**ETL Project - IMDb Data**

Jen Mattingly

Richard Nwokoye

Tyler Motter

The aim of this project is to extract movies dataset from IMDB movies database, clean the data and insert into a Postgres database.

The technologies we used include Postgres Database, SQLalchemy library, Pandas, and user defined functions.

**EXTRACT**

Source data will be CSV file pulled from IMDb datasource.

**TRANSFORM**

We cleaned the data types by making the date type organized into one format and cleaned any null or erroneous data points. Also, we separated each actor’s name into first and last name columns.

**LOAD**

Then, we appended each of the cleaned sources to a Postgres database. We chose Postgres because the data was structured in an organized manner, so we did not need a non-SQL database.