

SOK-1004, høst 2022, Prosjektoppgave

19, 167, 219

Even Hvinden

Oppklarende spørsmål kan rettes til even.c.hvinden@uit.no. Jeg retter en stor takk til Jørgen Modalsli for å ha delt data på inntektsfordeling.

Instruksjoner

Prosjektoppgaven skal leveres som `.pdf` i WiseFlow. Det ferdige dokumentet lagres som `[kandidatnummer]_SOK1004_PO_H22.pdf`. Hvis du har kandidatnummer 43, så vil filen hete `43_SOK1004_PO_H22.pdf`. Det er ingen begrensning på hvilken programvare dere benytter til å lage besvarelsen. Men figurer og utregninger skal gjøres i RStudio og dokumenteres ved hjelp av en `.qmd`-fil som kompiles som en `.pdf` og inkluderes som vedlegg i besvarelsen.

Merk at prosjektoppgaven består av tre oppgaver. Vekttallene er oppgitt i prosent.

Prosjektoppgaven kan besvares alene eller i grupper på opptil tre personer. Det er ikke lov å samarbeide på tvers av grupper, og hvert gruppemedlem må levere besvarelsen i WiseFlow. Grupper anbefales å undertegne en [samarbeidsavtale](#).

NB! Besvarelser som ikke oppfyller disse kravene vil ikke bli vurdert.

Last inn pakker og data

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

```
-- Attaching packages ----- tidyverse 1.3.2 --  
v ggplot2 3.4.0      v purrr   0.3.4  
v tibble  3.1.8      v dplyr   1.0.10  
v tidyr   1.2.1      v stringr 1.4.1
```

```
v readr 2.1.2 v forcats 0.5.2
```

Warning: package 'ggplot2' was built under R version 4.2.2

```
-- Conflicts ----- tidyverse_conflicts() --  
x dplyr::filter() masks stats::filter()  
x dplyr::lag() masks stats::lag()
```

```
library(ggthemes)
```

Warning: package 'ggthemes' was built under R version 4.2.2

```
library(openxlsx)
```

Warning: package 'openxlsx' was built under R version 4.2.2

```
library(ggpubr)
```

Warning: package 'ggpubr' was built under R version 4.2.2

```
url <- "https://modals.li/downloads/gini_series_norway.xlsx"
```

```
df_gini <- url %>%  
  read.xlsx(sheet=1) %>%  
  as_tibble()
```

```
df_gini_navn <- df_gini %>%  
  names()
```

```
df_gini <- df_gini %>%  
  pivot_longer(!year,  
               names_to = "var",  
               values_to = "verdi")
```

```
url <- "https://github.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/blob/main/rkode/inn"
```

```

df_andeler <- url %>%
  read.xlsx() %>%
  as_tibble()

df_andeler_navn <- df_gini %>%
  names()

df_andeler <- df_andeler %>%
  pivot_longer(!year,
               names_to = "var",
               values_to = "verdi")

df_andeler <- df_andeler %>%
  filter(year>=1875)

url <- "https://github.com/uit-sok-1004-h22/uit-sok-1004-h22.github.io/blob/e36b270c865e07

df_bnp <- url %>%
  read.xlsx(sheet=1) %>%
  as_tibble()

navn_BNP <- df_bnp %>%
  names()

df_bnp <- df_bnp %>%
  rename(BNP = navn_BNP[2],
         year = navn_BNP[1])

df_bnp <- df_bnp %>%
  mutate(DLBNP = 100*(log(BNP)-lag(log(BNP))))

df_bnp <- df_bnp %>%
  filter(year >= 1875)

```

Bakgrunn

I norsk økonomisk historie blir årene 1920 - 1930 beskrevet som harde. Les avsnittene om arbeidsløshet og inflasjon, paripolitikken, streiker og arbeidskonflikter i [artikkelen om mellomkrigstiden](#) i Store norske leksikon. Se også [deres artikkel om paripolitikk](#).

I denne prosjektoppgaven skal vi analysere utviklingen i norsk økonomi gjennom historiske data

fra Norges Bank og Statistisk sentralbyrå. Vi vil også bruke data fra artikkelen [Estimating long-run income inequality from mixed tabular data: Empirical evidence from Norway, 1875-2017](#) av Jørgen Modalsli, Rolf Aaberge, og Anthony B. Atkinson, publisert i *Journal of Public Economics* i 2021.

Bedømmelse

Formålet med oppgave er å utfordre dere til å gi en deskriptiv analyse av sammensatte økonomiske data gjennom visualisering og grunnleggende økonomisk teori. En god besvarelse er kjennetegnet av følgende trekk.

