What creates a movie hit and what causes a movie flop?

The secret is in the data

Group Members:

Vanessa Caldera, Matthew Hill, Timothy Tieng

Project Summary Quadrant Questions:

1. What was accomplished since the last update (or since the project started) – *these should be highlighted in the Kanban board*
   1. Discussion about data points in the dataset and what could be used to analyze the data.
   2. Discussion about using data to look at the corresponding popularity of the movie.
2. What is working well for the team?
   1. The whole team is on MS Teams now and communication about where to go next on the project.
   2. The ability to look at the different variables of the dataset and place them with one another determine interesting possible correlating elements.
3. Plans for the next update.
   1. Further discussion of the main data points and where we can make connections between the data points and create hypothesis for future movie entries.
   2. Determine if there is any correlation between movie genres and movie popularity and/or voting average.
   3. Determine if runtime affects popularity.
   4. How does vote count skew voting average? Does it skew it?
4. Issues/ what is not working well.
   1. Concern on how to manipulate/ clean the data as there is strings of text and it contains multiple areas of dictionary-like data in the dataset.

Link to dataset: <https://www.kaggle.com/datasets/tmdb/tmdb-movie-metadata>

Kanban Board Visuals

Graphical user interface, application, table

Description automatically generated

Kanban Board To-Dos

Graphical user interface, application

Description automatically generated