

SOK-1004, høst 2022, Mappeoppgave 1

246

Oppgave 1.

- A) Hva menes med inntektsulikhet i et land? Inntektsulikheter er forskjellen mellom inntekten til befolkningen i et land. Inntektsulikhet er hvordan landets inntekter og ressurser er fordelt for befolkningen.
- B) Hva er mulige grunner for at inntekten er fordelt ulikt i samfunnet? Globalisering er en drivende faktor for hvorfor inntekten i et samfunn er ulikt. Det kan forklares med at produsenter flytter virksomheten sin til et annet land der varer og tjenester kan produseres billigere. En annen driver er teknologisk utvikling, ny teknologi gjør det mulig å automatisere oppgaver slik at arbeid som tidligere er blitt gjort av mennesker kan nå bli gjort av maskiner. Det kan føre til økt arbeidsledighet og flere som står uten inntekt. Fordelingspolitikk i et land vil også påvirke også inntektsulikheter. Velferdssystemer, skatt, pensjon og trygd er med på å redusere ulikheten.
- C) Hvilke konsekvens kan ulikhet i inntektsfordeling føre til? Ulikhet i inntektsfordeling kan føre til et splittet samfunn og redusert tillit, både mellom mennesker og myndigheten. Dersom det er høy tillit i et samfunn kreves det mindre ressurser til regulering og kontroll og det er lettere å komme til de beste løsningene i forhandlinger. Dersom det er mistillit vil det bidra til økt polarisering der det blir vanskeligere å komme til enigheter. Ved mindre tillit i et samfunn kan det føre til at man blir mindre åpen for endringer, og at man henter informasjon fra ulike kilder som støtter opp mot egen side av saken. Dersom et samfunn har tillit til hverandre påvirker det den økonomiske posisjonen til et samfunn på en positiv måte. Når aktørene stoler på at avtaler blir holdt og at lovverket beskytter dem dersom avtaler blir misligholdt, er det desto større sjanse for at bedrifter inngår avtaler med hverandre, som vil bidra til en økonomisk vekst.

Oppgave 2.

- A) Forklar hvordan Gini-koeffisientene beregnes og hvordan de bør tolkes. Hva måler P90/P10 og S80/S20? Gini-koeffisienten måler inntektsulikheter innenfor et land, man ser på hvor mye inntekt og formue det er totalt i landet for så og deretter hvordan det er fordelt. Dersom et land har en Gini-koeffisient på 0 betyr det at alle i landet har lik inntekt og formue, mens dersom Gini-koeffisienten er på 1 vil det bety at en person står med all inntekten og formuen. P90/P10 sorterer befolkningen etter størrelse på inntekt etter skatt, for deretter å dele befolkningen opp i ti like store grupper. P90/P10 tar den personen som er rett over 90% og den som er rett under 10% på inntektsfordelingen og sammenligner de mot hverandre. Ulikheten kommer til å være stor her, men sammenlignet med de andre målene vil den vise en jevnere utvikling over tid. S80/S20 måler gjennomsnittsinntekten til de 20% rikeste og 20% fattigste i landet. Forskjellen på S80/S20 og P90/P10 er at man tar med de aller rikeste og aller fattigste i landet, slik vil gapet på gjennomsnittslønnen vil være enda større.
- B) Diskuter svakheter og styrker mellom disse målene på ulikheter. I hvilke situasjon er hvert mål mest relevant? Formålet til Gini-koeffisienten, P90/P10 og S80/S20 er å måle inntektsulikhetene mellom de rike og fattige og hvordan inntektene er fordelt for befolkningen. Gini-koeffisienten og S80/S20 kan være ganske lik, men har klare svakheter dersom det skulle bli brukt til å finne ut hvor mye den gjennomsnittlige personen i Norge tjener siden de ikke tar høyde for at de aller rikeste og aller

