

---

title: "SOK-1004, høst 2022, Mappeoppgave 1"

author: "55, 122, 54"

format: pdf

echo: true

output: true

editor: visual

---

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

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

```
suppressMessages(library(tidyverse))  
#install.packages("cowplot") Fjern # for å installere cowplot om det mangler.  
library(cowplot)  
suppressMessages(library(cowplot))  
library(gglorenz)  
suppressMessages(library(gglorenz))  
#install.packages("gglorenz") #Fjern # om kode nedenfor ikke kjører.
```

```
#Nå henter vi inn data fra URL og Leser da det og lagrer som oecd_n

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

oecd_n <- url %>%
  read_csv(show_col_types = FALSE)

url3 <- "https://raw.githubusercontent.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/main/data/skatt_troms_2015.RDS"

skatt_troms_2015 <- url3 %>%
  url() %>%
  gzcon() %>%
  readRDS()

#Vi klarte ikke å få fjernet de meldingene fjernet som vi ønsket. Vi ønsket at utprint skulle
#være ryddigere.
```

# Oppgave 1. Årsaker og konsekvenser av ulikhet

## Oppgave 1a

Hva menes med inntektsulikhet i et land?

---

Inntektsulikhet er forskjellen mellom hvor mye rike tjener og fattige tjener. «Noen tjener lite, noen tjener mye»

Inntektsulikhet forteller om ulikhetene mellom hva folk tjener i løpet av en avgrenset tidsperiode, ofte ett år.

I Norge så bruker SSB i hovedsak 3 måter på å måle ulikhet med p90p10, s80s20 og Gini koeffisient.

---

## Oppgave 1b

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

«Den dominerende faktoren bak økte lønnsforskjeller, og dermed økte ulikheter, er på globalt nivå teknologi, ikke globalisering.» «NHO Verden og oss 2018, side 159.»

Teknologisk fremgang i kapitalisme kan ofte føre til store inntekter, mens de arbeiderne som nå fikk jobben sin forenklet fra teknologi vil da potensielt tjene mindre. Det kan også bli mindre arbeidsplasser. Et eksempel på dette er «Spinning jenny». Dette ledet til at mange arbeidere mistet jobbene sine siden arbeidet ble effektivisert, men det gjorde også slik at man kunne omfordele arbeidskraften fra det arbeidet. De landene uten den teknologien fikk da vanskeligheter rundt konkurranse som da økte ulikhetene ved at de landene med «Spinning jenny» tjente mer mens de uten tjente mindre. Videre eksempel er da de la til damp og da økte

effektiviteten igjen mer.

Politikk holder mange grunner til ulikhet, ved laging av institusjoner og retningslinjer kan gi stort utslag i å bare gi tilgang til institusjoner eller tilsvarende.

Skattepolitikk, gitt at man snakker om inntekt etter skatt. Inntekt kan være ulik etter hvor man jobber eller hvilket yrke osv. Dette leder til at noen tjener mye mens andre tjener mindre og skatt omfordeler inntekt, og vil normalt utjevne forskjeller. Verdt å nevne er jo da ting som såkalt koppskatt, der skatten er et fast kronebeløp, uavhengig av inntekt som vil gi dårligere utslag en annen skattepolitikk.

Et eksempel på forandring i skattepolitikk er at I Norge ble arveavgift fjernet og vi kan anta at dette økte ulikhetene siden rike sendte formue videre til unge.

Utdanning og helsevesen er gode eksempler på institusjoner som kan fungere som fordelingsmekanismer, men når disse har høye kostnader så vil det øke ulikheter siden det da bare er de med ressurser som kan få tilgang på det.

Deregulering er en faktor som kan lede til økt ulikhet siden deregulering er å fjerne reguleringer.

Et eksempel på at deregulering kan lede til mer ulikhet er når lover blir opphevd slik at arbeidere kan utnyttes til større grad av bedrifter til lav lønn. Det kommer også såkalt "Non standard work"

Non standard work er selvstendig næringsdrivende, deltidsansatte og midlertidig ansatte. De type jobber øker andelen sysselsatte i bunnen av inntektsskalaen og dermed øker også ulikheter.

Innvandring vil også øke ulikheter. Et eksempel kan være når personer immigrerer til Norge. De vil ofte ha lavere formue og inntekt enn Nordmenn. Dette kan også være motsatt.

Eksempel på hvordan arv vil påvirke ulikhet kan være at;

Personer har også forskjellige evner. Noen vil ha praktiske jobber og andre vil ha teoretiske jobber. Noen vil jobbe med mat og jobbe på restaurant hele livet mens andre kan være interessert i å bli advokater osv.

Lavt utdannede foreldre vil ofte få lavt utdannede barn. Det samme går for høyt utdannede foreldre som vil ofte få høyt utdannede barn.

---

## Oppgave 1c

Hvilke konsekvenser kan ulikhet i inntektsfordelingen føre til?

---

Når man har veldig store ulikheter i et land så kommer ofte misnøye blandt populasjonen og dårlig økonomisk vekst. Ulikheter leder til at du har et samfunn delt i deler der du har rike og fattige. De rike vil da ofte få bedre tilgang på tjenester og andre goder som de fattige ikke har tilgang på. En direkte konsekvens på ulikhet kan være tilgang på kapital. Om du har en god ide uten kapital så vil du ikke kunne like lett starte en bedrift og potensielt skape ressurser eller tjenester. På den andre siden så kan du ha en dårlig ide med mye kapital og dermed starte en bedrift med en dårlig ide og potensielt da gjøre det dårlig å bortkaste de ressursene du hadde. Dette er et eksempel på hvordan ulikhet kan lede til dårlig effektivitet i en økonomi.

Dårlig effektivitet i økonomien gir utslag i vekstevnen. Om ressursene i økonomien blir plassert dårlig i forhold til andre land så vil andre land potensielt ha en større vekst slik at vi blir fattigere i forhold til dem.

«Lav, varierende og usikker inntekt kan også påvirke tilgangen til kreditt, herunder til bolig. Dermed kan økende ulikheter gi lavere boligbygging og lavere boliginvesteringer.» "NHO Verden og oss 2018, side 163"

Med lavere boligbygging og lavere boliginvesteringer så vil du også få mindre folk som eier boliger og flere

leietakere. Dette gir også da negativ innvirkning på evne for folk til å bygge formue som igjen gir økte ulikheter.

«Det er rimelig å anta at fallet i tillit til andre er størst hos dem som på ulike måter føler seg utenfor de etablerte maktstrukturene.» “NHO Verden og oss 2018, side 164.”

Politisk polarisering og lav tillitt kan være en årsak til ulikhet.

