

# Mappeinnlevering2

Eov016

2 2 2022

## OPPGAVE1

*#Etter å ha funnet json filen fra NY-times ved å lese gjennom sidekilden, lastet jeg opp linken. Deretter gjorde jeg den om til df slik jeg ville ha den.*

```
url <- read_json("https://static01.nyt.com/newsgraphics/2021/12/20/us-coronavirus-deaths-2021/ff0adde21")
```

```
cov_json_res <- url
str(cov_json_res)
```

```
## List of 51
## $ :List of 6
## ..$ name           : chr "Alabama"
## ..$ geoid           : chr "USA-01"
## ..$ deaths_before   : int 10790
## ..
```

```
cov_json <- lapply(cov_json_res, function(x) {
  x[sapply(x, is.null)] <- NA
  unlist(x)
})
```

```
NYT_cov_data <- as.data.frame(do.call("rbind", cov_json))
str(NYT_cov_data)
```

```
## 'data.frame':   51 obs. of  6 variables:
## $ name           : chr "Alabama" "Alaska" "Arizona" "Arkansas" ...
## $ geoid           : chr "USA-01" "USA-02" "USA-04" "USA-05" ...
## $ deaths_before   : chr "10790" "318" "17153" "5699" ...
## $ deaths          : chr "5590" "634" "6663" "3308" ...
## ..
```

*#Konverterer til numeric, som etter as\_tibble() blir double (numeric)*

```
as_tibble(NYT_cov_data)
```

```
## # A tibble: 51 x 6
##   name      geoid deaths_before deaths deaths_per_100k fully_vaccinate~
##   <chr>    <chr>   <chr>          <chr>   <chr>          <chr>
## 1 Alabama USA-01  10790         5590    13.84706814898~ 0.4728
## 2 Alaska  USA-02   318         634     10.52622170652~ 0.5602
## ..
```

```
NYT_cov_data$fully_vaccinated_pct_of_pop <-
  as.numeric(as.character(NYT_cov_data$fully_vaccinated_pct_of_pop))
NYT_cov_data$deaths_per_100k <-
```

```

  as.numeric(as.character(NYT_cov_data$deaths_per_100k))
NYT_cov_data$deaths_before <-
  as.numeric(as.character(NYT_cov_data$deaths_before))
NYT_cov_data$deaths <-
  as.numeric(as.character(NYT_cov_data$deaths))
str(NYT_cov_data)

## 'data.frame':   51 obs. of  6 variables:
##  $ name          : chr  "Alabama" "Alaska" "Arizona" "Arkansas" ...
##  $ geoid          : chr  "USA-01" "USA-02" "USA-04" "USA-05" ...
##  $ deaths_before  : num  10790 318 17153 5699 61034 ...
##  $ deaths         : num  5590 634 6663 3308 15117 ...
....

```

```
as_tibble(NYT_cov_data)
```

```

## # A tibble: 51 x 6
##   name          geoid deaths_before deaths deaths_per_100k fully_vaccinate~
##   <chr>         <chr>      <dbl>   <dbl>      <dbl>         <dbl>
## 1 Alabama      USA-01      10790   5590         13.8         0.473
## 2 Alaska       USA-02        318    634         10.5         0.560
....