1. Egenskapene til data visualiseres i gode figurer. En god figur har presist definerte enheter, fornuftig skalering, og illustrerer tydelig relevante trekk ved dataene. Bedømmelsen vil vektlegge en tydelig beskrivelse av figurene. Figurer bør ikke overbelastes med innhold og kan gjerne være vakre.
2. Besvarelsen svarer konkret på oppgaven med et konsist og presist språk. Påstander om årsakssammenhenger underbygges med argumenter og kildehenvisninger.
3. Koden som lagde figurene er korrekt og godt dokumentert.

Oppgave I - 20%

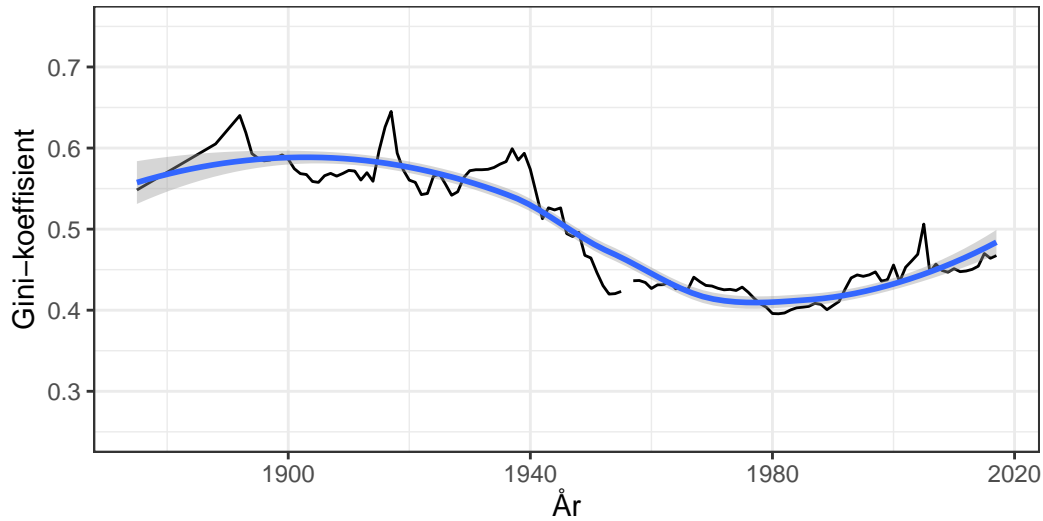
Betrakt følgende figurer.

```
df_gini %>%  
  filter(year >= 1875) %>%  
  filter(var %in% df_gini_navn[4]) %>%  
  ggplot(aes(x=year,y=verdi)) +  
  geom_line() +  
  geom_smooth() +  
  ylim(0.25,0.75) +  
  theme_bw() +  
  labs(title = "Gini-koeffisienter for inntektsfordelingen i Norge",  
        subtitle = "1875-2017. Gjennomsnitt av øvre og nedre grense.",  
        y="Gini-koeffisient",  
        x="År",  
        caption = "Kilde: Modalsli, Aaberge og Atkinson (2021)")
```

```
`geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).

Gini-koeffisienter for inntektsfordelingen i Norge 1875–2017. Gjennomsnitt av øvre og nedre grense.



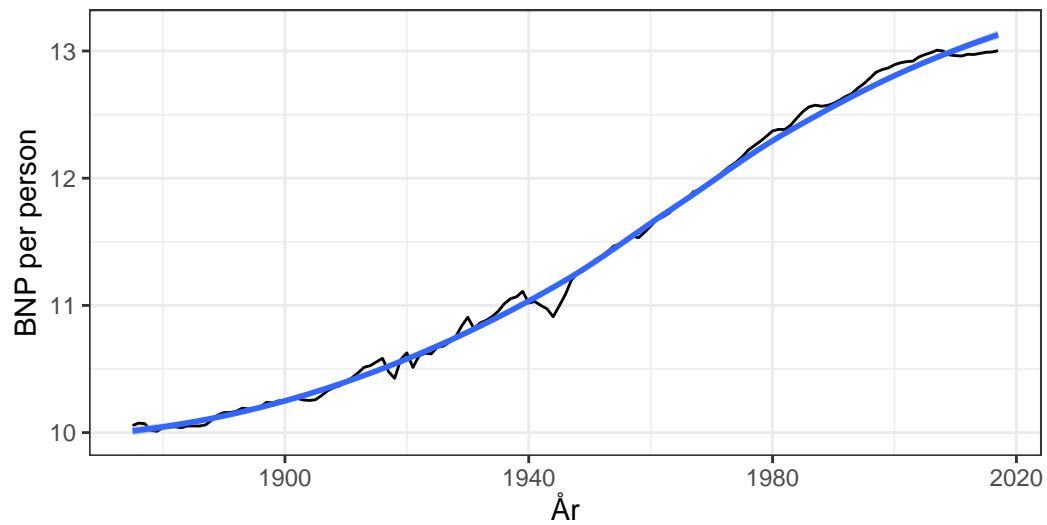
Kilde: Modalsli, Aaberge og Atkinson (2021)

