

## Refleksjonsnotat

Vi har underveis i prosjektet lært mye om datainnsamling, databehandling, dataanalyse og visualisering og generelt om samarbeid når det kommer til bruk av branches, og commits. Alt når det kommer til datainnsamling var nytt for oss, så det var en del å sette seg inn i. Det ble brukt en god del tid på å skjønne bruk av API kall og innhenting av data manuelt. Vi tok i bruk begge metodene for innhenting av data og endte opp med to json filer fra Frost (Metrologisk institutt) og en csv fil fra NILU (norsk klima- og miljøforskningsinstitutt). Disse inneholder henholdsvis homogenisert middelverdi for lufttemperatur og nedbørsmengde og historisk data for luftkvalitet. Vi prøvde i utgangspunktet å hente ut data for hver dag fra 1960-tallet frem til i dag, men det oppsto flere problemer da det ble for mye å hente ut. Vi tok oss dermed til nøye med en måling i måneden fra 1960 til i dag for å vise best mulig utvikling over tid.

Vi slet en god del med å finne gode dater for nedbørsmengde og luftkvalitet og, mener at disse dataene ikke reflekterer noen funn i særlig stor grad. Lufttemperatur fant vi god data på, som vi mener viser en god utvikling, da den inneholder mange data over en lang tidsperiode. Derfor mener vi at denne dataen også viser best utvikling ved visualisering og prediktiv analyse.

Når det kommer til databehandling fikk vi bruk for en del tillært kunnskap, sånn som blant annet iterering. Også her er det en del nye funksjoner som kommer godt med, som ikke vi har tatt i bruk før, blant annet bruken av Pandas SQL som egnet seg godt i dette prosjektet. Vi valgte å slette nullverdier og manglende verdier. En alternativ måte kunne vært å ta gjennomsnitt av den nærmeste eksisterende verdien før og etter.

Datavisualisering har vi hatt en del om fra før av, men vi benyttet oss nå av noen nye biblioteker som vi ikke har brukt før. Både Seaborn og Bokeh var helt nytt for oss. Seaborn var ganske rett frem å lære seg, da mye bygde på matplotlib, men Bokeh var litt mer krevende. Det var en helt annen måte å plote funksjoner på, hvor man fikk en interaktiv nettside, som illustrerte dataen en god del bedre enn mye annet. Dette økte vanskelighetsgraden, da det var en litt annerledes måte å kode på enn vi skulle gjort tidligere. Introduksjonen av nye funksjoner innad i Bokeh måtte vi ta hensyn til og dermed angripe denne delen av oppgaven på en helt annen måte.

Alt av repository og branches, commits osv. var helt nytt for oss det også. Vi brukte en god del tid på å lage et velfungerende repository, da det antagelig hadde skjedd en liten feil underveis når vi lagde SSH key. Det slo ikke gjennom med noen feilmelding så det var vanskelig å finne problemet. Det endte opp med at vi oppsøkte IT avdeling for hjelp.

Samarbeidet har funket utmerket mellom oss. Vi har jobbet i egne branches utover oppdraget, men likevel kommunisert og informert hverandre av hvordan de ulike mappene fungerer alene og hvordan disse benyttes sammen i en main funksjon for å lage et fullstendig produkt.

Vi er meget fornøyde med hvordan det endelige resultatet endte opp, både hva det gjelder av dataanalyse og behandling via pandas SQL til visualisering av dataene ved hjelp av Seaborn og Bokeh.

Eventuelle forbedringer som kunne bli satt i lyset kunne innebære en «oppgradering» av den interaktive visualiseringen, ved å innføre flere funksjoner og alternativer for brukeren, blant annet flere data på en graf og muligheten til å laste opp/velge egen data på nettsiden. Mappestruktur og generelt oppsett av prosjektet kunne også vært en god del forbedret, da vi endte opp med diverse problemer til tider når det gjaldt commits, pushing, pulling og merging av branches. Dette kunne vi unngått, hadde vi hatt mer erfaring med VS Code eller programmering generelt.

Det ville falt naturlig at utvikling av prosjektet blandt annet kan innebære å ta i bruk en regresjon som ikke er lineær. Vi føler at en lineær modell kanskje ikke passer dataen vår fullstendig. En annen funksjon som kunne ha forbedret prosjektet er å ha muligheten til å endre datatype for «Interactive\_Plot.py» på en mer effektiv måte, blant annet å ha en input i main som spør brukeren om ønskede filtype.

Viktige læringspunkter og erfaringer vi har erfart gjennom prosjektet er blant annet hvordan Github og VS Code kan integreres og benyttes til vitenskapelig forskning og analysering av data. Prosjektet har gitt god innsikt i hva som ligger bak enhver nettside og dens informasjon, altså baksystemet, og hvordan digitale verktøy benyttes til å få dette til å samhandle.

Prosjektet har gitt svært god innsikt og erfaring om hvordan arbeidsoppgaver i et arbeidsliv senere i livet vil kunne se ut. Erfaringene om hvordan en kan samarbeide

innad i et slikt prosjekt vil kunne anvendes uansett anledning, selv om det er programmering eller ei.

Ut ifra analysen av de ulike dataene er det en viktig trend som fanger øyet. Nemlig det faktum at lufttemperaturen stadig øker, noe man ser da regresjon viser en stigende trend. Noe som kan tyde på at global oppvarming har hatt en stor effekt de siste 80 årene.