

“SOK-1004, høst 2022, Mappeoppgave 1”

“277”

Forord

Alle oppgave tekstene er skrevet i vanlig skrift og er hentet fra oppgaven (kilder). Alle mine egne ord og svar på oppgavene er skrevet i kursiv eller som kommentarer i kodene.

Oppgave 1. Årsaker og konsekvenser av ulikhet

Oppgave 1A.

Hva menes med inntektsulikhet i et land?

Inntektsulikheten i et land er ett mål på hvor stor forskjell i inntekt det er mellom de rikeste og de fattigste i ett gitt land. Man kan også måle inntektsulikhet mellom forskjellige grupper i landet, f.eks. arbeiderklassen mot middelklassen, bosted, alder, osv. Inntektsulikhet måles kun i inntekter som f.eks. lønn, avkastning på investeringer, overføringer og ikke formue som f.eks. verdipapirer og eiendom.

Oppgave 1B.

Hva er mulige grunner for at inntekten er fordelt ulikt i samfunnet?

Den mest grunnleggende grunnen, som de andre grunnene bygger på til at inntekten er ulikt fordelt er fordi folk er forskjellige, med ulik kompetanse og muligheter. I mange land er det så grunnleggende at inntektene blir ulikt fordelt ut ifra hvilke kjønn man er. To andre grunner til at inntektene er ulikt fordelt i ett samfunn kan være globalisering og teknologi. Jeg skal nå kort gå gjennom disse to grunnene til at de kan skape ulikhet i inntekt.

Globalisering: Økningen i globaliseringen de siste tiårene har gjort det slik at produksjonen av varer og tjenester blir flyttet dit det er billigst. Som en konsekvens av dette er det ikke behov for arbeiderne som en gang produserte disse varene innad i landet. Samtidig har de som sitter som eiere tjent mer penger på denne overgangen. Dersom vi ser på dette som ett lukket samfunn ville arbeiderne i verste fall mistet jobben og eierne tjent mer. Her kan vi se en klar

inntekts ulikhet. Vi kan også se på det slik at selskaper flytter produksjonen sin til fattigere land og skaper jobber i disse landene. Det øker da inntekten til de fattigste i det aktuelle landet. Da kan vi se at de fattigste får det litt bedre mens de rike blir rikere.

Arbeidsinnvandringen til de rike landene har også økt. Dette vil si at arbeidere med samme kompetanse kommer til ett land og tar jobbene til de tidligere arbeiderne for lavere lønn. Dette igjen øker inntektene til eierne, og derav økt inntektsulikheten. Det er ikke bare arbeidsinnvandringen som spiller inn på inntektsulikhetene. Generell innvandring fra f.eks. krig og dårlig levevilkår, kan også skape inntektsulikheter hvis de som kommer til et land ikke klarer å integrere seg og komme på arbeidsmarkedet.

Teknologi: Teknologien har gjort store framskritt de siste 20 årene, som en konsekvens av dette har en stor andel av jobbene med lavt kompetansenivå blitt automatisert. Dette fører til at de som en gang var arbeidere i fabrikker ble byttet ut med maskiner som kun krevde regelmessig vedlikehold og kunne jobbe døgnet rundt. Igjen kan vi se at det går ut over arbeiderne og de med lavest inntekt i samfunnet. Samtidig som arbeiderne blir byttet ut, skaper det nye jobber for høyt utdannende og kompetente mennesker som kan drifte disse maskinen. Vi kan se at IT og finanssektoren har hatt en betydelig inntektsvekst av dette (NHO, 2018).

Teknologi er ikke bare maskiner og roboter. Vi kan også se at utviklingen og verdiskapningen av apper har hatt en betydelig vekst. Dette betyr at f.eks. ett selskap som utvikler en app kun trenger høyt utdannende og kompetente mennesker på ett kontor for å kunne levere til forbrukeren. Selskapet krever ikke f.eks. lastebilsjåfører, fabrikk arbeidere, håndverkere, osv.

Hvis vi ser på de samlede konsekvensene til globalisering og teknologi kan vi se at i moderniserte land blir det mindre lav kompetanse jobber og mer høy kompetanse jobber på grunn av at fabrikk jobbene blir automatisert eller flyttet til andre land. Mens i mindre teknologiske land blir det flere lav kompetanse jobber fordi selskapene flytter fabrikk produksjonen til fattigere land med lavere lønninger. Uten å trekke noen konklusjoner kan vi si at utviklingen går mest utover middelklassen på ett globalt nivå? Dette er bare to av mange mulige grunner til at det er inntektsulikhet i et land.

Oppgave 1C.

Hvilke konsekvenser kan ulikhet i inntektsfordelingen føre til?

Først skal jeg skrive litt om hva ulikhet i inntektsfordelingen har å si for enkeltpersonen, også skal jeg skrive litt om hva det har å si for samfunnet man lever i.

For enkeltpersonen har det mye å si hvilken side av inntektsfordelingen man er på. En person med lav inntekt har lavere levestandard og kortere liv. De vil heller ikke ha muligheten til å skape det livet de ønsker og drømmer om. En person med høy inntekt har høyere levestandard og lever lengere. Disse har også større frihet i hva de ønsker å gjøre med livet sitt. Som vi så i oppgave 1b var det de med lav utdanning og kompetansenivå som er taperne i dette. (NHO, 2018)

En PISA-undersøkelse (NHO, 2018) viser at skoleresultat kan avhenge av inntektsfordelingen. Barna med foreldre som hadde lav inntekt gjorde det generelt sett dårligere på skolen og barn med foreldre som hadde en høy inntekt gjorde det generelt sett bedre på skolen. Som ett resultat av dette kan vi si at inntektsulikheten går i arv.

For samfunnet i ett land har også inntektsfordelingen stor betydning. IMF (2016) gjennomførte en undersøkelse der de så på inntektsulikhet og tillit (styresmakten i landet og folk flest). Der så de at mennesker i samfunn med stor ulikhet har mindre tillit til hverandre og myndighetene sammenlignet med samfunn med høyere likhet. Lavere tillit kan bremse opp økonomisk vekst i ett land, på grunn av at aktører ikke stoler på hverandre eller styringsmakten hvis det skal skje noe med en avtale. Lav tillit mellom folk og styringsmaktene kan også gjøre at landet blir oppfattet som ustabilt og aktører ikke tørr å investere i landet.