fattigste. Forskjellen mellom de rikeste og fattigste i Norge kommer til å være såpass stor at det gir stort utslag for gjennomsnittet og derfor vil målingen ikke ha samsvar med virkeligheten. P90/P10 derimot tar høyde for nettopp det og måler kun den personen som er rett over 90% og rett under 10% av inntektsfordelingen i Norge, det vil gi et mer relevant bilde på hva gjennomsnittet ligger på. Dersom man skulle måle utviklingen i ulikheter ved inntekt til befolkningen vil det være mer relevant å bruke Gini-koeffisienten eller S80/S20 da de tar høyde for de rikeste og fattigste i landet. Det vil gi et klarere bilde på hvor stor forskjellen fra de med minst og størst inntekt er.

```
rm(list=ls())
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr  0.3.4
## v tibble  3.1.8      v dplyr  1.0.9
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.1
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
#Laster inn datasett
```

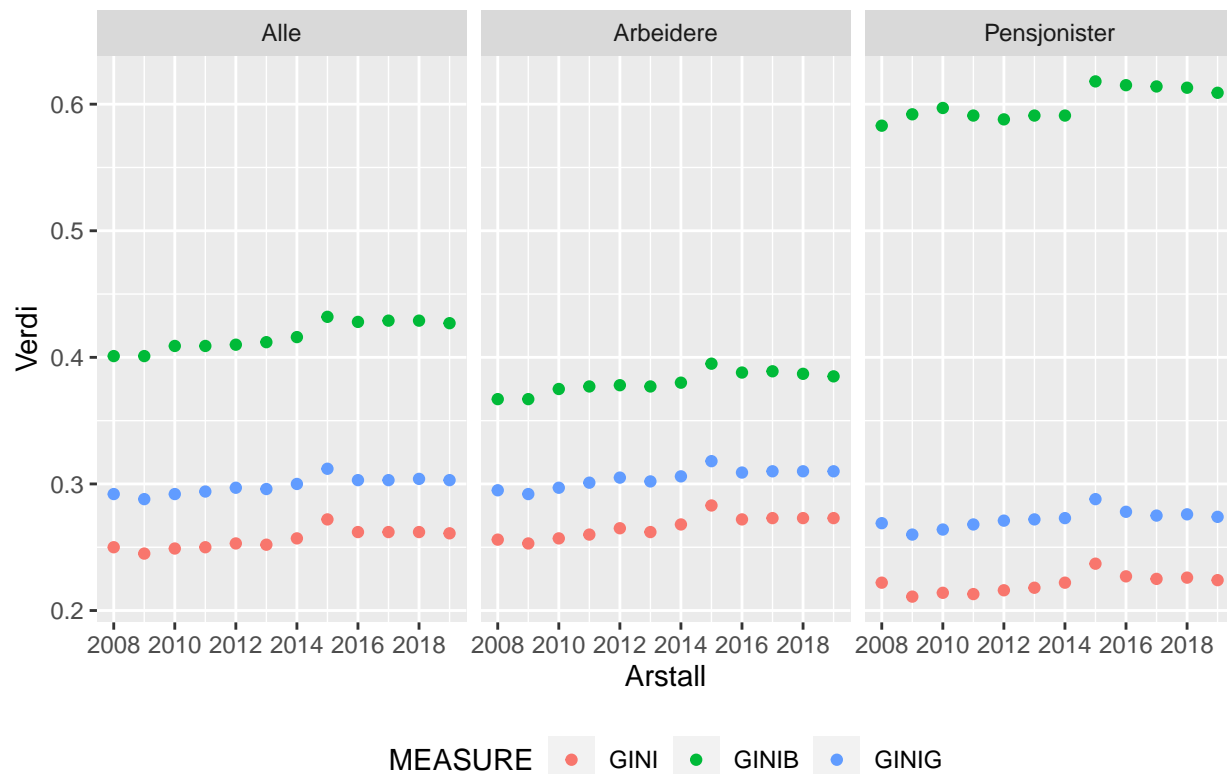
```
url <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/main/data/oecd_data.csv"
```

```
oecd_n <- url %>%
  read_csv()
```

```
## Rows: 216 Columns: 7
## -- Column specification -----
## Delimiter: ","
## chr (5): MEASURE, Measure, AGE, Age group, Unit
## dbl (2): Year, Value
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
oecd_n %>%
  filter(Unit == "0-1 scale") %>% #Velger hva jeg vil male
  ggplot(aes(x = Year, y=Value, color=MEASURE)) + #Definerer aksene
  geom_point() + #Velger scatterplot
  scale_x_continuous(breaks=c(2008, 2010, 2012, 2014, 2016, 2018)) + #Velger årstall i grafen
  labs(title = "Gini-koeffisient fra 2008-2019", y = "Verdi", x = "Årstall") + #Lager titler
  facet_wrap(~ factor(AGE, (levels=c("TOT", "WA", "OLD"))),
    labeller = as_labeller(
      c("TOT" = "Alle",
        "WA" = "Arbeidere",
        "OLD" = "Pensjonister"))) + #Lager titler på alle grafene i riktig rekkefølge
  theme(legend.position = "bottom")
```

