**ETL Project**

**Project Title:**

GDP vs Olympic Medals

**Team Members:**

Samantha van Wyngaarden

Stephanie Salvona

**Dataset to be used:**

* Olympics
* GDP by country

To be joined on Country/Country Code

**Data Clean up & Analysis**

The sources of data that you will extract from. - data.world

<https://data.world/nilanshuramteke/gdp-by-country/workspace/file?filename=GDP+By+Country.csv>

<https://data.world/johayes13/summer-winter-olympic-games>

The type of transformation needed for this data (cleaning, joining, filtering, aggregating, etc).

The type of final production database to load the data into (SQL).

The final tables or collections that will be used in the production database (Gold vs high GDP country, Low GDP Country vs no medal, Country’s with medals)

**Breakdown of tasks**

1. **Clean data in Pandas**
2. **Export to CSV**
3. **Import to SQL**
4. **Relate tables & ERD**