

Sentralitet, sysselsetting og struktur: Der ingen skulle tru at nokon kunne bu.

En analyse av norske demografiske mønstre fra en regional-økonomisk vinkling.

Harald Bjarne Vika, Henriette Hausken Kjærandsen, Karin Liang, Katinka Beate Valvatne,
Kristoffer Tufta
HVL

Tuesday 17 Feb, 2026

Abstract I denne studien undersøker vi systematiske variasjoner i folketallsutvikling, sysselsetting, arbeidsmarkedsstatus, boligpriser, demografiske kjennetegn, utdanningsnivå, inn/utflytting og inn/utpendling for norske kommuner av ulik sentralitet. Gjennom beskrivende analyser med kart og kurver, samt korrelasjons- og multiple regresjonsanalyser, dokumenterer vi betydelige forskjeller mellom sentralitetskategorier. Analysene viser at sentrale kommuner har høyere boligpriser, sterkere folketallsvekst, høyere andel unge voksne (20-29 år) og innbyggere med universitetsutdanning, samt positiv netto innflytting og innpendling. Mindre sentrale kommuner viser motsatt mønster med utflytting, utpendling, aldrende befolkning og lavere utdanningsnivå. Boligpriser og utdanningsnivå fremstår som særlig sterke prediktorer for befolkningsvekst i regresjonsanalysene, mens arbeidsledighet har begrenset forklaringskraft. Samlet sett er funnene i studien konsistente med regional vekstteori og tidligere forskning, og bekrefter at sentralitet er en grunnleggende dimensjon for å forstå regionale forskjeller i Norge. Studien demonstrerer også viktigheten av å kontrollere for multikollinearitet ved analyse av regionale utviklingsmønstre.

Innholdsfortegnelse

1 Innledning	2
2 Datainnhenting	3
2.1 Folketall	3
2.2 Demografiske data	3
2.3 Sentralitetsindeksen for norske kommuner	3
2.4 Prosentvis arbeidsledighet	4
2.5 Sysselsetting	4
2.6 Gjennomsnittlig kvadratmeterpris for bolig	4
2.7 Utdanningsnivå	4
2.8 Ut- og innflyttinger pr 1000 innbyggere	4

2.9 Inn- og utpendling	4
3 Resultater	5
3.1 Kart	5
3.1.a Folketallsvekst	5
3.1.b Andel befolkning i alder 20-29	6
3.1.c Arbeidsledighet og sysselsettingsvekst	7
3.1.d Andel med universitetsutdanning	8
3.1.e Netto utflytting	9
3.1.f Netto innpendling og kvm. pris for eneboliger	10
3.2 Kurver	11
3.2.a Gjennomsnittlig folketall per sentralitetsklasse	11
3.2.b Andel innbyggere 20-29 år etter sentralitetskategorier	12
3.2.c Arbeidsledighet etter sentralitetskategorier	12
3.2.d Sysselsettingsvekst etter sentralitetskategorier	13
3.2.e Kvadratmeterpris bolig	13
3.2.f Utvikling i utdanning	14
3.2.g Utvikling i netto innflytting	15
3.2.h Utvikling i netto innpendling	15
3.3 Korrelasjoner	15
3.4 Paneldata	21
3.4.a Demeaning	24
3.5 Pearson vs Spearman	24
3.6 Visualisering av noen sammenhenger	24
3.7 Systematiske forskjeller mellom kommuner i ulike sentralitetskategorier	25
3.8 Forskjeller mellom kommuner i ulike kategorier hos Andersson et al. (2019)	26
3.9 Sammenhengen mellom befolkning og sysselsetting	26
3.10 Relevante hypoteser basert på deskriptiv statistikk	27
3.11 En multippell regresjonsmodell for folketallsutvikling	28
3.11.a Start-år-modell	28
3.11.b Periode-gjennomsnitt-modell	29
3.11.c FE-modell	30
3.12 En multippell regresjonsmodell for sysselsettingsvekst	31
3.12.a Start-år-modell	31
3.12.b Periode-gjennomsnitt-modell	32
3.12.c FE-modell	33
4 Refleksjon	34
5 Konklusjon	34
6 Litteraturliste	35

1 Innledning

I denne studien vil vi se på systematiske variasjoner i folketallsutvikling, sysselsetting, arbeidsmarkedsstatus, boligpriser, demografiske kjennetegn, utdanningsnivå, inn/utflytting og inn/utpendling for kommuner av ulik sentralitet. For å gjøre dette, kommer vi til å samle dataene inn i kart, kurver, ulike korrelasjonsundersøkelser, og multiple regresjonsanalyser. Vi vil også formulere noen hypoteser om sammenhengene til de ulike variablene og til slutt komme med en konklusjon basert på funnene og foreslå elementer til videre forskning.

2 Datainnhenting

Denne studien benytter kommunevise paneldata for å analysere sammenhenger mellom sentralitet, demografi, arbeidsmarked, boligmarked og flyttemønstre i norske kommuner. Nedenfor gis en kort redegjørelse for de enkelte datasettene og deres rolle i de videre deskriptive og analytiske analysene.

2.1 Folketall

Folketall er en sentral variabel i analysen av regional utvikling og brukes både som utfallsvariabel og som skaleringsgrunnlag for andre størrelser. I den deskriptive delen analyseres prosentvis folketallsvekst etter 2000 gjennom kartframstillinger og tidsserier for ulike sentralitetskategorier. Videre inngår folketallsutvikling som avhengig variabel i regresjonsanalysene, der formålet er å undersøke hvordan arbeidsmarked, boligpriser, utdanningsnivå, flytting og pendling samvarierer med endringer i befolkningen.

2.2 Demografiske data

Demografiske data, særlig alderssammensetningen i befolkningen, brukes til å analysere selektiv flytting og regionale livsløp. Spesiell vekt legges på aldersgruppen 20–29 år, som ofte er mest mobil. Andelen unge voksne inngår i kartframstillinger, tidsserier og korrelasjonsanalyser.

2.3 Sentralitetsindeksen for norske kommuner

Sentralitetsindeksen brukes som en overordnet strukturell forklaringsvariabel i analysen. Indeksen ligger til grunn for gruppering av kommuner i den deskriptive delen og brukes videre i korrelasjonsanalyser.

Sentralitetsindeksen er en kode som beskriver hver kommunens sentralitet. Beregningen av sentralitetsindeksen er basert på reisetid til arbeidsplasser og servicefunksjoner fra alle beboerde grunnkretser. Indeksen er sammensatt av to del-indekser:

- Antall arbeidsplasser de som bor i den enkelte grunnkrets kan nå med bil i løpet av 90 minutter.
- Hvor mange ulike typer servicefunksjoner (varer og tjenester) de som bor i den enkelte grunnkrets kan nå med bil i løpet av 90 minutter.

Antallet vektes, slik at en arbeidsplass eller servicefunksjon som ligger nært bostedet teller mer enn en som ligger lengre bort. Indeksen fordeles på en kontinuerlig skala fra 0 (minst sentral) til 1000 (mest sentral), og fordeles så inn i 6 forskjellige sentralitetsklasser, basert på kommunens sentralitetsverdi.

Endringer fra 2017 -> 2020 og 2024 skyldes (i følge SSB selv) i stor grad endringer i kommunegrensene. I perioden har det vært flere kommuneoppløsninger/sammenslåinger, og indeksen vil påvirkes ved hver slik grensejustering (Høydahl, n.d., s. 22). I 2017 (ved første kalkulering av indeksen), ble det også gjort noen feil i databehandlingen som senere har blitt rettet opp i. Ved én kjøring av dataekspport ble det satt en cutoff på antall arbeidsplasser den enkelte kunne nå innen 60 minutter i stedet for de 90 som indeksen skulle baseres på. Dette påvirket resultatet svært mye for de aktuelle områdene (Hordaland, Sogn og Fjordane, Møre og Romsdal, Trøndelag og Nord-Norge).

2.4 Prosentvis arbeidsledighet

Arbeidsledighet benyttes som indikator på arbeidsmarkedets tilstand i kommunene. Variabelen inngår i kartframstillinger, tidsserier og korrelasjonsanalyser.

Arbeidsledighet kan defineres som andelen av et lands arbeidstilskre som ikke er i arbeid, men som aktivt søker jobb og er tilgjengelig for å starte. I Norge publiserer både NAV og Statistik sentralbyrå (SSB) tall for arbeidsledighet, men bruker ulike metoder når det kommer til å samle inn data (Horgen, 2025).

NAVs tall omfatter personer som er registrert som helt ledigh hos NAV. Dette er en fulltelling av alle som faktisk har meldt seg som arbeidssøkere i NAVs systemer. SSB derimot, måler arbeidsledighet gjennom Arbeidskraftundersøkelsen (AKU), som er en intervjuundersøkelse basert på et representativt utvalg av hele befolkningen.

Begge benytter de samme grunnleggende kriteriene for hva som regnes som arbeidsledighet. Forskellen i tall skyldes først og fremst innsamlingsmetoden. SSB inkluderer også personer som er uten arbeid og aktivt søker jobb. Flere av disse kan være personer som ikke er registrert hos NAV. Dette fører til at SSB rapporterer et høyere ledighetstall enn NAV.

NAVs tall gir derimot et bedre bilde av hvor mange som faktisk mottat og har behov for oppføring fra et statlig organ, mens SSBs tall gir et bredere og mer helhetlig bilde av arbeidsledigheten i befolkningen som helhet (Horgen, 2025).

2.5 Sysselsetting

Sysselsettingsdata brukes til å analysere lokal næringsutvikling og arbeidsplassvekst, og inngår både som avhengig og forklaringsvariabel i analysene.

2.6 Gjennomsnittlig kvadratmeterpris for bolig

Boligpriser benyttes som mål på etterspørsel og attraktivitet i regionale boligmarkeder. Prisene sier også noe om tilbuddet i regionale boligmarkeder. Variabelen inngår i kart, tidsserier og regresjonsanalyser.

2.7 Utdanningsnivå

Utdanningsnivå brukes som indikator på kommunenes humankapital og analyseres i sammenheng med sentralitet, sysselsetting og boligpriser.

2.8 Ut- og innflyttinger pr 1000 innbyggere

Flyttedataene brukes til å analysere demografisk dynamikk og kommunenes attraktivitet, inkludert netto flytting og aldersspesifikke flytterater.

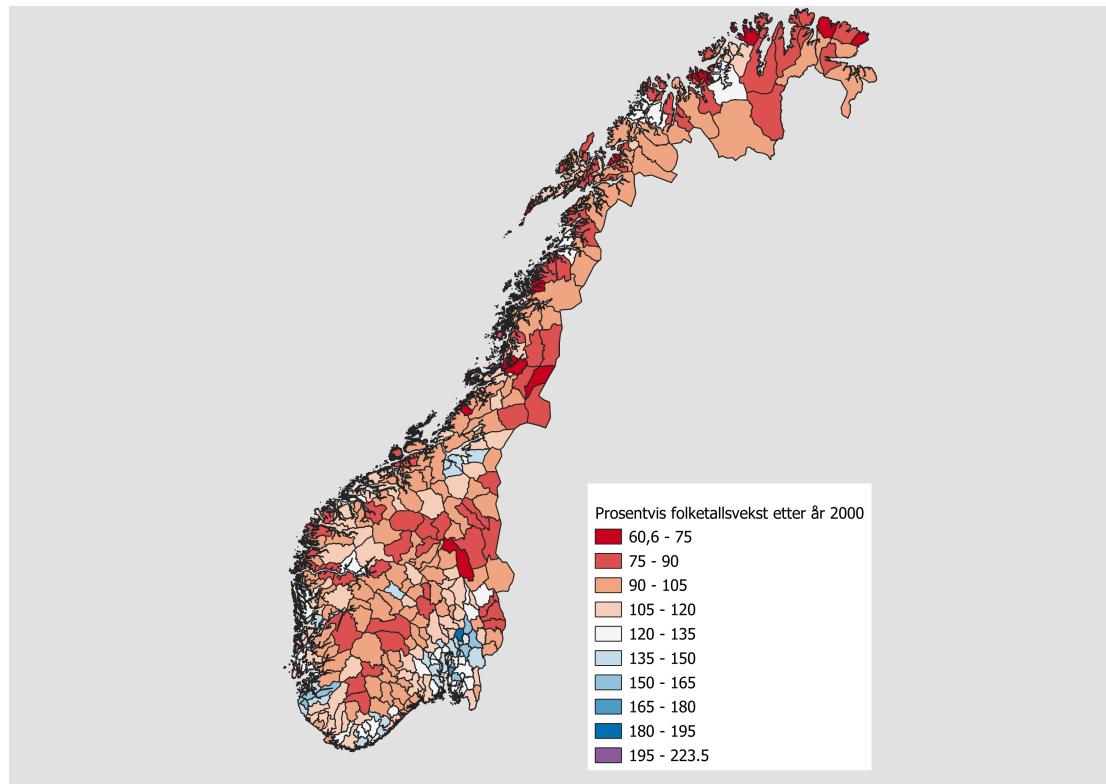
2.9 Inn- og utpendling

Pendlerdata brukes til å analysere funksjonelle arbeidsmarkedsregioner og samspillet mellom bosted og arbeidssted.