Oppgave 2. Måling av ulikhet

I datasettet fra OECD finner man 5 mål på inntektsulikhet: tre ulike Gini-koeffisienter, P90P10 og S80S20.

Oppgave 2A.

Forklar hvordan Gini-koeffisientene beregnes og hvordan de bør tolkes. Hva måler P90P10 og S80S20?

Gini-koeffisientene er ett mål på inntektsulikheter i ett land der man ser på hvor mye inntekt det totalt sett er i ett land og deretter hvordan det er fordelt i befolkningen. Resultatet ligger på ett sted mellom 0 og 1, der 0 (perfekt likhet) er at inntektene er helt likt fordelt og 1 (perfekt ulikhet) er at inntektene er hos en enkelt person. Gini indeksen er forskjellen mellom perfekt likhet og Lorenz kurven som viser hvordan inntektene faktisk er fordelt i ett samfunn.

I datasettet er det tre ulike Gini mål. GINI tar utgangspunkt i inntekt etter skatt og overføringer (netto inntekt). GINIB tar utgangspunkt i markedsinntekt før skatt. GINIG tar utgangspunkt i inntekt før skatt og overføringer (brutto inntekt).

P90P10 er et mål der man fordeler befolkningen etter størrelse på inntekt etter skatt inn i 10 “desiler”. I denne fordelingen utelukkes de 10 prosentene med høyest og lavest inntekt. Dette er fordi de rikeste ikke skal trekke snittet drastisk opp og de fattigste trekke snittet drastisk ned. Dette gjør at denne måten å måle ulikhet på viser en jevnere utvikling over tid. Det vil si at personer i “desilene” 1 og 2 (P10) og personer i “desilene” 9 og 10 (P90) ikke er med i utregningen. Som ett eksempel kan vi si at hvis P90P10 forholdstallet kommer ut som 2 betyr det at befolkningen med høyere inntekt enn 90 prosent av resten har 2 ganger så høy inntekt som befolkningen med lavere inntekt enn 90 prosent av resten.

S80S20 er et mål som viser forholdstallet mellom gjennomsnittsinntektene til befolkningens 20 prosent rikeste og gjennomsnittsinntekten til befolkningens 20 prosent fattigste etter skatt.

Denne måten å måle på tar med seg de rikeste og fattigste i befolkningen til motsetning fra P90P10 målingen. Her skjer da altså det motsatte at de rikeste trekker snittet drastisk opp mens de fattigste trekker snittet drastisk ned. Som ett eksempel kan vi si at hvis forholdstallet til S80S20 er 4 betyr det at de 20% med høyest inntekt i gjennomsnitt har 4 ganger høyere inntekt enn de 20% med lavest inntekt.

Oppgave 2B.

Diskuter svakheter og styrker med disse målene på ulikhet. I hvilke situasjoner er hvert mål mest relevant?

Alle de tre forskjellige Gini målene i datasettet måler på lik måte, men måler med ulike verdier. Når man skal måle med Gini-koeffisienter må man måle med de verdiene som er mest relevant for der man vil se hvor stor ulikhet det er. Gini målene er en forenkling av virkeligheten så det vil aldri være ett “bra” mål av ulikhet, men det gir oss muligheten til å sammenligne land siden vi bruker samme variabler. Gini fungerer ikke like bra til å finne ut hvor mye rikere en del av befolkningen er i forhold til en annen del av befolkningen.

P90P10 er ett godt mål når man vil se på ulikheter innad i ett land over lengere tid. Det gir deg også en indikasjon på hvor mye rikere de rikeste er i forhold til de fattigste. Hensikten med P90P10 er å utelukke de største variasjonene i befolkningens inntekt og få ett stabilt mål på ulikhet. Dette gir oss også en pekepinn på ulikhetene mellom “den normale person i gata”.

S80S20 har en tendens til å være veldig lik Gini-koeffisienter (SSB, 2019). Men her får man også en indikasjon som man kan bruke til å se hvor mye rikere de rikeste er i forhold til de fattige. S80S20 passer bedre hvis man ønsker å se innad i ett land på en kortere periode eller ett spesielt år siden man får en gjennomsnittsverdi (rikeste / fattigste). Dette er fordi det utelukker den “midterste” delen av befolkningen, og inkludert når f.eks. de rikeste tar ut utbytte.

Oppgave 3. Inntektsulikhet i Norge 2008-2019

Her skal det brukes et datasett fra OECD som gir ulike mål på ulikhet (Variabel: MEASURE eller Measure), for årene 2008-2019 (Year), for hele befolkningen (AGE=TOT), gruppa som er utgjør arbeidsstyrken (AGE=WA), samt de som er pensjonert (AGE=OLD). Datasettet inneholder også antall personer i hver av disse gruppene. Kolonne “Value” angir alle verdiene, og “Unit” forteller hvordan disse er målt.

Slett minne, last inn tidyverse.

```
# Sletter minnet og laster inn nødvendige pakker.  
rm(list=ls())  
library(tidyverse)
```

Last inn så datasettet, som jeg her kaller `oecd_n` (`df`).

```
show_col_types = FALSE
```

```
# Laster inn datasett som jeg kaller df.
lenke <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/m

df <- lenke %>%
  read_csv()
```

Oppgave 3A.

Bruk datasettet til å lage en figur med tre paneler ved siden av hverandre. Jeg anbefaler å bruke funksjonen `facets`, se eksempelet under. I hver figur skal du tegne et spredningsplott (*scatterplot*) ved hjelp av `ggplot` og `geom_point()`, med årstall på den horisontale x-aksen, og verdien på Gini på den vertikale y-aksen. Tegn inn alle tre definisjoner av Gini som du har i datasettet.

I det første panelet i figuren velger du hele befolkningen, i det andre panelet tar du arbeidere, og i det tredje pensjonister. Gi hele figuren og hvert panel passende titler, og husk benevning av aksene, samt legende til figuren (plasser den gjerne under figuren).

