

SOK-1004 Høst 2021

Prosjektoppgave for studenter på kull 2021.

Leveres i WiseFlow innen 15. desember 2021, kl. 16.00

Prosjektoppgaven er basert på inntektsdata i Troms fra 2015. Alle oppgaver skal besvares. Du skal benytte ggplot til alle figurer. Alle figurer skal ha meningsfulle titler, 'labels', benevninger og enheter på aksene.

Skattetall for Troms fylke 2015

Datasettet «skattetall.RData» inneholder skattetall fra Troms fylke i 2015. Det er totalt 138521 personer med i datasettet som inneholder følgende variabler:

"kommnr" angir kommunenummer. Tromsø er 1902, mens Harstad er 1903.

"aldersgruppe" er en grupperingsvariabel for alder.

"fodt" er fødselsår.

"kjoenn" er kodet «F» for kvinne og «M» for mann. Dataene inneholder også noen personer uten registrering, disse er kodet som «N».

"formue" angir formue i NOK i 2015.

"inntekt" angir inntekt i NOK i 2015.

"skatt" angir utlignet skatt i NOK i 2015.

Følgende R kode laster «skattetall» filen: `load(url("https://bit.ly/2YBntjg"))`

Benytt R pakken gglorenz for å lage figurer av ulikheter. Du finner pakken på CRAN:

<https://cran.r-project.org/web/packages/gglorenz/index.html>

Du finner repositoriet på github: <https://github.com/jichern/gglorenz>

For å tegne Lorenz-kurven under 45 graders linja må du benytte `stat_lorenz(desc = FALSE)`. Alle oppgaver skal besvares. Du skal benytte ggplot til alle figurer. Alle figurer skal ha meningsfulle titler, 'labels', benevninger og enheter på aksene. For å beregne Gini kan du for eksempel benytte pakken ineq (<https://cran.r-project.org/web/packages/ineq/>).

- a) Lag en Lorenz-kurve for inntekt i Troms i 2015 der Gini indeksen vises i figuren.

Fyll ut følgende tabell, der du regner Gini om til prosent ($\times 100$), og Inntekt er gjennomsnittlig inntekt.

Kommune	Gini (%)	Inntekt	Personer
Troms			138521

Kommenter dine funn.

- b) Lag en ny variabel «kommune», der Tromsø er kommunenr 1902, Harstad 1903, og Omegn er resten.

Fyll ut følgende tabell, der du regner Gini om til prosent ($\times 100$), Inntekt er gjennomsnittlig inntekt og Andel er prosentandelen av skattebetalerne i kategorien Kommune.

Kommune	Gini (%)	Inntekt	Personer	Andel
Harstad				
Omegn				
Tromsø				

Kommenter dine funn, og sammenlign dem med det du fant i a).

- c) Mange økonomer mener at det blir mer riktig å beregne Gini på disponibel inntekt, dvs inntekt fratrullet skatt. Beregn denne størrelsen. Dersom skatt er større enn inntekt vil disponibel inntekt bli negativ. Sett alle negative verdier lik 0. Lag en Lorenz-kurve for disponibel inntekt i Troms i 2015 der Gini indeksen vises i figuren. Kommenter dine funn, og sammenlign dem med det du fant i a).
- d) Lag et plot av inntekt på y-aksen og skatt på x-aksen. Kommenter dine funn.
- e) Lag et plot av formue på y-aksen og skatt på x-aksen. Kommenter dine funn, og sammenlign dem med det du fant i d).
- f) Fyll ut følgende tabell, der du regner Gini om til prosent ($\times 100$), Formue er gjennomsnittlig formue og Andel er prosentandelen av skattebetalerne i kategorien Kommune.

Kommune	Gini (%)	Formue	Personer	Andel
Harstad				
Omegn				
Tromsø				
Troms			138521	100%

Kommenter dine funn, og sammenlign dem med det du fant i a) og b).

- g) Benytt variabelen «aldersgruppe» til å beregne Gini per aldersgruppe. Lag et plot der Gini er på y-aksen og aldersgruppe er på x-aksen. Kommenter dine funn.
- h) Benytt variabelen «aldersgruppe» og «kjønn» til å beregne Gini per aldersgruppe per kjønn. Ta bort kategorien «N» i kjønn før du gjør din analyse. Lag et plot der Gini er på y-aksen og aldersgruppe er på x-aksen, sett farge etter kjønn. Kommenter dine funn.

Skriv et kort sammendrag om inntekt, formue og ulikheter i Troms fylke, basert på dine analyser.