```
df_bnp %>%  
  filter(year >= 1875) %>%  
  mutate(BNP = log(BNP)) %>%  
  ggplot(aes(x=year,y=BNP)) +  
    geom_line() +  
    geom_smooth() +  
    theme_bw() +  
    labs(title = "BNP per person",  
         subtitle = "Log faste-2005 kroner. 1875-2017.",  
         y="BNP per person",  
         x="År",  
         caption = "Kilde: Norges Bank")
```

`geom_smooth()` using method = 'loess' and formula = 'y ~ x'

BNP per person

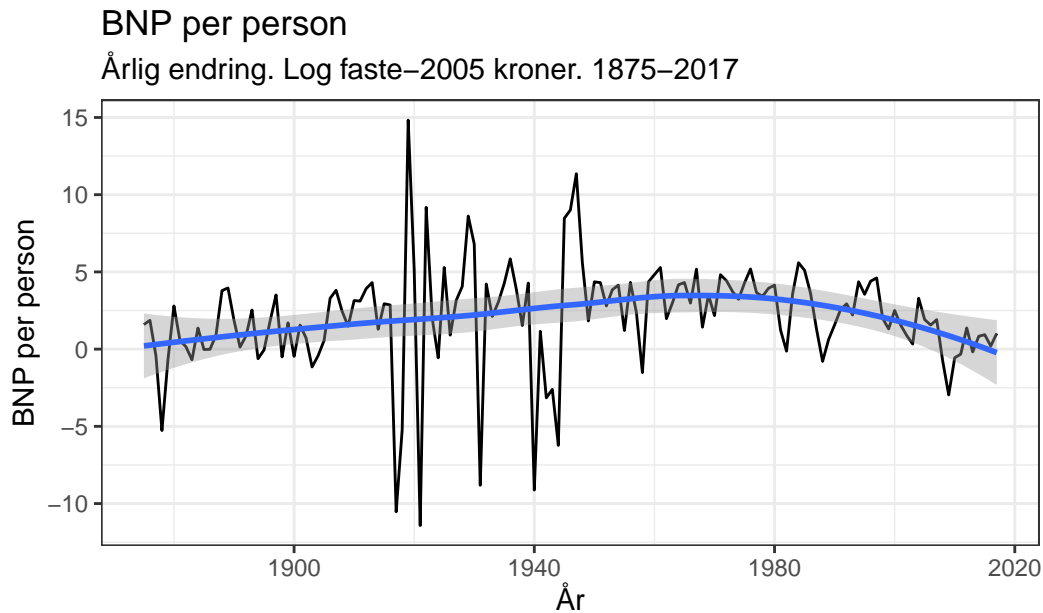
Log faste-2005 kroner. 1875-2017.



Kilde: Norges Bank

```
df_bnp %>%  
  filter(year >= 1875) %>%  
  ggplot(aes(x=year,y=DLBNP)) +  
    geom_line() +  
    geom_smooth() +  
    theme_bw() +  
    labs(title = "BNP per person",  
         subtitle = "Årlig endring. Log faste-2005 kroner. 1875-2017",  
         y="BNP per person",  
         x="År",  
         caption = "Kilde: Norges Bank")
```

`geom_smooth()` using method = 'loess' and formula = 'y ~ x'



Spørsmål: Beskriv hvordan ulikheten og veksten i inntekter utviklet seg i 1920-årene sammenlignet med øvrige perioder.

Hint. Det kan være nyttig å endre koden for å se nærmere på bestemte tidsintervaller.

Merknad: Anta at estimatet av BNP og Gini-koeffisienten i figuren er sann. Det er ikke nødvendig å diskutere metoden Norges Bank og Modalsli, Aaberge og Atkinson (2021) har brukt til å estimere BNP og Gini-koeffisienten eller den statistiske usikkerheten rundt estimatet.

Besvarelse oppgave I

Mellom-krigstiden var en økonomisk volatil periode for Norge med høy inflasjon og arbeidsløshet, samtidig som det var en økende mengde med tvangsauksjoner og bankkrise hvor mange banker og kommuner med gjeld ble satt under administrasjon.

Dette var følger av at under den første verdenskrig så økte ulikhetene kraftig og for å deretter falle raskt ned igjen etter krigen. Dette kan vi se i den ene grafen, at BNP per person hadde et dropp etter krigen (den kraftige inflasjonen var over, priser falt kraftig, produksjonsinnskrenkninger, omsetninger sank, mye arbeidsløshet.)

Så i grunn har ikke inntekten hatt noe særlig stor vekst i 1920-tallet, sammenlignet med de øvrige periodene. Samtidig ser vi i grafen for Gini-koeffisienter at inntektsfordelingene minsker over tid, noe som er med på å senke ulikhetene. Men begynner ikke å få en jevn nedgang før

1930-tallet der det går over til en noe mer aggressiv nedgang. Dette kommer av den spesielle økonomiske situasjonen under mellom-krigstiden.

Oppgave II - 20%

I `df_andeler` er det lagret inntektsandeler fordelt på desil og år. Merk at dette er tallene bak Figur A9 i vedleggene til MAA, tilgjengelig [her](#). For eksempel viser `share_10_20` hvor stor andel av den totale inntekten som gikk til intervallet mellom tiende- og tyvende persentil. I følgende figur kombinerer jeg BNP-tallene fra oppgave I med inntektsandelen til å beregne gjennomsnittlig inntekt per person, fordelt på desil og år.

Betrakt følgende figurer.

Merknad. Inntektene må skaleres når vi beregner andeler. La $a_{x-y,t}$ være inntektsandelen til intervallet mellom x og y i år t . BNP per person i år t er gitt ved bnp_t og er lagret i `df_bnp`. For eksempel så vil $a_{50-60,2005}$ og bnp_{2005} være gitt ved `df_andeler$share_50_60` og `df_bnp$BNP` når `year` er lik 2005. Da vil gjennomsnittlig inntekt i intervallet være gitt ved

$$a_{x-y,t} \times \frac{100}{x-y} \times bnp_t.$$

