Title: Movie Loops

Group Members: Carlos Chafloque-Siu, Avni Patankar, Sarah Johnson

Project Description/Outline: What are factors that affect movie release? Can these known trends be quantitatively shown in movie data? Things like genre and production tend to run in cycles that can usually be affected by Box Office Revenue, but are there other factors that have a strong enough correlation to say that changes in release schedule for the year? What conclusions can be drawn from the publicly available data for films released from 1969 to 2019?

Research Questions to Answer: Are there correlations between genre, year MPAA rating, run time, and user/critic rating? Does box office revenue affect popularity of genre or the time of year a particular genre gets released?

Datasets to be Used: Box Office Mojo library for Python and the Omdb Api

Breakdown of Tasks:

Creating dataframe and cleaning the data

Getting rid incomplete data or supplementing from Box Office Mojo

Collating the pertinent data to be analyzed

Plotting the data

Creating scatter plots and coefficient of correlation to determine if there is a relationship between the aspects being analyzed.

Creating tables to group certain variables together in order to determine cycles or a relationship.

Creating charts to look at distribution of the sample into different categories

Creating the presentation and providing film history and industry context for conclusions