3 Resultater

3.1 Kart

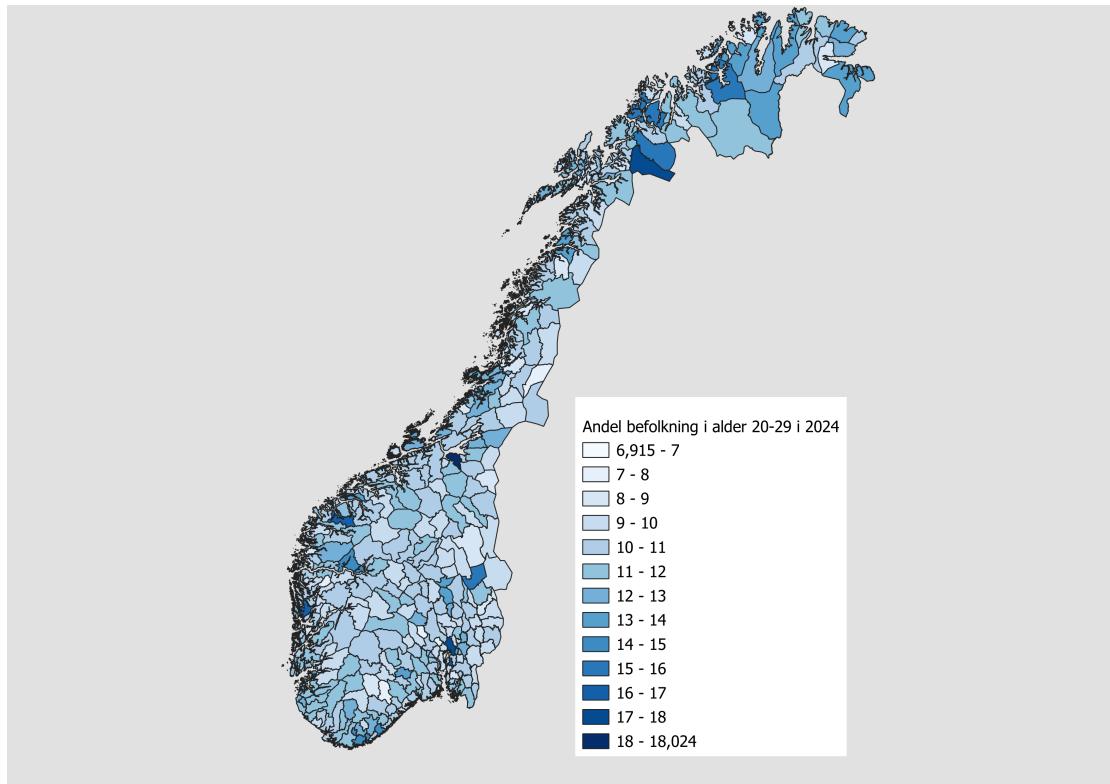
3.1.a Folketallsvekst



Figur 1: Prosentvis folketallsvekst etter år 2000

Den prosentvise folketallsveksten etter år 2000, til 2024, kan ses i figur 1. Her kan man se at det er ett fåtall kommuner som har klasse 6 eller høyere. Der er flere kommuner i Akershus med forholdsvis høy befolningsvekts, for eksempel Gjerdrum, Nannestad og Eidsvoll. Ullensaker har hatt høyest vekst i landet. Kystområdene sør i landet har generelt sett hatt høyere vekst enn innlandet og Nord-Norge. De kommunene med lavest vekst ligger for det meste i Trøndelag og lengre nord i landet. Det at hele landet har hatt en vekst i befolkning kan tyde på at det har vært en del innvandring, eller en økning i samlet fruktbarhet. Samlet fruktbarhet er ett mål på gjennomsnittlig mengde barn kvinner vil få i løpet av livet (Sørlien, 2025).

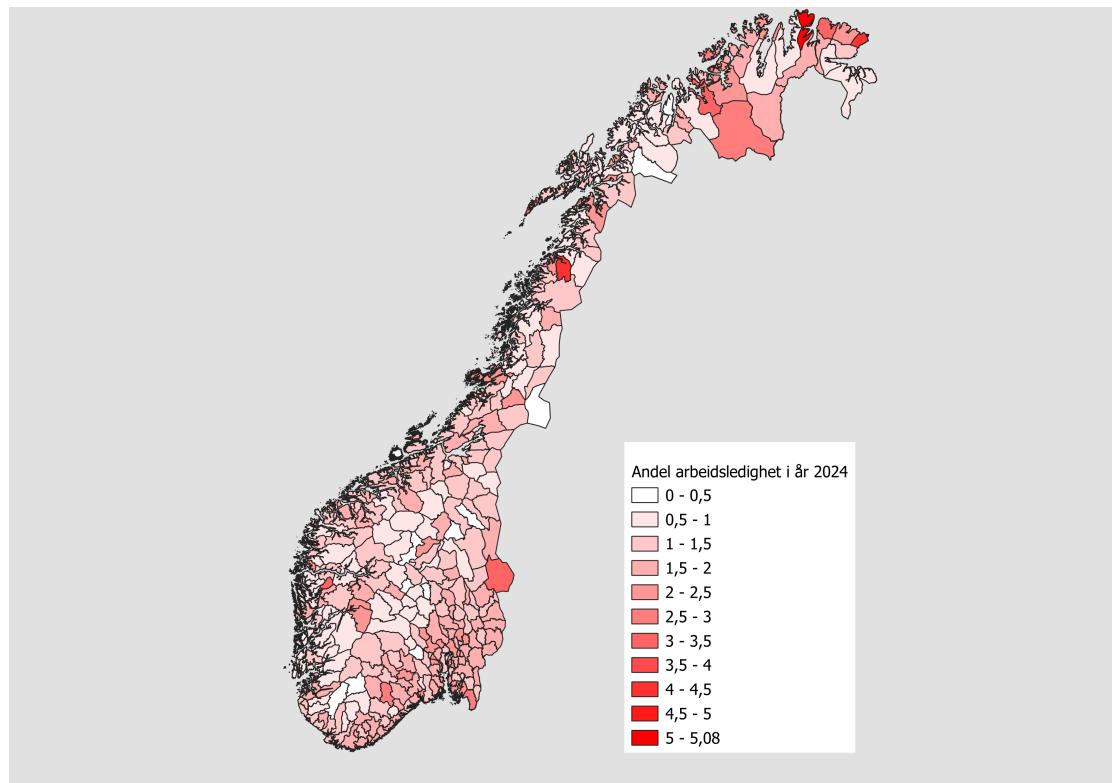
3.1.b Andel befolkning i alder 20-29



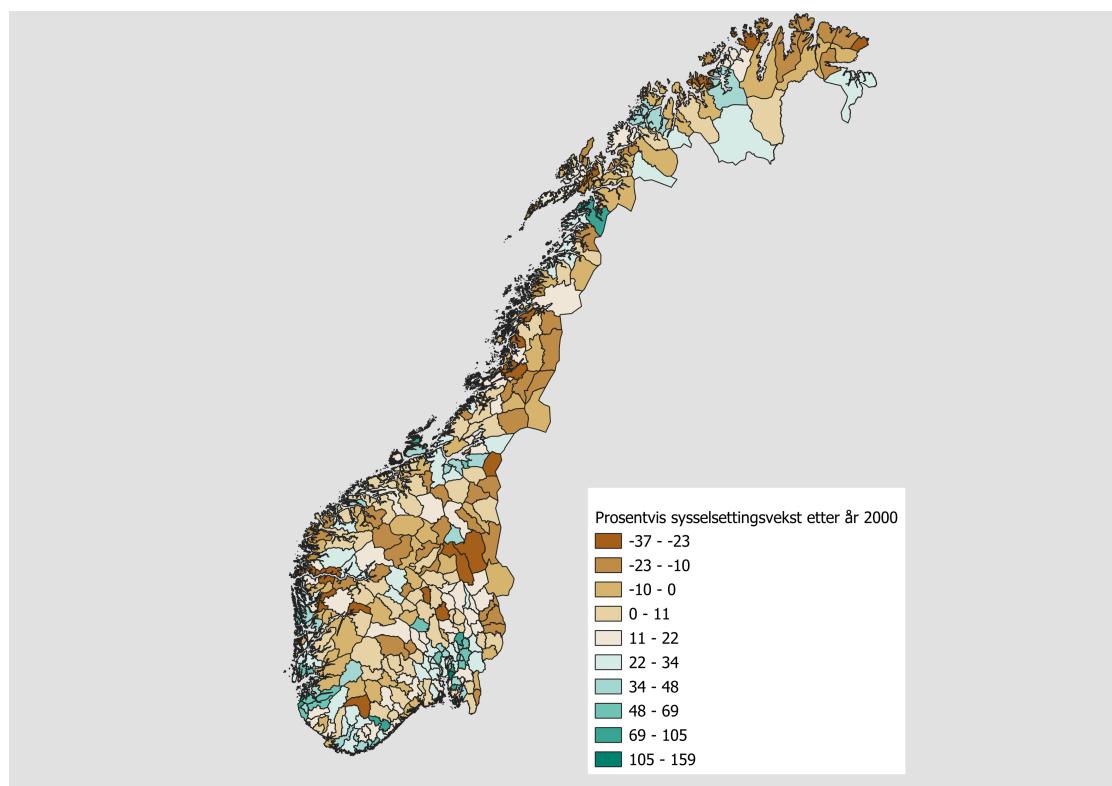
Figur 2: Andel befolkning i alder 20-29 i 2024

I figur 2 vises andel av befolkningen i aldersgruppen 20-29 i 2024. Her er det høyest andel i Troms, og en del plasser i Finnmark har høy andel i aldersgruppen. Noen kommuner i landet har en spesielt høy andel (18%). Det er spesielt de kommunene som er populære studiesteder, e.g. Oslo, Trondheim og Bergen, som har en høy andel i denne aldersgruppen.

3.1.c Arbeidsledighet og sysselsettingsvekst



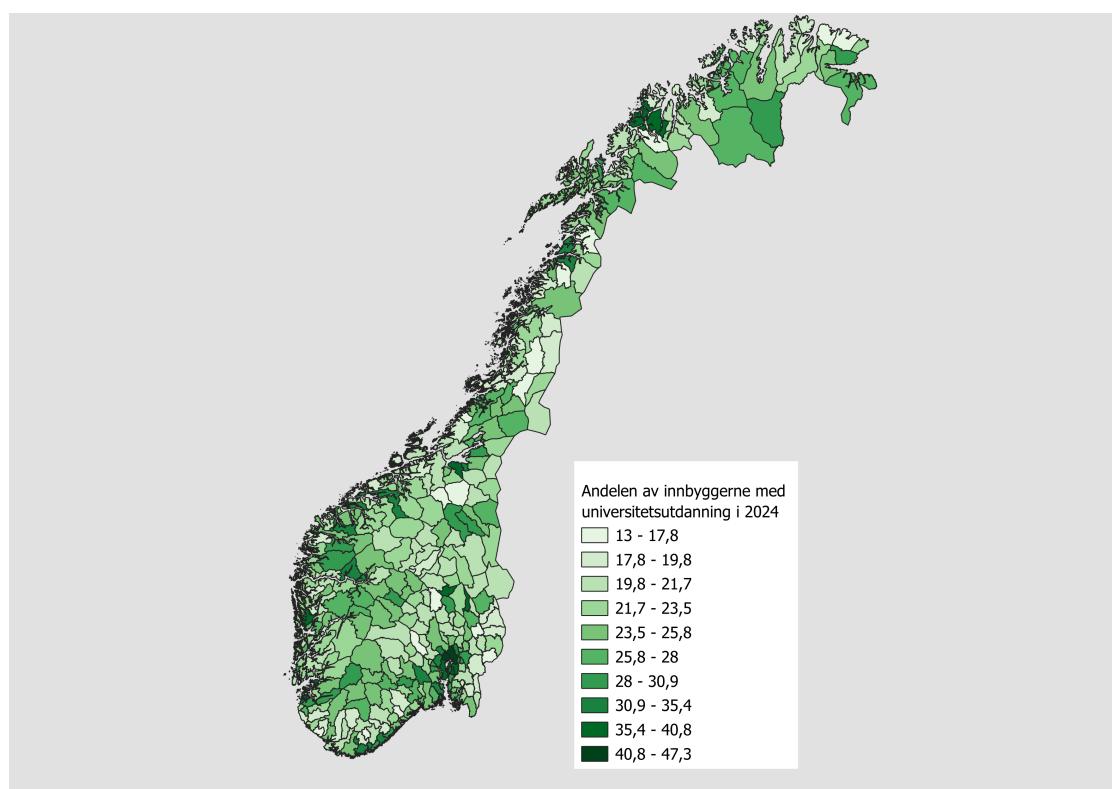
Figur 3: Andel arbeidsledighet i år 2024



Figur 4: Prosentvis sysselsettingsvekst etter år 2000. De tre første klassene viser til negativ vekst, resten viser positiv vekst

Figur 3 viser et kart over arbeidsledigheten i Norge i 2024. Kartet viser få uteliggere med høy arbeidsledighet. Høyest arbeidsledighet er i Finnmark, men man kan finne en uteligger i Nordland. Det er også mer arbeidsledighet i området rundt og i Akershus enn vest i landet. Arbeidsledighet er størst i urbane områder, og lavest i rurale. Dette kan komme av «amenities», eller «lokale goder», som gjør at det er mer ettertraktet å være boende i urbane områder. For eksempel kan det være mer ettertraktet å bo en plass med muligheter for permanent høyere inntekt. Personer vil prioritere å bo urbant, dersom man kan tjene vesentlig mer enn i rurale områder, selv om det er vanskelig å få seg en jobb (Todaro, 1969, s. 147). Videre i figur 4 kan man se prosentvis sysselsettingsvekst etter år 2000, til 2024. Her kan man se at det har både vært negativ og positiv vekt i landet. Igjen kommer kystområdene i Sør-Norge, samt områdene rundt og i Oslo godt ut, og har en høy klasse vekst. Innlandet og nordover i landet har en del negativ vekst. Hamarøy skiller seg spesielt ut og har klasse 10 vekst, selv om nabo kommunene har svak eller negativ vekst. Kartet tyder på at det er mer vekst i sysselsetting i urbane enn rurale områder, noe som kan komme av en økning av urbanisering i form av nye bedrifter og dermed arbeidsplasser.

