

SOK-1004 H24 - Case 3

Andrea Karlsen

Instruksjoner

Denne oppgaven er laget opprinnelig av Even S. Hvinden. Den er endret og oppdatert av Derek J. Clark. Sistnevnte er ansvarlig for eventuelle feil og mangler.

Oppgaven skal løses interaktivt i RStudio ved å legge inn egen kode og kommentarer. Det ferdige dokumentet lagres med kandidatnummeret som navn `[kandidatnummer]_SOK1004_C3_H24.qmd` og lastes opp på deres GitHub-side. Hvis du har kandidatnummer 43, så vil filen hete `43_SOK1004_C3_H24.qmd`. Påse at koden kjører og at dere kan eksportere besvarelsen til pdf. Lever så lenken til GitHub-repositoriet i Canvas.

Bakgrunn

Prisveksten har vært høy i Norge, og som [denne overskriften fra 2023](#) viser kan en del av prisveksten skyldes en historisk stor vekst i matpriser.

[Denne saken fra Dagens Næringsliv](#) viser at en reduksjon i matpriser kan også bidra til at inflasjonen faller:

Hvor mye har matpriser bidratt til prisveksten? I denne oppgaven skal vi benytte prisdata fra SSB til å besvare dette spørsmålet. Jeg anbefaler dere å lese [Konsumprisindeksen - en levekostnadsindeks](#) av Randi Johannesen, Økonomiske analyser 5/2014.

Oppgave I: Tolk vekstbidraget

For å forstå øvelsen skal vi først se nærmere på hvordan en prisindeks bygges opp. La P_t være verdien til konsumprisindeksen i tid t , gitt av et vektet gjennomsnitt av $n \geq 1$ priser eller prisindekser

$$P_t = \sum_{i=1}^n v_{i,t} p_{i,t} \quad (1)$$

hvor vektene summerer til én i hver periode t , $\sum_{i=1}^n v_{i,t} = 1$. Vektene viser hvor stor andel av budsjettet en representativ konsument bruker på vare eller vare gruppe i . Vi vil se på månedlig KPI på hovedgruppenivå, hvor $n = 12$ og t løper fra januar 1979 til april 2024 (som var siste måned tilgjengelig da dette dokumentet ble skrevet).

Vi betegner endringen over tolv måneder i KPI ved $P_t - P_{t-12} := \Delta P_t$, eller

$$\Delta P_t = \sum_{i=1}^n v_{i,t} p_{i,t} - \sum_{i=1}^n v_{i,t-12} p_{i,t-12} = \sum_{i=1}^n \Delta(v_{i,t} p_{i,t}). \quad (2)$$

Merk at både vektene og prisene kan endre seg fra $t - 12$ til t . Vekter endres i januar hvert år, og er fast resten av året. I praksis vil vektene endre seg lite. For å forenkle fremstillingen vil vi anta at $v_{i,t} = v_{i,t-12}$. I tillegg så deler vi uttrykket på P_{t-12} , og ganger med 100. Da har vi

$$100 \times \frac{\Delta P_t}{P_t} = 100 \times \frac{\sum_{i=1}^n v_{i,t-12} \Delta p_{i,t}}{P_{t-12}}. \quad (3)$$

På venstre side av likhetstegnet har vi prosentvis tolv måneders endring i konsumprisindeksen, eller inflasjon. På høyre side har vi en sum med n argumenter. Vi fokuserer nå på et vilkårlig element i ,

$$100 \times \frac{v_{i,t-12} \times \Delta p_{i,t}}{P_{t-12}}. \quad (4)$$

Tolk ligning (4). Gi en konkret forklaring på hva tallet representerer.

Svar på oppgave 1

Ligning 4 brukes til å regne ut vekstbidraget til én varegruppe (i) til den totale endringen i konsumprisindeks. Endringen er målt fra år til år. Den viser hvor mye endringen av prisen på varegruppen påvirker inflasjon (med tanke på hvor mye vekt varegruppen har i budsjettet).

v i, t-12 viser vekten for varegruppe i for 12 måneder siden

delta P i, t står for endringen i prisindeksen over de siste 12 månedene

P t-12 er kpi-en for 12 måneder siden

Så ganges det med 100 slik at vi kan se tallet i prosent

Oppgave II: Rydd i data

Vi begynner med å rydde og laste inn pakker.

