

SOK-1005: Prosjektoppgave Vår 2022

Prosjektoppgaven kan gjennomføres alene, eller i en gruppe på inntil 3 personer. Målet med prosjektet er at du tar oss gjennom hele data science prosessen med å lese inn, bearbeide og presentere din analyse av dataene som svarer på problemstillingene i denne oppgaven.

Filene som utgjør leveransen på prosjektet skal ligge på et åpent github repository. Lenken til dette repositoryet skal leveres på Wiseflow innen fristen. Er det en gruppeinnlevering skal medlemmene i gruppa levere samlet. Utfyllende veiledning kommer.

Prosjektet skal gjennomføres i R, og det skal genereres en html fil fra en Rmarkdown fil, som er å anse som selve besvarelsen. Dersom prosjektet består av flere filer må det leveres en readme fil som forklarer hvordan prosjektet er bygget opp.

Kvalitet og drøfting teller fremfor kvantitet. Dersom du/dere lager en figur eller tabell for et produkt/produktgruppe er det begrunnelsen og beskrivelsen av denne som er viktig. Det er ikke nødvendig å gjenta denne figuren/tabellen flere ganger.

Det er ingen begrensninger på lengde, men tenk deg at besvarelsen skal leses av personer som ville vurdert om ditt firma fikk oppdraget som er beskrevet i teksten. Ikke for langt, ikke for kort, du/dere bestemmer lengden.

Bakgrunn

Artikkelen “Sandwich analytics: A dataset comprising one year's weekly sales data correlated with crime, demographics, and weather” av Spaulding et. al. (2019) gir oss et innblikk i salgsdata fra en hurtigmatkjede for 10 av deres utsalgssteder i USA.

Artikkelen finner du her:
<https://www.sciencedirect.com/science/article/pii/S2352340919306067#!>

eller i mappen “prosjektoppgave” i Canvas. Les artikkelen grundig før dere starter på prosjektet.

Dataene kan lastes ned her: <https://data.mendeley.com/datasets/6htjnf578b/1>

Bakgrunn

Du/ditt team er nå ansatt som analytiker i ACME analytics. Selskapet som “eier” dataene har lyst ut en anbudskonkurranse. De ønsker å få utviklet et system som rapporterer salg for alle utsalgsstedene i sitt konsern. Sandwich dataene skal benyttes for å vise hva du/dere kan levere på.

Oppgaver

Hver oppgave teller ¼ av endelig karakter.

Oppgave 1

Den første oppgaven er å skrive R kode som slår sammen de 6 datasettene til et stort datasett. Du må benytte de variablene som de ulike datasettene har til felles for å gjøre dette. Denne prosessen skal kort dokumenteres og kommenteres.

Oppgave 2

Dataene skal benyttes til en ukentlig salgsrapport til et enkelt utsalg. Gi noen eksempler på hva innholdet i en slik kortsiktig individuell rapport bør inneholde. Begrunn dine valg og tankegangen bak figurer og eventuelle tabeller.

Oppgave 3

Dataene skal benyttes til en månedlig salgsrapport på aggregert nivå til konsernledelsen. Gi noen eksempler på hva innholdet i en slik langsiktig konsernrapport bør inneholde. Begrunn dine valg og tankegangen bak figurer og eventuelle tabeller.

Oppgave 4

Kan dataene benyttes til å planlegge nye utsalg? Dersom konsernledelsen ønsker å etablere et nytt utsalg, hvordan kan de benytte disse dataene til å finne den beste lokasjonen?

Prosjektpresentasjon

Det skal gjennomføres en 10 minutters presentasjon av prosjektet fredag 13 mai. I denne presentasjonen skal du/dere fokusere på de “innovative” analytiske grepene som gjør at konsernet skal velge ACME analytics som leverandør av denne tjenesten. Godkjent prosjektpresentasjon er nødvendig for å få vurdert prosjektoppgaven.

Bedømmelse

Formålet med prosjektoppgaven er å gi dere trening i å:

1. ... skrive godt og tydelig
2. ... anvende data, lage figurer og tabeller
3. ... levere i henhold til en beskrivelse

Innleveringen bedømmes i hovedsak etter innholdet i rapporten, men det forventes også at R-koden er godt dokumentert, ryddig og oversiktlig. Husk at leveransen skal kunne leses både av daglig leder av en bedrift, samt av konsernledelsen. En god besvarelse svarer konkret på oppgaven, gir tydelige definisjoner av relevant data og økonomiske begrep, drøfter sammenhenger med utgangspunkt i økonomisk teori, og forklarer utfyllende innholdet i figurer og tabeller. Husk at selvforklarende titler og kategorier på figurer er essensielt.

Innleveringsfrist

Frist for innlevering av prosjektoppgaven er tirsdag 7 juni kl 13:00.

Husk at filene som utgjør leveransen på prosjektet skal ligge på et åpent github repository. Lenken til dette repositoryet skal leveres på Wiseflow innen fristen. Er det en gruppeinnlevering skal medlemmene i gruppa levere samlet. Utfyllende veiledning kommer.