## Introduksjon

(Generell informasjon om nettsiden)

## Hvordan kommer data inn?

Databasen vi har valgt tar utgangspunktet i et datasett som bruker data fra en undersøkelse på studenter fra to skoler i Portugal. For at ny data skal bli lagt inn i databasen, må det gjøres enda en undersøkelse ved enten de samme skolene eller en ny skole. Dataen må deretter legges inn via nettskjemaet med følgende felt på nettsiden:

* Skole
* Kjønn
* Alder
* Mors-utdanning
* Fars-utdanning
* Reisetid
* Studietid
* Stryk
* Skolestøtte
* Ekstratimer (Betalt)
* Aktiviteter utenom faget
* Internett
* Romantisk forhold
* Familie Relasjon
* Fritid
* Gå ut med venner
* Alkohol i ukedagene
* Alkohol i helgen
* Helse
* Fravær
* Karakter semester 1
* Karakter semester 2
* Sluttkarakter

Totalt er det 22 felt som må fylles ut, men når dette er sagt kommer mange av feltene til å ha standardverdier. I tillegg til nettskjemaet vil det være mulig å laste opp datasett i form av CSV eller JSON som har de samme kolonnene. Rader som er duplikater, vil ikke bli tatt hensyn til siden det er mulig å produsere de samme verdiene. Dette vil si at noen må manuelt legge til data via nettsiden for at dataen skal dukke opp i databasen, den vil ikke lytte etter endringer i datasettet via API-er.

# Riak eller ETCD?

I prosjektet vårt har vi bestemt oss for å bruke ETCD, siden ETCD prioriterer konsistent data over tilgjengelighet. Dette er viktig, siden for dette prosjektet er det mye viktigere at dataen stemmer, enn å ha rask respons-tid. Hele hensikten med nettsiden er å forstå studiene som er tatt ved de forskjellige skolene, derfor må dataen være oppdatert. Dessuten gjelder CAP-teoremet i praksis kun når det oppstår en nettverksfeil, så det vil ellers være rask responstid.

ETCD ytelse - <https://etcd.io/docs/v3.4/op-guide/performance/>