```

*#Forsøkte å finne en måte å forkorte navnene på, men konkluderte med å bare endre  
#på de en etter en før ggplot.*

```

NYT_cov_data$name[NYT_cov_data$name == "Wyoming"] <- "Wyo."
NYT_cov_data$name[NYT_cov_data$name == "West Virginia"] <- "W.Vi."
NYT_cov_data$name[NYT_cov_data$name == "Montana"] <- "Mont."
NYT_cov_data$name[NYT_cov_data$name == "Idaho"] <- "Ida."
NYT_cov_data$name[NYT_cov_data$name == "Alabama"] <- "Ala."
NYT_cov_data$name[NYT_cov_data$name == "Mississippi"] <- "Miss."
NYT_cov_data$name[NYT_cov_data$name == "Louisiana"] <- "Lou."
NYT_cov_data$name[NYT_cov_data$name == "Tennessee"] <- "Tenn."
NYT_cov_data$name[NYT_cov_data$name == "Indiana"] <- "Ind."
NYT_cov_data$name[NYT_cov_data$name == "North Dakota"] <- "N.Da."
NYT_cov_data$name[NYT_cov_data$name == "North Carolina"] <- "N.Ca."
NYT_cov_data$name[NYT_cov_data$name == "Kansas"] <- "Kan."
NYT_cov_data$name[NYT_cov_data$name == "Hawaii"] <- "Haw."
NYT_cov_data$name[NYT_cov_data$name == "Illinois"] <- "Ill."
NYT_cov_data$name[NYT_cov_data$name == "Wisconsin"] <- "Wis."
NYT_cov_data$name[NYT_cov_data$name == "California"] <- "Calif."
NYT_cov_data$name[NYT_cov_data$name == "New Hampshire"] <- "N.H."
NYT_cov_data$name[NYT_cov_data$name == "Maryland"] <- "Mar."
NYT_cov_data$name[NYT_cov_data$name == "South Dakota"] <- "S.Da."
NYT_cov_data$name[NYT_cov_data$name == "Colorado"] <- "Col."
NYT_cov_data$name[NYT_cov_data$name == "Oregon"] <- "Ore."
NYT_cov_data$name[NYT_cov_data$name == "New Mexico"] <- "N.Me."
NYT_cov_data$name[NYT_cov_data$name == "Virginia"] <- "Vir."
NYT_cov_data$name[NYT_cov_data$name == "Maine"] <- "Mai."
NYT_cov_data$name[NYT_cov_data$name == "Florida"] <- "Flo."
NYT_cov_data$name[NYT_cov_data$name == "Oklahoma"] <- "Okl."
NYT_cov_data$name[NYT_cov_data$name == "Kentucky"] <- "Ken."
NYT_cov_data$name[NYT_cov_data$name == "Missouri"] <- "Miss."
NYT_cov_data$name[NYT_cov_data$name == "Michigan"] <- "Mich."
NYT_cov_data$name[NYT_cov_data$name == "Nevada"] <- "Nev."