```
# ny tabell hvor vi skal beregne inntekter
df <- df_andeler

# lagrer en liste med årene hvor vi har observasjoner med inntektsandeler
years_list <- df$year %>%
  unique()

# wide-format er enklere
df <- df %>%
  pivot_wider(names_from = var, values_from = verdi)

# ny tabell med bnp-tall for årene hvor vi har observasjoner
df_temp <- df_bnp %>%
  filter(year %in% years_list)

# anstendige navn
df <- df %>%
  rename("0-10" = "share_0_10",
         "10-20" = "share_10_20",
         "20-30" = "share_20_30",
         "30-40" = "share_30_40",
```



```

    "40-50" = "share_40_50",
    "50-60" = "share_50_60",
    "60-70" = "share_60_70",
    "70-80" = "share_70_80",
    "80-90" = "share_80_90",
    "90-100" = "share_90_100",
    "90-99" = "share_90_99",
    "99-100" = "share_99_100")

# lagrer navn
df_navn <- df %>%
  names()

df <- df %>%
  mutate(`0-10` = 10 * `0-10` * df_temp$BNP,
    `10-20` = (100/10) * `10-20` * df_temp$BNP,
    `20-30` = (100/10) * `20-30` * df_temp$BNP,
    `30-40` = (100/10) * `30-40` * df_temp$BNP,
    `40-50` = (100/10) * `40-50` * df_temp$BNP,
    `50-60` = (100/10) * `50-60` * df_temp$BNP,
    `60-70` = (100/10) * `60-70` * df_temp$BNP,
    `70-80` = (100/10) * `70-80` * df_temp$BNP,
    `80-90` = (100/10) * `80-90` * df_temp$BNP,
    `90-100` = (100/10) * `90-100` * df_temp$BNP,
    `90-99` = (100/9) * `90-99` * df_temp$BNP,
    `99-100` = (100/1) * `99-100` * df_temp$BNP,)

df <- df %>%
  pivot_longer(!year,
    names_to = "var",
    values_to = "verdi")

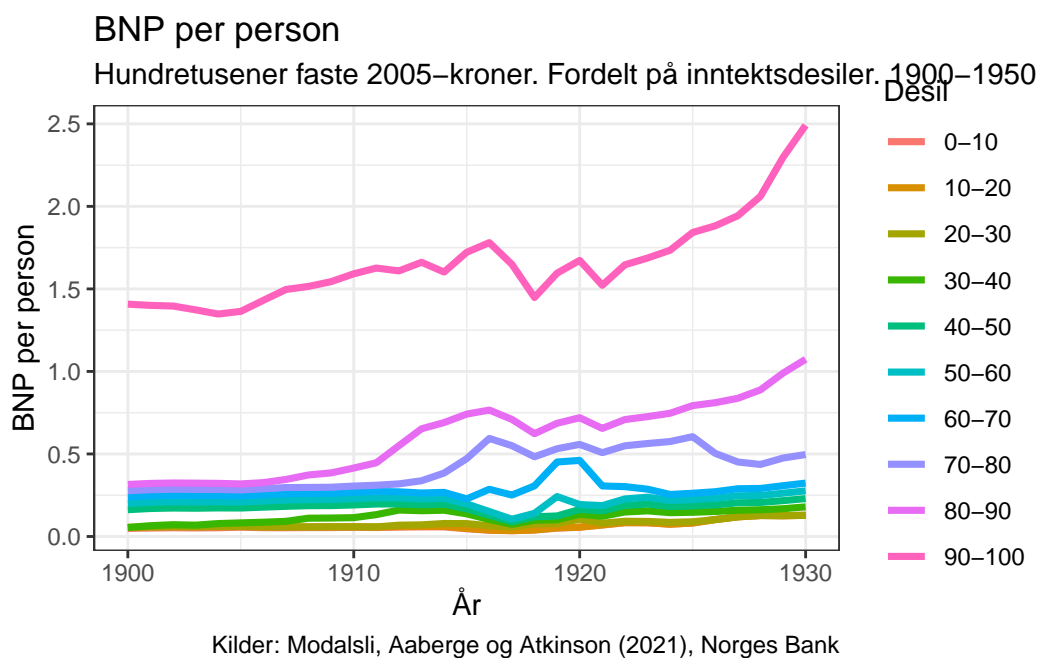
andeler <- df$var %>%
  unique()