Du får en splittelse i samfunnet når du har veldig rike og veldig fattige. Tilgang på kapital gir deg tilgang på politisk innflytelse i mange former som bruk av media osv. Du ender opp med å ha mange som da ikke har tillitt til myndighetene eller hverandre. En negativ effekt av denne splittelsen for vekst er politisk uro. Hvordan kan vi forvente at folk skal ta risiko og starte bedrifter når landet du er i plutselig kan gå en annen retning og forby det du gjør. For eksempel I Storbritannia når det var mye uro og snakk om å forlate EU, så var det mye høyere risiko å starte en bedrift som driver handel med EU. Dette viser da en effekt som bremser vekst fra politisk polarisering og lav tillitt. Den lave tillitten ledet til «Brexit» og da forlot Storbritannia EU. Da vil den bedriften ikke kunne leve lengere. Britene mistet i stor grad tilgang på EU sitt marked, og «vice versa». Dette har ledet til en mindre effektiv økonomi.

Det er korrelasjon mellom veldig lav BNP per person og lav levealder. Land som Norge kan betale mye for gode medisiner og behandlinger mens sør afrikanske land vil slite med å gjøre det samme, så her kan vi se en sammenheng mellom rikdom og fattigdom i at vi lever lengere og har bedre kvalitet på mat, vann, helsevesen, utdanning, jobber, osv.

<https://ourworldindata.org/grapher/life-expectancy-vs-gdp-per-capita> (<https://ourworldindata.org/grapher/life-expectancy-vs-gdp-per-capita>)

Livstilfredshet på selvrapporterte studier så kommer land med høy BNP svært godt ut. Vi kan anta at dette kommer fra at vi får bedre muligheter til å gjøre det vi vil i stedet for arbeide hele dagen for å overleve. Vi har også enklere tilgang til underholdning og bedre utdanning.

Kilde til livstilfredshet er <https://ourworldindata.org/grapher/gdp-vs-happiness?xScale=linear&time=2020..latest> (<https://ourworldindata.org/grapher/gdp-vs-happiness?xScale=linear&time=2020..latest>)

Siden vesten har hatt det godt økonomisk sett så har vi hatt muligheten til å bygge fabrikker og drive produksjon i utlandet. Kina har vært et billig land for vesten og produsere i siden arbeidere der har tatt forholdsvis veldig lite betalt og vi har fått veldig mange varer for lite kostnad. Effekten av dette har vært todelt. Mennesker i vesten har hatt luksusen til å velge bort arbeid som industriell produksjon, og Kina har fått industrialisert landet og hentet veldig mange mennesker ut av fattigdom. Vesten har endt opp med mye teknologisk utvikling og Asia har endt opp med å produsere varer som ofte kommer fra den teknologiske utviklingen.

Samfunnets ressurser omfordes til der det antas å gi størst avkastning.

I industrialiserte land så kan en av de positive konsekvensene av ulikhet være at flere velger å ta høyere utdanning.

Med økonomisk mobilitet så er det flere insentiver til å jobbe mer og hardere for å tjene mer. Personer som starter suksessfulle bedrifter, motiverer andre til og også starte nye bedrifter som igjen kan lede til ny teknologisk utvikling.

Med ulikheter i et samfunn så får du konkurranse om ressursene, du vil få mennesker som vil bygge verdier og da kan fort ny teknologi komme. Teknologisk fremgang i et land vil kunne gi konkurranse fortrinn for det landet på global skala som derved kan gi større inntekter til det landet. Dette kan da lede til at samtidig som eiere av den nye teknologien tjener og øker ulikheter i landet, så vil de også da skatte over ny inntekt og da betale for tjenester til fattige. Med den muligheten til økonomisk mobilitet i en økonomi så kan det fort komme stor fremgang i økonomien som løfter alle selv om det kan komme mer ulikhet.

Stor ulikhet kan også lede til at land får politisk motivasjon til å bedre stats finansierte institusjoner som utdanning.

Og for å snu på spørsmålet: hva ville skjedd dersom alle tjente det samme?

De som studerer i årevis finansiert av studielån ville tape penger på utdanningen. Vi ville hatt færre insentiver til å jobbe hardere for entreprenørskap og innovasjon.

## 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-koeffisienten måler inntektsulikheter. Bør sammenlignes mellom variabler siden gini tallet alene sier svært lite. Gini-koeffisient beregnes kort forklart ved se på all inntekt og formue det er totalt sett i ett land, deretter se på hvordan dette er fordelt mellom folk.

Markedets inntekt – totale inntekter fra markedet, inkludert formue

Brutto inntekt - betyr markedsinntekt pluss verdien av overføringer fra trygdeordninger. Det subtraherer verdien av overføringer gjort til arbeidsrelaterte og pensjonsordninger. kilde for definisjon av brutto inntekt: <https://www.oecd.org/els/soc/IDD-ToR.pdf> (<https://www.oecd.org/els/soc/IDD-ToR.pdf>)

Disponibel inntekt – Totale inntekter etter betaling av inntektsskatt og utbetalinger av støtte ordninger fra stat. Tar ikke med indirekte skatter. Ekskluderer gratis tilbud fra stat og offentlige tjenester.