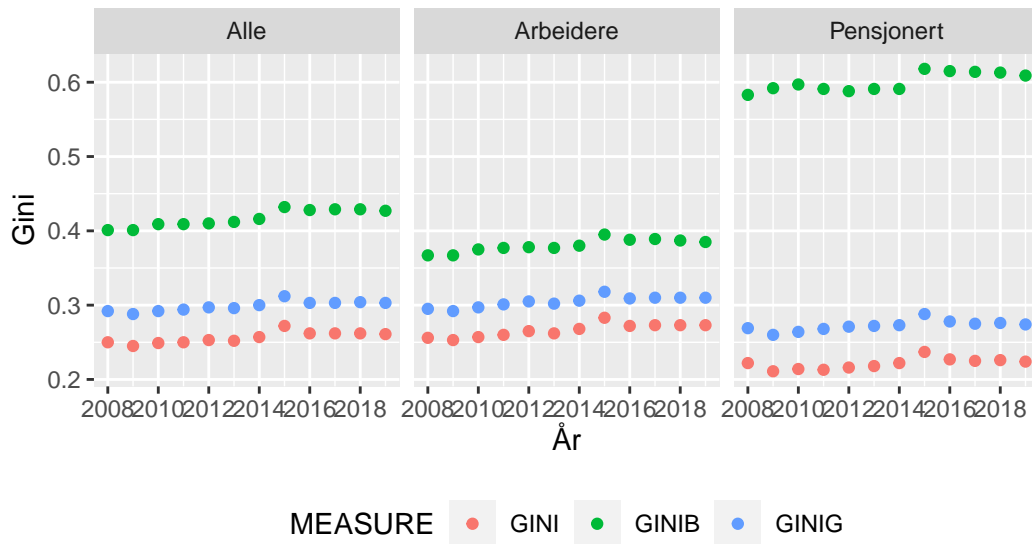
For å hjelpe kan dere tilpasse denne koden som viser befolkningstall fra OECD datasettet:

```
# Lager en liste som inneholder de tre definisjonene av Gini.
list_gini <- c("GINI", "GINIB", "GINIG")

# Lager figuren med hjelp fra koden som er hentet i oppgaven.
df %>%
  filter(MEASURE %in% list_gini) %>%
  ggplot(aes(x = Year, y=Value, color=MEASURE))+
  geom_point()+
  scale_x_continuous(breaks=c(2008, 2010, 2012, 2014, 2016, 2018))+ #Velger annet hvert år
  labs (title = "Mål på ulikhet i Norge, 2008-2019", subtitle = "Tre forskjellige Gini mål",
  facet_wrap(~ factor (AGE, (levels=c("TOT", "WA", "OLD"))),
  labeller = as_labeller(
    c("TOT" = "Alle",
      "WA" = "Arbeidere",
      "OLD" = "Pensjonert")))+
  theme(legend.position = "bottom")
```

Mål på ulikhet i Norge, 2008–2019

Tre forskjellige Gini mål på ulikhet i Norge



Gi en nøyaktig beskrivelse av figuren med de tre Gini-målene. Bruk figuren til å svare på følgende spørsmål:

Figuren viser *GINI*, *GINIB* og *GINIG* indeksen på y-aksen og årstall (annenhvert år) på x-aksen. Hvert punkt representerer verdien til *GINI*, *GINIB* og *GINIG* det aktuelle året.

- Hva viser Gini-målene om utviklingen av ulikhet i Norge? Hvordan forklarer man eventuelle forskjeller mellom de tre Gini-koeffisientene?

Hvis vi først ser på hele befolkningen (alle), kan vi se at Norge har i perioden 2008 - 2019 hatt en stabil *GINI* (netto inntekt) verdi på ca. 0.25 til 0.3. I 2014 gikk *GINI* (netto inntekt) verdien litt opp før den gikk ned og stabiliserte seg igjen. Vi kan også se at *GINI* (netto inntekt) verdien fra 2008 til 2019 har hatt en liten økning generelt sett. Dette tyder på at ulikheten på netto inntekt i landet har økt litt.

Vi ser at *GINIG* (brutto inntekt) har en veldig lik utvikling som *GINI* (netto inntekt), bare at ulikheten er større. Dette gir mening ettersom at de med høyest inntekt ikke har ett "tak" på hva de kan tjene mens de med lavere inntekt har "normale" (8 -16 jobber) jobber som betaler "normalt" (tariff, fastlønn, osv.). Hvis vi sammenligner *GINI* (netto inntekt) og *GINIG* (brutto inntekt) kan vi se at siden *GINI* (netto inntekt) verdien er lavere enn *GINIG* (brutto inntekt) verdien og det vil si at de med høyest inntekt betaler prosentvis mer av sin inntekt i skatt og overføringer. Som skaper mindre ulikhet.

Til slutt skal vi ta for oss *GINIB* (markedsinntekt før skatt). Her kan vi se at ulikheten ligger på ca. 0,4 til 0,45 mellom perioden 2008 - 2019. At *GINIB* har en større ulikhet

enn GINI (netto inntekt) og GINIG (brutto inntekt) gir mening fordi den regner med inntekter som kommer fra å eie noe som f.eks. utbytte og leieinntekter før skatt. Vi kan anta at det er de med høyest inntekt fra før, er også de som har høyest inntekt fra å eie noe. Det vil si at de med høyest inntekt får muligheten til å kjøpe mer, som igjen øker inntekten enda mer i forhold til de med lavere inntekt.

- Er det forskjeller i ulikhet mellom arbeidere og pensjonister?

Hvis vi ser på forskjellene mellom arbeidere og pensjonister kan vi fort se at ulikheten på GINI (netto inntekt) og GINIG (brutto inntekt) er størst hos arbeiderne mens GINIB (markedsinntekt før skatt) forskjellene er størst hos pensjonistene. Det er mange av de samme faktorene som jeg har diskutert over som går inn på å forklare ulikhetene i de to gruppene. Men GINIB (markedsinntekt før skatt) forskjellene for pensjonister skiller seg veldig ut blant den andre dataen. Dette kan være fordi grupper av pensjonistene (sannsynligvis de som hadde høy inntekt før) har gjennom mange år investert i eiendommer og andre verdier som gir de inntekt. Mens resten av pensjonistene (sannsynligvis de som hadde lavere inntekt før) ikke har investert å kun lever på det de får utbetalt i pensjon pr. mnd. En annen viktig faktor når vi ser på “dagens” pensjonister er at før jobbet ikke damene, og dermed ikke sparte opp / investerte til pensjon i store deler av livet.

- Hvilken rolle spiller samfunnets institusjoner, for eksempel skatte- og pensjonssystemet, for utjevning av ulikhet mellom arbeidere og pensjonister?

For å se på hva skattesystemet gjør for å jevne ut ulikhetene kan vi se på GINI (netto inntekt) og GINIG (brutto inntekt). Her kan vi se at GINI (netto inntekt) er lavere enn GINIG (brutto inntekt). Dette tyder på at ulikhetene er større før skatt og at de med høyere inntekt betaler en større prosentandel av inntekten sin i skatt. De sitter igjen med en større inntekt enn de som skatter mindre, men ulikhetene synker. Dette reflekteres både hos arbeiderne og hos pensjonistene.

