

Refleksjonsnotat

I denne mappen startet vi med å lage et utviklingsmiljø. Dette var første gangen vi har arbeidet med GIT, noe førte til en bratt læringskurve. Derfor er noen av commit-meldingene forskjellig fra oppstarten av prosjektet, før vi lagde en generell standard der vi skrev først hvilken fil vi er i, og deretter hva som ble gjort. Det er derfor litt rot i starten med tanke på hva som har blitt commitet og slettet.

Valg av datasett

Vi hadde ingen forkunnskap om datainnsamling, men med en gang vi forsto nettsiden til Metrologisk institutt, var det lett å samle inn data. Vi har lastet ned datasettet fra Metrologisk institutt ved bruk av api. Vi har lært viktigheten av .env med tanke på datasikkerhet, og lagret client.id i env. Vi valgte å bruke API i stedet for å laste ned en statisk csv-fil, så vi ikke trenger å manuelt endre datasettet ved endringer av innholdet. Fordelen med datasettet vi valgte var at datasettet hadde tre forskjellige datasett, slik at vi kunne fordele arbeidet ved dataanalyse, visualisering og prediktiv analyse.

Behandling av datasett

Vi valgte å ikke ta med alle attributtene for å få data som var lett å forstå, men likevel inneholdte mest mulig nyttig informasjon. Dette gjorde det lettere å visualisere, ettersom datasettet hadde konkrete attributter. Vi har lært hva vi gjør med manglende verdier og lært hvilke verdier det er lurt å ha med i datasettet.

Framstilling av data

Vi har lært at de ulike datasettene måtte visualiseres med forskjellige grafer for å få fram dataen best. Dette er fordi dataen inneholder forskjellig informasjon, som f.eks. at noen har negative og positive verdier, mens andre bare har positive. Vi har lært viktigheten av et godt filoppsett og at man må kunne både main-fila og de individuelle notebookene for å at alle de nødvendige bibliotekene skal kunne brukes.

Prediktiv analyse

I den prediktive analysen fikk vi utformet grafer og resultater som skapte større forståelse ovenfor våre ulike datasett. I datasettet til vind og nedbør kunne vi

konkludere med at begge regresjonene passet tilnærmet like bra, noe vi ikke kan si om temperatur.

Etter å ha analysert hvert datasett separat, prøvde vi å sammenligne de ulike datasettene for å se om vi kunne finne noen sammenhenger. Ettersom verdiene i de ulike datasettene er veldig forskjellige, var det ikke så lett, eller hensiktsmessig.

Dersom vi hadde hatt litt bedre tid kunne vi nok dykket dypere i sammenligningen, og utforsket bedre måter å sammenstille dataene på.

Utfordringer

Det var utfordrende å forstå hvordan GIT fungerte i sin helhet og hvordan vi skulle bruke den på best mulig måte for å unngå for mye dobbeltarbeid, eller å ødelegge for hverandre.

Hva vi har lært

De viktigste læringspunktene for oss har vært:

- Hvordan lage et godt filoppsett for å lett finne tilbake til filene
- Hvordan kommunisere godt i GIT for å jobbe mer effektivt
- Viktigheten av pålitelig kilder ved innhenting av data

Vi har lært hvordan det er å bli kastet inn i et prosjekt uten mye tidligere erfaring, der man må bruke kunnskapen man har for å gjennomføre prosjektet. Vi har fått et innblikk i datainnhenting, databehandling og sortering av data i praksis, noe som vi som arkivstudenter vil få nytte av foregår når det kommer til arbeid med bevaring og sikring av data over tid.

Vi er fornøyde med det endelige resultatet totalt sett. Vi har lært å jobbe i gruppe, men også selvstendig. Vi har fått prøvd oss frem ved å prøve ulike visualiseringer. Vi har lært ulike visualisering passer til ulike analyser.

KI

Vi har ikke brukt KI til å lage kode eller produsere resultater som figurer, tabeller eller tallverdier i vår mappe. I tilfeller hvor koden har vært ferdig, men ikke fungert og vi har ikke funnet løsninger andre steder har vi brukt KI for å finne småfeil i koden.

Tilbakemelding til gjennomføring av prosjektet

Noe som skapte utfordringer for oss med tanke på gjennomføringen av prosjektet var at emnet hadde lagt opp mye tid til øvingene og lite tid til mappeinnleveringene. I tillegg ble innleveringsdatoene endret midt i prosjektet flere ganger, noe vi ikke synes var greit.

Andre utfordringer var at enkelte forelesninger ikke var relevante for mappeinnleveringen, eller kom sent i forhold til tiden avsatt til arbeidet med mappen.

Vurdering av samarbeidet og arbeidsfordeling

Vi har hatt et godt samarbeid med mappen. Oppgavene i mappe del 1 bygget på hverandre, så vi valgte derfor å arbeide sammen, mens Sofie skrev og commitet til git. I mappe del 2 lagde vi tre branches til datasettet; wind, temperature og precipitation. Slik kunne vi alle arbeide med hvert vårt datasett uten å risikere krøll med git og overskriving.

Marte var sykemeldt i perioden mellom 27. april og 10. mai, derfor er det ingen aktivitet i denne perioden.