Gini-koeffisient fra 2008–2019



```
#Grønn strek måler markedsinntekt før skatt og overføringer.
#Blå strek måler brutto inntekt før skatt
#Rød strek måler disponibel inntekt etter skatt
```

Gini-koeffisienten viser at det har generel økning i ulikheter fra 2008-2019 med den disponible inntekten. Man ser at det enkelte år har vært oppgang og nedgang i forhold til året før men at det totalt sett har blitt mer ulikheter over årene totalt sett. Tar man pensjonister alene har de den samme Gini-koeffisienten som de hadde i 2008.

I grafen har jeg målt tre forskjellige Gini-koeffisienter. Den grønne grafen måler den totale markeds inntekten til befolkningen før man har trukket fra skatt og andre overføringer. Den blå grafen måler brutto inntekt, altså før skatt er trukket fra, mens den røde grafen måler netto inntekt.

Som man kan se på panelet er det forskjell mellom arbeiderne og pensjonister når det kommer til Gini-koeffisienten, det er større ulikheter hos arbeiderne enn hos pensjonistene.

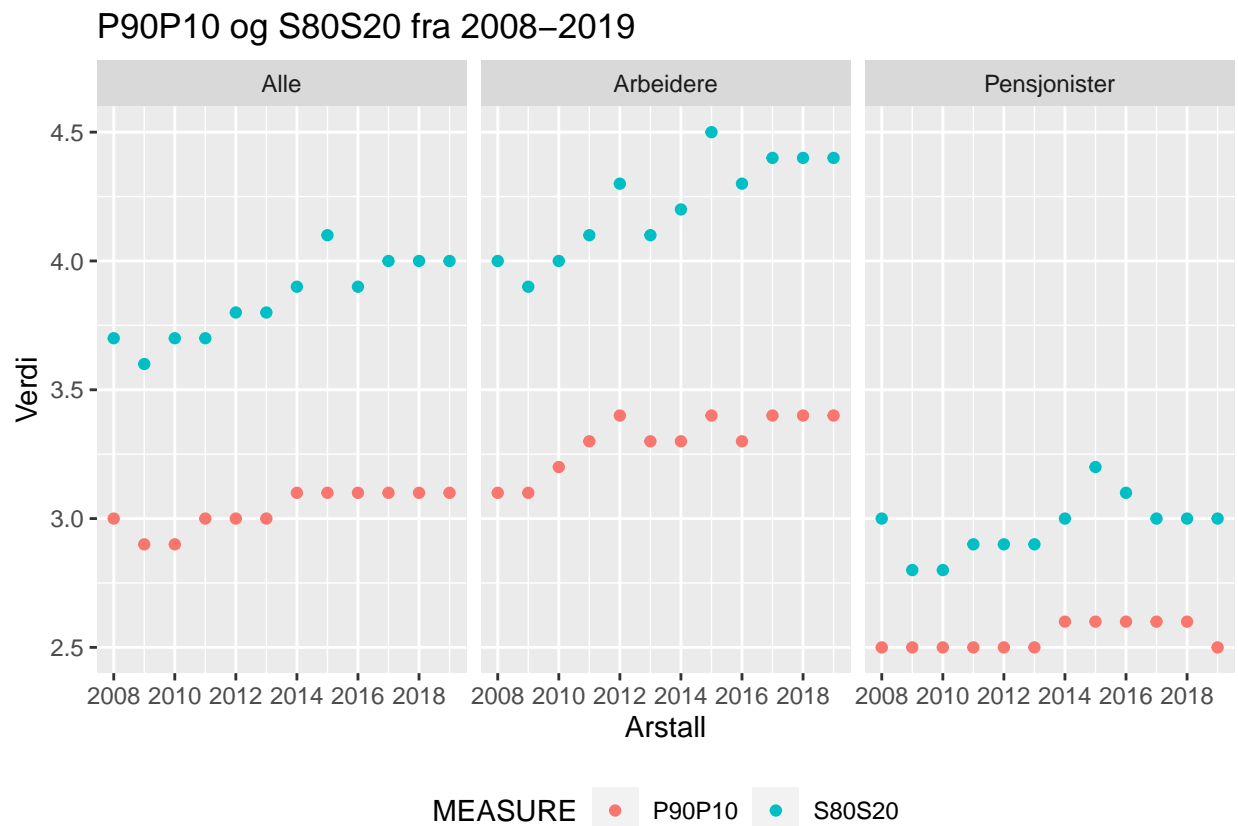
Noen pensjonister har store formuer før skatt og overføringer er trukket i fra, noe som gjør at det er store ulikheter for pensjonister, men når både pensjon er kommet inn og formueskatt og vanlig skatt er trukket i fra kan