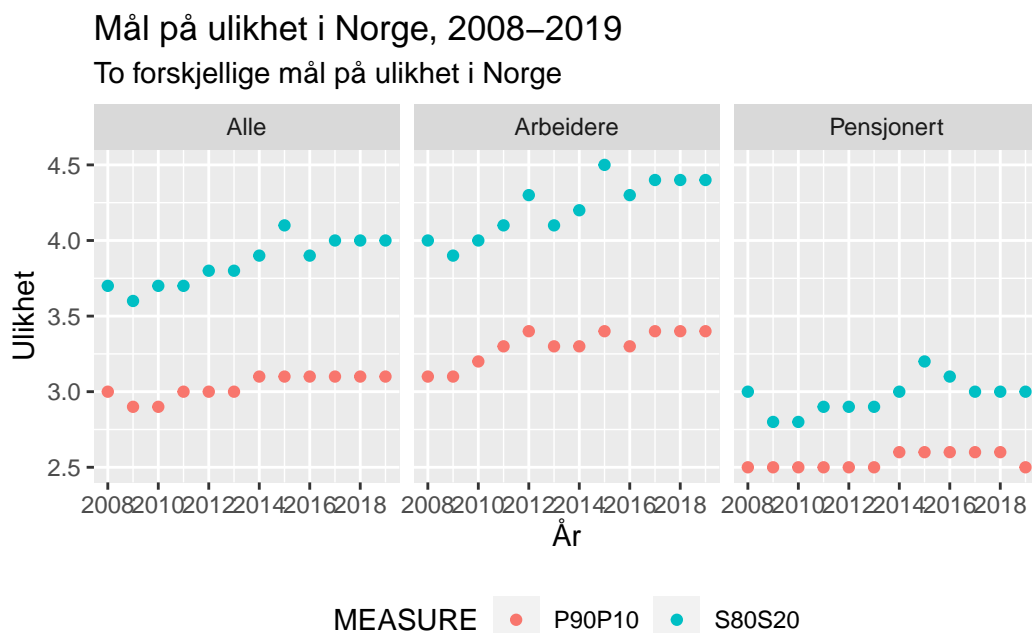
Vi kan se at inntektsulikheten mellom pensjonistene generelt er lavere enn hos arbeiderne (sett bort i fra GINIB (markedsinntekt før skatt)). Dette kan tyde på at pensjonssystemet i Norge har en innvirkning på inntektsulikheten blant pensjonistene.

Oppgave 3B.

Lag en ny figur hvor du bruker de andre ulikhetsmålene (P90P10 og S80S20) til å undersøke ulikhet mellom arbeidere og pensjonister i tidsperioden 2008-2019. Her kan du for eksempel tegne et spredningsplott for hver gruppe. Bruk figuren som grunnlag for å beskrive inntektsulikhet. Er svarene dine det samme som i 3a? Hvorfor/hvorfor ikke?

```
# Lager en liste som inneholder de andre ulikhetsmålene.
list_p9_s8 <- c("P90P10", "S80S20")

# Lager figuren
df %>%
  filter(MEASURE %in% list_p9_s8) %>%
  ggplot(aes(x = Year, y=Value, color=MEASURE))+
  geom_point()+
  scale_x_continuous(breaks=c(2008, 2010, 2012, 2014, 2016, 2018))+
  labs (title = "Mål på ulikhet i Norge, 2008–2019", subtitle = "To forskjellige mål på ulikhet i Norge"),
  facet_wrap(~ factor (AGE, (levels=c("TOT", "WA", "OLD"))),
  labeller = as_labeller(
    c("TOT" = "Alle",
      "WA" = "Arbeidere",
      "OLD" = "Pensjonert")))+
  theme(legend.position = "bottom")
```



Bruk figuren som grunnlag for å beskrive inntektsulikhet. Er svarene dine det samme som i 3a? Hvorfor/hvorfor ikke?

Som beskrevet tidligere i oppgaven er de 10% rikeste og fattigste etter skatt utelukket av P90P10 målet. Dette gir ett bedre mål på inntektsulikhetene til den “normale mannen i gata”. Som vi

kan se av figuren over er ulikhetene minst hos pensjonistene og størst hos arbeiderne. Hvis vi ser på året 2008 (alle) var det en P90P10 verdi på ca. 3. Dette betyr at vi kan anslå at de 10% rikeste var 3 ganger så rik som de 10% fattigste.

Når vi ser på S80S20 målet vet vi fra tidligere oppgaver at det tar med de 20% rikeste og 20% fattigste etter skatt. Her utelukkes altså alle mellom. Vi ser at S80S20 målet kommer ut med høyere verdier enn P90P10 målet. Igjen ser vi at ulikhetene her er størst hos arbeiderne. Dette kan tyde på at de rike trekker gjennomsnitte på inntekt lengere opp enn de fattigste drar det ned. Når vi ser på året 2016 (alle) har det en S80S20 verdi på ca. 4. Dette betyr at de 20% rikeste har en gjennomsnitts inntekt som er ca. 4 ganger høyere enn de 20% fattigste.

Hvis vi skal sammenligne disse observasjonene med figuren i oppgave 3A ser vi at ulikhetene er større hos arbeiderne enn hos pensjonistene etter skatt i begge oppgavene. Vi observerer også i begge oppgavene at det åpenbart er ulikheter i inntekt i samfunnet. Til forskjeller fra oppgavene får vi en bedre indikasjon på hvor stor forskjeller det er i befolkningen av oppgave 3B siden vi får ett tall som indikerer akkurat hvor store forskjell det er. Dette er selvfølgelig en forenkling av virkeligheten og vi kan ikke sette ett tall på hvor stor inntekts forskjell det er.