```

```

NYT_cov_data$name[NYT_cov_data$name == "Arizona"] <- "Ari"
NYT_cov_data$name[NYT_cov_data$name == "Texas"] <- "Tex."
NYT_cov_data$name[NYT_cov_data$name == "Alaska"] <- "Ala."
NYT_cov_data$name[NYT_cov_data$name == "Minnesota"] <- "Min."
NYT_cov_data$name[NYT_cov_data$name == "Colorado"] <- "Col."
NYT_cov_data$name[NYT_cov_data$name == "Pennsylvania"] <- "Pen."
NYT_cov_data$name[NYT_cov_data$name == "Georgia"] <- "Geo."
NYT_cov_data$name[NYT_cov_data$name == "Arkansas"] <- "Ark."
NYT_cov_data$name[NYT_cov_data$name == "South Carolina"] <- "S.Ca."
NYT_cov_data$name[NYT_cov_data$name == "Washington"] <- "Was."
NYT_cov_data$name[NYT_cov_data$name == "Nebraska"] <- "Neb."
NYT_cov_data$name[NYT_cov_data$name == "Washington, D.C."] <- "W.D.C."
NYT_cov_data$name[NYT_cov_data$name == "New Jersey"] <- "N.J."
NYT_cov_data$name[NYT_cov_data$name == "New York"] <- "N.Y."
NYT_cov_data$name[NYT_cov_data$name == "Massachusetts"] <- "Mas."
NYT_cov_data$name[NYT_cov_data$name == "Connecticut"] <- "Conn."
NYT_cov_data$name[NYT_cov_data$name == "Rhode Island"] <- "R.Is."
NYT_cov_data$name[NYT_cov_data$name == "Vermont"] <- "Ver."

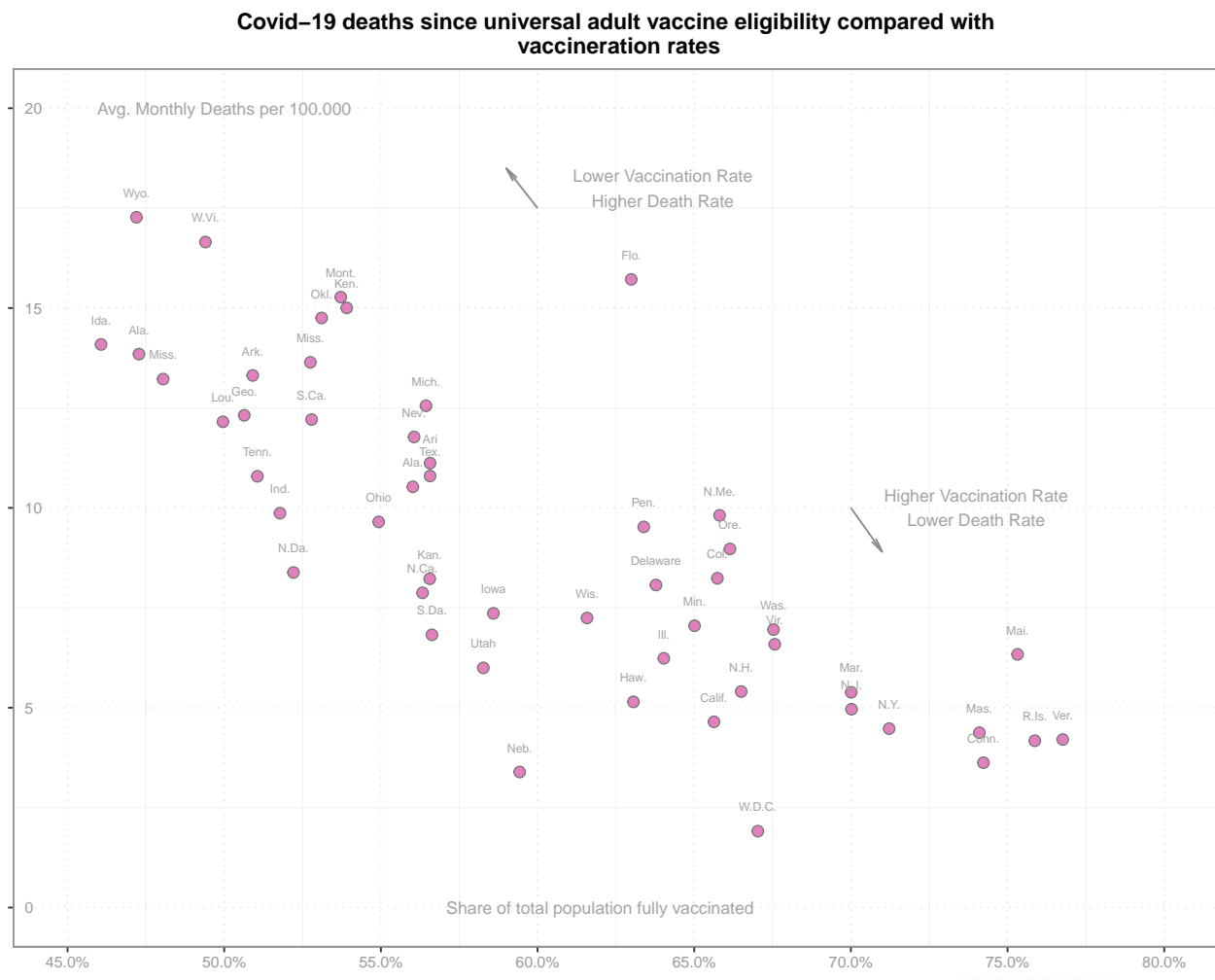
```

*#Fulgte NY-times sin graf, forsøkte å replisere den. Fikk med alt og tok meg noen friheter.*

```

p <- ggplot(NYT_cov_data, aes(x=fully_vaccinated_pct_of_pop,
                             y = deaths_per_100k, label = name))
p + geom_point(position = "identity", fill = "mediumvioletred", alpha= 0.60, color = "black", size = 2.5) +
  geom_text(vjust = 0, nudge_y = 0.5, size = 2.5, color = "gray60")+
  scale_x_continuous(labels = scales::percent, breaks=seq(0.45,1.00,0.05), limits = c(0.45, 0.80))+
  scale_y_continuous(name = "",limits = c(0, 20, 4))+
  labs(title = "Covid-19 deaths since universal adult vaccine eligibility compared with\n vaccination",
       x = "Share of total population fully vaccinated",
       y = "")+
  annotate(geom="text", x=0.50, y=20, label="Avg. Monthly Deaths per 100.000",
         color="gray60", size = 3.5)+
  annotate(geom = "segment", x = 0.60, xend = 0.59, y = 17.5, yend = 18.5, color ="gray50",
         arrow = arrow(angle = 10, length = unit(.3,"cm")))+
  annotate(geom="text", x=0.64, y=18, label="Lower Vaccination Rate\nHigher Death Rate",
         color="gray60", size = 3.5)+
  annotate(geom = "segment", x = 0.70, xend = 0.71, y = 10, yend = 8.9, color ="gray50",
         arrow = arrow(angle = 10, length = unit(.3,"cm")))+
  annotate(geom="text", x=0.74, y=10, label="Higher Vaccination Rate\nLower Death Rate",
         color="gray60", size = 3.5)+
  annotate(geom="text", x=0.62, y=-0.0, label="Share of total population fully vaccinated",
         color="gray60", size = 3.5)+
  theme(plot.title = element_text(face = "bold",hjust = 0.5),
        axis.title.x = element_text(hjust = 0.9, size =5, color = "gray98"),
        axis.text = element_text(colour = "gray60"),
        axis.text.y = element_text(hjust = 0, margin = margin(0, -0.7, 0, 0, 'cm')),
        panel.background = element_rect(fill = "white", colour = "grey85"),
        panel.grid.major = element_line(colour = "gray92", linetype = "dotted"),
        panel.grid.minor = element_line(colour = "gray94", linetype = NULL),
        panel.border = element_rect(fill = alpha("white", 0.10),color = "gray60", size = 1))