Som vist i figuren nedenfor ser vi at gini-koeffisienten beregnes med formelen:  $\text{Gini Koeffisienten} = A / (A + B)$ . A er området som befinner seg over Lorenz kurven og B er området under Lorenz kurven. Resultatet er ett tall som vil ligge mellom 0 og 1. Gini-koeffisienten måler inntektsulikheter mellom land.

```
#Her Legges figur på lorenz
tromso <- filter(skatt_troms_2015, kommnr == "1902")

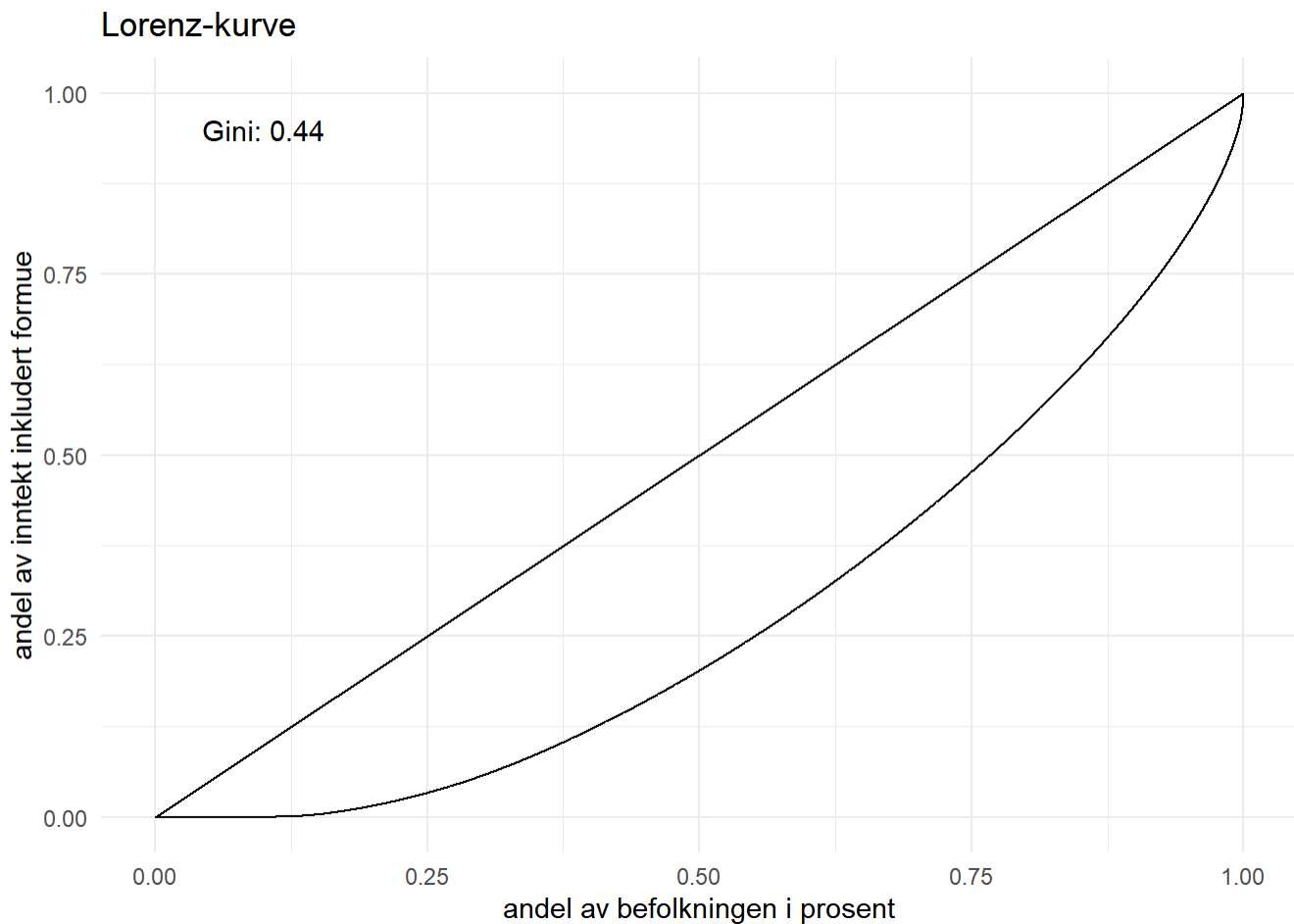
tromso <- tromso %>%
  mutate(netto = inntekt-skatt)

tromso[tromso < 0] <- 0

Eksempel_lorenz <- ggplot(tromso, aes(netto))+
  stat_lorenz()+ # tegner Lorenz
  annotate_ineq(tromso$netto)+ # beregner Gini
  labs (x="andel av befolkningen i prosent", y= "andel av inntekt inkludert formue", title =
"Lorenz-kurve")+
  theme_minimal()+

  geom_segment(aes(x=0,xend=1,y=0,yend=1))

Eksempel_lorenz
```



I P90P10 deler man opp befolkningen i ti like store deslier ut ifra inntekt. Forholdstallet beregnes ved å se på persentilen som har høyere inntekt enn 90 prosent av befolkningen og persentilen som har lavere enn 90 prosent av befolkningen. Dette resulterer ofte til at det er lettere å måle ulikhet over tid enn ved andre ulikhetsmål.

S80S20 er ett mål som viser forholdstallet mellom gjennomsnittsinntekten til de 20 prosent av befolkningen som har høyest og lavest inntekt. Det er gjerne små forskjeller mellom Gini-koeffisienten og dette ulikhetsmålet når man ser på utvikling i ulikhet. Produktet til dette målet vil gjerne være høyere enn å bruke bare gini.

## Oppgave 2b

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

---

Gini-koeffisienten er at det kun er ett tall. Det beskriver hvordan inntekt er fordelt i ett land eller en befolkning. En av de negative sidene med dette er at det ikke viser fordelingen innad i befolkningen.

Det som er positivt med gini er at man enkelt kan sammenlikne land mot hverandre fordi det er enkelt å samle mye data.

En av de bra tingene med å bruke P90P10 for å måle ulikhet er at den måler mere nøyaktig for personer som ligger mellom den andre persentil og den niende persentil. Minuset med denne metoden er at vi ser vekk fra de aller fattigste og de aller rikeste.

S80S20 er en god metode for å sammenlikne de riskeste og fattigste i en befolkning. Den negative siden med dette målet er at det ikke ser på hele befolkningen, men bare topp og bunn persentilene sammenlagt.

Alle disse måtene å måle ulikhet på viser forskjellige data om samme befolkning. Det vil derfor være positivt å bruke alle metodene for å se ett større bilde på ulikhet i en befolkning ved å bruke alle metodene å sammenlikne mot hverandre.

---

## Oppgave 3. Inntektsulikhet i Norge 2008-2019.

### Oppgave 3a

Bruk datasettet til å lage en figur med tre paneler ved siden av hverandre. Jeg anbefaler å bruke funksjonen `facets` (<https://r4ds.had.co.nz/data-visualisation.html#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.