Oppgave 4. Casestudie: Tromsø kommune 2015

Jeg starter med å slette minnet og laste inn nødvendige pakker så jeg starter oppgaven med blanke ark.

```
# Sletter minnet og laster inn nødvendige pakker.  
rm(list=ls())  
library(tidyverse)
```

Her skal vi se på skattetall for Troms fylke fra 2015:

```
# Laster inn datasett som jeg kaller df_troms.  
lenke <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/main/data/df_troms.csv"  
  
df_troms <- lenke %>%  
  url() %>%  
  gzcon () %>%  
  readRDS()
```

Vi skal kun bruke Tromsø kommune, så vi lager en ny dataframe med observasjoner fra Tromsø (kommune 1902):

```
# Lager en ny dataframe med observasjoner kun fra Tromsø.
df_tromsø <- filter(df_troms, kommnr == "1902")
```

Lag en ny variabel ved hjelp av `mutate()` som viser inntekt fratrasket skatt. Noen av disse verdiene er negative, og disse kan du sette lik 0, for eksempel ved `data[data < 0] <- 0` dersom dataframe heter “data”).

```
# Lager en ny variabel som er inntekt etter skatt som jeg kaller "netto_inntekt".
df_tromsø <- df_tromsø %>%
  mutate(netto_inntekt = inntekt - skatt)

# Erstatte alle negative tall i df_tromsø med 0.
df_tromsø[df_tromsø < 0] <- 0
```

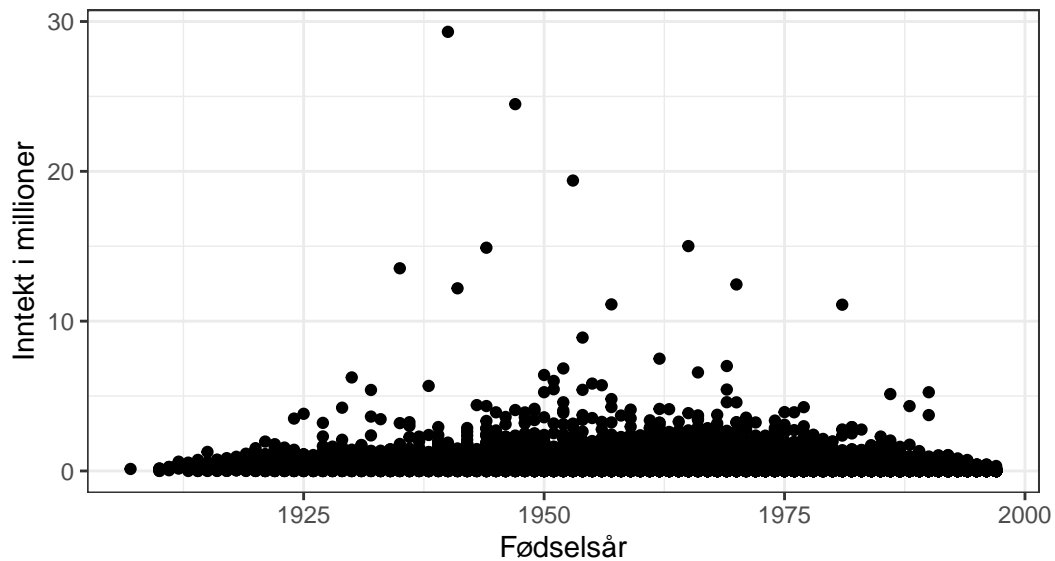
Oppgave 4A.

Tegn et spredningsplott med fødselsår på x-aksen og inntekt (før og etter skatt) på y-aksen.

```
# Lager spredningsplottet før skatt
df_tromsø %>%
  ggplot(aes(x = fodd, y = inntekt/10^6)) +
  geom_point() +
  labs(title = "Inntekt fordelt på alder (før skatt) i Tromsø i 2015",
        subtitle = "Inntekt fordelt på fødselsår i en periode fra 1907 til 1997",
        x = "Fødselsår",
        y = "Inntekt i millioner") +
  theme_bw()
```

Inntekt fordelt på alder (før skatt) i Tromsø i 2015

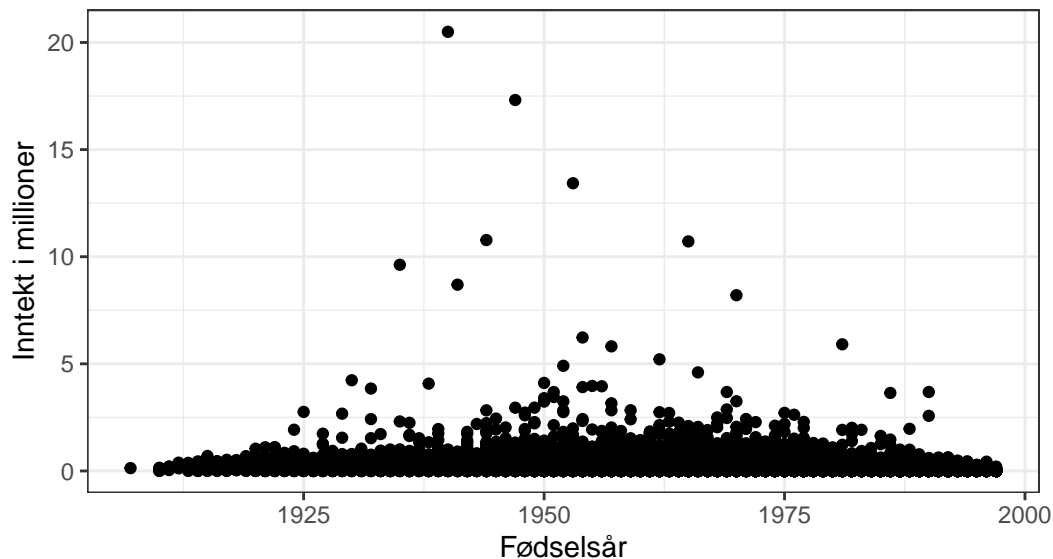
Inntekt fordelt på fødselsår i en periode fra 1907 til 1997



```
# Lager spredningsplot etter skatt
df_tromsø %>%
  ggplot(aes(x = fodd, y = netto_inntekt/106)) +
  geom_point() +
  labs(title = "Inntekt fordelt på alder (etter skatt) i Tromsø i 2015",
        subtitle = "Inntekt fordelt på fødselsår i en periode fra 1907 til 1997",
        x = "Fødselsår",
        y = "Inntekt i millioner") +
  theme_bw()
```

Inntekt fordelt på alder (etter skatt) i Tromsø i 2015

Inntekt fordelt på fødselsår i en periode fra 1907 til 1997



Hvordan kan vi tolke dette bildet i forhold til inntektsulikhet i Tromsø kommune, og hvilken rolle spiller skattesystemet her?

