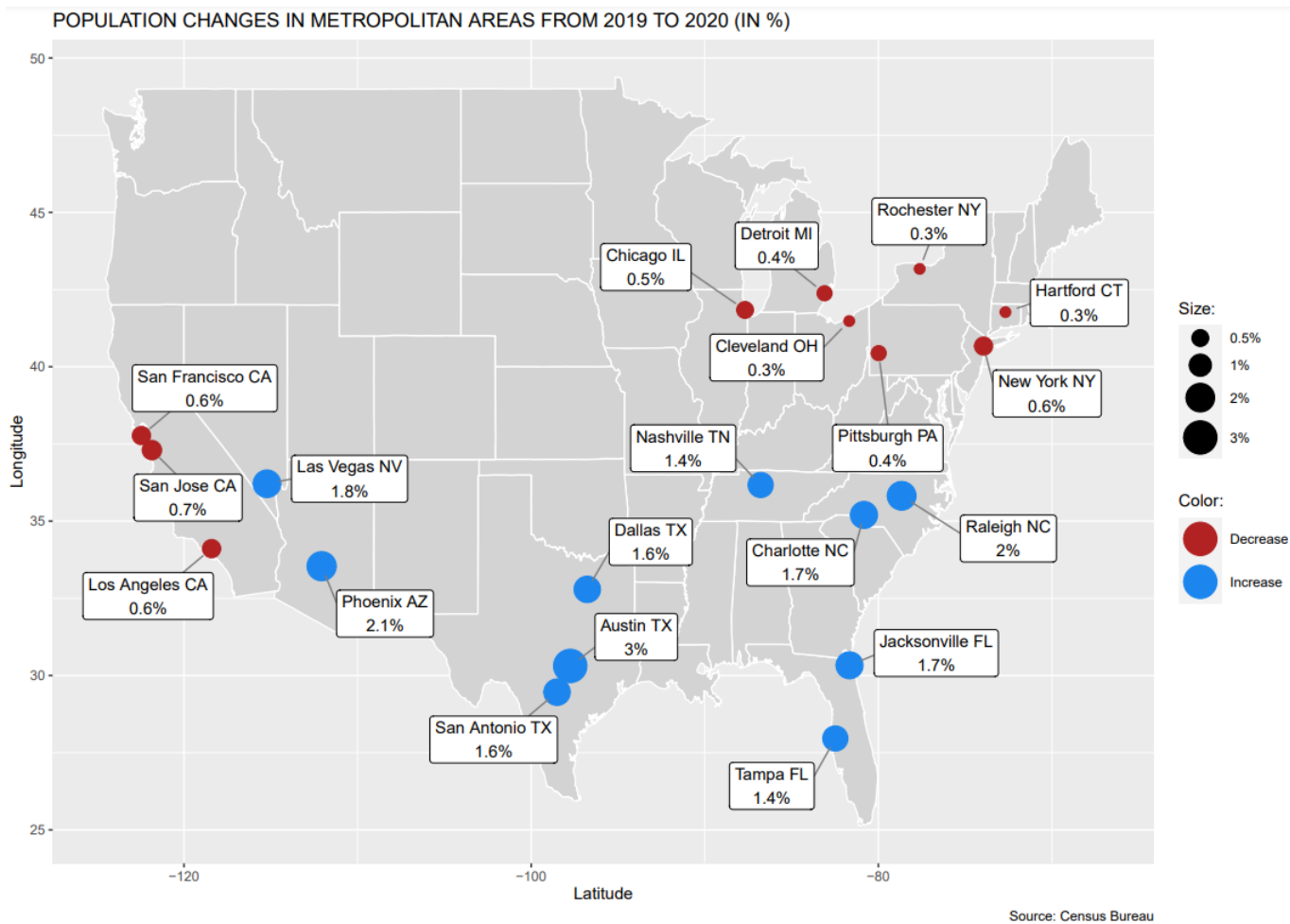
November 1, 2021

1 Poprawiana wizualizacja



źródło: <https://www.politico.com/news/2021/10/21/covid-american-cities-charts-516295>

2 Utworzona wizualizacja



3 Kod

```
library(ggplot2)
library(maps)
library(dplyr)
library(ggrepel)

poprawiony <- us.cities %>%
  filter(name %in% c("Austin TX","Phoenix AZ","Raleigh NC","Las Vegas NV","Charlotte NC","Jacksonville FL","San Antonio TX",
                    "Dallas TX","Nashville TN","Tampa FL","Cleveland OH","Hartford CT","Rochester NY","Detroit MI",
                    "Pittsburgh PA","Chicago IL","Los Angeles CA","New York NY","San Francisco CA","San Jose CA")) %>%
  arrange(name) %>%
  mutate(popdiff=c(3,1.7,-0.5,-0.3,1.6,-0.4,-0.3,1.7,1.8,-0.6,1.4,-0.6,2.1,-0.4,2,-0.3,1.6,-0.6,-0.7,1.4)
)

MainStates <- map_data("state")

ggplot() +
  geom_polygon(data=MainStates, aes(x=long, y=lat, group=group),
              color="white", fill="lightgrey" ) +
  geom_point(data=poprawiony, aes(x=long, y=lat, size = abs(popdiff),color=ifelse(popdiff>0,"firebrick","dodgerblue2")))+
  geom_label_repel(poprawiony,mapping=aes(x=long,y=lat,label = paste(paste(name,abs(popdiff),sep="\n"),"%",sep="")),
                  box.padding = 1,
                  point.padding = 1,
                  segment.color = 'grey50')+
  guides(color = guide_legend(override.aes = list(size = 10)))+
  scale_color_manual(labels = c("Decrease", "Increase"), values = c("firebrick", "dodgerblue2"))+
  scale_size_continuous(range = c(5, 10),labels=c("0.5%","1%","2%","3%"),breaks = c(0.5,1,2,3))+
  labs(title="POPULATION CHANGES IN METROPOLITAN AREAS FROM 2019 TO 2020 (IN %)",
       caption = "Source: Census Bureau",
       color="Color:",
       size="Size:",
       x="Latitude",
       y="Longitude")
```

4 Uzasadnienie

Poprawiony wykres jest lepszy od początkowego ze względu na 3 aspekty:

1. Ze względu na to, że strzałki zaczynają się na różnej wysokości są one dużo trudniejsze do porównania niż wielkość kropek.
2. Z tego samego powodu, na wykresie początkowym trudniej jest się odnieść do skali niż na poprawionym.
3. Kropka na poprawionym wykresie znajduje się dokładnie w miejscu, które opisuje. Natomiast na pierwszym wykresie problem może sprawić dopasowanie początku strzałki, do miejsca któremu odpowiada, szczególnie w bardziej zagęszczonych obszarach.

W wizualizacji samych danych o zmianie populacji lepiej sprawdziłby się wykres kolumnowy lub słupkowy, jednak nie zawierają one informacji o położeniu opisywanych obszarów na mapie USA, dlatego nie zostały użyte.