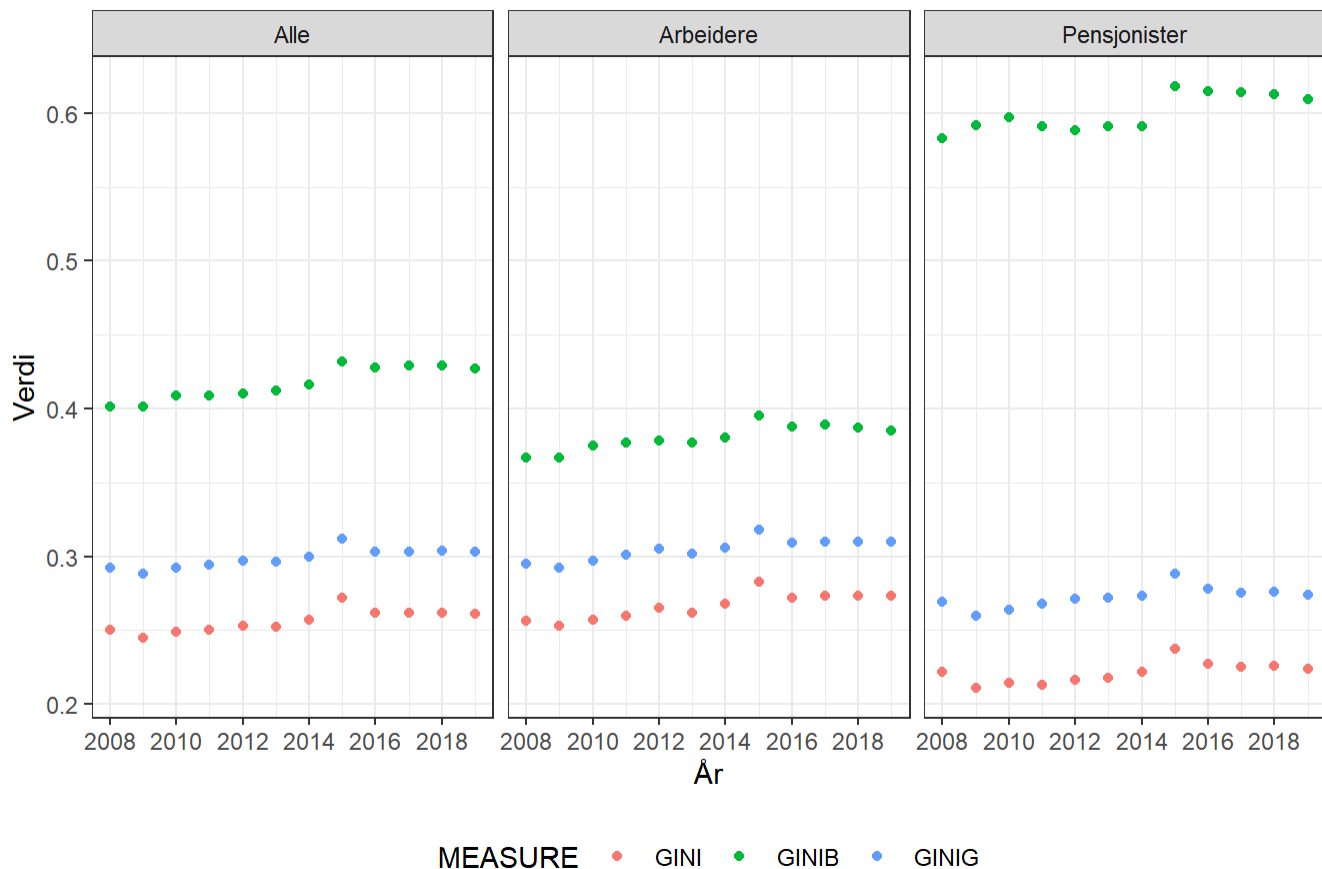
I det første panelet i figuren velger du hele befolkningen, i det andre panelet tar du arbeidere, og i det tredje pensjonister. Gi hele figuren og hvert panel passende titler, og husk benevnning av aksene, samt legende til figuren (plasser den gjerne under figuren).

```
list_gini = c("GINI", "GINIB", "GINIG")

oecd_n %>% # data

  filter(MEASURE %in% list_gini) %>% # velger ut de observasjonene vi vil ha
  ggplot(aes(x = Year, y=Value, color=MEASURE))+ # definer aksene
  geom_point()+ # plott som punkt
  theme_bw() + #endrer bakgrunnsfarge
  scale_x_continuous(breaks=c(2008, 2010, 2012, 2014, 2016, 2018))+ # unngår at årstall blir
gjort om til desimal
  labs (title = " GINI, GINIB og GINIG mål for Norge, 2008-2019", y = "Verdi", x= "År")+ # ti
ttel på figuren og benevning av aksene
  facet_wrap(~ factor (AGE, (levels=c("TOT", "WA", "OLD"))), # lag tre paneler etter alder i
fornuftig rekkefølge med passende titler
            labeller = as_labeller(
              c("TOT" = "Alle",
                "WA" = "Arbeidere",
                "OLD" = "Pensjonister")) + # tilpass tittel på panelene
  theme(legend.position = "bottom") # flytt Legende for at størrelsen på plottene blir bra
```

GINI, GINIB og GINIG mål for Norge, 2008-2019



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

---



Figuren viser gini verdiene hos aldersgruppene 18-65(arbeidere), 65+(pensjonerte), og hele befolkningen i hvert sitt panel. Vi ser utviklingen fra 2008 til 2019 med markedetsinntekt, disponibel inntekt og endelig inntekt i spredningsplott.

---

### **Hva viser Gini-målene om utviklingen av ulikhet i Norge?**

---

Gini-målene vises på figuren å være relativt stabil, det har vært en liten økning i perioden 2008-2019.

---

### **Hvordan forklarer man eventuelle forskjeller mellom de tre Gini-koeffisientene?**

---

GINIB tar markedsinntekt og formue før skatt, uten å inkludere støtteordninger.

GINIG tar markedsinntekt pluss verdien av overføringer fra trygdeordninger. Det subtraherer verdien av overføringer gjort til arbeidsrelaterte og pensjonsordninger

GINI tar markedsinntekt og formue etter skatt med støtteordninger.

---

### **Er det forskjeller i ulikhet mellom arbeidere og pensjonister?**

---

Figuren viser en distinkt forskjell mellom arbeidere og pensjonerte. Ser vi på GINI og GINIG er inntektsulikheten mindre i aldersgruppen "pensjonister" enn i aldersgruppen "arbeidere" ettersom det er færre i arbeid og dermed flere på støtteordninger.

Ser vi på GINIB har pensjonistgruppen en mye større inntektsulikhet igjen fordi det ikke beregner støtteordninger som pensjon som inntekt.

---

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

---

Institusjoner som fordeler midler i Norge gjør at forskjellene jevner seg ut ved at personer som ikke ellers ville hatt en inntekt nå får utbetalinger. Vi vil også se at personer som ellers ikke har penger til å ta utdanning nå kan gjøre dette fordi dette dekkes av staten.