man se at det utjevner seg slik at ulikheten ikke er så stor lenger. For arbeiderne sin del er ikke ulikheten like stor. Samfunnets institusjoner slik som trygd og pensjon påvirker Gini-koeffisienten slik at ulikhetene blir mindre blant befolkningen, uten statlig hjelp ville differansen vært mye høyere.

```
#Laster inn datasett
url <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/main/data/oecd_da

oecd_n <- url %>%
  read_csv()
```

```
## Rows: 216 Columns: 7
## -- Column specification -----
## Delimiter: ","
## chr (5): MEASURE, Measure, AGE, Age group, Unit
## dbl (2): Year, Value
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

oecd_n %>%
  filter(MEASURE == "P90P10" | MEASURE == "S80S20") %>% #Velger hva jeg vil måle
  ggplot(aes(x = Year, y=Value, color=MEASURE)) +
  geom_point() + #Velger scatterplot
  scale_x_continuous(breaks=c(2008, 2010, 2012, 2014, 2016, 2018)) + #Velger årstall til panelene
  labs(title = "P90P10 og S80S20 fra 2008-2019", y = "Verdi", x = "Årstall") + #Lager titler
  facet_wrap(~ factor(AGE, (levels=c("TOT", "WA", "OLD"))),
    labeller = as_labeller(
      c("TOT" = "Alle",
        "WA" = "Arbeidere",
        "OLD" = "Pensjonister"))) + #Gir navn på alle panelene
  theme(legend.position = "bottom")
```



