

HJEMMEEKSAMEN MET4



Vår 2024

Start: 22. april 2024 kl. 09:00

Slutt: 24. april 2024 kl. 14:00

BESVARELSEN SKAL LEVERES I WISEFLOW

På våre nettsider finner du informasjon om hvordan du leverer din besvarelse:

<https://www.nhh.no/for-studenter/eksamen/heimeeksamen-og-innlevering/>

Kandidatnummer blir oppgitt på StudentWeb i god tid før innlevering. Kandidatnummer skal være påført på alle sider øverst i høyre hjørne (ikke navn eller studentnummer). Ved gruppeinnlevering skal alle gruppemedlemmers kandidatnummer påføres.

UTFYLLENDE BESTEMMELSER OM EKSAMEN

Du finner utfyllende bestemmelser under overskriften «Fulltidstudiene»:

<https://www.nhh.no/for-studenter/forskrifter/>

Se avsnitt om gjennomføring av eksamen uten tilsyn, kap. 4.0 om vurderingsformer.

Antall sider, inkludert forside: 7

Antall vedlegg: 1 (heks_v24.Rdata)

Introduksjon

Å forstå og måle livskvalitet er avgjørende for å få et mer helhetlig bilde av et lands velferd og innbyggernes velvære. Bruttonasjonalprodukt (BNP) er en viktig økonomisk indikator som måler et lands økonomiske ytelse gjennom summen av alle varer og tjenester som produseres innenfor et land i løpet av et år. Det er mange aspekter som er essensielle for menneskers lykke og tilfredshet med livet som ikke fanges opp av økonomiske indikatorer som BNP. I denne oppgaven skal dere utforske akkurat disse sammenhengene.

Studier har vist at høyere BNP per innbygger ikke nødvendigvis korrelerer med høyere livstilfredshet. For eksempel kan land med moderat BNP ha svært høye lykkenivåer på grunn av sterk sosial samhörighet, tilgang til natur og fritidsmuligheter, og et godt balansert arbeidsliv. Samtidig er det slik at økonomisk velstand ofte henger sammen med selvrapportert lykke.

World Happiness Report er en rapport som skrives hvert år basert på Gallup gjennomførte spørreundersøkelser av ca 1000 innbyggere i hvert land. Disse dataene gir innsikt i hvordan ulike faktorer bidrar til følelsen av lykke og velvære på tvers av land. Vi har tilgang til et aggregert datasett på land-år.

Se en video om the world happiness report [her](#).

Last inn datasettene ved navn `whr`, `whr_balansert`, `whr_average` og `whr_ny` i R-minnet ved hjelp av følgende kommando, der vi antar at filen ligger i arbeidsmappen din:

```
load("heks_v24.Rdata")
```

Datasettet `whr` inneholder lykkenivåer (`lifeladder`) fra ulike land i perioden 2011 til 2022, i tillegg til en rekke andre variabler. Se Tabell 1 for den originale beskrivelsen av variablene på engelsk. Merk at variabelen `corruption` ikke er med i dette datasettet, men beskrivelsen er relevant for oppgave 4.

Det viser seg at datasettet `whr` ikke er **komplett**. Flere land mangler observasjoner for flere år. Altså er panel-datasettet ikke balansert. Datasettet `whr_balansert` er en balansert versjon av `whr` som er laget ved å filtrere ut (fjerne) de landene som fullstendig mangler observasjoner for enkelte år i perioden 2011 - 2022. For eksempel er Algerie observert i alle år utenom 2013 og 2015 (det finnes ingen rad i `whr` for Algerie for 2013 og 2015) og derfor blir ingen observasjoner fra Algerie inkludert i `whr_balansert`. Det er likevel noen NA verdier i `whr_balansert`.

Spesielt mange land blir fjernet fordi de mangler observasjoner for 2020. Ideelt sett er denne filtreringen helt tilfeldig i den forstand at det ikke er spesielle kjennetegn (for eksempel rike land fra en bestemt verdensdel) ved land som inkluderes/ekskluderes fra datasettet `whr_balansert`. Da vil det balanserte datasettet `whr_balansert` ennå være representativt for verden.