---

## **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?

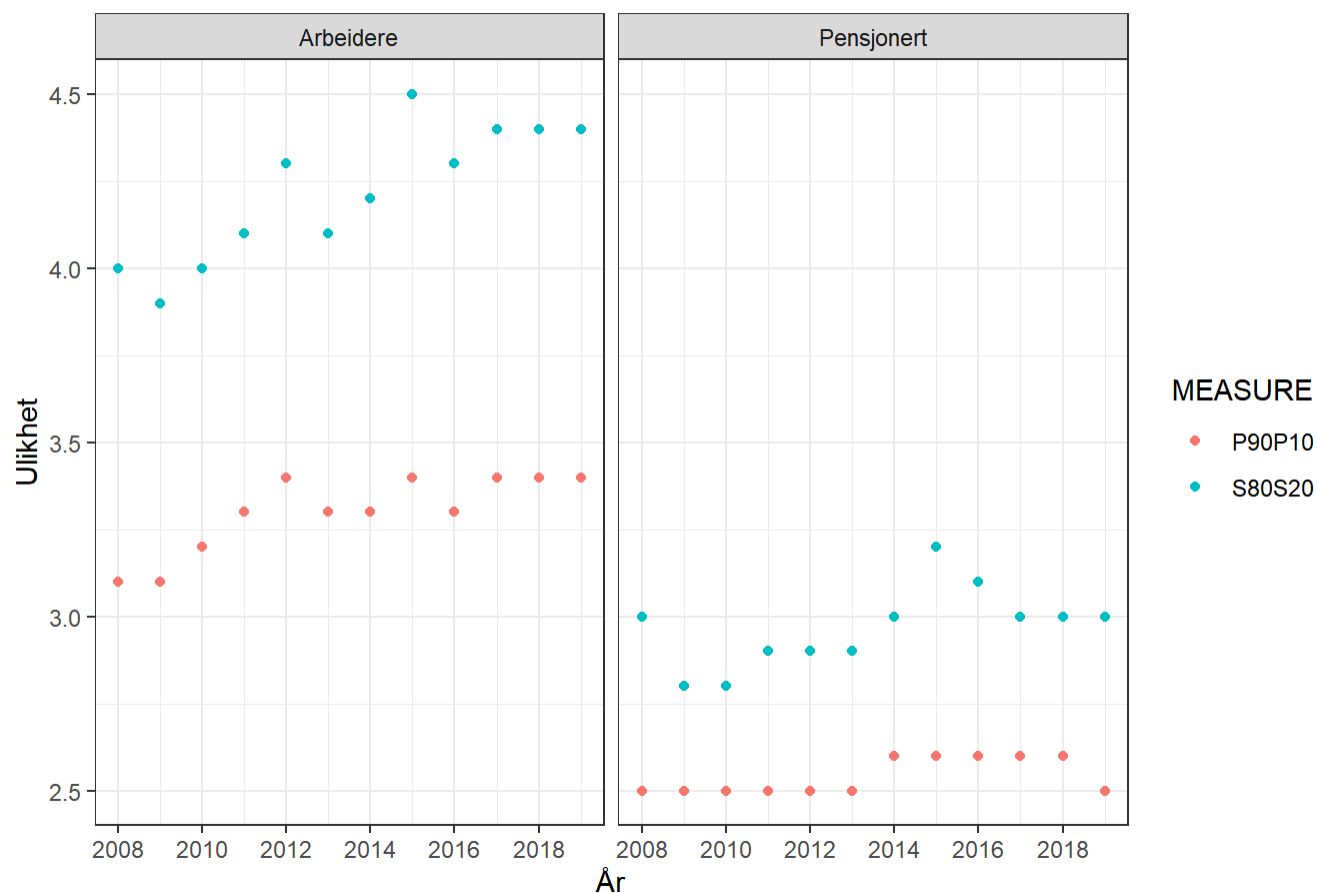
---

Her lager vi to figurer der vi bruker både P90P10 og S80S20 slik at vi ser forskjellen som vises mellom målene.

```
#Lager lister slik at det enkelt kan filtreres ut
list_P9010_S8020 = c("S80S20", "P90P10")
list_wa_old = c("WA", "OLD")

oecd_n %>% # data
  filter(MEASURE %in% list_P9010_S8020) %>% # velger ut de observasjonene vi vil ha
  filter(AGE %in% list_wa_old) %>%
  ggplot(aes(x = Year, y=Value, color=MEASURE))+ # definer aksene
  geom_point()+ # plott som punkt
  theme_bw() + #Endrer bakgrunnsfarge
  scale_x_continuous(breaks=c(2008, 2010, 2012, 2014, 2016, 2018))+ # unngår at årstall blir
gjort om til desimal
  labs (title = " P90P10 og S80S20 for Norge, 2008-2019", y = "Ulikhet", x= "År")+ # tittel p
å figuren og benevnning av aksene
  facet_wrap(~ factor (AGE, (levels=c("WA", "OLD"))), # lager to paneler etter alder i fornuf
tig rekkefølge med passende titler
            labeller = as_labeller(
              c("WA" = "Arbeidere",
                "OLD" = "Pensjonert"))) # tilpass tittel på panelene
```

P90P10 og S80S20 for Norge, 2008-2019



Er svarene dine det samme som i 3a? Hvorfor/hvorfor ikke?

---

Hvis vi sammenligner figurene ser vi at de følger hverandre parallelt med oppgang og nedgang fra perioden 2008-2019, spesielt i P90P10. Det er mer stabilitet for pensjonister enn arbeidere. I S80S20 ser vi at ulikhetene øker mer hos arbeidere ettersom den ser på ekstremiteter, da kan vi anta at de rikere har blitt betydelig mer rike enn de fattigste i landet over nevnt periode.

Det å sammenligne S80S20 P90P10 og de forskjellige GINI koeffisientene vi har brukt gir et godt sammensatt blikk på hvordan ulikhetene har forandret seg i Norge

