Journey to the Center of Movies

James Bryant II, Amee Yang, Rebecca Lubera, Greg Spahlinger (aka ‘The Brat Pack’)

**Project description**

The goal of this project is to identify correlations between user rating and various factors that might influence user enjoyment. Our first goal is to find several data sets that contains the data we are going to analyze. We want a few different variables that could track with enjoyment including: Budget, Gross Revenue, Runtime, Release Year, Release Month, and possibly others. Our metric for how well people enjoy each film will be IMDb user ratings. Once these sets have been cleaned and merged into a manageable dataset we will explore the data, look for correlations, and plot figures. Then we will write the analysis up and present our findings.

**Research question:**

What variables influence how much people enjoy movies?

**Sub questions:**

Does the budget of a movie affect how much people enjoy it?

Does the amount of money it makes indicate enjoyment?

Does the time of release predict user enjoyment?

Does the length of a film predict enjoyment?

**Data sets:**

Opening weekend box office (Kaggle)

IMDb Movies Extensive Dataset (Kaggle)

The Movie Dataset (Kaggle)

**Tasks:**

Data cleaning and assembly – All, in concerted fashion

Data exploration and visualization – All, but independently, Amee takes point

Conclusion editing and selection – All, together

Report writing – Greg

Report editing – All

Presentation assembly - Rebecca

Presentation editing – All

Presentation speaker – James