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

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --  
v dplyr      1.1.4      v readr      2.1.5  
v forcats    1.0.0      v stringr    1.5.1  
v ggplot2    3.5.1      v tibble     3.2.1  
v lubridate  1.9.3      v tidyr      1.3.1  
v purrr      1.0.2  
-- Conflicts ----- tidyverse_conflicts() --  
x dplyr::filter() masks stats::filter()  
x dplyr::lag()     masks stats::lag()  
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(lubridate)  
library(rjstat)
```

Attaching package: 'rjstat'

The following object is masked from 'package:dplyr':

id

```
library(janitor)
```

Attaching package: 'janitor'

The following objects are masked from 'package:stats':

chisq.test, fisher.test

```
library(gdata)
```

Attaching package: 'gdata'

The following objects are masked from 'package:dplyr':

combine, first, last, starts_with

The following object is masked from 'package:purrr':

keep

The following object is masked from 'package:tidyr':

starts_with

The following object is masked from 'package:stats':

nobs

The following object is masked from 'package:utils':

object.size

The following object is masked from 'package:base':

startsWith

```
library(httr)
```

Vi bruker dataene fra [Tabell 0313: Konsumprisindeksen fra SSB](#). Jeg laster ned ved hjelp av API. Se [brukerveiledningen](#) her.

```
url <- "https://data.ssb.no/api/v0/no/table/03013/"

query <- '{
  "query": [
    {
      "code": "Konsumgrp",
      "selection": {
        "filter": "vs:CoiCop2016niva2",
        "values": [
          "01",
```

```

        "02",
        "03",
        "04",
        "05",
        "06",
        "07",
        "08",
        "09",
        "10",
        "11",
        "12"
    ]
  }
},
{
  "code": "ContentsCode",
  "selection": {
    "filter": "item",
    "values": [
      "KpiIndMnd",
      "KpiVektMnd"
    ]
  }
}
],
"response": {
  "format": "json-stat2"
}
}'

hent_indeks.tmp <- url %>%
  POST(body = query, encode = "json")

df <- hent_indeks.tmp %>%
  content("text") %>%
  fromJSONstat() %>%
  as_tibble()

```

Følgende kode benytter kommandoen `ymd` fra `lubridate` til å lage en anstendig tidsserie.

```
df <- df %>%
  separate(måned,
           into=c("year", "month"),
           sep="M") %>%
  mutate(dato = ymd(paste(year, month, "1")))
```

Nå er det deres tur til å rydde. Slett variablene `year` og `month`. Gi variablene formålstjenlige navn. Påse at variablene har riktig format. Fjern data fra før år 2011, slik at vi kan beregne tolv måneders endring fra 2012. Løs oppgaven slik at du ekstraherer navnene på variablene og verdiene ved hjelp av kode.

Hint. Bruk `as.Date()` for å filtrere på datoer.

```
# Svar på oppgave II

# Konverterer "dato" til et date-object
df$dato <- as.Date(df$dato)

# Sletter variablene year og month, gir variabler nye navn, filtrerer bort data fra før 2011
df <- df %>%
  select(-year) %>%
  select(-month) %>%
  rename("vekt" = "value") %>%
  rename("kons_indeks" = "statistikkvariabel") %>%
  filter(dato >= as.Date("2011-01-01"))
```

Oppgave III: Beregn et vektet gjennomsnitt

Vi skal nå beregne KPI som et vektet gjennomsnitt av konsumgruppene og sammenlign med totalindeksen.

Oppgave IIIa: Endre verdi på vektene

Del vektene i `df` på 1000, og sjekk at de summerer seg til om lag 1 for hver måned. Hvor store avvik får du?

```
# løs oppgave IIIa her

df <- df %>%
  mutate(vekt = vekt/1000)
```

Jeg observerer ikke spesielt store avvik. De fleste observasjonene ligger mellom 0,9 til 1,1. Det er også en del observasjoner rundt 0,8 og 1,3.

Verdier innad samme varegruppe har minimale avvik. For eksempel verdien for gruppen matvarer og alkoholfrie drikkevarer finner vi på et spenn fra 1,331 til 1,325.

Oppgave IIIb: Beregn prisindeksen fra gruppene med vekter

Beregn en totalindeks hvor dere bruker vektene og verdiene på prisindeks i formel (1) fra oppgave I. Hvordan kan du vite om beregningen er riktig?