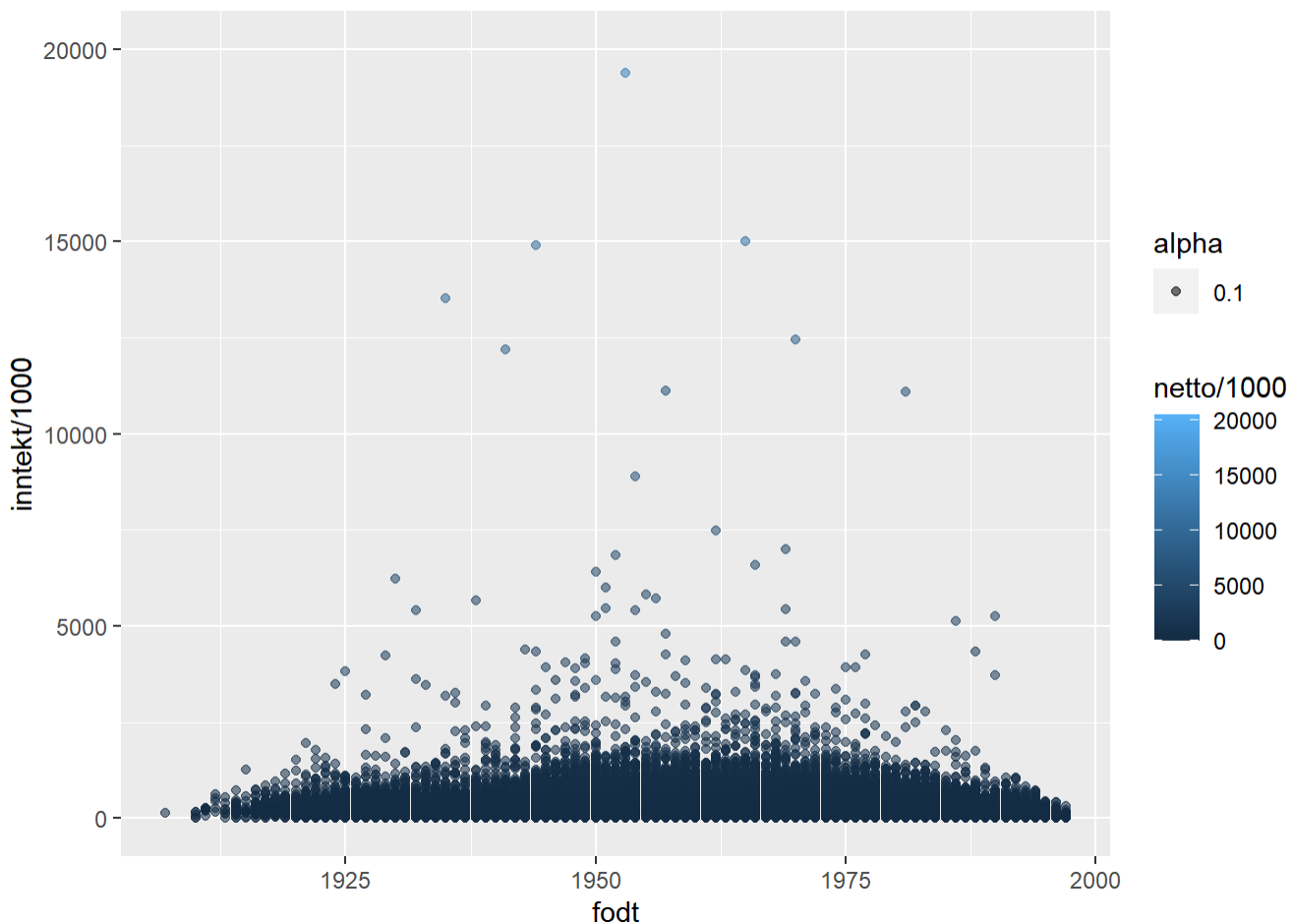
## Oppgave 4. Casestudie: Tromsø kommune 2015

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

### Oppgave 4a

Tegn et spredningsplott med fødselsår på x-aksen og inntekt (før og etter skatt) på y-aksen. Hvordan kan vi tolke dette bildet i forhold til inntektsulikhet i Tromsø kommune, og hvilken rolle spiller skattesystemet her?

```
tromso %>%  
  
  ggplot(aes(x = fodd, y = inntekt/1000, alpha = 0.1, color = netto/1000))+  
  
  geom_point()+  
  coord_cartesian(ylim = c(0, 20000))
```



I figuren ovenfor ser vi ett spredningsplott med inntekt og netto-inntekt delt på tusenvis av kroner i 2015 i y-aksen og hvilke år folk er født i x-aksen. Det vi kan lese av figuren er at majoriteten av personer i Tromsø har relativt lik inntekt men at den del personer i jevnt fordelt alder har betydelig høyere inntekt.

Skattesystemet setter sitt preg på figuren ved å gi oss inntekt etter skatt, kalt netto-inntekt. Skattesystemet gjør også at de rikeste betaler mest for at de fattigste kommer høyere ved å betale mere skatt, slik at institusjonen fordeler midlene på de svakeste i samfunnet.

## 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

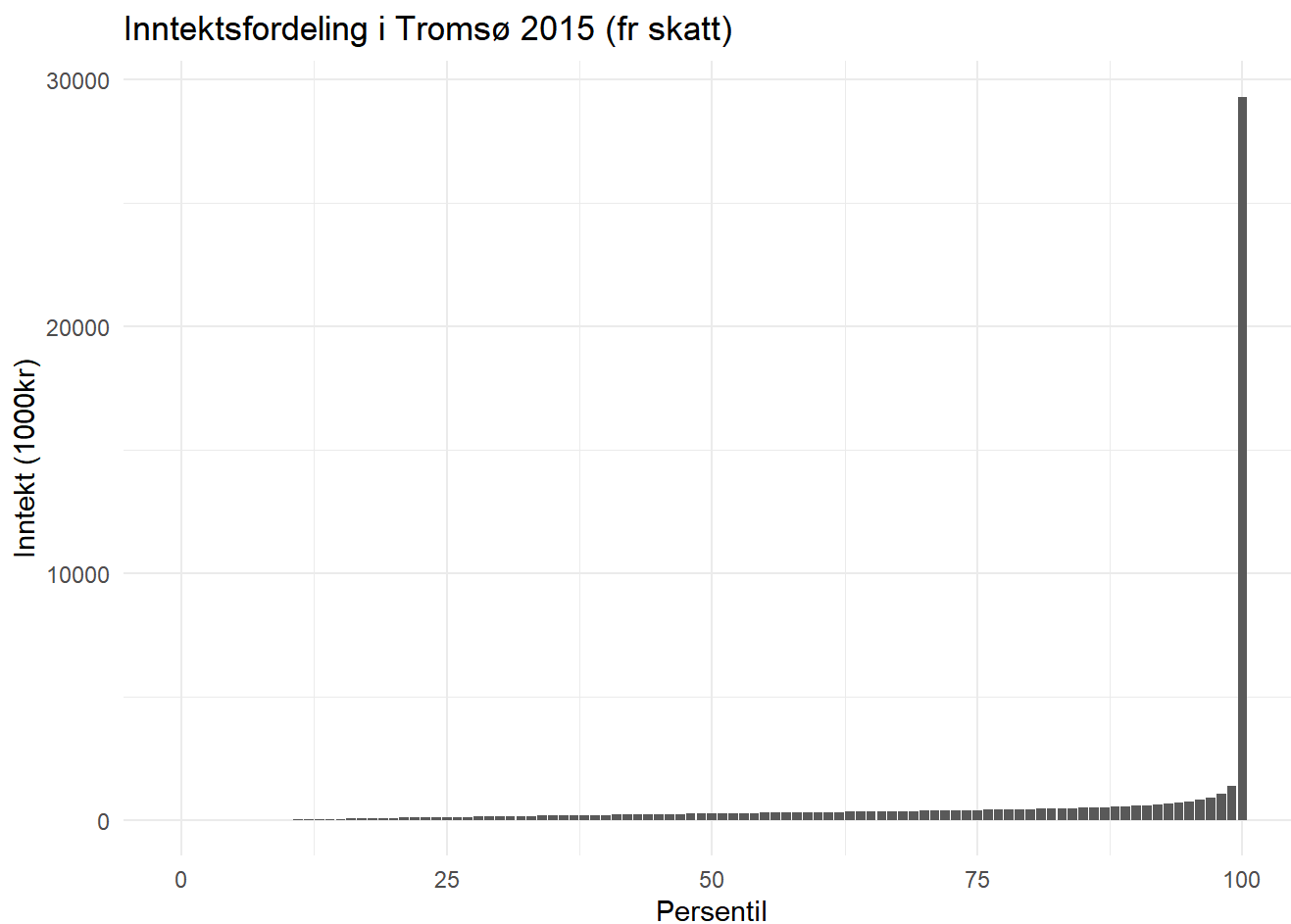
```
df <- tromso %>%
  select(inntekt) %>%
  quantile(probs = seq(0,1, .01), na.rm=TRUE)

df <- as_tibble(df)

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

df <- df %>%
  rename(inntekt_fr_skatt=value)

df %>%
  ggplot(aes(x=Persentil,y=inntekt_fr_skatt/1000)) +
  geom_col()+
  labs (title = "Inntektsfordeling i Tromsø 2015 (fr skatt)",
        y = "Inntekt (1000kr)", x= "Persentil")+
  theme_minimal()
```



