

# Covid-oppgaver

Louisa Boulaziz - louisabo@uio.no

September 3, 2020

Her jobber dere med data fra John Hopkins University som jeg lastet ned på overnevnte dato. Last in datasettet vha denne lenken:

`https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse\_covid\_19\_data/csse\_covid\_19\_daily\_reports/09-01-2020.csv`

For å løse oppgavene trenger du pakken Tidyverse.

## Oppgaver

1. Hvor mange land har dødsfall?
2. Hvor mange kinesiske provinser er det i datasettet?
3. Hva er gjennomsnittlig lengdegrad (longitude) for provinser med dødsfall == 0?
4. Opprett tre nye variabler: en som viser andel av smittede som har blitt friske, en som viser andel smittede som har dødd, og en som viser andel smittede som lever men fortsatt er syke.
5. Lag et distribusjonsplott (density plot) for hver av de nye variablene.
6. Hvor langt fra ekvator er den gjennomsnittlige staten/provinsen i datasettet? Bruk `abs()` for å finne absolutt verdi.
7. Lag et spredningsdiagram (scatter plot) over distanse fra ekvator og andel smittede som har dødd. Legg så på en lineær regresjon.

Figure 1:

Hubei	China
	Italy

Figure 2:

Hubei	China
Italy	Italy

8. Se på datasettet. For alle land som ikke i datasettet er delt opp i stater eller provinser står det feltet blankt. Fyll ut denne kolonnen: de som nå er blanke skal i stedet ha navnet til landet. Se eksempel ovenfor.