Tabell 1: Variabler i datasettet

Variabel	Beskrivelse
<code>lifeladder</code>	National average response to Gallup World Poll's life evaluation question, using a 0 to 10 ladder metaphor, 10 being the highest possible life satisfaction.
<code>logGDP</code>	Logged GDP (in dollars) per capita data from World Development Indicators.
<code>lifeexpectancy</code>	Life expectancy (in years) based on WHO data.
<code>socialsupport</code>	National average of binary responses to having support in times of trouble, with 1 being the highest possible support.
<code>freedom</code>	National average of responses to satisfaction with freedom to choose life activities, with 1 being the highest possible freedom.
<code>generosity</code>	Residual of regressing national average of charity donation responses on GDP per capita. Larger values corresponds to more generosity.
<code>percep corruption</code>	Average national perception of corruption based on GWP survey responses regarding government and business corruption. "Is corruption widespread throughout the government or not" and "Is corruption widespread within businesses or not?" The overall perception is just the average of the two 0-or-1 responses.
<code>region</code>	Classification of countries into world regions.
<code>corruption</code>	UNODC official statistics on the number of corruption cases per 100,000 inhabitants.
<code>year</code>	Year of observation.
<code>country</code>	Country of observation.

Oppgaver

Svar på følgende oppgaver. Prosentvis omfang er veiledende og trengs ikke følges absolutt.

1: Deskriptiv statistikk og utforskende dataanalyse (20%)

- a) Utfør en utforskende dataanalyse med deskriptiv statistikk basert på det balanserte datasettet `whr_balansert`. Fokuser spesielt på variabelen `lifeladder`. Presenter en tabell og tre figurer dere mener er relevante, og forklar hva de viser.
- b) Koronaåret 2020 var et år mange ser tilbake på som tøft. Finner vi grunnlag for dette i datasettet `whr_balansert`? Formuler en hypotese og test den.

2. Vurdering av populasjon (30%)

Datasettet `whr_average` inneholder gjennomsnittsverdier av variablene i `whr` per land. I tillegg inkluderer `whr_average` en dummyvariabel `included_in_balanced` som tar verdien `TRUE` dersom landet er med i det filtrerte datasettet `whr_balanced` og `FALSE` ellers.

- a) Bruk deskriptiv statistikk basert på datasettet `whr_average` til å utforske om filtreringen vi gjør for å lage det balanserte datasettet `whr_balansert` er tilfeldig. Fokuser spesielt på variablene `lifeladder` og `logGDP`.
 - b) Estimer en eller flere logistiske regresjonsmodeller basert på datasettet `whr_average` som kan brukes til å utforske om filtreringen vi gjør for å lage det balanserte datasettet `whr_balansert` er tilfeldig. Diskuter hvilken populasjon `whr_balansert` eventuelt representerer.
 - c) Bruk først datasettet `whr` til å se på gjennomsnitt av `logGDP` og `lifeladder` over tid. Er det noen overraskende funn? Sammenlign med tilsvarende analyse for det balanserte datasettet `whr_balansert`. Kommenter i lys av funnene dine i oppgave 2a) og 2b).
- R-hint oppgave c): Dersom man har et datasett `df` og ønsker å lage et datasett bestående av gjennomsnittet av variabelen `x` per `year` kan man bruke følgende R-kode:

```
df_average <- df %>%  
  group_by(year) %>%  
  summarise(average_x = mean(x, na.rm = TRUE))
```

argumentet `na.rm = TRUE` i `mean()` funksjonen brukes dersom det er NA verdier i `df`.

Argumentet `na.rm = TRUE` i `mean()`-funksjonen vert brukt dersom det er NA-verdiar i `df`. ## 3: Regresjonsanalyse (30%)

- a) Utfør en eller flere regresjonsanalyser for å undersøke forholdet mellom `lifeladder` (som avhengig variabel) og ulike forklaringsvariabler. Diskuter resultatenes tolkning i lys av antagelsen om eksogenitet. Bruk datasettet `whr_balansert`.
- b) Diagnostiser en av regresjonsmodellene fra oppgave a).