I figuren over ser vi ulikhetene i inntekt for Tromsø kommune i 2015 med Y-verdi tusenvis av kr. Sett fra grafen tjener den 100 persentil særdeles mer enn resten av befolkningen, som gjør det vanskelig å lese av figuren.

Under kommer vi til å prøve å få ett mer nøyaktig bildet av inntekten til resten av befolkningen i Tromsø kommune. Det gjør vi ved å sette parameteret på Y-aksen lavere.

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

NettoTromso <- as_tibble(NettoTromso)

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

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

plottskattetter <- NettoTromso %>%
  ggplot(aes(x=Persentil,y=inntekt_etter_skatt/1000)) +
  geom_col()+
  labs (title = "Inntektsfordeling i Tromsø 2015 (etter skatt)",
        y = "Inntekt (1000kr)", x= "Persentil")+
  theme_minimal()+
  coord_cartesian(ylim = c(0, 2000))

#####

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

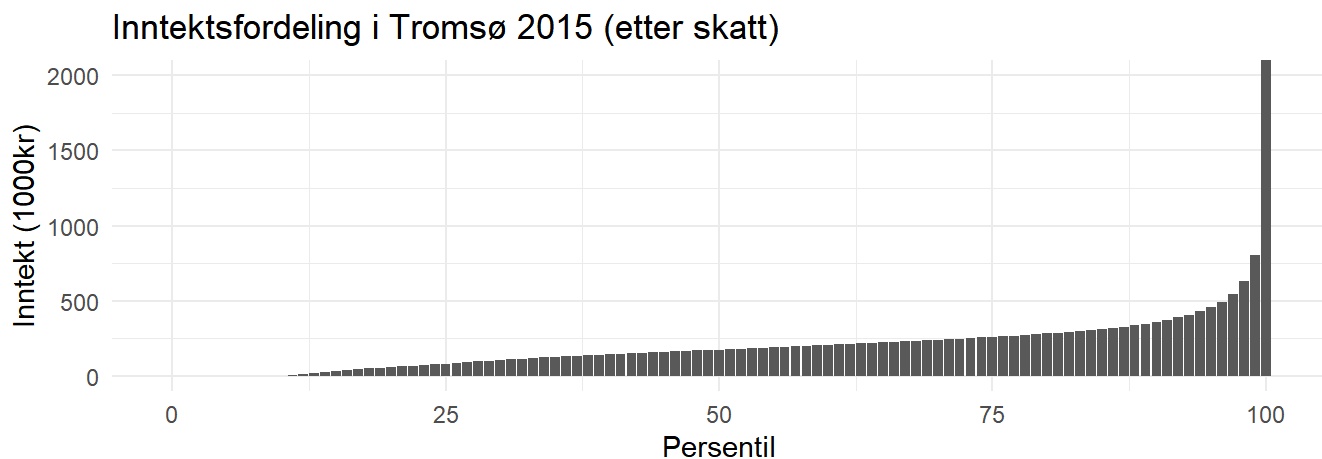
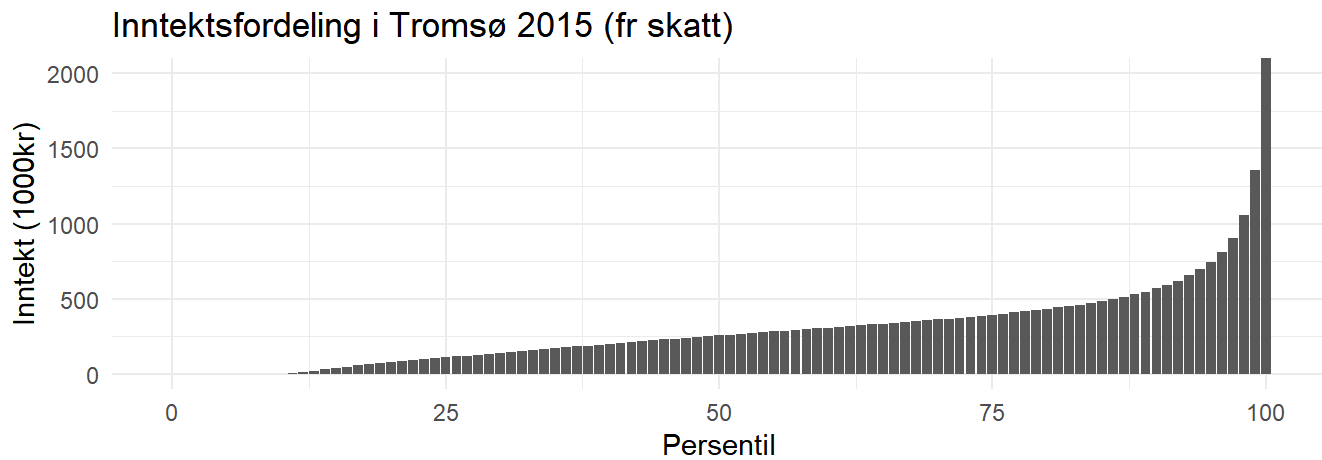
SkattTromso <- as_tibble(SkattTromso)