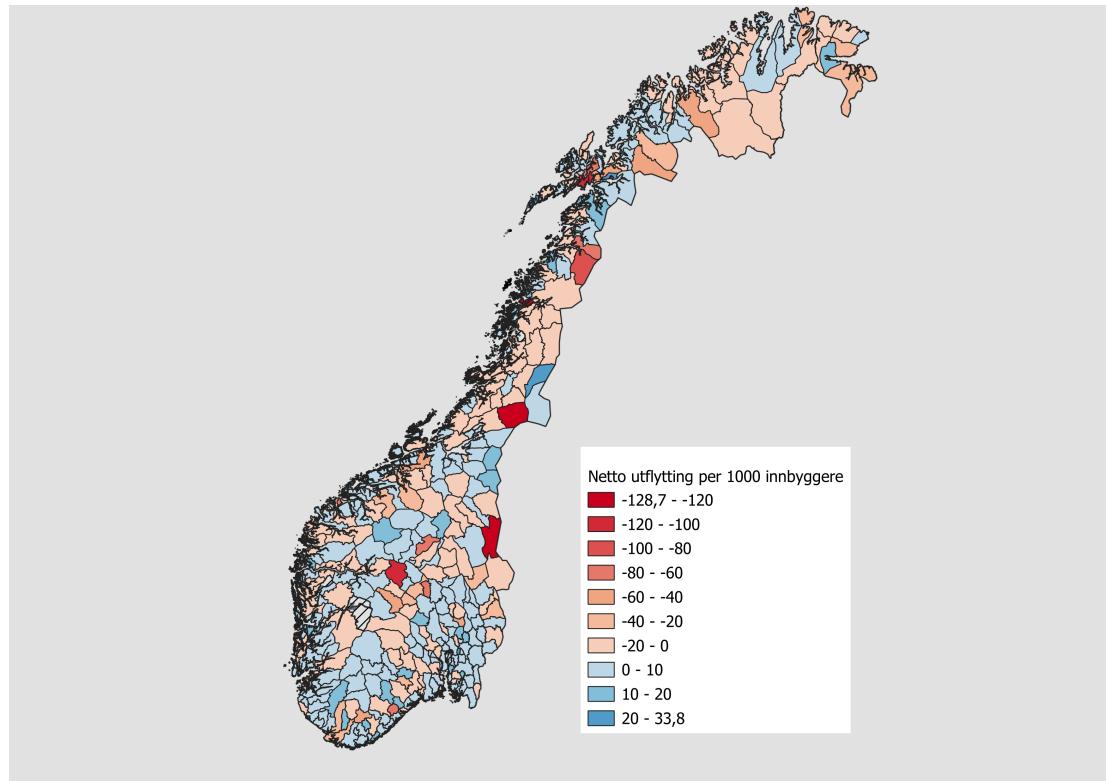
3.1.d Andel med universitetsutdanning



Figur 5: Andelen av innbyggerne med universitetsutdanning i 2024

Andelen av innbyggere med universitetsutdanning i 2024 kan ses i figur 5. Her er det blant annet en høy andel i området i og rundt Oslo. Det er flere kommuner med forholdsvis lav andel (under 17,8%), men de fleste kommunene har en andel over den første klassen. Urbane områder har høyere andel med universitetsutdanning enn i mer rurale områder, for eksempel i indre strøk.

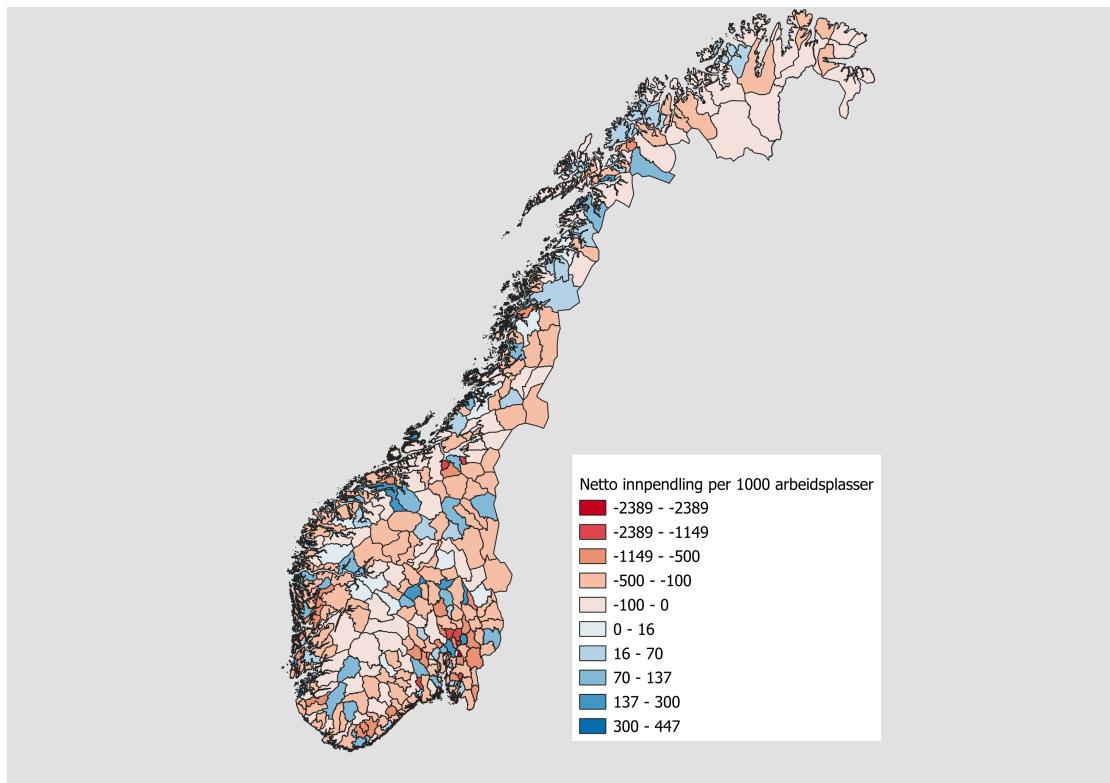
3.1.e Netto utflytting



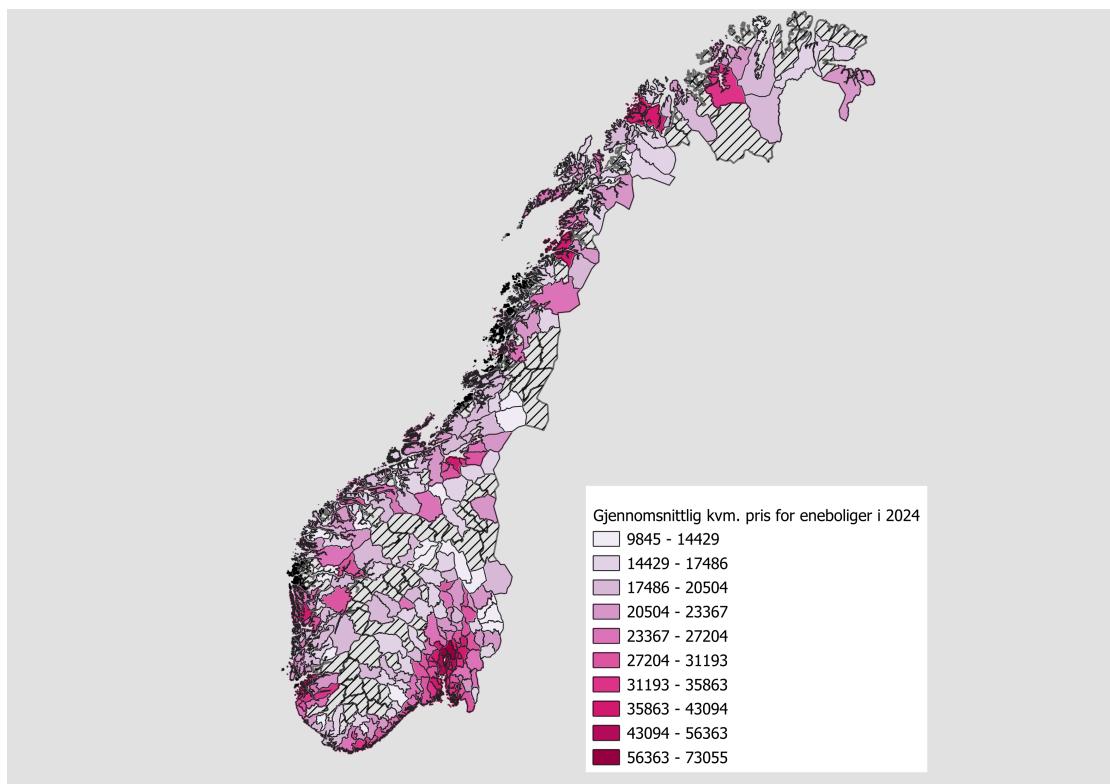
Figur 6: Netto utflytting per 1000 innbyggere. Røde kategorier viser til netto utflytting, mens blåe viser til netto innflytting. Skravete kommuner mangler data

I figur 6 kan man se netto utflytting per 1000 innbyggere. Det er mer kommuner med netto innflytting enn utflytting. Noen kommuner har spesielt høy utflytting, for eksempel Engerdal, Vang og Snåsa. Det at det er mer kommuner med en netto innflytting kan komme av innvandring. Dette kunne det ha vært interessant å se nærmere på i videre studier. Videre hadde det vært interessant å dele opp kartet i ulike aldersgrupper for å se om vi kan se noen generelle mønstre, for eksempel studenter sine flyttingsmønstre.

3.1.f Netto innpendling og kvm. pris for eneboliger



Figur 7: Netto innpendling per 1000 arbeidsplasser. Røde verdier viser til netto utpendling, og blåe til netto innpendling

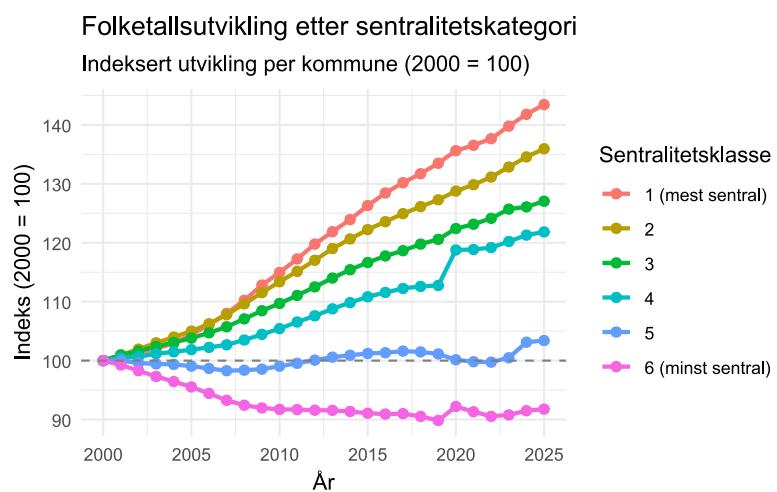


Figur 8: Gjennomsnittlig kvm. pris for eneboliger i 2024. Kommuner uten verdier er skravert

Netto pendling per 1000 arbeidsplasser vises i figur 7. Det er mest høyest grad av utpendling i Østlandet, spesielt i Akershus. Det er minst kommuner med mer innpendling enn utpendling, disse kommunene er spredt over hele landet. Noen av disse kommunene er derimot rett vedsiden av kommunene med mest utpendling, som tyder på at arbeiderne pendler mye til nabokommunene. Figur 8 viser den gjennomsnittlige kvadratmeter prisene for eneboliger i 2024. Det er en del manglende data for priser. I og rundt Oslo er det spesielt høy pris per kvm. Det samme gjelder blant annet for noen kommuner nord i Norge. Man kan se at det er høyere priser ved kysten enn midt i landet. Det er en generell trend med høyere boligpriser i urbane enn rurale områder. Den høye innpendlingen i Oslo kan være ett resultat av den høye kvadratmeterprisen for eneboliger.

3.2 Kurver

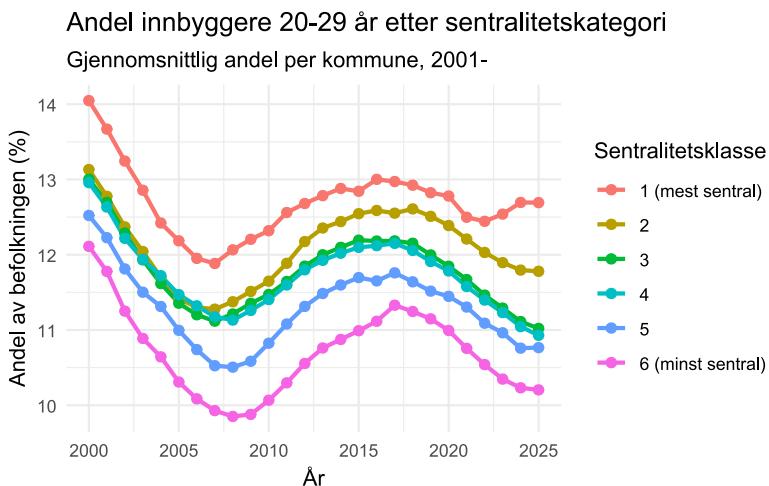
3.2.a Gjennomsnittlig folketall per sentralitetsklasse



I figuren over ser vi gjennomsnittlig folketallsutvikling i de ulike sentralitetsklassene, indeksert for å vise prosentvis utvikling. Startåret på grafen er år 2000.