I denne grafer måler jeg inntekts ulikhetene for arbeidere og pensjonister i årene 2008-2019 med bruk av P90P10 og S80S20. Grafene er henholdsvis ganske lik, men ved bruk av S80S20 så er differansen en del større enn ved bruk av P90P10. Ved å bruke P90P10 får man over tid en mer konkret beskrivelse over hvor stor inntektsforskjellen er i da man ikke tar i betraktning de rikeste og fattigste i landet. Det er en del likhet

i den her grafen som den jeg lagde ved bruk av Gini-koeffisienten, men det er høyere verdier.

```
#Laster inn datasett
url <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/main/data/skatt_t_

skatt_troms_2015 <- url %>%
  url() %>%
  gzcon() %>%
  readRDS()

#Lager en ny variabel
tromso <- filter(skatt_troms_2015, kommnr == "1902") %>%
  mutate(disponibel_inntekt=inntekt - skatt) %>%
  mutate(disponibel_inntekt=lag(disponibel_inntekt/10^3)) %>%
  mutate(inntekt=lag(inntekt/10^3))

tromso[tromso < 0] <- 0

tromso_long <- tromso %>%
  pivot_longer(c("inntekt", "disponibel_inntekt"), names_to = "var", values_to = "inntekt")

#Lager grafen
tromso_long %>%
  ggplot(aes(x = aldersgruppe, y = inntekt, color = var)) +
  geom_point() +
  geom_boxplot(alpha = 0.2) +
  scale_y_log10(labels = scales::comma) +
  labs(title = "Inntektsfordeling i Tromsø",
        subtitles = "Inntekt i Tromsø for og etter skatt",
        x = "Alder",
        y = "Inntekt") +
  theme_bw()
```

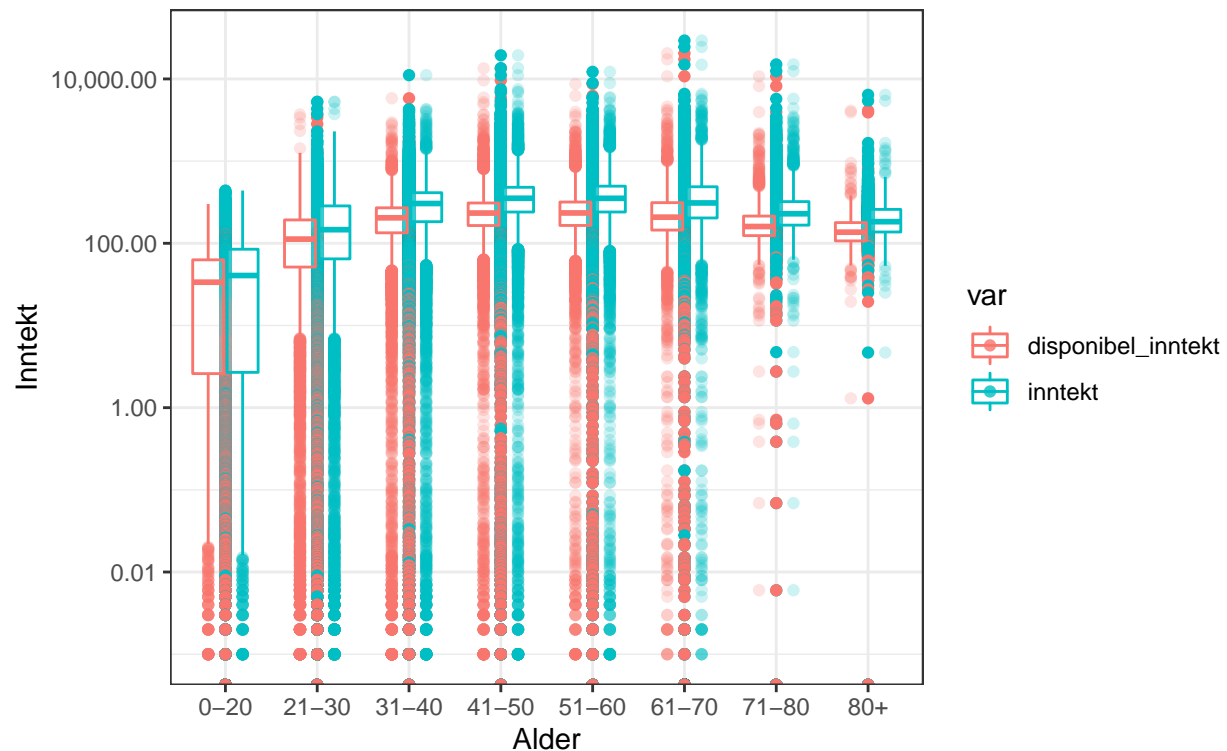
```
## Warning: Transformation introduced infinite values in continuous y-axis
## Transformation introduced infinite values in continuous y-axis
```

```
## Warning: Removed 9498 rows containing non-finite values (stat_boxplot).
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```

Inntektsfordeling i Tromsø

Inntekt i Tromsø for og etter skatt



Grafen viser inntektsulikheter i Tromsø kommune før og etter skatt til de følgende aldersgruppene. Ved bruk av boxplot vises median i tillegg. Her kan man se at det er enkelte som trekker gjennomsnittet opp og noen som trekker ned.