df %>%
  filter(year %in% c(1900:1930)) %>%
  filter(var %in% andeler[1:10]) %>%
  mutate(verdi=verdi/(10^5)) %>%
  ggplot(aes(x=year,y=verdi,color=var)) +
  geom_line(size=1.25) +

```

```
labs(title = "BNP per person",
      subtitle = "Hundretusener faste 2005-kroner. Fordelt på inntektsdesiler. 1900-1950",
      y = "BNP per person",
      x = "År",
      color = "Desil",
      caption = "Kilder: Modalsli, Aaberge og Atkinson (2021), Norges Bank") +
theme_bw()
```

Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
 i Please use `linewidth` instead.

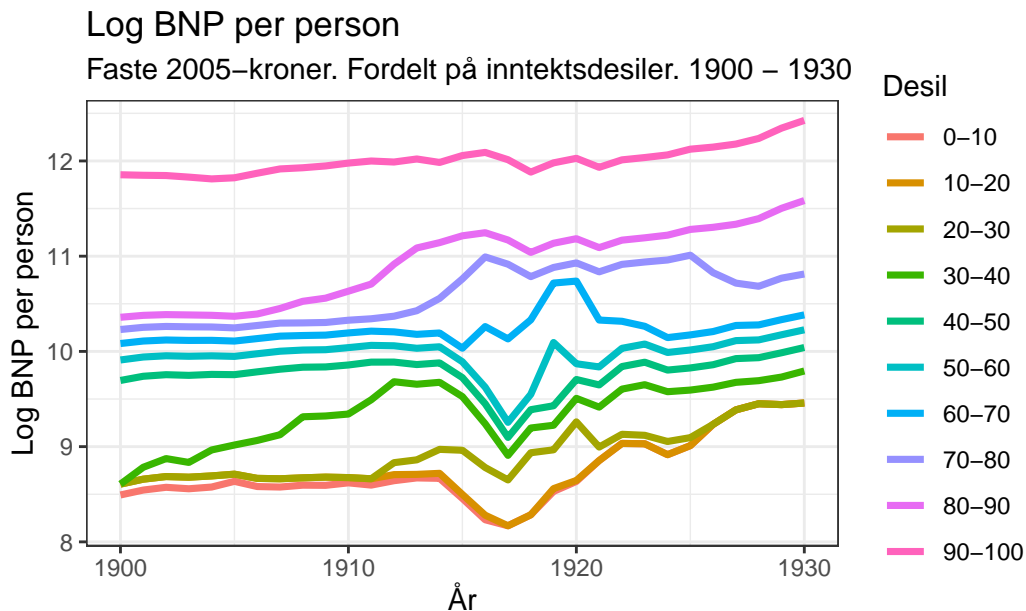


```
df %>%
  filter(year %in% c(1900:1930)) %>%
  mutate(verdi=log(verdi)) %>%
  filter(var %in% andeler[1:10]) %>%