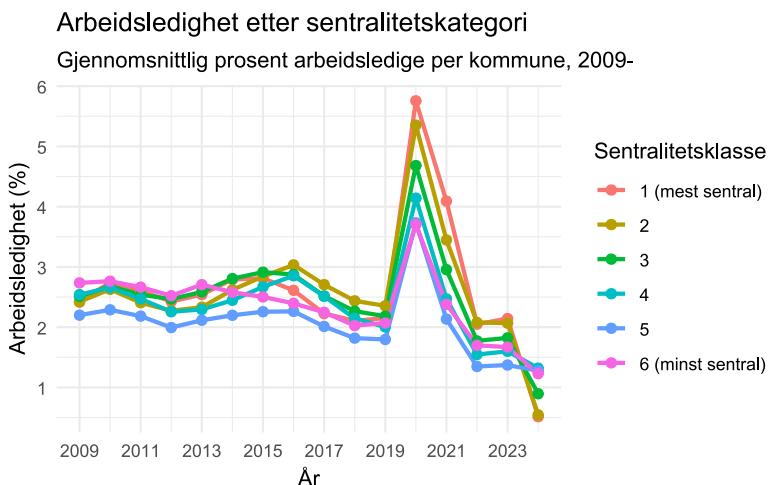
Folketallsutviklingen har vært stigende for de fire mest sentrale sentralitetskategoriene. I klasse 4 ser vi også et betraktelig hopp i folketall fra 2019 til 2020. I klasse 5 var det befolkningsnedgang fra år 2000, før det snudde til positiv utvikling rundt år 2007. Denne utviklingen snudde til det negative igjen rundt 2016, før vi igjen ser et hopp i år 2020. De aller minst sentrale kommunene har konsekvent hatt befolkningsnedgang siden 2000, foruten et hopp i 2020 og en mild oppgang de siste 2-3 årene. Hoppene i 2020 må antas å skyldes kommunereformen. De minste kommunene (i folketall) ble her slått sammen for å spare på administrative utgifter, og det er derfor vi ser størst hopp i folketall i klasse 4 og 6 i denne perioden. Kommunereformen ser også ut til å ha en lignende effekt på klasse 5, men noe stort hopp i folketall ser vi ikke før i 2024. Den (svake) oppgangen i tiden etter er også interessant. Kan det være normalisering av hjemmekontor og økt fokus på distriktsutvikling som driver denne desentraliseringen? Presser høye boligpriser folk ut fra de mest sentrale kommunene? Ønsker folk større boliger, noe som er mer tilgjengelig i periferien(Capello, 2016, s. 54)? Det finnes mange mulige sentrifugale krefter som fungerer desentralisende på befolkningen, men vi har ikke grunnlag til å trekke konklusjoner på dette ennå.

3.2.b Andel innbyggere 20-29 år etter sentralitetskategori



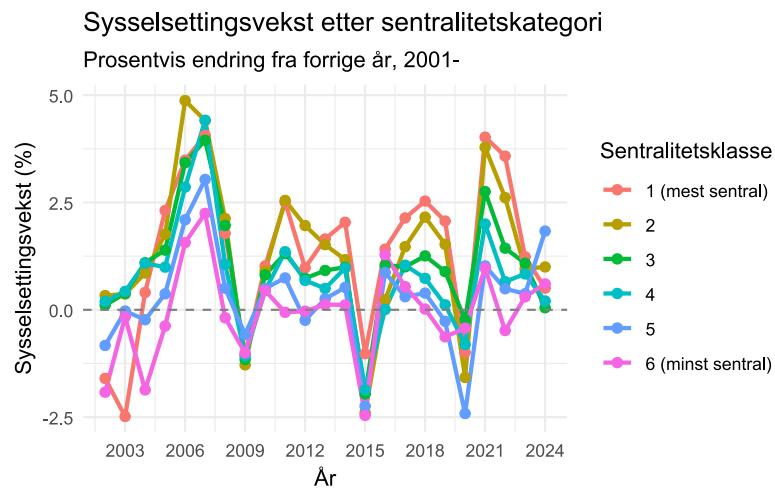
Andel innbyggere i 20-årene ser ut til å følge en noenlunde lik utvikling nasjonalt, unntatt for sentralitetskasse 1 og 2 som har hatt en svakere nedgang i unge innbyggere enn resten av landet. Dette kan skyldes at unge mennesker reagerer bedre på de positive effektene av urbanisering, som et bredere kulturtibud og flere ameniteter (Capello, 2016, p. s.187). En annen hypotese kan være at de i aldersgruppen 20-29 er på det livsstadiet at de vil ta høyere utdanning - «learning by schooling» - og at de derfor i større grad trekkes mot steder der det finnes universiteter, som ofte er lagt til de større, mer sentrale byene (Göcke, 2004).

3.2.c Arbeidsledighet etter sentralitetskategori



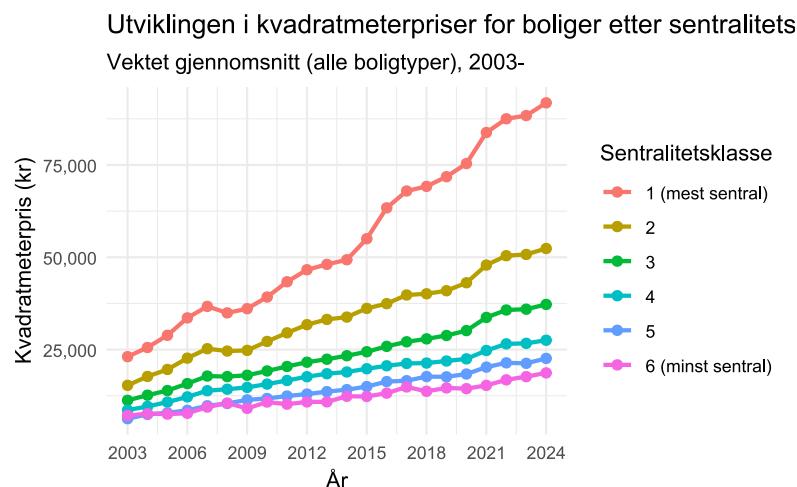
I figuren over ser vi en grafisk framstilling av gjennomsnittlig arbeidsledighet (i prosent) i hver sentralitetskategori, beregnet ut fra data hentet fra SSB. Arbeidsledigheten ser ut til å være relativt lik over hele landet, med mindre enn ett prosentpoeng variasjon frem til 2020. I koronaåret 2020 ser vi en eksplosjon i ledighet for alle klassene, men med størst økning i sentrale strøk. Sjokket fra COVID-19 ser ut til å ha gitt seg i 2022, og grafen trender mot mer typiske verdier igjen i 2023. Datagrunnlaget viser at flere kommuner har oppgitt 0% arbeidsledighet i 2024, og vi antar at dette er en feil i statistikken.

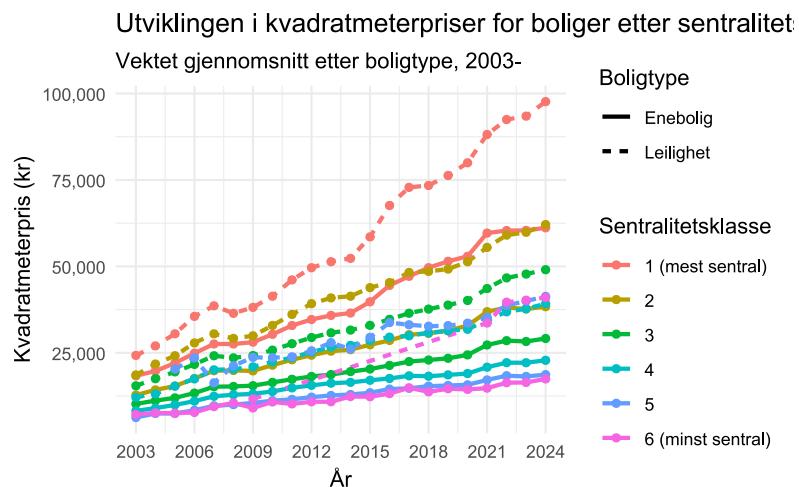
3.2.d Sysselsettingsvekst etter sentralitetskategori



Figuren over viser sysselsettingsvekst. Vi ser positiv utvikling i de fleste årene, unntatt negativ utvikling i 2009, 2015 og 2020. Dette kan ses i forbindelse med internasjonale økonomiske trender. Nedgangen i 2009 kan forklares som en konsekvens av «the great recession» i 2008. Nedgangen i 2015 kan vi forbinde med oljekrisen i 2014 da tusenvis av nordmenn mistet jobbene sine i oljeindustrien som følge av et ras i oljeprisen. Kyst-Norge er sterkt knyttet opp til oljeindustrien, og ringvirkningene av dette oljeprisfallet kunne dermed føles sterkt også i andre industrier i landet. Nedgangen i 2020 må ses i forbindelse med koronapandemien.

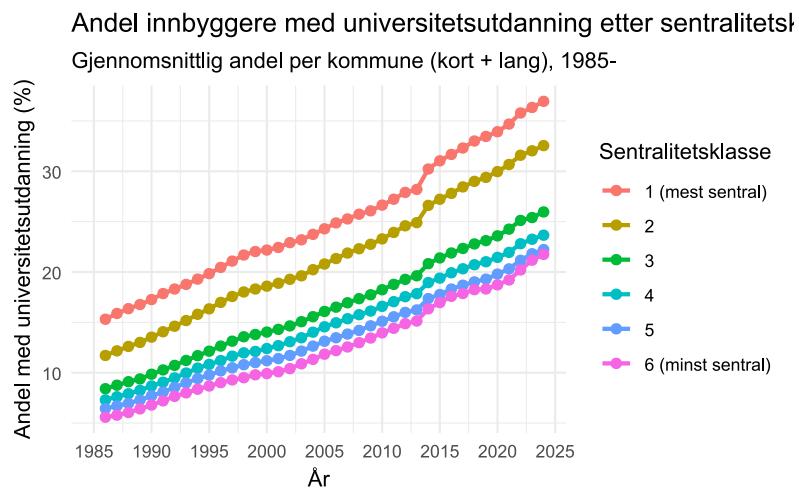
3.2.e Kvadratmeterpris bolig





Kvadratmeterprisen på bolig har hatt en positiv utvikling over hele landet, men den har vært aller sterkest i de mest sentrale kommunene. Det er ellers interessant å se at kvadratmeterprisen ser ut til å korreleres med sentralitet.

3.2.f Utvikling i utdanning

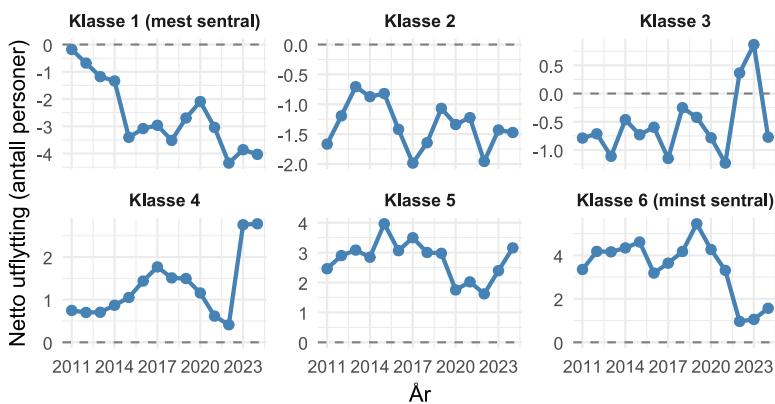


Utviklingen i andel høyt utdannede innbyggere har vært positiv nasjonalt, men også her ser det ut til å korreleres med sentralitet. De mest sentrale kommunene har også høyest andel høyt utdannede innbyggere. Dette kan skyldes at befolkningen i bynære miljø har enklest tilgang til universiteter og høyskoler, men det kan også være en indikasjon på at det i bykommunene det er best tilgang på arbeid som krever høye kvalifikasjoner.

3.2.g Utvikling i netto innflytting

Netto utflytting etter sentralitetskategori

Gjennomsnittlig netto utflytting per kommune, 2011-

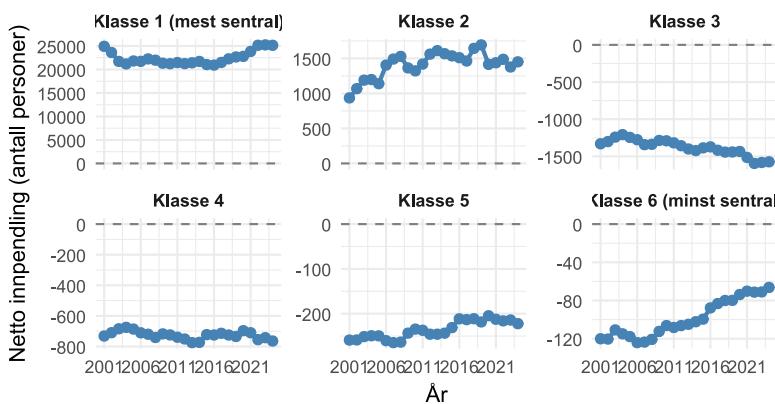


Grafene over må tydes med at negative tall viser positiv flyttebalanse (altså mer innflytning enn utflytning).