```
rm(list=ls())
library(tidyverse)

#URL
url <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/main/data/skatt_tromso_2015.csv"

#Laster inn datasett
skatt_troms_2015 <- url %>%
  url() %>%
  gzcon() %>%
  readRDS()

#Filtrerer Tromsø kommune
tromso <- filter(skatt_troms_2015, kommnr == "1902")

#Lager ny variabel
tromso <- tromso %>%
  mutate(inntekt_etter_skatt=inntekt-skatt)

tromso[tromso < 0] <- 0

#muterer slik at inntekt blir vises i tusen
tromso <- tromso %>%
```

```

mutate(inntekt_etter_skatt=inntekt_etter_skatt/1000)

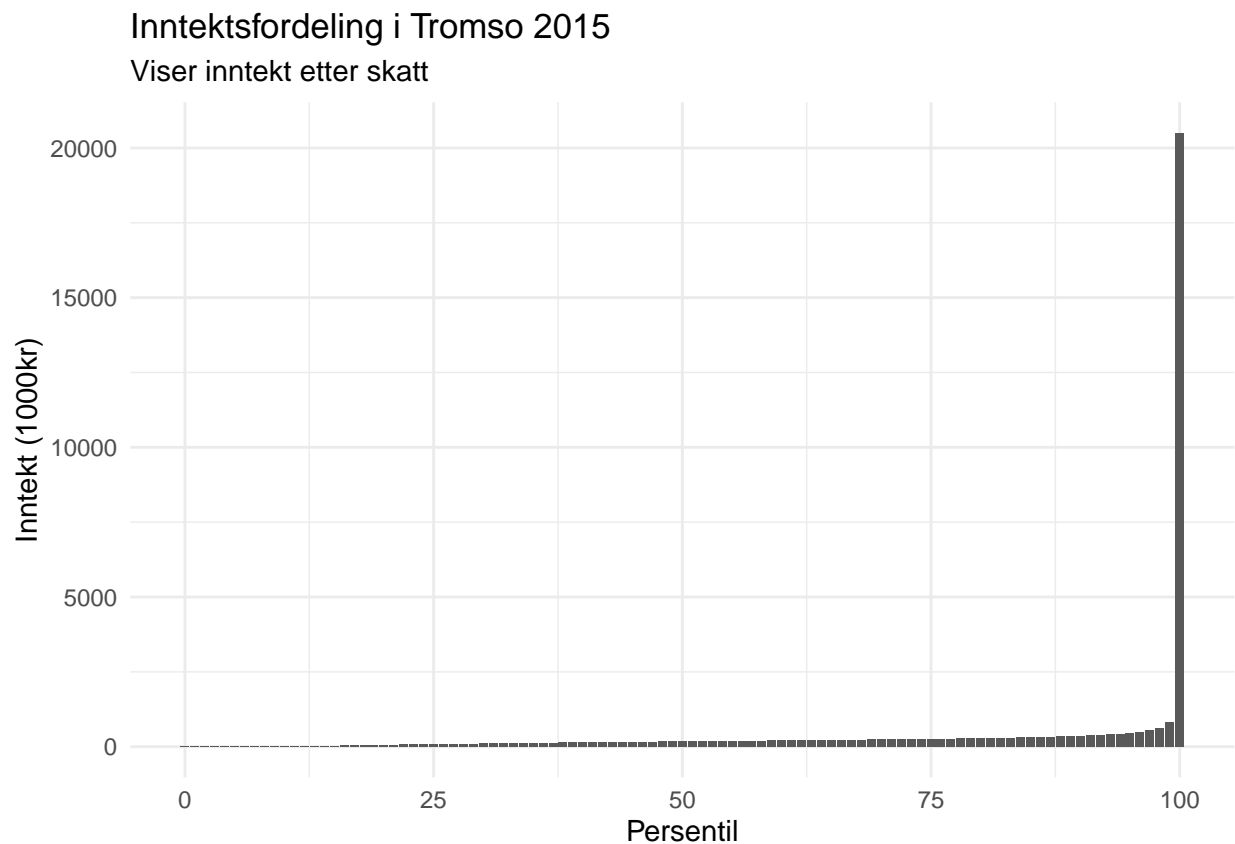
df <- tromso %>%
  select(inntekt_etter_skatt) %>%
  quantile(probs = seq(0,1, .01), na.rm=TRUE)

