Successful Movie Analysis- Project Proposal

**Summary:**

If we wanted to make our own movie, what aspects would we need to have the best chance of success? To determine this there are a few factors to consider. First we would need to answer questions relating to movie themes. Once this is done we will need to collect multiple datasets, transform them into key aspects/fields, and load it into a singular readable dataset for comparing and analyzing.

**Questions to answer:**

-What are the most successful movie genres?

-How much did they gross

-How much did they cost to make?

-Do movies typically do better when they are released at a certain time?

-What company makes the more successful movies?

**How we going to extract the data:**

-CSV files from Kaggle

* <https://www.kaggle.com/datasets/narayan63/most-popular-movies-dataset>
* <https://www.kaggle.com/code/arnavr10880/topboxoff-eda>
* <https://www.kaggle.com/datasets/sanjeetsinghnaik/top-1000-highest-grossing-movies>

-Web scraping from BoxOfficeMojo.com

* <https://www.boxofficemojo.com/year/world/2021/?ref_=bo_cso_table_1>

These data sources will be used to compile a list of successful movies and collect key aspects from the list to determine what makes a successful movie

**How we will transform the data:**

By dropping unnecessary aspects

Limiting the datasets to the most successful movies (top 100) to compare their top aspects

**Loading the data**

We will load it into a singular database for easy accessing and readability