**Project Title:** The higher the budget the more gross a movie produces.

**Team Members:** Amandeep Brar | Arpine Bankikyan | Cassie Folkers | Lindsey Giron | Ricardo Negrete | Sriven Ankam

**Project Description/Outline:**

* Reviewing all genres
* Profit recognition
* Awards won
* Budget to gross
* 5 years range

**Research Questions to Answer:**

1. Does a higher budget produce a more profitable movie?
2. What genres yield the biggest bang for the buck based on budget to gross?
3. Does a larger budget yield a higher popularity?

**Datasets to Be Used:**

* The Movie Database API –
* IMDb
* OMDb
* Kaggle – Netflix regarding original titles

**Rough Breakdown of Task:**

Documentation – everyone will keep track and contribute

Identify datasets to be used and assure it contains the data we need for analysis – was done as a group in class

Pull dataset – lead Cass (we are all going to pull from API)

Clean datasets – Arpi/Cass

Create script to extract data for analysis – Sriven/Ricardo

Outliners, linear regression, correlation – Lindsey

Create plots for visualization – Amandeep

Analysis – Lindsey/Apri

Presentation – Ricardo

Ultimately assist each other as a group!

**NOTES**

\*\*\* The Movie DataBase Statement:  
“You shall use the TMDb log to identify your use of the TMDb’s API’s.

You shall place the following notice prominently on your application: ‘This product uses the TMDb API but is not endorsed or certified by TMDb.’

Any us of the TMDb log in your application shall be less prominent than the logo or mark that primarily describes the application and your use of the TMDb logo shall not imply any endorsement by TMDb,”\*\*\*

\*\*\*OMDb

**Visualization**

Budget, gross, year.

Netflix for popularity.

Friday

Netflix

568 titles we were only data available 44 titles.

OMDb

Needed to remove OMDb due to little data from metascores 9043 out 84294

Budget vs profit -

Budget vs rating

Profit vs genre – bar chart

IMDb

TMDb

Budget is defined –

Revenue is defined

## Quartile results for Movie Profit (%)[¶](http://localhost:8888/notebooks/Import%20Data/lgiron_outliners_linear_regression_coorelation.ipynb#Quartile-results-for-Movie-Profit-(%))

* The lower quartile of movie profits is: -26%
* The upper quartile of movie profits is: 289%
* The interquartile range of movie profits is: 315%
* The median of movie profits is: 82%
* Values below -500% could be outliers.
* Values above 763% could be outliers.
* Needed to remove outlier as it was over 40,000% from the upper quartile Secret
* The top 3 most genres being made are Drama, Action, and Comedy