ggplot(aes(x=year,
            y=verdi,
            color=var)) +
```

```
geom_line(size=1.25) +

labs(title = "Log BNP per person",
      subtitle = "Faste 2005-kroner. Fordelt på inntektsdesiler. 1900 - 1930",
      y = "Log BNP per person",
      x = "År",
      color = "Desil",
      caption = "Kilder: Modalsli, Aaberge og Atkinson (2021), Norges Bank") +
theme_bw()
```



Kilder: Modalsli, Aaberge og Atkinson (2021), Norges Bank

Spørsmål: Beskriv hvordan endringen i reelle inntekter fordelte seg på inntektsdesiler mellom 1920 og 1930.

Besvarelse oppgave II

Man ser at endringene i de øverste to desilene skiller seg ut fra resten der de har hatt en mye større økning mellom 1920 og 1930 i forhold til hva de andre har hatt. Alle har hatt vekst, men de laveste åtte desilene har hatt en relativ lavere vekst. Noe som kan pekes ut er at de lavere desilene er mer sårbare for økonomiske svingninger i motsetning til de øverste.

Vi ser at de laveste desilene under den første verdens krig ble mye kraftigere påvirket av den økonomiske situasjonen samtidig som det tok dem mye lengre tid å hente seg inn igjen i forhold

til de øvrige. Noe vi konkluderte med tidligere med tanke på økonomisk sårbarhet. De øverste har heller ikke hatt særlig stor endring og holdt seg stødig både før og etter krigen, noe som viser oss at de har nyttet godt av økonomisk stabilitet.

Måtene desilene er fordelt på sier oss at den øverste desilen er den eneste som kontinuerlig hatt en BNP på over 100 000 siden 1900. Dette endrer seg ikke før i slutten av 1920-tallet da den nest øverste desilen har hatt noe vekst og krysser 100 000 grensen rundt ca. 1927/1928. Annet enn dette ser vi at den tredje høyeste desilen har noe vekst som igjen faller i midten av 1920-tallet og smålig begynner å oppnå vekst igjen mot midten av tiåret.

Oppgave III - 60%

Dere skal nå presentere og beskrive utviklingen i to ytterligere økonomiske forhold som kjennetegnet de harde 1920-årene. Jeg anbefaler å bruke følgende kilder til å finne data:

- [Historisk statistikk fra Norges Bank](#)
- [Historisk statistikk fra SSB](#)

Her følger fire eksempler på tema dere kan, men ikke nødvendigvis må velge. Lenkene leder til definisjonen i SNL, og dataene finnes i de ovennevnte kildene.

- [Realrente](#)
- [Utlån \(kreditt\) fra banker](#)
- [Arbeidsledighet](#)
- [Arbeidskonflikter og streik](#)

I besvarelsen skal dere forklare hvorfor dataene er relevante og gi en grundig beskrivelse av deres utvikling. Lengden på besvarelsen skal være på 750 - 1500 ord.

Besvarelse Oppgave III

I denne oppgaven har vi valgt ut temaene arbeidsledighet og nominell årlig inntekt. For sistnevnte har vi valgt ut tall fra noen industrier. Dette er sentrale industrier som gir en god oversikt uten at det blir rotete i figuren. Disse to temaene mener vi gir et godt innblikk i den økonomiske tilstanden i den 20 års perioden vi har tatt for oss.

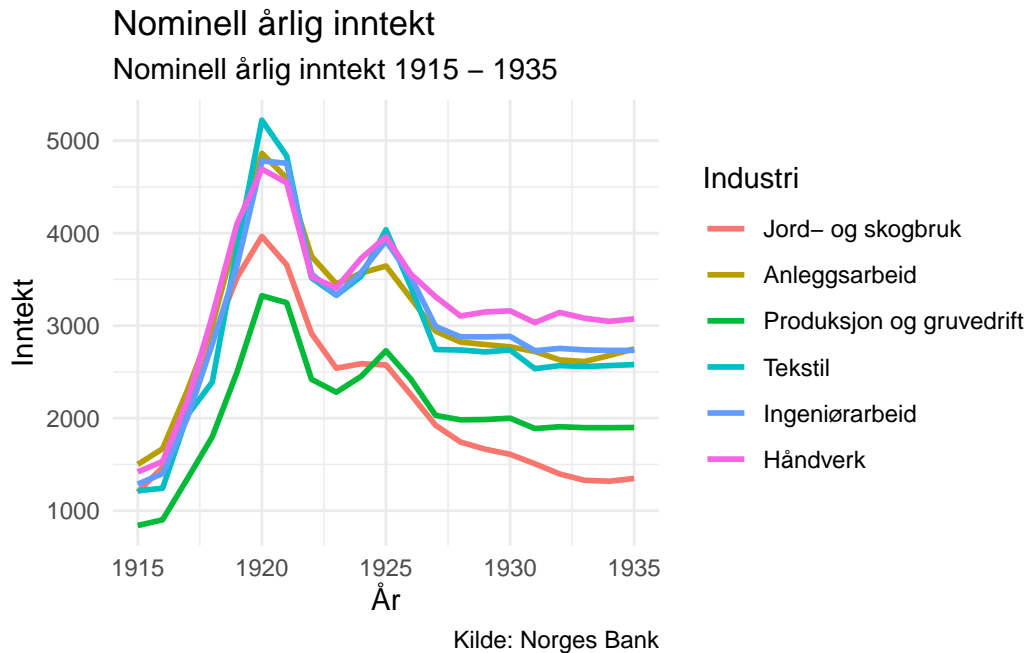
Nominell årlig inntekt

```
#Henter inn xlsx fila fra Norges Bank
url <- "https://www.norges-bank.no/globalassets/upload/hms/data/wages_ind.xlsx"