3.2.h Utvikling i netto innpendingning

Netto innpendingning etter sentralitetskategori

Gjennomsnittlig netto innpendingning per kommune, 2001-



Grafene over viser netto innpendingning etter sentralitetskategori. Vi ser at for alle sentralitetsklassene utenom 1 og 2 har vi negativ gjennomsnittsinnpendling per kommune, altså vil det være flere arbeidere som pendler ut fra kommunene enn pendler inn. Ett av målene for sentralitetsindeksen er nemlig antall arbeidsplasser innenfor rimelig pendleavstand (90 minutter) vektet etter nærhet. Det er dermed logisk at det pendles mer inn til de mest sentrale kommunene.

3.3 Korrelasjoner

I dette delkapittelet analyseres sammenhengen mellom utvalgte variabler fra datasettene som ble hentet inn i oppgave 1. For å undersøke disse sammenhengene benyttes korrelasjonsanalyse. En korrelasjonsberegnning gir en korrelasjonskoeffisient, et tall mellom -1 og 1 , som uttrykker styrken og retningen på den lineære sammenhengen mellom to variabler. En positiv verdi indikerer at variablene beveger seg i samme retning, mens en negativ verdi innebærer at

de beveger seg i motsatt retning. En verdi nær 0 tyder på svak eller ingen lineær sammenheng (Frøslie, 2025; Salomão, 2024).

I analysen benyttes Pearsons korrelasjonskoeffisient, som måler graden av lineær samvariasjon mellom to kvantitative variabler. Det er viktig å presisere at korrelasjon ikke innebærer kausalitet; en høy korrelasjon betyr ikke nødvendigvis at den ene variabelen forårsaker endringer i den andre, men kun at de varierer systematisk sammen.

Korrelasjonsberegningene gjennomføres på tvers av norske kommuner for utvalgte år. For de fleste variablene benyttes data fra 2014 og 2023, slik at vi kan sammenligne sammenhenger før og etter en periode med betydelige økonomiske og demografiske endringer. For sentralitetsindeksen er det kun tilgjengelige data for 2018, 2020 og 2023. I disse analysene benyttes derfor 2018 og 2023 for å sikre sammenlignbarhet over tid.

I de tilfellene hvor korrelasjonskoeffisientene viser særlig sterke sammenhenger, tydelige endringer over tid eller uventede resultater, vil disse bli diskutert nærmere. Det skal ikke bare presentere tallverdier, men vi skal også tolke dem.

- Folketall mot andel i aldersgruppen 20-29

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014  0.414
2 2023  0.386
```

Korrelasjonsberegningene viser en moderat positiv sammenheng mellom kommunestørrelse og andelen unge voksne (20–29 år). Dette indikerer at større kommuner relativt sett har en yngre befolkningssammensetning. Sammenhengen er noe svakere i 2023 enn i 2014, noe som kan indikere en svak reduksjon i graden av aldersselektiv urbanisering

- Sentralitet mot andel i aldersgruppen 20-29

```
# A tibble: 2 × 2
  år     cor
  <dbl> <dbl>
1 2018  0.273
2 2023  0.285
```

- Sentralitet mot arbeidsledighet

```
# A tibble: 2 × 2
  år     cor
  <dbl> <dbl>
1 2018  0.225
2 2023  0.225
```

Får lik beregning, noe som kan si at sentraliteten er lite bevegelig.

- Sentralitet mot enebolig kvm.pris

```
# A tibble: 2 × 2
  år     cor
  <dbl> <dbl>
1 2018  0.721
2 2023  0.715
```

- Sentralitet mot andel med høyere utdanning

```
# A tibble: 2 × 2
  år     cor
  <dbl> <dbl>
1 2018  0.590
2 2023  0.523
```

- Sentralitet mot utflytting

```
# A tibble: 2 × 2
  år     cor
  <dbl> <dbl>
1 2018  0.188
2 2023 -0.0218
```

Her får vi en motsetning av korrelasjonsberegningen, uansett om den er liten. I 2018 får vi en positiv korrelasjonskoeffiseint på 0.188 mens i 2023 får vi -0.0218 som kan sees på som null. Dette viser en svakt sammenheng til ingen mellom sentralitet og utflytting. Sentralitet forklarer i liten grad hvor mye folk flytter ut av en kommune.

- Sentralitet mot innpendling

```
# A tibble: 2 × 2
  år     cor
  <dbl> <dbl>
1 2018  0.387
2 2023  0.362
```

Sentralitet mot utpendling

```
# A tibble: 2 × 2
  år     cor
  <dbl> <dbl>
1 2018  0.620
2 2023  0.610
```

Her har vi en godt positiv korrelasjoner mellom sentralitet og utpendling men beveger seg lite fra 2018 til 2023, noe som kan ha med at sentralitet har en påvirkning på utpendling men varierer ikke noe over årene. Ledig mot andel i aldersgruppen 20-29

```
# A tibble: 2 × 2
  år     cor
```

```
<int> <dbl>
1 2014 0.111
2 2023 0.0759
```

- Ledig mot enebolig kvm.pris

```
# A tibble: 2 × 2
  år      cor
  <int>   <dbl>
1 2014  0.0196
2 2023  0.169
```

- Ledig vs andel med høyere utdanning

```
# A tibble: 2 × 2
  år      cor
  <int>   <dbl>
1 2014 -0.00976
2 2023  0.0651
```

- Ledig mot utflytting per 1000

```
# A tibble: 2 × 3
  år cor_brutto cor_netto
  <int>   <dbl>     <dbl>
1 2014    0.210     0.0868
2 2023    0.0714   -0.0220
```

- Ledig mot pendling per 1000

```
# A tibble: 2 × 3
  år cor_brutto cor_netto
  <int>   <dbl>     <dbl>
1 2014    0.228     0.0160
2 2023    0.228     0.0418
```

- Bolig kvm.pris mot andel med høyere utdanning

```
# A tibble: 2 × 2
  år      cor
  <int>   <dbl>
1 2014  0.795
2 2023  0.741
```

- Bolig kvm.pris mot andel i aldersgruppen 20-29

```
# A tibble: 2 × 2
  år      cor
  <int>   <dbl>
```

```
1 2014 0.832
2 2023 0.617
```

Her får vi en sterk korrelasjon mellom bolig pris per kvadrat meter mot andel i aldersgruppen 20-29. Korrelasjonen mellom boligpriser og andel 20–29-åringer er svært høy i 2014 (0,83) og fortsatt tydelig positiv i 2023 (0,62). Dette indikerer en sterk strukturell sammenheng mellom urbane boligmarkeder og konsentrasjonen av unge voksne. Sentrale og attraktive kommuner med høye boligpriser har også systematisk høyere andel unge innbyggere. Nedgangen i korrelasjonen kan tyde på en viss desentralisering eller mer utjevnet boligprisvekst etter pandemien.

- Bolig kvm.pris mot innpendling

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014 0.641
2 2023 0.701
```

Korrelasjonen mellom boligpriser og innpendling er høy og stabil i både 2014 og 2023. Dette indikerer at kommuner med høye boligpriser også er sentrale arbeidsmarkeder med betydelig innpendling. Sammenhengen har blitt noe sterkere over tid, noe som kan reflektere økt konsentrasjon av arbeidsplasser i sentrale regioner.

- Bolig kvmpris mot utpendling

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014 0.736
2 2023 0.786
```

Korrelasjonen mellom boligpriser og antall utpendlere er høy i både 2014 og 2023. Dette indikerer at kommuner med høye boligpriser også har omfattende pendlestrømmer ut av kommunen. Sammenhengen reflekterer trolig at sentrale og folkerike kommuner inngår i større, funksjonelle arbeidsmarkedsregioner med betydelig krysspendling.

- Bolig kvm.pris mot utflytting

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014 0.310
2 2023 -0.0286
```

I 2014 var det en moderat positiv sammenheng mellom boligpriser og utflytting per 1000 innbyggere, noe som trolig reflekterer høy mobilitet i sentrale og kostbare kommuner. I 2023 er sammenhengen tilnærmet null. Dette indikerer at boligprisnivå i seg selv ikke systematisk

forklarer bruttoutflytting, og at eventuelle sammenhenger i 2014 sannsynligvis var drevet av underliggende strukturelle forskjeller mellom kommuner.

- Bolig kvm.pris mot innflytting

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014  0.676
2 2023  0.174
```

I 2014 var det en sterk positiv sammenheng mellom boligpriser og innflytting per 1000 innbyggere, noe som indikerer at sentrale og kostbare kommuner også var de mest attraktive målt ved innflytting. I 2023 er sammenhengen betydelig svakere. Dette kan reflektere endrede flyttemønstre etter pandemien, økt bruk av hjemmekontor og mer geografisk utjevning i boligprisveksten.

- Nettoflytting mot andel i aldersgruppen 20-29

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014  0.0640
2 2023 -0.0588
```

Her har vi en positiv sammenheng i 2014 til en negativ i 2023, men korrelasjonene mellom disse variablene er nesten lik null og kan ses på som svært lite påvirkning på hverandre.

- Nettoflytting mot netto utpendling

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014  0.0327
2 2023 -0.00212
```

Her får vi også en liten sammenheng mellom de to variablene, selv om årene 2014 og 2023 går fra en positiv til en negativ korrelasjon.

- Netto utpendling og bolig kvm.pris i tilgrensende kommuner

```
# A tibble: 2 × 2
  år     cor
  <int> <dbl>
1 2014 -0.106
2 2023 -0.151
```

Korrelasjonen mellom netto pendling og gjennomsnittlig boligpris i nabokommunene er svak, men negativ i både 2014 og 2023. Dette indikerer at kommuner som grenser til kommuner med høye boligpriser i større grad fungerer som bostedskommuner med netto utpendling.

Resultatet er konsistent med en sentrum–omland-struktur, der høye boligpriser i sentrale kommuner bidrar til økt bosetting og utpendling fra omlandskommuner.

3.4 Paneldata

- Årlig endring i folketall og sysselsetting

```
# A tibble: 357 × 3
  kommunenummer n_obs cor_folk_syss
  <chr>        <int>      <dbl>
1 0301          24       -0.0889
2 1101          24        0.0797
3 1103          24       -0.153
4 1106          24       -0.113
5 1108          24       -0.106
6 1111          24       -0.180
7 1112          24        0.0569
8 1114          24       -0.197
9 1119          24        0.0467
10 1120         24        0.509
# i 347 more rows
```

```
# A tibble: 1 × 7
  n_komm   mean median    sd    min    max n_na
  <int>  <dbl>  <dbl> <dbl> <dbl> <dbl> <int>
1 357  0.0157 0.0153 0.226 -0.648 0.995     0
```

Den gjennomsnittlige korrelasjonen mellom årlige endringer i sysselsetting og folketall er tilnærmet null, noe som indikerer at kortsigte endringer i arbeidsplasser i liten grad fører til umiddelbar befolkningsvekst. Dette tyder på at arbeidsmarkedets tilpasning i større grad skjer gjennom pendling enn gjennom flytting. En slik tolkning er i tråd med regionaløkonomisk teori om funksjonelle arbeidsmarkedsregioner, hvor arbeidskraftens daglige mobilitet gjør det mulig å tilpasse seg endringer i sysselsetting uten at husholdninger nødvendigvis endrer bosted (Capello, 2016).

Årlig endring i prosentvis arbeidsledighet og årlig prosentvis endring i sysselsetting

```
# A tibble: 357 × 5
  kommunenummer n_obs sd_ledig sd_syss cor_ledig_syss
  <chr>        <int>    <dbl>    <dbl>      <dbl>
1 0301          16      0.798    1.66      -0.855
2 1101          16      0.781    1.78      -0.506
3 1103          44      0.577    1.82      -0.837
4 1106          16      0.966    2.00      -0.734
5 1108          32      0.643    1.96      -0.398
6 1111          16      0.576    3.58      0.0591
7 1112          16      0.487    3.47      0.289
8 1114          16      0.514    2.62      -0.297
9 1119          16      0.482    1.48      -0.263
10 1120         16      0.742    2.27      -0.286
# i 347 more rows
```

```
# A tibble: 1 × 7
  n_komm   mean median    sd    min    max  n_na
  <int>  <dbl>  <dbl> <dbl> <dbl> <dbl> <int>
1     357 -0.303 -0.307 0.279 -0.880  0.777     1
```

Den gjennomsnittlige korrelasjonen mellom årlig endring i arbeidsledighet og årlig sysselsettingsvekst er moderat negativ (-0.303). Dette indikerer at arbeidsmarkedet reagerer relativt raskt på endringer i sysselsetting, ved at økt sysselsetting samvarierer med fallende ledighet. Resultatet tyder på at arbeidsledighet fungerer som en kortsiktig justeringsmekanisme i regionale arbeidsmarkededer, i tråd med regionaløkonomisk teori om faktorjustering (Capello, 2016).

Årlig endring i boligpriser og årlig netto innflytting

```
# A tibble: 192 × 5
  kommunenummer n_obs sd_kvpris sd_flytt cor_bolig_nettoflytt
  <chr>        <int>    <dbl>    <dbl>            <dbl>
1 0301          14     2399.    1.25           -0.591
2 1101          14     1062.    1.73            0.157
3 1103          14     1568.    2.25           -0.408
4 1106          14     796.    1.31           -0.368
5 1108           4     1233.    2.00           -0.102
6 1111          14     1796.    3.03           -0.0138
7 1112           11     1253.    4.12           -0.0224
8 1114           11     2408.    4.05           -0.245
9 1119          14     1267.    2.98            0.0793
10 1120          14     1545.    1.91            0.208
# i 182 more rows
```