I disse figurene har vi inntekt i millioner på y-aksen og fødselsår på x-aksen. Hvert punkt representerer en person i Tromsø. Figurene viser oss hvilke fødselsår som tjente mest i Tromsø i 2015 før og etter skatt. Vi kan se at personene som tjente best både før og etter skatt var de som var født mellom 1925 og 1975.

Figurene forteller oss at i Tromsø i 2015 var det noen enkelte individer som tjente veldig mye mer enn resten av befolkningen. Det var fem personer som tjente over 10 millioner etter skatt. I figurene klarer vi å se inntektsulikhetene mellom de som tjener veldig godt og de som tjener mindre. Men figurene gir ikke ett veldig godt bilde av ulikheten mellom den "normale person i gata". Det er noen få personer som sitter med veldig høy inntekt.

Når vi sammenligner før og etter skatt ser vi at verdiene blir mindre, men figurene er veldig lik. Dette tyder på at skattesystemet ikke har veldig stor innvirkning på inntektsulikhetene. Som vi har diskutert i tidligere oppgaver så vi at de med høyest inntekt skatter prosentvis mer av sin inntekt. Dette er nokk også det som skjer her, men det er vanskeligere å se siden figuren blir såpass strekt ut av de med høyest inntekt.

Når vi skal se på inntekt ulikhet mellom fødselsår kan vi se at de som er født tidlig og seinest er de med dårligst inntekt både før og etter skatt (med unntak av noen unge enkeltindivider). Dette er sannsynligvis fordi de eldste har gått av med pensjon og de yngste studerer eller nettopp har startet i arbeidsmarkedet.

Hadde vi fjernet de med høyest inntekt og lavest inntekt (P90P10) hadde vi fått et bedre bilde på inntektsulikhetene mellom “den normale person i gata”.

Oppgave 4B.

Vi skal nå tegne en figur hvor vi deler befolkningen i stigende rekkefølge etter inntekt (før skatt), og hvor vi tar gjennomsnittsinntekten til hver persentil i kommunen. Kjør følgende kode for å få dette til:

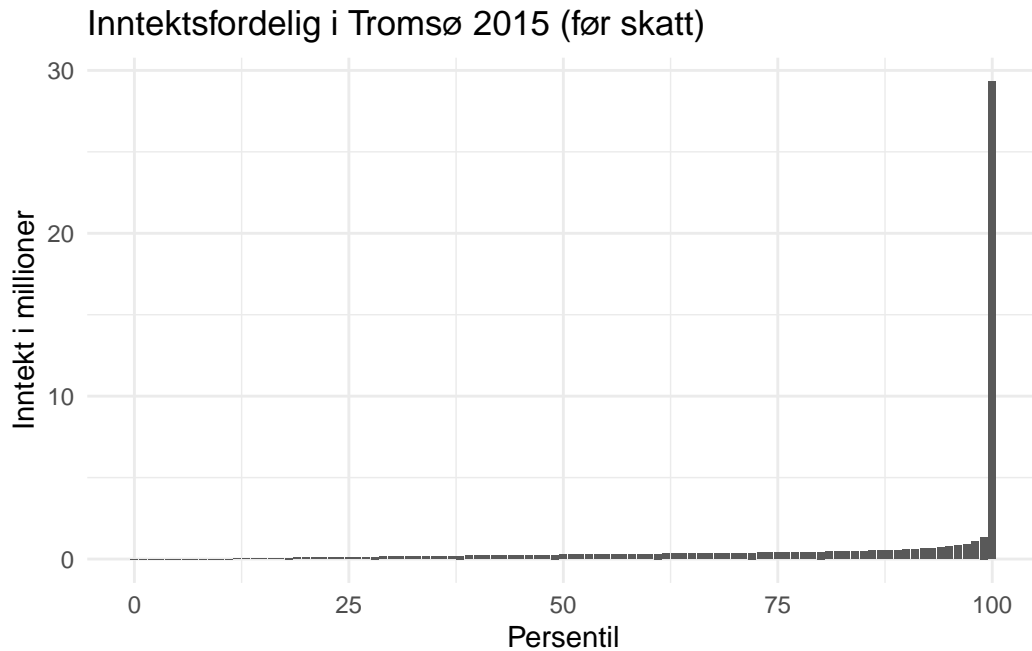
```
# Kjøreren koden som er hentet fra oppgaven (kildeliste) og tilpasser til min oppgave for å
df_tromsø_1 <- df_tromsø %>%
  select(inntekt) %>%
  quantile(probs = seq(0,1, .01), na.rm = TRUE)

tibble_tromsø_1 <- as_tibble(df_tromsø_1)

tibble_tromsø_1 <- tibble_tromsø_1 %>%
  add_column(Persentil = c(0:100))

tibble_tromsø_1 <- tibble_tromsø_1 %>%
  rename(inntekt_før_skatt = value)

tibble_tromsø_1 %>%
  ggplot(aes(x = Persentil, y = inntekt_før_skatt/10^6)) +
  geom_col() +
  labs(title = "Inntektsfordeling i Tromsø 2015 (før skatt)",
       y = "Inntekt i millioner",
       x = "Persentil") +
  theme_minimal()
```



Lag en tilsvarende figur for inntekt etter skatt og kommentér og forklar eventuelle forskjeller.