#Lagrer de ønskede radene/kolonnene i et datasett.
nom_annual_wages <- url %>%
  read.xlsx(sheet=2, rows=c(3,193:213), cols=c(1,3,5:9)) %>%
  as_tibble() %>%
  rename(year = Year)

#Lager et nytt datasett med pivot_tall slik at vi kan bruke alle kolonnene inne på samme f
nom_annual_wages_tall <- nom_annual_wages %>%
  pivot_longer(!year, names_to="industri", values_to="verdi")

#Plotter inn figuren
nom_annual_wages_tall %>%
  ggplot(aes(x=year,y=verdi,color=industri)) +
  geom_line(size=1) +
  scale_color_discrete(labels = c("Jord- og skogbruk","Anleggsarbeid","Produksjon og gruvene")) +
  labs(title = "Nominell årlig inntekt",
        subtitle = "Nominell årlig inntekt 1915 - 1935",
        y = "Inntekt",
        x = "År",
        color = "Industri",
        caption = "Kilde: Norges Bank") +
  theme_minimal()
```



Etter å ha importert den nominelle årlige inntekten for industrien i Norge fra Norgesbank, så har vi valgt og vist den i årene 1915 til 1935. Dette er for å få med litt data fra før og etter 1920-tallet, for å sette ting mer i perspektiv. Dette gir oss mer data å jobbe med samtidig som vi får et bredere perspektiv på hvordan den nominelle årlige inntekten for de forskjellige industriene i Norge var på 1920-tallet.

Vi kan starte med å se på årene litt før 1920, fra 1915 av så ser vi at den nominelle årlige inntekten for industriene i Norge er i kraftig vekst, selv om at det er krig i verden så fortsetter den. Den fortsetter å være i vekst helt fram til etter krigen og begynner svakt å avta i 1920. Når vi kommer til 1921 og 1922 så begynner den å synke kraftig. Dette droppet i inntektsvekst henter seg noe inn og det begynner å øke i vekst for noen industrier igjen, men dette er kortvarig. For i 1925 igjen, er det et kraftig fall i inntekten til de ansatte over alle de forskjellige industriene. Den årlige inntekten til de fleste i de forskjellige industriene vil ikke stoppe å falle før omtrent halvveis inn i 1925. Alle utenom én industri, agrikultur og skogbruks- industrien er den store taperen etter dette økonomiske kaoset.

Grafene setter jo ting veldig godt i perspektiv på hvor hardt den årlige nominelle lønnen til arbeiderne i disse industriene ble rammet. Under denne paragrafen har vi tatt tallene fra 1920 og trukket de fra med tallene fra 1930 for å finne differansen fra fallet i denne tiårs perioden. Deretter har vi tatt differansen og brukt den til å finne den prosentvise endringen i den årlige nominelle inntekten.

I de respektive industriene så er dette de prosentvise endringene i årlig nominell inntekt fra høyest til lavest endring:

Jord- og skogbruk har et fall på: **-59.41 %**

Ingeniørarbeid har et fall på: **-47.59 %**

Konstruksjon har et fall på: **- 43.02 %**

Tekstil har et fall på: **- 39.82 %**

Produksjon og gruvedrift har et fall på: **-39.66 %**

Håndverk har et fall på: **- 32.66 %**

I gjennomsnitt så har den årlige nominelle inntekten for arbeiderne i disse industriene gått ned med **-43.69%**

Disse prosentvise endringene er regnet ut med de årlige nominelle inntektene til arbeiderne i de respektive industriene nevnt ovenfor og viser oss den faktiske endringen i prosent for deres lønn. Så gjennomsnittlig har deres lønn nesten halvert på dette tiåret. Justert opp mot inflasjonen så er nok tapet i verdi enda høyere enn tallene som ligger i datasettet viser.

Arbeidsledighet

```
#Ettersom tallene vi skal bruke er fra et digitalisert gammelt dokument må vi skrive inn t
year <- c(1915, 1916, 1917, 1918, 1920, 1921, 1922, 1923, 1924, 1924, 1925, 1926, 1927, 19

unemployment <- c(2.1, 0.8, 0.9, 1.4, 1.6, 2.3, 17.6, 17.1, 10.6, 8.5, 13.2, 24.3, 25.4, 1