3: Analyse med nytt mål på korrupsjon (20%)

Vi skal nå se på datasettet `whr_ny`. Dette datasettet inkluderer de balanserte WHR dataene fra `whr_balansert` og en ny variabel `corruption`. Denne variabelen er samlet inn av et FN organ (UNODC) og angir offisiell statistikk på antall korrupsjonssaker per 100,000 innbygger. Det vil si at vi nå har to mål på korrupsjon; oppfattet korrupsjon målt via spørreundersøkelser gjennomført av Gallup (`percep corruption`) og offisiell statistikk rapportert til FN om antall korrupsjonssaker per 100,000 innbygger (`corruption`).

- a) Hvilke sammenhenger finner vi mellom de to målene på korrupsjon? Illustrer med figur eller tabell. Hvilke inntrykk får dere av de to målene? Hva overrasker dere, og hva kan det bety?
- b) Analyser sammenhengen mellom lykke (som avhengig variabel) og målene på korrupsjon, samt andre forklaringsvariabler ved bruk av regresjonsmodeller både med og uten faste effekter. Tolk resultatet.

Administrative bestemmelser

- Gruppeeksamen i MET4 må leveres i grupper på 2, 3, eller 4 studenter.
- Det er ikke tillatt å diskutere eksamen med studenter utenfor din gruppe etter at oppgavesettet er frigitt.
- Besvarelsene vil bli rettet i henhold til retteskjema postet på Canvas.
- Du kan besvare eksamen på norsk eller engelsk.
- Send en e-post til både Geir Drage Berentsen (geir.berentsen@nhh.no) og Julie Bjørkheim (julie.bjorkheim@nhh.no) ved spørsmål til oppgaven.
- Ingen informasjon utover eksamensteksten vil bli gitt. Ingen spørsmål av typen “Er det meningen at vi skal ... ?” vil bli besvart. Det er opp til dere å gjøre fornuftige presiseringer eller innskrenkninger dersom det er nødvendig.
- Kun spørsmål vedrørende eventuelle feil oppgaveteksten vil bli besvart mens eksamen pågår. Rettelser av betydning vil bli postet på Canvas.
- Rapporten skal ikke være lenger inn 10 sider. Tabeller, figurer og referanser er inkludert i de 10 sidene. Dersom rapporten har en forside uten noen form for svar på oppgavene kan forsiden komme i tillegg til de 10 sidene. Innholdsfortegnelse teller med i sidetallet, men er ikke nødvendig. Prioriter hva dere tar med i rapporten!
- Rapporten skal skrives med fonten Times New Roman, størrelse 12 og linjeavstand 1.15. Tekst i figurer og tabeller kan ha font ned til størrelse 9.
- Eksamen administreres i Wiseflow. Besvarelsen må leveres som en enkelt .pdf-fil. Andre format (f.eks. .doc, .docx eller .R) er ikke akseptert.

Referanser og andre kilder

Helliwell, J. F., Layard, R., Sachs, J. D., Aknin, L. B., De Neve, J.-E., & Wang, S. (Eds.). (2023). World Happiness Report 2023 (11th ed.). Sustainable Development Solutions Network. <https://worldhappiness.report/ed/2023/>

Helliwell, J. F., Huang, H., Wang, S., & Norton, M. (2020). Statistical appendix for chapter 2 of world happiness report 2020. *New York: Sustainable Development Solutions Network*. <https://happiness-report.s3.amazonaws.com/2021/Appendix1WHR2021C2.pdf>

Data UNODC - Metadata Information Corruption and Economic Crime https://data.unodc.un.org/sites/dataunodc.un.org/files/metadata_corruption_and_economic_crime.pdf

World Happiness Report 2024 Launch YouTube <https://www.youtube.com/playlist?list=PLfe8F4t8Q4aaN4dhpYTgquuHz2KJB9qtp>

Link til Kaggle for å laste ned data: <https://www.kaggle.com/code/lukestarnes/world-happiness-report-2023/notebook> <https://www.kaggle.com/datasets/mayzannilarthein44/world-happiness-report-2015-to-2022>

Nrk «Ny rapport: Unge ser meir negativt på framtida»: https://www.nrk.no/norge/fns-lykkerapport-viser-at-noreg-er-lykkelege_-men-at-mange-unge-er-misnogde-1.16811390 20.03.2024

United Nations Office on Drugs and Crime <https://dataunodc.un.org/dp-crime-corruption-offences>. Se metadata for mer informasjon om disse dataene.