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

SkattTromso <- SkattTromso %>%
  rename(inntekt_fr_skatt=value)

plottskattfr <- SkattTromso %>%
  ggplot(aes(x=Persentil,y=inntekt_fr_skatt/1000)) +
  geom_col()+
  labs (title = "Inntektsfordeling i Tromsø 2015 (fr skatt)",
        y = "Inntekt (1000kr)", x= "Persentil")+
  theme_minimal()+
  coord_cartesian(ylim = c(0, 2000))

plot_grid(plottskattfr, plottskattetter, nrow=2)
```



Hvilket bilde får du av inntektsulikhhet i Tromsø kommune? Lag en tilsvarende figur for inntekt etter skatt og kommentér og forklar eventuelle forskjeller.

---

I grafene ovenfor kan vi se inntektsulikheten i Tromsø kommune i 2015 før og etter skatt tusenvis av kroner. I disse grafene er Y-aksen satt ned for å kunne se nærmere på inntektsulikheten. Merk at søylen for dem som tjente mest går forbi høyeste verdi i Y-aksen og kan ikke leses av i disse grafene. Når vi ser på disse grafene kan vi se at det er en jevn stigning mellom persentilene. Vi kan også se at det er i overkant av ti prosent av befolkningen som er uten inntekt. Når vi sammenlikner grafene kan vi se at inntektsfordelingen er parallell med hverandre, men verdien etter skatt er lavere enn før skatt.

---

## 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](https://cran.r-project.org/web/packages/gglorenz/gglorenz.pdf) (<https://cran.r-project.org/web/packages/gglorenz/gglorenz.pdf>). Her er et eksempel hvor jeg har brukt Harstad kommune som grunnlag (kommunenr 1903):

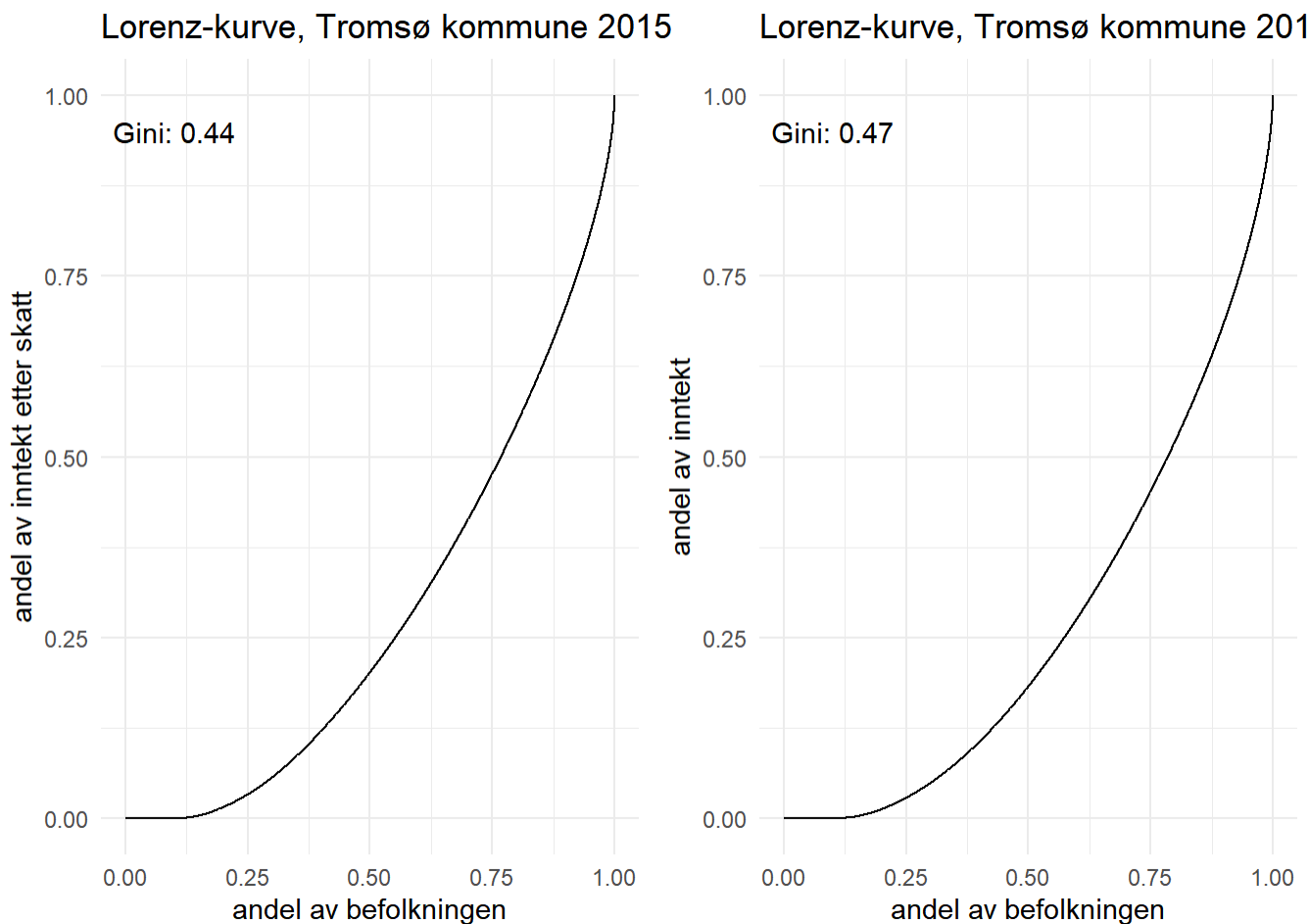
Tegn to Lorenz-kurver for Tromsø kommune (før og etter skatt), samt tilhørende Gini-koeffisienter. Sammenlikn med de relevante Gini-koeffisientene for hele Norge fra 2015 fra OECD datasettet. Er inntektsfordelingen i Tromsø mer eller mindre lik sammenliknet med landet som helhet? Gi mulige forklaringer.

```
gini_for_skatt <- ggplot(tromso, aes(inntekt))+
  stat_lorenz()+ # tegner Lorenz
  annotate_ineq(tromso$inntekt)+ # beregner Gini
  labs (x="andel av befolkningen", y= "andel av inntekt", title = "Lorenz-kurve, Tromsø kommu
ne 2015")+
  theme_minimal()

##

gini_etter_skatt <- ggplot(tromso, aes(netto))+
  stat_lorenz()+ # tegner Lorenz
  annotate_ineq(tromso$netto)+ # beregner Gini
  labs (x="andel av befolkningen", y= "andel av inntekt etter skatt", title = "Lorenz-kurve,
Tromsø kommune 2015")+
  theme_minimal()

plot_grid(gini_etter_skatt, gini_for_skatt)
```



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



---

Inntektsulikheten er mindre lik i Tromsø sammenliknet med hele Norge. Som vi kan se i figuren over er gini tallene før og etter skatt ganske lik hverandre, men i datasettet til oecd for Norge kan vi se at ginitallet er 0.272, som er betydelig lavere enn for Tromsø.

En av grunnen til dette er at, befolkningstallet er ganske lite i Tromsø sammenlignet med hele Norge, det gjør at sammenlikningsgrunnlaget for gini-koeffisienten blir mer vannet ut.

---