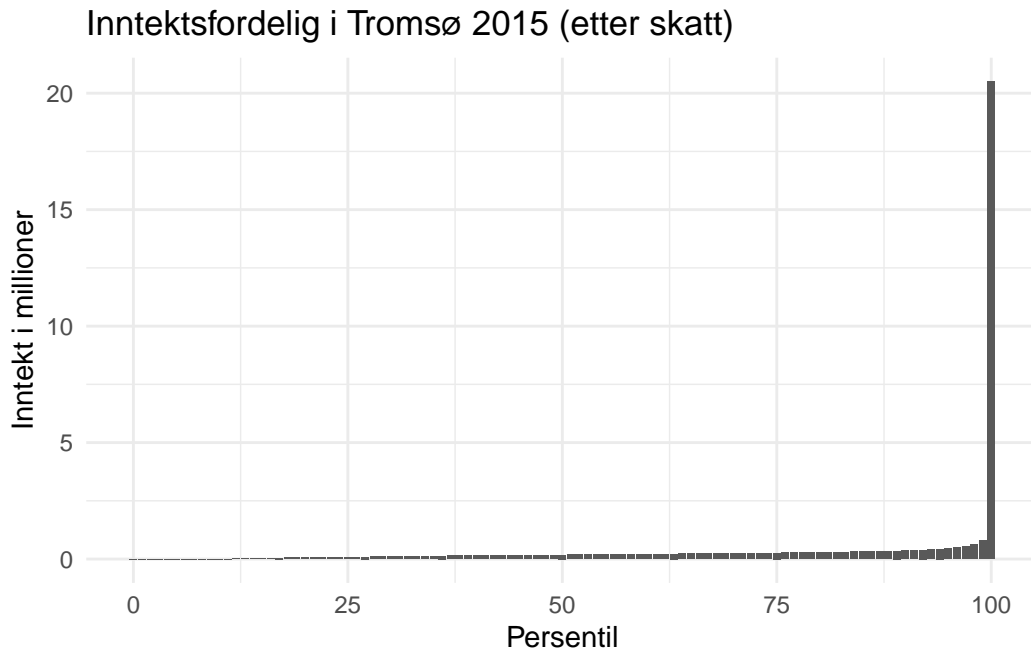
```
# Lager en tilsvarende figur for etter skatt.
df_tromsø_2 <- df_tromsø %>%
  select(netto_inntekt) %>%
  quantile(probs = seq(0,1, .01), na.rm = TRUE)

tibble_tromsø_2 <- as_tibble(df_tromsø_2)

tibble_tromsø_2 <- tibble_tromsø_2 %>%
  add_column(Persentil = c(0:100))

tibble_tromsø_2 <- tibble_tromsø_2 %>%
  rename(inntekt_etter_skatt = value)

tibble_tromsø_2 %>%
  ggplot(aes(x = Persentil, y = inntekt_etter_skatt/10^6)) +
  geom_col() +
  labs(title = "Inntektsfordelig i Tromsø 2015 (etter skatt)",
       y = "Inntekt i millioner",
       x = "Persentil") +
  theme_minimal()
```



Hvilket bilde får du av inntektsulikhet i Tromsø kommune?

I disse figurene har vi inntekt i millioner (brutto og netto) på y-aksen og hvilken persentil befolkningen med den inntekten tilhører på x-aksen.

Som vi observerer i begge figurene er det befolkningen som befinner seg i persentil 100 som helt klart har høyest inntekt. Forskjellene blir så store at vi ikke klarer å observere hvilken inntekt gruppene under persentil 100 har. Dette tyder selvfølgelig på at inntektene i Tromsø kommune er ulikt fordelt.

Som vi har diskutert tidligere i oppgaven skatter de med høy inntekt prosentvis mer av sin inntekt enn de med lavere inntekt. Men igjen kan vi se at de med høy inntekt sitter igjen med veldig mye mer enn de med lav inntekt selv om de skatter mer.

Hvis vi ser bort ifra persentil 100, blir det som nevnt tidligere vanskelig å observere noen data. Men ut ifra det vi kan observere ser det ut som et ulikhetene jevner seg ut når vi tar bort denne gruppen. Det vil si at det er noen få i befolkningen som har en svært mye høyere inntekt enn resten, men resten har en relativt lik inntekt.

Oppgave 4C.

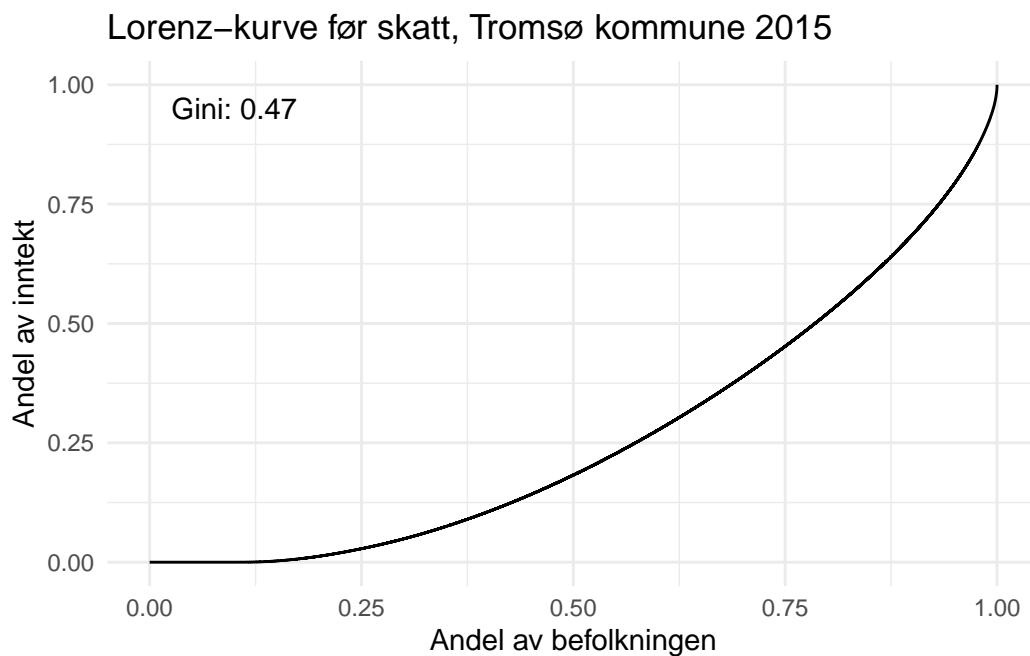
Vi skal nå tegne en Lorenz-kurve og beregne en Gini-koeffisient for Tromsø kommune. Først må du `install.packages("gglorenz")` , og laste pakken inn: `library(gglorenz)` . Hvordan pakken fungerer kan du finne her: [gglorenz](#).

```
# Laster inn nødvendige pakker.  
library(gglorenz)
```

Tegn to Lorenz-kurver for Tromsø kommune (før og etter skatt), samt tilhørende Gini-koeffisienter. Sammenlikn med de relevante Gini-koeffisientene for hele Norge fra 2015 fra OECD datasettet. Er inntektsfordelingen i Tromsø mer eller mindre lik sammenliknet med landet som helhet? Gi mulige forklaringer.