```
# A tibble: 1 × 7
  n_komm   mean median    sd    min    max  n_na
  <int>  <dbl>  <dbl> <dbl> <dbl> <dbl> <int>
1     192  0.0637  0.0897 0.529     -1      1      4
```

Den gjennomsnittlige korrelasjonen mellom årlige endringer i boligpriser og nettoflytting er svak og nær null (0.0637). Samtidig er variasjonen mellom kommuner betydelig. Dette indikerer at sammenhengen mellom boligmarked og flytteprosesser varierer sterkt mellom ulike typer kommuner, og at effektene ikke nødvendigvis opptrer samtidig innen samme år.

Årlig prosentvis endring i sysselsetting og årlig endring i boligpriser

```
# A tibble: 192 × 5
  kommunenummer n_obs sd_syss sd_kvpris cor_syss_bolig
  <chr>        <int>    <dbl>    <dbl>            <dbl>
1 0301          22     1.86    2099.           0.379
2 1101          22     2.05    1143.           0.702
3 1103          22     3.38    1718.           0.601
4 1106          22     2.50    772.            0.583
5 1108           4     1.45    1233.           0.576
```

```

6 1111          20    4.63    1614.      0.0927
7 1112          15    3.97    1331.      0.177
8 1114          14    3.13    2224.      0.172
9 1119          22    1.89    1184.      0.272
10 1120         22    2.24    1516.      0.512
# i 182 more rows

```

```

# A tibble: 1 × 7
  n_komm  mean median    sd   min   max n_na
  <int>  <dbl> <dbl> <dbl> <dbl> <dbl> <int>
1     192 0.101 0.166 0.564    -1     1     4

```

Den gjennomsnittlige korrelasjonen mellom årlig sysselsettingsvekst og årlig endring i boligpriser er svakt positiv (0.101), men variasjonen mellom kommuner er betydelig. Dette indikerer at boligmarkedets respons på arbeidsmarkedsendringer varierer systematisk mellom ulike typer kommuner, og at andre faktorer enn lokal sysselsettingsvekst spiller en viktig rolle for boligprisutviklingen.

Årlig endring i innpendling og årlig prosentvis endring i sysselsetting

```

# A tibble: 357 × 5
  kommunenummer n_obs sd_innpPEND sd_syss cor_innpPEND_syss
  <chr>        <int>    <dbl>    <dbl>           <dbl>
1 0301          24    4900.    2.03            0.933
2 1101          24     61.0    2.02            0.808
3 1103          24    1826.    3.36            0.948
4 1106          24     301.    2.41            0.948
5 1108          24     752.    2.84            0.955
6 1111          24     36.6    3.78            0.844
7 1112          24     31.2    3.99            0.894
8 1114          24     20.1    2.59            0.736
9 1119          24     66.1    2.06            0.826
10 1120         24     127.    2.51            0.900
# i 347 more rows

```

```

# A tibble: 1 × 7
  n_komm  mean median    sd   min   max n_na
  <int>  <dbl> <dbl> <dbl> <dbl> <dbl> <int>
1     357 0.719 0.751 0.184 -0.648 0.991     0

```

Den gjennomsnittlige korrelasjonen mellom årlig sysselsettingsvekst og endring i innpendling er høy og positiv (0,72). Dette indikerer at økning i antall arbeidsplasser i en kommune i stor grad absorberes gjennom økt innpendling. Sammenhengen er konsistent på tvers av kommuner og viser at lokale arbeidsmarkededer fungerer som integrerte, funksjonelle regioner snarere enn lukkede enheter.

3.4.a Demeaning

For enkelte relasjoner, som mellom boligpriser og flytting, er det rimelig å anta at virkninger oppstår med tidsforsinkelser. Vi lager derfor boligprisene ett år for å ta hensyn til tregheter i husholdningenes tilpasning.

```
# A tibble: 1 × 1
  cor
  <dbl>
1 -0.0290
```

Demeaning fjerner tidsinvariante kommuneegenskaper som geografisk beliggenhet, størrelse og historisk sentralitet. Dette gjør at korrelasjonene i større grad reflekterer samvariasjon over tid innen samme kommune, snarere enn strukturelle forskjeller mellom kommuner. Dette er særlig nyttig i paneldata, der en kan risikere uekte korrelasjoner som er drevet av permanente nivåforskjeller.

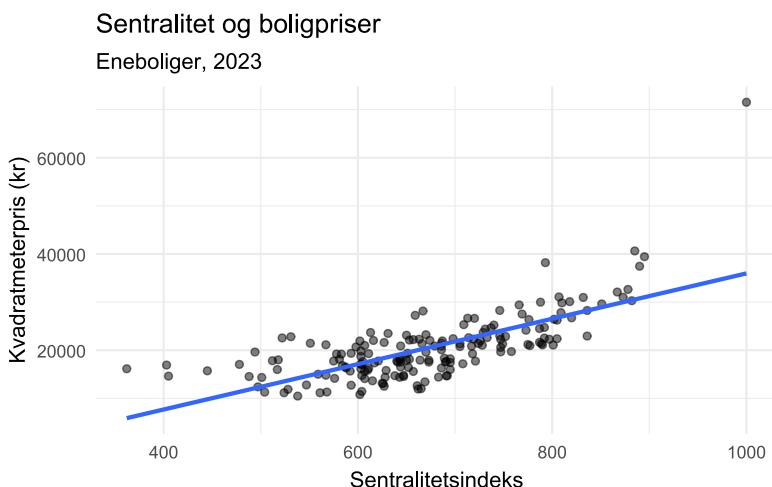
3.5 Pearson vs Spearman

Vi har beregnet en Spearman korrelasjonskoeffisient og en Pearson korrelasjonskoeffisient. Spearman-korrelasjon måler styrken og retningen til variablene som beveger seg i samme retning, uavhengig av om sammenhengen er lineær, mens Pearson-korrelasjon måler den lineære styrken og retningen til variablene mens de beveger seg kontinuerlig. Vi tar eksempelet mellom sentralitet og kvadratmeterpris ("Korrelasjon," n.d.).

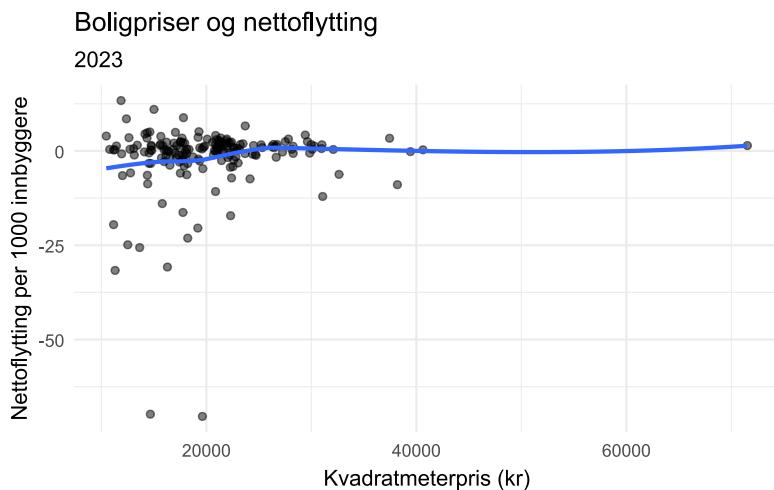
```
# A tibble: 1 × 2
  pearson spearman
  <dbl>     <dbl>
1 0.715     0.710
```

Pearson- og Spearman-korrelasjonene er svært like i størrelse (0,715 og 0,710). Dette indikerer at sammenhengen mellom variablene er tilnærmet lineær, og at resultatet ikke i vesentlig grad påvirkes av uteliggere eller skjev fordeling. Rangeringen av kommunene samsvarer godt med nivåforskjellene i variablene, noe som tyder på en stabil og systematisk samvariasjon.

3.6 Visualisering av noen sammenhenger



Spredningsdiagrammet viser en klart stigende og tilnærmet lineær sammenheng, noe som støtter bruk av Pearson-korrelasjon.



Figuren indikerer en monoton, men ikke-lineær sammenheng, med betydelig spredning. Dette tilsier at Spearman-korrelasjonen gir et mer robust mål enn Pearson.

Korrelasjonskoeffisientene gir kun informasjon om samvariasjon, som er et statistiks mål på strken og retningen, men kan ikke helt tolke årsaksbestemmelsen. Sammenhengene kan være påvirket av simultanitet, utelatte variabler og strukturelle forskjeller mellom kommuner. Videre kan aggregert kommunedata skjule betydelig intern ulikheter. Resultatene bør derfor tolkes som deskriptive, og primært brukes som grunnlag for videre analyse.

3.7 Systematiske forskjeller mellom kommuner i ulike sentralitetskategorier

	# A tibble: 6 × 3	cor	n
	‘Klasse 2023’	<dbl>	<dbl>
		<dbl>	<int>
1	1	-0.418	19
2	2	0.135	67
3	3	-0.0697	234
4	4	0.128	436
5	5	0.136	596
6	6	0.0861	214

Korrelasjonsanalysene viser at sammenhengen varierer systematisk mellom kommuner i ulike sentralitetskategorier. For de fleste sentralitetsklasser (2–6) er korrelasjonen mellom variablene svak og svakt positiv. I de mest sentrale kommunene (klasse 1) er sammenhengen derimot klart negativ. Dette kan indikere at svært høye boligpriser i de største bykommunene bidrar til å begrense netto innflytting, blant annet ved at flere husholdninger velger å bosette seg i omlandskommuner og pendle inn til byene. Samtidig er antall observasjoner i denne gruppen lavt, noe som tilsier at resultatene bør tolkes med forsiktighet. Korrelasjonen mellom boligpriser og nettoflytting er sterkere i store byer enn i distriktskommuner, noe som kan indikere at boligmarkedet spiller en mer aktiv rolle i flyttebeslutninger i urbane strøk. Videre kan en forvente at sammenhengene varierer systematisk mellom ulike typer

kommuner, slik de er klassifisert hos (Andersson et al., 2019). For eksempel kan relasjonen mellom sysselsetting, boligpriser og flytting være sterkere i funksjonelle byregioner enn i perifere distriktskommuner. Dette gir derimot ikke et helt grunnlag for kausale tolkninger men kan indikere systematiske samvariasjoner som danner utgangspunkt for videre analyser og hypoteser.

3.8 Forskjeller mellom kommuner i ulike kategorier hos Andersson et al. (2019)

# A tibble: 4 × 3			
andersson_type	cor	n	
<chr>	<dbl>	<int>	
1 Forstad / omland	0.135	67	
2 Perifere distrikter	0.117	810	
3 Regionale sentra	0.164	670	
4 Storbykjerne	-0.418	19	

Analysene viser tydelige forskjeller mellom kommuner gruppert etter en forenklet typologi. For storbykjerne er korrelasjonen mellom boligpriser og nettoflytting klart negativ, noe som indikerer at svært høye boligpriser i de største bykommunene kan virke begrensende på netto innflytting. For regionale sentra, forstads- og omlandskommuner samt perifere distriktskommuner er sammenhengen derimot svakt positiv. Dette kan tolkes som at boligpriser i disse områdene i større grad reflekterer attraktivitet og etterspørsel, uten de samme kapasitetsbegrensningene som i storbykjerne. Her er også antall observasjoner for storbykjerne relativt lavt, noe som tilslirer at resultatene bør tolkes med en viss varsomhet. Resultatene understøtter dermed antakelsen om at sammenhenger mellom boligmarked og flytteprosesser varierer systematisk med kommunenes funksjonelle rolle i det regionale systemet (Andersson et al., 2019).

3.9 Sammenhengen mellom befolkning og sysselsetting

	Pooled OLS	Kommune + år FE
(Intercept)	-2254.159*** (47.010)	
befolkning	0.676*** (0.001)	0.589*** (0.021)
Num.Obs.	8925	8925
R2	0.976	0.999
R2 Within		0.871
Std.Errors	by: kommunenummer	

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Når faste effekter for både kommuner og år inkluderes, kontrollerer modellen for kommunespesifikke faktorer som er konstante over tid, samt felles nasjonale sjokk i hvert år. I denne spesifikasjonen identifiseres sammenhengen mellom befolkning og sysselsetting utelukkende

ved hjelp av variasjon innen samme kommune over tid. Koeffisienten på 0,601 innebærer at en økning i befolkningen på en person innen en kommune i gjennomsnitt er assosiert med om lag 0,601 flere sysselsatte.

At koeffisienten i modellen med faste effekter er lavere enn i den pooled regresjonen, indikerer at en del av den positive sammenhengen i pooled modellen skyldes varige strukturelle forskjeller mellom kommuner som kontrolleres for når faste effekter inkluderes.

Regresjonsmodellene gir en mer presis beskrivelse av sammenhengen mellom befolkning og sysselsetting enn korrelasjonskoeffisientene i deloppgaven foran. Mens korrelasjonskoeffisientene kun måler graden av samvariasjon mellom variablene, gir regresjonsmodellene en strukturert analyse der sammenhengen estimeres eksplisitt. Dette gjør det mulig å skille mellom ulike kilder til variasjon og gir en klarere forståelse av hvordan befolkning og sysselsetting henger sammen.

Resultatene fra regresjonsmodellene kan i større grad enn en enkel korrelasjonskoeffisient tolkes kausalt. En korrelasjonskoeffisient viser kun om to variabler samvarierer. Den kontrollerer ikke for andre forhold. Regresjonsmodellene kontrollerer derimot for flere viktige faktorer. Når faste effekter for kommuner og år inkluderes, tas det hensyn til både varige forskjeller mellom kommuner og felles sjokk som rammer alle kommuner i samme år. Estimatene bygger da på endringer innen samme kommune over tid. Dette gir et bedre grunnlag for kausal tolkning. Likevel kan resultatene ikke tolkes som fullt ut kausale. Det kan fortsatt finnes omvendt kausalitet og tidsvarierende forhold som ikke er kontrollert for i modellen.

3.10 Relevante hypoteser basert på deskriptiv statistikk

Den deskriptive statistikken viser tydelige sammenhenger mellom sentralitet, befolkningsutvikling, utdanningsnivå, boligpriser og pendlemønstre. Hypotesene som formuleres nedenfor bygger på disse observerte mønstrene og kan forankres i etablert urban- og regionaløkonomisk teori.

For det første viser utviklingskurvene at mer sentrale kommuner har hatt sterkere befolningsvekst enn mindre sentrale kommuner. Dette gir grunnlag for hypotesen om en positiv sammenheng mellom sentralitet og befolningsvekst. Ifølge agglomerasjonsteorien oppstår byvekst fordi konsentrasjon av økonomisk aktivitet gir produktivitetsgevinster gjennom spesialisering og kunnskapsspredning (Brueckner, 2011). I regional vekstteori beskrives dette som kumulative og selvforsterkende prosesser, hvor økonomisk aktivitet tiltrekker ytterligere aktivitet (Capello, 2016).

Når det gjelder netto utflytting og andelen innbyggere i aldersgruppen 20–29 år, viser materialet at sentrale kommuner har en høyere andel unge voksne. Hypotesen er derfor at sentrale kommuner fungerer som mottakere av unge i utdannings- og etableringsfasen. Dette samsvarer med teorier om arbeidsmarkedets tykkelse og jobbmatching, hvor større og mer differensierte arbeidsmarkeder gir høyere forventet avkastning på humankapital (Brueckner, 2011).

Sammenhengen mellom prosentvis arbeidsledighet og netto utflytting fremstår svakere. Teoretisk skulle høy arbeidsledighet øke utflyttingen, men i integrerte arbeidsmarkedsregioner kan husholdninger i stedet velge å pendle. Den monosentriske bymodellen viser hvordan bosted og arbeidssted kan skilles når pendlekostnader balanseres mot boligpriser (Brueckner, 2011).

Videre viser dataene en positiv sammenheng mellom andelen med høyere utdanning og sysselsetting per 1000 innbyggere. Innen endogen vekstteori fremheves humankapital som en sentral drivkraft for produktivitet og sysselsetting (Capello, 2016). Kommuner med høy kompetansekonstrasjon forventes derfor å ha høyere sysselsettingsintensitet.

Sammenhengen mellom sentralitet og boligpriser er sterk. Ifølge budrenteteorien faller boligpriser med økende avstand til sentrum fordi husholdninger må kompenseres for høyere pendlekostnader (Brueckner, 2011). Den sterke positive korrelasjonen mellom sentralitet og kvadratmeterpris er dermed i tråd med byøkonomisk teori.

Når det gjelder netto utpendling og boligpriser i nabokommuner, kan dette forstås gjennom teorien om funksjonelle arbeidsmarkedsregioner, hvor sentrum fungerer som arbeidsplasskjerne og omkringliggende kommuner som bostedskommuner (Capello, 2016).

Samlet sett er de deskriptive funnene konsistente med sentrale bidrag innen urban og regional økonomi, herunder agglomerasjonsteori og regional vekstteori (Brueckner, 2011; Capello, 2016).

3.11 En multippel regresjonsmodell for folketallsutvikling

3.11.a Start-år-modell

