

## OPPGAVE – Data Engineer

**Kunde:** DyrWatt AS

**Bransje:** Strømproduksjon

**Mål:** Optimalisere produksjonen ved hjelp av data

### Bakgrunn

DyrWatt AS vil planlegge og optimalisere strømproduksjonen sin bedre.

For å gjøre dette trenger de å se:

- Fyllingsgrad i vannmagasinene
- Værdata, spesielt nedbør

De vet at det finnes historiske data for både magasinfylling og vær, og ønsker at du:

1. Henter disse dataene
2. Lager dem strukturer
3. Gjør dem tilgjengelige for analyser og rapporter

### Hva kunden ønsker av deg

Som første leveranse ønsker de et sett med rapporter som viser:

- Trender over tid for:
  - fyllingsgrad i vannmagasin
  - nedbør
- Maksimums- og minimumsverdier
- Sammenligning med året før
- Perioder med størst endring i fyllingsgrad eller nedbør
- Forslag til hvordan dette best kan visualiseres (grafer, tabeller, dashboards osv.)

### Datakilder

#### Magasinstatistikk (NVE)

Bruk NVE sine åpne API-er:

- Dokumentasjon: <https://biapi.nve.no/magasinstatistikk/swagger/index.html>

Her finner du data om fyllingsgrad i vannmagasin.

#### Værdata (MET / Frost API)

Bruk Frost-API-et fra MET:

- Dokumentasjon: <https://frost.met.no/index.html>

Du skal:

- Velge én eller flere værstasjoner i samme el-område som magasinene
- Hente spesielt data om nedbør

Eksempel på kall:

[https://frost.met.no/observations/v0.jsonld?sources=SN18700&referencetime=2022-01-01%2F2022-02-01&elements=accumulated\(precipitation\\_amount\)](https://frost.met.no/observations/v0.jsonld?sources=SN18700&referencetime=2022-01-01%2F2022-02-01&elements=accumulated(precipitation_amount))

## Teknisk løsning

### Henting av data

Du velger selv hvordan du vil hente data fra API-ene, for eksempel:

- Python
- PowerShell
- Andre verktøy du foretrekker

### Lagring og modellering

- Dataene skal lagres i en database
- Transformasjon av data skal gjøres i SQL
  - For eksempel med view eller prosedyrer
- Målet er å lage en dimensjonsmodell (stjernemodell eller lignende) som er enkel å bruke i rapporteringsverktøy

### Database

Hvilken database du benytter for lagring av data er helt opp til deg.

Det flere gratis databaser du kan benytte, to alternativer er:

#### freesqldatabase.com

<https://www.freesqldatabase.com/freemysqldatabase/>

NB! For å koble f.eks. Power BI til denne databasen må du bruke en MySQL-driver:

- Last ned her: <https://downloads.mysql.com/archives/c-net/>
- Bruk versjon 9.1 (nyere versjoner fungerer foreløpig ikke mot freesqldatabase.com)

#### duckDB

<https://duckdb.org/>

ODBC driver for duckDB: <https://duckdb.org/docs/stable/clients/odbc/windows>

### Rapportering

Rapportene skal lages i et rapporteringsverktøy, for eksempel Power BI. Bruk dimensjonsmodellen du har laget som grunnlag for rapportene.

### Spørsmål?

Hvis du har spørsmål til oppgaven, send e-post til:

[hans.frode.andersen@fraktal.no](mailto:hans.frode.andersen@fraktal.no)