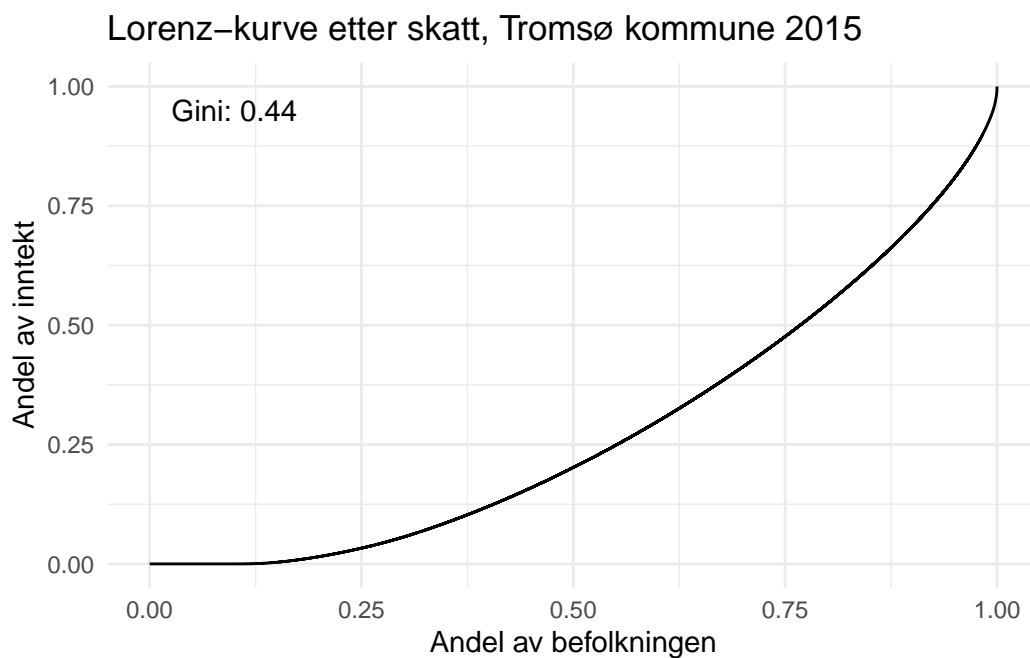
Jeg tegner først opp begge lorenz-kurvene for Tromsø kommune (før og etter skatt), samt tilhørende Gini-koeffisienter. Under disse kurvene svarer jeg på spørsmålene.

```
# Figur 1 for Tromsø kommune (før skatt).  
# Bruker en tidligere datafram fra oppgave 4A (df_tromsø)  
  
# Lager figuren  
ggplot(df_tromsø, aes(inntekt)) +  
  stat_lorenz() +  
  annotate_ineq(df_tromsø$inntekt) +  
  labs (x="Andel av befolkningen",  
        y= "Andel av inntekt",  
        title = "Lorenz-kurve før skatt, Tromsø kommune 2015") +  
  theme_minimal()
```




```
# Figur 2 for Tromsø kommune (etter skatt)
# Bruker en tidligere datafram fra oppgave 4A (df_tromsø)

# Lager figuren
ggplot(df_tromsø, aes(netto_inntekt)) +
  stat_lorenz() +
  annotate_ineq(df_tromsø$netto_inntekt) +
  labs (x="Andel av befolkningen",
        y= "Andel av inntekt",
        title = "Lorenz-kurve etter skatt, Tromsø kommune 2015") +
  theme_minimal()
```



Er inntektsfordelingen i Tromsø mer eller mindre lik sammenliknet med landet som helhet? Gi mulige forklaringer.

Vi har tidligere i oppgaven sett at Norge i 2015 har en GINI verdi på 0.31 før skatt og 0.27 etter skatt. Når vi ser på figurene over kan observere at Tromsø i 2015 har en GINI verdi på 0.47 før skatt og 0.44 etter skatt. I oppgave 2A fant vi ut at en GINI verdi på 0 er total likhet og en GINI verdi på 1 er total ulikhet. Det vil si at GINI indeksen indikerer at Tromsø i 2015 hadde en større inntektsulikhet enn resten av landet både før og etter skatt.

En mulig grunn til at inntektsulikhetene er større i Tromsø enn resten av landet, kan være at Tromsø har universitet. Det vil si at det er en del studenter i byen med lav inntekt som trekker

ned total inntekten i kommunen. En annen virkningen av dette er at nesten halvparten av befolkningen har høyere utdanning (SSB 2021). Mens den andre halvparten har videregående eller grunnskole utdanning. Jeg har tidligere i oppgaven (oppgave 1b) skrevet om at det blir flere jobber med høyere inntekt til de med høyere utdanning mens de uten høyere utdanning får færre jobber og en lavere inntekt.

En annen mulighet til at inntektsulikheten er større i Tromsø enn resten av landet, kan være at Tromsø har nesten halvparten av sin arbeidsstokk i offentlig sektor (iTromsø, 2018). Offentlig sektor har generelt sett en lavere inntekt enn privat sektor (SSB, 2011). Det vil si at du har en stor del av befolkningen med inntekter som er styrt av det offentlige og en stor del av befolkningen som har inntekter styrt av det private. Siden den private sektoren ikke har ett "tak" på hva de kan tjene blir ulikhetene naturlig større når vi sammenligner med den offentlige sektoren som har ett slikt "tak".

Dette er bare noen få mulige grunner av hvorfor inntektsulikhetene er større i Tromsø enn resten av landet. Det er antakeligvis en kombinasjon av flere grunner til hvorfor det er økt inntektsulikhet i Tromsø i forhold til resten av landet.

Kildeliste

Oppgaven: https://uit-sok-1004-h22.github.io/eksamen/mappe_1_h22_endelig.html

NHO, (2018), Økende ulikhet, NHO, hentet 26.09.2022 <https://www.nho.no/publikasjoner/p/naringslivets-perspektivmelding/okende-ulikhet/>

Epland, J., Tuv, N., (09.05.2019), Slik måler SSB ulikhet, SSB, hentet 28.09.2022 <https://www.ssb.no/inntekt-og-forbruk/inntekt-og-formue/statistikk/inntekts-og-formuesstatistikk-for-husholdninger/artikler/slik-maler-ssb-ulikhet>

Grini, K., Lien, H. (05.12.2011), Økende lønnsforskjeller det siste tiåret, SSB, hentet 06.10.2022 <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/okende-lonnsforskjeller-det-siste-tiaaret>

Bowles, S., Carlin, W., Stevens, M. (2017). *Core*. Oxford University Press.

SSB, (2021), Kommune Tromsø, SSB, hentet 09.10.2022 <https://www.ssb.no/kommunefakta/tromso>

Pedersen. C. (23.08.2018), Tromsø kommune er mest næringsvennlig i Nord-Norge, iTromsø, hentet 09.10.2022 <https://www.itromso.no/nyheter/i/8BdaAd/tromso-kommune-er-mest-naeringsvennlig-i-nord-norge>