```
# løs oppgave IIIb her

df_indeks <- df %>%
  filter(kons_indeks %in% first("Konsumprisindeks (2015=100)")) %>%
  select(dato, kons_indeks, vekter) %>%
  arrange(dato, kons_indeks)

df_vekter <- df %>%
  filter(kons_indeks %in% first("Konsumprisindeks (vekter)")) %>%
  select(dato, kons_indeks, vekter) %>%
  arrange(dato, kons_indeks)

df_total <- df_indeks %>%
  select(dato, kons_indeks) %>%
  mutate(verdi = df_indeks$vekter*df_vekter$vekter)
```

Oppgave IV: Beregn matprisens bidrag til vekst

Lag en figur som illustrerer vekstbidraget til konsumgruppen “Matvarer og alkoholfrie drikkevarer”. Gi figuren en anstendig tolkning.

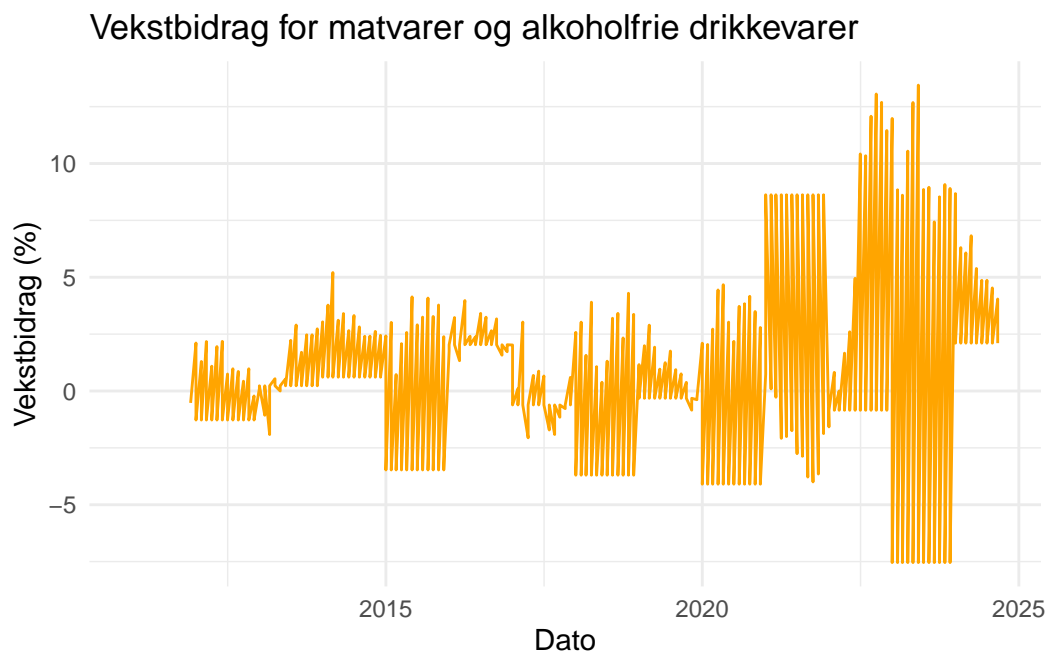
```
# løs oppgave IV her

df_filtered <- df %>%
  filter(konsumgruppe == "Matvarer og alkoholfrie drikkevarer")

df_filtered <- df_filtered %>%
  mutate(vekstbidrag = 100 * (vekter - lag(vekter, 12)) / lag(vekter, 12))
```

```
df_filtered %>%
  ggplot(aes(x = dato, y = vekstbidrag)) +
  geom_line(color = "orange") +
  labs(title = "Vekstbidrag for matvarer og alkoholfrie drikkevarer",
       x = "Dato",
       y = "Vekstbidrag (%)") +
  theme_minimal()
```

Warning: Removed 1 row containing missing values or values outside the scale range (`geom_line()`).



Grafen viser vekstbidraget for konsumgruppen “Matvarer og alkoholfrie drikkevarer”. Når linjen er positiv, betyr det at prisindeksen for denne gruppen har økt sammenlignet med samme periode forrige år. Når linjen er negativ, har prisindeksen falt.