```



## OPPGAVE2

*#Gitt metode for lm()*

```
lm(deaths_per_100k ~ fully_vaccinated_pct_of_pop, data = NYT_cov_data)
```

```
##
```

```
## Call:
```

```
## lm(formula = deaths_per_100k ~ fully_vaccinated_pct_of_pop, data = NYT_cov_data)
```

```
##
```

```
## Coefficients:
```

```
##              (Intercept)  fully_vaccinated_pct_of_pop
```

```
##              31.15              -36.66
```

Jeg tolker dette som at ca. 63.34% av befolkningen er full-vaksinert, altså er -36.66 % ikke full-vaksinert. Det er veldig tydelig at med høyere vaksinasjons rate er det lavere dødstall pr. 100K. Basert på disse data, tolker jeg det som at tallene "31.15" er antall døde pr 100K statene samlet. Altså, er tallet på antall døde i dette tilfelle ca. 31 150 per 100K den perioden.

*#La til geom\_smooth(method = lm) i ggplot fra oppg1.*

```
p <- ggplot(NYT_cov_data, aes(x=fully_vaccinated_pct_of_pop,  
                             y = deaths_per_100k, label = name))
```

```
p + geom_point(position = "identity", fill = "mediumvioletred", alpha= 0.60, color = "black", size = 2.5,  
              geom_text(vjust = 0, nudge_y = 0.5, size = 2.5, color = "gray60")+  
              geom_smooth(method = lm, lwd = 0.6, color = alpha("lightseagreen", 0.70), fill = alpha("lightseagreen", 0.70)))
```

```

scale_x_continuous(labels = scales::percent, breaks=seq(0.45,1.00,0.05), limits = c(0.45, 0.80))+
scale_y_continuous(name = "",limits = c(0, 20, 4))+
labs(title = "Covid-19 deaths since universal adult vaccine eligibility compared with\n vaccination",
      x = "Share of total population fully vaccinated",
      y = "")+
annotate(geom="text", x=0.50, y=20, label="Avg. Monthly Deaths per 100.000",
         color="gray60", size = 3.5)+
annotate(geom = "segment", x = 0.60, xend = 0.59, y = 17.5, yend = 18.5, color ="gray50",
         arrow = arrow(angle = 10, length = unit(.3,"cm")))+
annotate(geom="text", x=0.64, y=18, label="Lower Vaccination Rate\nHigher Death Rate",
         color="gray60", size = 3.5)+
annotate(geom = "segment", x = 0.70, xend = 0.71, y = 10, yend = 8.9, color ="gray50",
         arrow = arrow(angle = 10, length = unit(.3,"cm")))+
annotate(geom="text", x=0.74, y=10, label="Higher Vaccination Rate\nLower Death Rate",
         color="gray60", size = 3.5)+
annotate(geom="text", x=0.62, y=-0.0, label="Share of total population fully vaccinated",
         color="gray60", size = 3.5)+
theme(plot.title = element_text(face = "bold",hjust = 0.5),
      axis.title.x = element_text(hjust = 0.9, size =5, color = "gray98"),
      axis.text = element_text(colour = "gray60"),
      axis.text.y = element_text(hjust = 0, margin = margin(0, -0.7, 0, 0, 'cm')),
      panel.background = element_rect(fill = "white", colour = "grey85"),
      panel.grid.major = element_line(colour = "gray92", linetype = "dotted"),
      panel.grid.minor = element_line(colour = "gray94", linetype = NULL),
      panel.border = element_rect(fill = alpha("white", 0.10),color = "gray60", size = 1))

## `geom_smooth()`` using formula 'y ~ x'

```

**Covid-19 deaths since universal adult vaccine eligibility compared with vaccination rates**