#Gjør det om til en tibble
df <- as_tibble(df)

#lager en rad med persentiler
df <- df %>%
  add_column(Persentil=c(0:100))

#Plotter grafen
df %>%
  ggplot(aes(x=Persentil,y=value)) +
  geom_col()+
  labs (title = "Inntektsfordeling i Tromso 2015",
        subtitle = "Viser inntekt etter skatt",
        y = "Inntekt (1000kr)", x= "Persentil")+
  theme_minimal()

```



Grafen over viser inntektsfordelingen i Tromsø etter skatt i 2015. Man ser en klar likhet mellom denne grafen og den før skatt, der den siste persentilen er langt høyere enn resten. Sammenligner vi de to grafene finner vi ut at skatten i Norge ikke gjør så mye for å jevne ut inntektsfordelingen.

```

#Laster inn pakker
rm(list=ls())
library(tidyverse)
library(gglorenz)
library(patchwork)

#URL
url <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/main/data/skatt_t.

#laster inn datasett
skatt_troms_2015 <- url %>%
  url() %>%
  gzcon() %>%
  readRDS()

#Filtrerer Tromsø kommune
tromso <- filter(skatt_troms_2015, kommnr == "1902")

#Muterer slik at ny variabler vises med inntekt etter skatt
tromso <- tromso %>%
  mutate(disponibel_inntekt = inntekt - skatt)