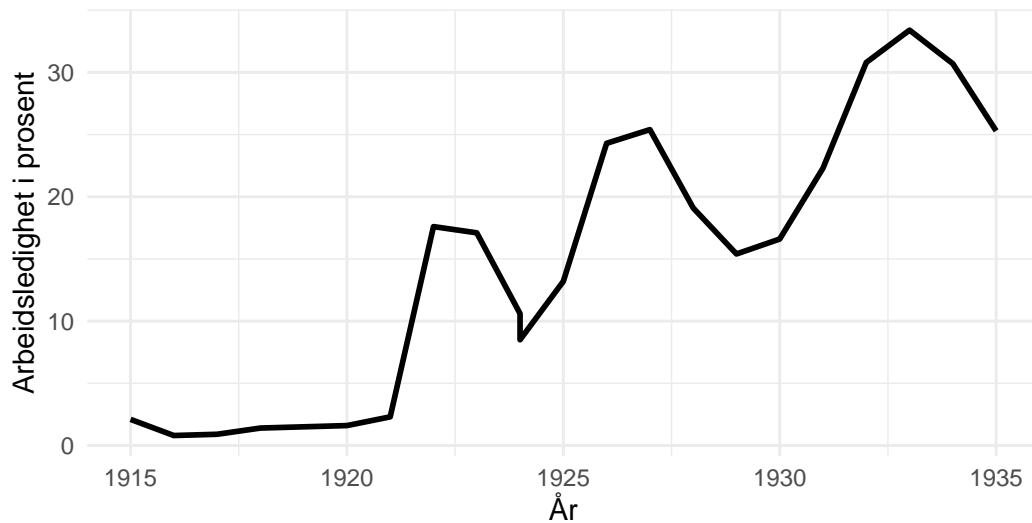
#Lager et datasett med de nye variablene
unemployment <- data.frame(year, unemployment)

#Plotter inn figuren
unemployment %>%
  ggplot(aes(x=year,y=unemployment)) +
  geom_line(size=1) +
  labs(title = "Arbeidsledighet",
        subtitle = "Arbeidsledighet i prosent. 1915 - 1935",
        y = "Arbeidsledighet i prosent",
        x = "År",
        caption = "Kilde: Norges offisielle statistikk XII 291. Tabell 47") +

theme_minimal()
```

Arbeidsledighet

Arbeidsledighet i prosent. 1915 – 1935



Kilde: Norges offisielle statistikk XII 291. Tabell 47

På samme måte som vi har gjorde med den Nominelle årlige inntekten til industriarbeidere så har vi igjen valgt og vist den i årene 1915 til 1935.

Før 1920, så nøy Norge godt av lav arbeidsløshet og god økonomisk tilstand, noe som raskt ville endre seg på 20- tallet. Krisen som rammet Norge og verden på 1920-tallet kom for fullt og skapte store problemer som prisfall, over investering, og problemer med å selge varer. Norges store handelsflåte ble kuttet ned med en tredjedel etter å ha blitt lagt til opplag. Næringer og individer slet med å få betalt ned renter og avdrag, som igjen førte til store tap for bankene. Reserver ble brukt opp og de fleste banker ble tatt på senga og var på brinken av konkurs.

I grafen ser vi at arbeidsledigheten økte med omtrent 15% fra 1921-1922, altså fra ca. 2% til 17%. Arbeidsløsheten ville igjen gå ned i 1923, men denne nedgangen ville være kortvarig ettersom at det traff en ny bølge i 1924 hvor arbeidsløsheten igjen økte og ville fortsette fram til 1927. Økningen i arbeidsløshet fra 1924-1927 var på 14.8% og var i 1927 på 25.4%. Årene etter ville tallet gå ned og synke til 15.4% i 1929. Nedgangen i arbeidsløshet ville deretter stoppe og begynne å øke igjen inn i 1930- tallet.

Her har vi plottet det samlede gjennomsnittet for industriene og plottet det over den prosentvise arbeidsledigheten

Hvis man sammenligner grafene så ser vi at den økonomiske tilstanden på 20- tallet traff flere forskjellige økonomiske og sosiale områder. Noe som førte til en svekket nasjonal økonomisk

handlekraft. Mye av dette stammer fra egenproduserte problemer med regjeringens paripolitikk. Som omhandlet at myndighetene ville knytte pengeverdien til den norske kronen til gull, for å få kronen tilbake til førkrigsverdi og samtidig oppnå en stabil valuta.

Dataene her er relevant da det forteller oss hvor hard den økonomiske situasjonen i Norge var på denne tiden, samtidig som det gir en pekepinnen til hvordan den var for verden på 1920-tallet. Denne langvarige økonomiske krisen ville sette den Norske økonomiske veksten på pause og holde den tilbake i ca. 15 år.

Referanseliste

Tvedt, Knut Are. (2009, 14. februar). *Norges historie fra 1905 til 1939*. Store Norske Leksikon. https://snl.no/Norges_historie_fra_1905_til_1939

Dørum, Knut. (2014, 22. juli). *Norsk industrihistorie før 1940*. Store Norske Leksikon. https://snl.no/norsk_industrihistorie_f%C3%B8r_1940

<http://www.kildenett.no/portal/artikler/2008/1225702044.27> (Siden ble tatt ned/var ikke tilgjengelig da vi skulle gjøre oss ferdig med selve referanselisten. Vi innser at dette er et problem, men den har gitt oss info og vi føler det er bedre å inkludere kilden i tilfelle siden skulle komme opp igjen.)