```
Call:
lm(formula = folketallsvekst ~ sentralitet + arbeidsledighet_pct +
    hoyere_utdanning_andel + sysselsetting_per1000 +
    netto_innpending_per1000 +
    netto_utflytting_per1000 + ln_boligpris + andel_20_29 +
    factor(landsdel),
    data = df_6_csA)

Residuals:
    Min      1Q  Median      3Q     Max 
-12.2214 -1.9365 -0.4556  1.3411 18.4082 

Coefficients:
                                         Estimate Std. Error t value
(Intercept)                  -22.68322610  7.36548663 -3.080
sentralitet                      0.01872326  0.00276729  6.766
arbeidsledighet_pct            -0.73477563  0.25566456 -2.874
hoyere_utdanning_andel        -0.29479627  0.04790442 -6.154
sysselsetting_per1000          -0.00001328  0.00207137 -0.006
netto_innpending_per1000        0.00013866  0.00003324  4.172
netto_utflytting_per1000       -0.14800636  0.03524131 -4.200
ln_boligpris                     1.02422914  0.89035638  1.150
andel_20_29                      67.82872472 12.04418432  5.632
factor(landsdel)sør             1.99289961  0.50692117  3.931
factor(landsdel)vest              4.11452354  0.47631365  8.638
factor(landsdel)Trøndelag        2.25121239  0.59629797  3.775
factor(landsdel)nord              2.51294555  0.63766844  3.941
                                         Pr(>|t|)    
(Intercept)                           0.002181 **
```

```

sentralitet           0.0000000000357 ***
arbeidsledighet_pct      0.004219 **
hoyere_utdanning_andel 0.000000015077 ***
sysselsetting_per1000    0.994888
netto_innpendinging_per1000 0.0000353855993 ***
netto_utflytting_per1000 0.0000314140496 ***
ln_boligpris          0.250523
andel_20_29            0.0000000291983 ***
factor(landsdel)sør     0.0000958466449 ***
factor(landsdel)vest     < 0.0000000000000002 ***
factor(landsdel)Trøndelag 0.000178 ***
factor(landsdel)nord      0.0000922512419 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.684 on 522 degrees of freedom
Multiple R-squared:  0.3935,   Adjusted R-squared:  0.3796
F-statistic: 28.22 on 12 and 522 DF,  p-value: < 0.0000000000000022

```

Denne modellen er en start-år-modell startåret angir det tidligste året med komplette forklaringsvariabler for hver kommune. Sluttåret angir det siste året med tilgjengelige folketall. Den avhengige variablene er kommunens samlede prosentvise folketallsvekst over hele observasjonsperioden. Alle forklaringsvariabler er målt i startåret. Sentralitet er representert med verdien i 2023 som et tilnærmet tidsinvariant strukturelt mål. Øst fungerer som referansekategori for landsdel. Koeffisientene for øvrige landsdeler viser gjennomsnittlige forskjeller relativt til Østlandet, kontrollert for øvrige forklaringsvariabler.

3.11.b Periode-gjennomsnitt-modell

```

Call:
lm(formula = folketallsvekst ~ sentralitet + arbeidsledighet_pct +
    hoyere_utdanning_andel + sysselsetting_per1000 +
    netto_innpendinging_per1000 +
    netto_utflytting_per1000 + ln_boligpris + andel_20_29 + landsdel,
    data = df_6_csB)

Residuals:
    Min      1Q      Median      3Q      Max 
-12.6748 -2.6629 -0.6332  1.9477 17.7698 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -41.05309693 13.87074272 -2.960   0.003351 ***
sentralitet   0.01783805  0.00518823  3.438   0.000677 ***
arbeidsledighet_pct -0.43702274  0.61251725 -0.713   0.476157    
hoyere_utdanning_andel -0.31960682  0.08524760 -3.749   0.000217 ***

```

```

***  

sysselsetting_per1000 -0.00112992 0.00345739 -0.327  

0.744061  

netto_innpending_per1000 0.00013067 0.00004161 3.140 0.001873  

**  

netto_utflytting_per1000 -0.11551935 0.05912694 -1.954  

0.051755 .  

ln_boligpris 2.76513828 1.67533458 1.650  

0.099995 .  

andel_20_29 87.42684340 22.75080892 3.843 0.000151  

***  

landsdelsør 2.13636058 0.90642567 2.357 0.019137  

*  

landsdelvest 5.60620310 0.81542857 6.875 0.0000000000423  

***  

landsdelTrøndelag 3.18265930 1.07132375 2.971 0.003236  

**  

landsdelnord 2.94233937 1.11890264 2.630 0.009034  

**  

---  

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.574 on 272 degrees of freedom
Multiple R-squared: 0.3922, Adjusted R-squared: 0.3654
F-statistic: 14.63 on 12 and 272 DF, p-value: < 0.0000000000000022

```

Den avhengige variabelen er kommunens samlede prosentvise folketallsvekst, identisk definert som i Start-år-modellen. Forklaringsvariablene er definert som gjennomsnitt over perioden fra startåret til sluttåret, og reflekterer langsiktige strukturelle forhold. Sentralitet er representert med verdien i 2023 og er konstant over perioden, slik at periodegjennomsnittet tilsvarer denne faste verdien. Denne modellformuleringen kan tolkes som en sammenheng mellom kommunens gjennomsnittlige kjennetegn og langsiktig befolkningsutvikling, men gir ikke grunnlag for kausale tolkninger.

3.11.c FE-modell

```

OLS estimation, Dep. Var.: dlog_folketall
Observations: 1,964
Fixed-effects: kommunenummer: 282, år: 15
Standard-errors: Clustered (kommunenummer)

Estimate Std. Error t value Pr(>|t|)  

arbeidsledighet_pct 0.000402380 0.00054810 0.734139 0.4634764  

hoyere_utdanning_andel 0.000495563 0.00077970 0.635579 0.5255680  

sysselsetting_per1000 0.000032864 0.00002220 1.480209 0.1399376  

netto_innpending_per1000 -0.000000544 0.00000101 -0.541228 0.5887796  

netto_utflytting_per1000 0.000566743 0.00006454 8.781164 < 2.2e-16 ***
ln_boligpris 0.001265786 0.00324323 0.390286 0.6966210  

andel_20_29 0.194743092 0.06506961 2.992843 0.0030095 **  

---  

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

RMSE: 0.00779 Adj. R2: 0.408679
 Within R2: 0.096905

FE-modellen utnytter kun variasjon innen samme kommune over tid og identifiserer hvordan årlige endringer i forklaringsvariablene samvarierer med årlig befolkningsvekst. Tidsinvariante forhold som sentralitet og landsdel fanges opp av kommune-faste effekter og kan derfor ikke estimeres eksplisitt. Sammenlignet med tverrsnittsmodellene (start-år-modell og periodegjennomsnitt-modellen) mister flere variabler statistisk signifikans i FE-modellen. Dette er forventet og reflekterer at FE-modellen kontrollerer for uobserverte, tid-invariante forskjeller mellom kommuner. Netto utflytting per 1000 innbyggere har en negativ og signifikant sammenheng med årlig befolkningsvekst innen samme kommune. Andel 20–29 år har en positiv og signifikant sammenheng med årlig befolkningsvekst innen kommuner.

3.12 En multippel regresjonsmodell for sysselsettingsvekst

3.12.a Start-år-modell

```
Call:
lm(formula = sysselsettingsvekst ~ folketallsvekst + sentralitet +
    arbeidsledighet_pct + hoyere_utdanning_andel + sysselsetting_per1000 +
    netto_innpending_per1000 + netto_utflytting_per1000 + ln_boligpris +
    andel_20_29 + factor(landsdel), data = df_7_csA)

Residuals:
    Min      1Q  Median      3Q     Max 
-27.490 -2.475 -0.176  2.436 31.834 

Coefficients:
                                         Estimate Std. Error t value
(Intercept)                 -51.016618639 16.113013991 -3.166
folketallsvekst                0.926928130  0.088780841 10.441
sentralitet                  -0.016309904  0.005919227 -2.755
arbeidsledighet_pct            1.514245924  0.545816970  2.774
hoyere_utdanning_andel        0.116715463  0.112903140  1.034
sysselsetting_per1000          0.000825499  0.004599957  0.179
netto_innpending_per1000       0.000001053  0.000062296  0.017
netto_utflytting_per1000      -0.198829842  0.129980618 -1.530
ln_boligpris                   5.868627959  1.953153641  3.005
andel_20_29                     8.183514948 31.419159634  0.260
factor(landsdel)sør           -0.662665315  1.240768283 -0.534
factor(landsdel)vest           -0.903816654  1.008854458 -0.896
factor(landsdel)Trøndelag     -1.454352411  1.172113935 -1.241
factor(landsdel)nord           1.807183376  1.421858749  1.271
                                         Pr(>|t|)    
(Intercept)                 0.00168 **  
folketallsvekst              < 0.000000000000002 ***
sentralitet                  0.00617 **  
arbeidsledighet_pct           0.00584 **  
hoyere_utdanning_andel        0.30197    
sysselsetting_per1000          0.85768
```

```

netto_innpendingning_per1000          0.98652
netto_utflytting_per1000              0.12702
ln_boligpris                          0.00285 **
andel_20_29                            0.79466
factor(landsdel)sør                  0.59363
factor(landsdel)vest                  0.37094
factor(landsdel)Trøndelag            0.21553
factor(landsdel)nord                  0.20459
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.275 on 343 degrees of freedom
Multiple R-squared:  0.4289,    Adjusted R-squared:  0.4073
F-statistic: 19.81 on 13 and 343 DF,  p-value: < 0.0000000000000022

```

Dette er en tverrsnittmodell der sysselsettingsvekst over hele perioden er avhengig variabel. Forklарingsvariablene (unntatt folketallsvekst) mles i startåret. Folketallsvekst inkluderes som egen forklaringsvariabel (vekstrate over hele perioden). Sentralitet er representert med 2023-verdien (tilnærmet tidsinvariant). Øst er referansekategori for landsdel.

