



UiO : Universitetet i Oslo

# OpenRefine



# Hva er OpenRefine

- Tidligere Google Refine
- Open Source verktøy
- Jobbe med uryddig data
- Rense data
- Endre format

# Importere data

- Flere typer formater
- TSV (tab-separated values), CSV (comma-separated values), Excel, JSON (javascript object notation), XML, Google Spreadsheet

## Lage et prosjekt

- Create Project
- Get Data from This Computer
- Choose Files

Prosjekter lagres automatisk mens du jobber med dem  
For å åpne eksisterende prosjekter velf 'Open' eller klikk på  
OpenRefine ikonet øverst til venstre

# Oppgave 1

- Prøv å importere en CSV fil

# Layout

- Layout som i Excel. Tabular format med rader og kolonner
- Endre kolonne navn og rearrengere kolonner
- Klikk på nedtrekks knappen i første kolonne der det står 'All'

## Sortere data

- Velg nedtrekks knappen ved kolonnen du ønsker å sortere og velg 'Sort'
- Sortering er ikke permanent og dataen kan endres tilbake til opprinnelig tilstand

## Oppgave 2

- Sorter data på en kolonne



# Fasetter

- Nyttig funksjon
- Tekst fasett
- <http://127.0.0.1:3333/preferences>

## Oppgave 3

- Lag en fasett
- Endre navn på fasetten

## Filtre: Regular expression & Booleans

- Edit cells, Transform
- `value.replace("Argentina", "ARG")`
- `if(value.contains("Argentina"),"ARG",value)`
- `value.match(/[^\@]+.*)[0]`

# Importere fra URL

# Oppgave 4

- Importere fra URL
- <http://folk.uio.no/ahmedmo/sloyd/letras.csv>