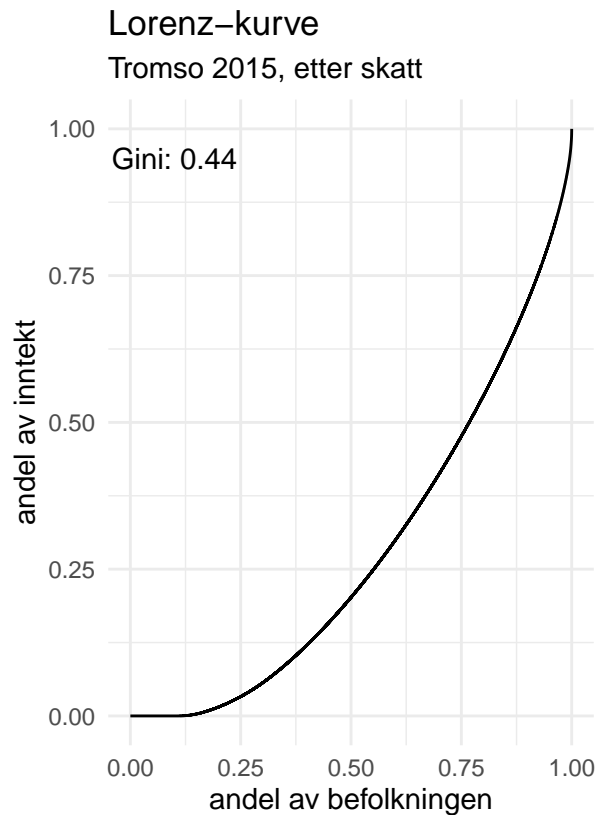
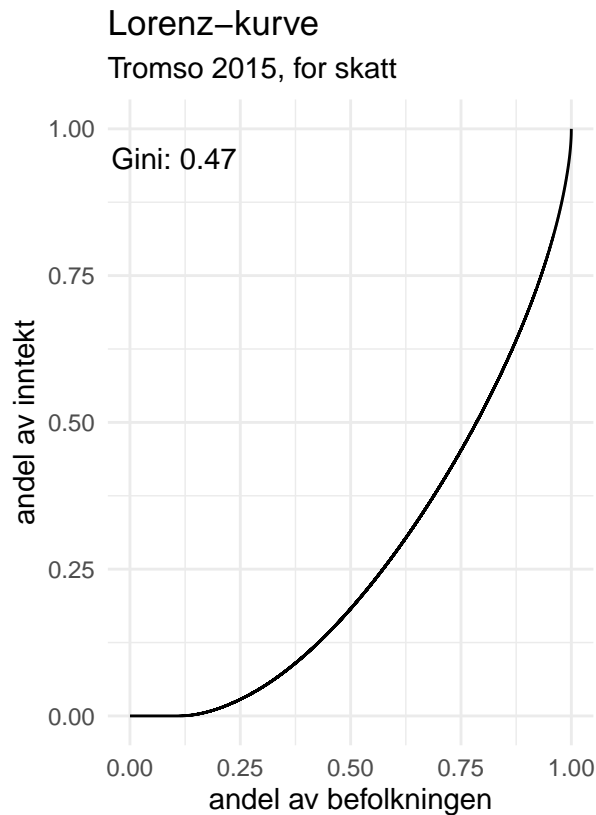
#Gjør negative verdier om til 0
tromso[tromso < 0] <- 0

#Plotter graf om inntekt før skatt
lorenzkurve_for_skatt <- tromso %>%
  ggplot(aes(inntekt)) +
    stat_lorenz() + # tegner Lorenz
    annotate_ineq(tromso$inntekt) + # beregner Gini
    labs(x="andel av befolkningen", #Lager titler
         y= "andel av inntekt",
         title = "Lorenz-kurve",
         subtitle = "Tromso 2015, for skatt") +
    theme_minimal()

#Plotter graf om inntekt etter skatt
lorenzkurve_etter_skatt <- tromso %>%
  ggplot(aes(disponibel_inntekt)) +
    stat_lorenz() + # tegner Lorenz
    annotate_ineq(tromso$disponibel_inntekt) + # beregner Gini
    labs(x="andel av befolkningen", #Lager titler
         y= "andel av inntekt",
         title = "Lorenz-kurve",
         subtitle = "Tromso 2015, etter skatt") +
    theme_minimal()

lorenzkurve_for_skatt + lorenzkurve_etter_skatt

```

Hvis vi tar utgangspunkt fra oppgave 3a så ser vi at Gini-koeffisienten til landet ligger på 0.26 mot Tromsø sin Gini-koeffisient på 0.44 (etter skatt), det vil si at Tromsø har en større ulikhet i inntekt i forhold til resten av landet. Slik vi så tidligere på oppgave 4B er det tydelig enkelte i Tromsø som har svært mye høyere inntekt enn resten, det vil bidra til at Gini-koeffisienten blir større og ulikheten mellom Tromsø og resten av landet forverres.