3.12.b Periode-gjennomsnitt-modell

```

Call:
lm(formula = sysselsettingsvekst ~ folketallsvekst + sentralitet +
    arbeidsledighet_pct + hoyere_utdanning_andel + sysselsetting_per1000 +
    netto_innpendingning_per1000 + netto_utflytting_per1000 + ln_boligpris +
    andel_20_29 + factor(landsdel), data = df_7_csb)

Residuals:
    Min      1Q   Median      3Q     Max 
-25.8024 -3.3918 -0.1199  3.0578 27.0748 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept) -59.30601250 29.97423135 -1.979 0.04938  
folketallsvekst          0.85377007 0.11185631 7.633 0.0000000000127
sentralitet        -0.01506838 0.01007074 -1.496 0.13633  
arbeidsledighet_pct       1.44469902 1.23041606 1.174 0.24188  
hoyere_utdanning_andel    0.06406474 0.19007930 0.337 0.73648  
sysselsetting_per1000      0.01863085 0.00702242 2.653 0.00869  
netto_innpendingning_per1000 -0.00003746 0.00007407 -0.506 0.61366  
netto_utflytting_per1000     -0.38273385 0.17412183 -2.198 0.02921  

```

```

ln_boligpris           6.41049561   3.59944387   1.781
0.07659
andel_20_29            -24.33424496  54.74012449  -0.445
0.65718
factor(landsdel)sør    0.20632221   2.12435502   0.097
0.92274
factor(landsdel)vest    0.36811      -1.47916336  1.63936488  -0.902
0.20774
factor(landsdel)Trøndelag -2.46159915  1.94698271  -1.264
0.98309
factor(landsdel)nord     -0.04767269  2.24631505  -0.021

(Intercept)          *
folketallsvekst       ***
sentralitet
arbeidsledighet_pct
hoyere_utdanning_andel
sysselsetting_per1000  **
netto_innpendinging_per1000
netto_utflytting_per1000 *
ln_boligpris          .
andel_20_29
factor(landsdel)sør
factor(landsdel)vest
factor(landsdel)Trøndelag
factor(landsdel)nord
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.409 on 181 degrees of freedom
Multiple R-squared:  0.4573,    Adjusted R-squared:  0.4183
F-statistic: 11.73 on 13 and 181 DF,  p-value: < 0.0000000000000022

```

Periode-gjennomsnitt-modellen bruker periodegjennomsnitt av forklaringsvariabler (startår-sluttår). Sysselsettingsvekst er samlet vekst fra startåret til sluttåret. Folketallsvekst inkluderes som egen forklaringsvariabel (samme periode). Sentralitet er i praksis tidsinvariant (fast 2023-verdi), og gjennomsnittet tilsvarer dette.

3.12.c FE-modell

```

OLS estimation, Dep. Var.: dlog_sysselsetting
Observations: 1,964
Fixed-effects: kommunenummer: 282,  år: 15
Standard-errors: Clustered (kommunenummer)
Estimate Std. Error      t value      Pr(>|t|)
dlog_folketall          0.32531359  0.08528247  3.814542  0.00016777681060
***                      *
arbeidsledighet_pct     -0.00977435  0.00234186 -4.173763  0.00003999230587
***                      *
hoyere_utdanning_andel  0.00184426  0.00198452  0.929324

```

```

0.35351893460208
sysselsetting_per1000      0.00056485 0.00008769 6.441608 0.00000000051111
***
netto_innpending_per1000 -0.00000146 0.00000102 -1.425005
0.15526554201819
netto_utflytting_per1000 0.00031320 0.00016747 1.870189
0.06249716244416 .
ln_boligpris             -0.02098985 0.00946443 -2.217760 0.02737100025999
*
andel_20_29               -0.44528050 0.17564886 -2.535061 0.01178540012032
*
---
Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.020758   Adj. R2: 0.241932
Within R2: 0.185363

```

Avhengig variabel er årlig sysselsettingsvekst, målt som log-differanse i sysselsetting. Befolkningsutvikling inngår som forklaringsvariabel, representert ved årlig log-differanse i folketall. Modellen inkluderer kommune- og årsfaste effekter og utnytter kun variasjon innen samme kommune over tid. Tidsinvariante forhold som sentralitet og landsdel fanges dermed opp av kommune-faste effekter. Standardfeil er klustrert på kommunenivå.

4 Refleksjon

Noen av metodene som er blitt brukt gir ikke grunnlag for kausale tolkninger i seg selv. Disse resultatene må bli sett sammen med andre undersøkelser for å se om det faktisk finnes en forbindelse mellom ulike variabler. Vi har benyttet oss av flere metoder, som gir oss en innsikt i mulige forbindelser, men vi kunne ha gått enda dypere inn i undersøkelsen og brukt flere metoder for å få en dypere innsikt i forholdene, og om det da finnes endogenitetsutfordringer. Det finnes flere ulike måter for å håndtere mulige endogenitetsutfordringer, noe som Zaefarian et al. (2017) har samlet i en artikkel. Artikkelen sier for eksempel at man kan bruke instrumentvariabler og benytte seg av høyere orden moment eller generalisert metode av momenter (Zaefarian et al., 2017).

Videre i undersøkelsen vår, kunne vi ha delt inn variablene i ulike grupper, for eksempel lav/middels/høy andel pendling, ulike aldersgrupper eller tatt hensyn til kjønn. Da kunne vi ha sett på hvordan forholdene endres dersom vi har hensyn til disse ulike gruppene. Dette hadde vært interessant å se nærmere på til videre forskning. Som tidligere nevnt, kunne det også vært lurt å se på ulike årsaker til at kommuner flest har mer innflytting enn utflytting.

5 Konklusjon

Denne studien har analysert systematiske forskjeller mellom norske kommuner med utgangspunkt i sentralitet, demografi, arbeidsmarked, boligmarked, flytting og pendling. Samlet sett viser resultatene konsistente mønstre i tråd med regionaløkonomisk teori.

De mest sentrale kommunene har hatt sterkere befolkningsvekst, høyere andel unge voksne og høyere andel høyt utdannede innbyggere enn mindre sentrale kommuner. Samtidig er boligprisene klart høyest i de mest sentrale områdene, og sammenhengen mellom sentralitet og kvadratmeterpris er både sterk og stabil. Dette er konsistent med urban økonomisk teori,

hvor agglomerasjonsfordeler, arbeidsmarkedets tykkelse og budrentemekanismer bidrar til konsentrasjon av både arbeidsplasser og befolkning i sentrale områder.

Pendleanalysene gir særlig sterke resultater. Endringer i sysselsetting samvarierer i liten grad med årlige endringer i folketall innen samme kommune, men har en svært sterk positiv sammenheng med endringer i innpendling. Dette indikerer at arbeidsmarkedets kortsliktige tilpasning primært skjer gjennom daglig mobilitet, ikke gjennom flytting. Resultatet støtter forståelsen av norske kommuner som deler av funksjonelle arbeidsmarkedsregioner, hvor bosted og arbeidssted i betydelig grad er adskilt. Arbeidsledighet reagerer også relativt raskt på sysselsettingsendringer, noe som viser at lokale arbeidsmarkeder har interne justeringsmekanismer før mer langsiktige demografiske tilpasninger finner sted.

Boligmarkedets rolle fremstår mer sammensatt. På tvers av kommuner er sammenhengen mellom boligpriser og innpendling sterk, mens sammenhengen mellom boligpriser og nettoflytting varierer betydelig. I storbykjerner er sammenhengen mellom boligpriser og nettoflytting negativ, noe som kan tyde på at svært høye boligpriser virker begrensende på netto innflytting og bidrar til økt bosetting i omlandskommuner. I mindre sentrale kommuner er sammenhengen svakere og ofte svakt positiv. Dette peker mot en struktur hvor sentrale kommuner fungerer som arbeidsplasskjerner, mens omlandet i større grad fungerer som bostedsområder.

Regresjonsanalysene nyanserer bildet ytterligere. I pooled modeller fremstår sammenhengen mellom befolkning og sysselsetting sterk, men når det kontrolleres for kommune- og årsfaste effekter reduseres koeffisienten betydelig. Dette viser at en del av den observerte samvariasjonen skyldes varige strukturelle forskjeller mellom kommuner. Fixed effects-modellene viser at innen-kommunale endringer over tid gir svakere og mer selektive effekter enn tverrsnittsanalysene antyder. Netto utflytting og andel unge voksne har konsistente sammenhenger med befolkningsvekst innen kommuner, mens flere andre variabler mister signifikans når forskjeller over tid kontrolleres for.

Alt i alt viser analysen et helhetlig og sammenhengende mønster av et integrert og lagdelt regionalt system i Norge. Sentralitet strukturerer langsiktige nivåforskjeller mellom kommuner, mens kortsliktige tilpasninger i arbeidsmarkedet primært skjer gjennom pendling og endringer i ledighet snarere enn gjennom umiddelbar flytting. Boligmarkedet spiller en viktig rolle, men effektene varierer systematisk mellom ulike typer kommuner.

Resultatene gir ikke grunnlag for entydige kausale tolkninger, men de er i stor grad konsistente med sentrale teorier innen urban og regional økonomi, særlig agglomerasjonsteori, budrenteteori og teorien om funksjonelle arbeidsmarkedsregioner. Videre forskning bør i større grad ta hensyn til endogenitet, tidsforsinkelser og romlige avhengigheter, samt undersøke heterogenitet mellom ulike typer kommuner og befolkningsgrupper.

6 Litteraturliste

- Andersson, M., Håkansson, P. G., & Thorsen, I. (2019). Centralization and urbanization tendencies in norway. In P. G. Håkansson & H. Bohman (Eds.), *Investigating spatial inequalities* (pp. 31–54). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-78973-941-120191003>
- Arel-Bundock, V. (2022). modelsummary: Data and model summaries in R. *Journal of Statistical Software*, 103(1), 1–23. <https://doi.org/10.18637/jss.v103.i01>

- Arel-Bundock, V. (2025). *Modelsummary: Summary tables and plots for statistical models and data: Beautiful, customizable, and publication-ready*. <https://modelsummary.com>
- Bivand, R. (2025). *Spdep: Spatial dependence: Weighting schemes, statistics*. <https://github.com/r-spatial/spdep>
- Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R, second edition*. Springer, NY. <https://asdar-book.org/>
- Bivand, R., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, 27(3), 716–748. <https://doi.org/10.1007/s11749-018-0599-x>
- Brueckner, J. K. (2011). *Lectures on urban economics*. MIT Press.
- Capello, R. (2016). *Regional economics* (Second Edition). Routledge.
- Frøslie, K. F. (2025). *korrelasjon*. <https://snl.no/korrelasjon>
- Göcke, M. (2004). Learning-by-doing versus schooling in a model of endogenous growth. *Journal of Economics*, 83(1), 47–69. <https://doi.org/10.1007/s00712-003-0064-4>
- Gohel, D., & Skintzos, P. (2025). *Flextable: Functions for tabular reporting*. <https://ardata-fr.github.io/flextable-book/>
- Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3), 1–25. <https://www.jstatsoft.org/v40/i03/>
- Horgen, E. H. (2025). Hvorfor ulike arbeidsledighetstall? In *SSB*. <https://www.ssb.no/arbeid-og-lonn/sysselsetting/artikler/hvorfor-ulike-arbeidsledighetstall>
- Høydahl, E. (n.d.). *Sentralitetsindeksen. Oppdatering med 2020-kommuner*.
- Korrelasjon. (n.d.). In *Metodeguiden*. <https://www.metodeguiden.no/korrelasjon>
- Langsrød, Ø., & Bruusgaard, J. (2024). *PxWebApiData: PX-web data by API*. <https://statisticsnorway.github.io/ssb-pxwebapidata/>
- Ooms, J. (2025). *Writexl: Export data frames to excel xlsx format*. <https://ropensci.r-universe.dev/writexl>
- Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1), 439–446. <https://doi.org/10.32614/RJ-2018-009>
- Pebesma, E. (2025). *Sf: Simple features for r*. <https://r-spatial.github.io/sf/>
- Pebesma, E., & Bivand, R. (2023a). *Spatial Data Science: With applications in R*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429459016>
- Pebesma, E., & Bivand, R. S. (2023b). *Spatial data science with applications in R*. Chapman & Hall. <https://r-spatial.org/book/>
- Roger Bivand. (2022). R packages for analyzing spatial data: A comparative case study with areal data. *Geographical Analysis*, 54(3), 488–518. <https://doi.org/10.1111/gean.12319>
- Salomão, A. (2024). Pearson-korrelasjon: Forstå matematikken bak relasjoner - mind the graph-bloggen. In *Mind the Graph-blogg*. <https://mindthegraph.com/blog/nb/pearson-correlation/>

- Sørlien, K. (2025). *Fruktbarheten går opp*. <https://www.ssb.no/befolkning/fodte-og-dode/statistikk/fodte/artikler/fruktbarheten-gar-opp>
- Spinu, V., Grolemund, G., & Wickham, H. (2024). *Lubridate: Make dealing with dates a little easier*. <https://lubridate.tidyverse.org>
- Todaro, M. P. (1969). A model of labor migration and urban unemployment in less developed countries. *The American Economic Review*, 59(1), 138–148. <https://www.jstor.org/stable/1811100>
- Wickham, H. (2023). *Tidyverse: Easily install and load the tidyverse*. <https://tidyverse.tidyverse.org>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wickham, H., & Bryan, J. (2025). *Readxl: Read excel files*. <https://readxl.tidyverse.org>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A grammar of data manipulation*. <https://dplyr.tidyverse.org>
- Xie, Y. (2014). Knitr: A comprehensive tool for reproducible research in R. In V. Stodden, F. Leisch, & R. D. Peng (Eds.), *Implementing reproducible computational research*. Chapman; Hall/CRC.
- Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman; Hall/CRC. <https://yihui.org/knitr/>
- Xie, Y. (2025). *Knitr: A general-purpose package for dynamic report generation in r*. <https://yihui.org/knitr/>
- Zaefarian, G., Kadile, V., Henneberg, S. C., & Leischnig, A. (2017). Endogeneity bias in marketing research: Problem, causes and remedies. *Industrial Marketing Management*, 65, 39–46. <https://doi.org/10.1016/j.indmarman.2017.05.006>