

Nyregistreringer og konkurser blant norske foretak

Introduksjon

I denne obligatoriske mappeoppgaven skal du skrape data fra Brønnøysundregistrene, og se på nyregistreringer og konkurser blant norske foretak.

Data

På nettsiden til Brønnøysundregistrene kan du gjøre søk på kunngjøringer vedrørende nyregistreringer og konkurser på foretak.

<https://w2.brreg.no/kunngjoring/>

Du skal gjøre to søk på “Velg kunngjøringstype”. Ett på nyregistreringer og ett på konkurser. Se Figure 1 under for hvordan søkevinduet ser ut.

Oppgave

Du skal skrape dataene fra de to søkene og lage en tabell med følgende kolonner:

- Foretaksnavn
- Organisasjonsnummer (9 siffer), her skal du fjerne mellomrom og ta bort privatpersoner som er registrert med fødselsdato (6 siffer)
- Registreringsdato (dd.mm.åååå)
- Kunngjøringstype (Nyregistrering eller Konkurs/Tvangsavvikling)

For Kunngjøringstype på Konkurs/Tvangsavvikling skal du kun ta med kategorien “konkurs-åpning”.

Avansert søk i kunngjøringer

Fra dato * (påkrevd)	Til dato
<input type="text" value="01.01.2024"/>	<input type="text" value="15.03.2024"/>
Velg landsdel	Velg kunngjøringstype
<input type="text" value="Hele landet"/>	<input type="text" value="Konkurs/tvangsavvikling"/>
Velg fylke	Velg underkategori 1
<input type="text" value="---"/>	<input type="text" value="Konkurs-åpning"/>
Velg kommune	Velg underkategori 2
<input type="text" value="---"/>	<input type="text" value="---"/>
Velg hovedbransje	
<input type="text" value="Alle bransjer"/>	
Velg bransje	
<input type="text" value="---"/>	
Velg bransjegruppe	
<input type="text" value="---"/>	

Figure 1: Søkevindu på nettsiden

Utfordringer

Søket ditt benyttes til å lage en unik url. Denne url'en skal du "manipulere" og bruke til å skrape dataene. Du **skal** benytte deg av `purrr`-pakken og `map()` funksjonen når du skrapet dataene. Dette må du gjøre fordi hvert søk er begrenset til 5000 treff. Du må derfor skrape dataene i flere omganger, og deretter kombinere dataene.

Det enkleste er å manipulere tidsperioden. Du skal skrape data fra 1 januar 2017 frem til dagens dato (når du kjører koden din). Du kan også manipulere søket ved å velge fylke. Du kan også velge å skrape dataene per fylke, og deretter kombinere dataene.

Rensking

Du skal rense dataene slik at de er klare for analyse. Dette innebærer blant annet å fjerne duplikater og manglende verdier. Du skal også kategorisere dataene per fylke. Dersom du gjør et søk på "Hele landet", se Figure 2 under for hvordan resultatet ser ut (starter med Oslo).

I første linje for hvert fylke, ser du tittelen for hvert fylke. Så du har to alternativer, søke på alle fylker, eller søke per fylke.

Du søkte på:

Dato 01.01.2024 til 15.03.2024
Sted Hele landet
Kunngjøringstype Konkurs-åpning

Antall treff 1040

		Dato	Kunngjøringstype
Oslo			
A MILJØ GRUPPEN AS	923 174 265	06.02.2024	Konkursåpning
A R BISTRO AS	929 325 249	13.02.2024	Konkursåpning
ABDI ABDIWAHID HASSAN	210186	19.01.2024	Konkursåpning
ADETILOYE MARTIN	210188	13.02.2024	Konkursåpning
ADLER TAXISERVICE AS	928 033 147	11.01.2024	Konkursåpning
AKSJESELSKAPET HYTTE 547	827 068 322	14.03.2024	Konkursåpning
-			
ALMA BYGG	993 141 569	01.03.2024	Konkursåpning

Figure 2: Resultat av søk som html tabell

Analyse

Benytt dato til å aggregere til månedsdata for hele landet og per fylke. Du vil da ha en tidsserie per fylke og for hele landet for nyregistreringer og konkurser. Du skal deretter lage en tidsserie for differansen mellom nyregistreringer og konkurser, som du kaller “netto”.

Verdens helseorganisasjon erklærte 11. mars 2020 koronavirusutbruddet som en pandemi, og 5. mai 2023 erklærte de pandemien for over. La hele mars 2020 være en “pandemi-måned”, og hele mai 2023 være en “etter-pandemi-måned”. Du skal til slutt lage en figur som viser “netto” foretak per måned før, under og etter korona fra januar 2017 til februar 2024. Velg selv om du vil presentere analysen for hele landet, eller velg et spesifikt fylke. Figuren skal være selvforklarende.

Leveranse

Skriv en R, eller quarto fil som gjør denne analysen. Det forventes at du benytter egendefinerte funksjoner i denne prosessen, og at du har en ryddig og oversiktlig kode med kommentarer. Lagre kodefila og en .png fil av figuren fra din analyse i ditt github repository, og lever url til repositoryet i Canvas innen fristen.