ETL Project

For our project we will extract, transform and load several datasets to answer the all important question, for adaptations in films, books and video games which came first and at what times?

For instance, the terminator film series has been successful but what about the terminator video games, how did they fare? We will use beautiful soup to grab all Wikipedia HTML tables and use regex to use the list of films based on movies as a sort of key. The movies released and video games are both in .csv format and the columns will need to be selected and modified for matching keys, dates are mismatched, currencies are mismatched, etc..

Datasets to be used or generated:

* Movie released dataset : <<https://data.world/popculture/imdb-5000-movie-dataset>>
* Video games : <https://www.kaggle.com/ashaheedq/video-games-sales-2019/>
* List of films based on Movies (by beautifulsoup): <<https://en.wikipedia.org/wiki/List_of_films_based_on_video_games>>

Team Members:

Daniel Adam Cebula  
William Lim  
Rohan Chaudhari