ETL Project Proposal

Quinton Butterfield / Minami Imoto

* **Data Sources**
  + Average Wages from OECD

https://data.world/oecd/average-wages

* + Cost of Living per Country

https://www.kaggle.com/andresionek/cost-of-living-per-country/downloads/cost-of-living-per-country.zip/3

* + ISO3166 Countries with Regional Codes

https://github.com/lukes/ISO-3166-Countries-with-Regional-Codes/blob/master/all/all.csv

* **How you would connect the data**

Connect the country name in Cost of Living with the country name in ISO 3166 and connect their three letters of country codes to three letters of country codes in Average Wages

* What transformations do you think you have to do?
  + Drop all unnecessary columns
  + Drop N/A rows
  + Data aggregation
  + Filter 2018 in the data source Average Wages because the data source Cost of Living only has data in 2018