ETL Project – Final Report

\* \*\***E**\*\*xtract: Original datasets were pulled from Kaggle in csv format

\* \*\***T**\*\*ransform: Data was loaded into jupyter notebook for cleaning.

Tasks include:

* dropping unnecessary columns
* removing commas from integer values
* splitting ‘cells’ with comma separated values into multiple columns
* renaming columns
* rearranging columns

\* \*\***L**\*\*oad: After cleaning, I loaded my files into my SQL database, “movies.” I loaded each file into it’s own separate table for ease of access. The data is available to be joined and analyzed for reporting.