1. **Introduction- Patrick (3 mins)**

We chose movies as our subject of exploration for the machine learning project & final. Imagine yourself as a production office executive and needing information on new releases to determine the gross amount of revenue your company would make for a particular movie. This is the crystal ball that we created with our project. It is also fun to explore new releases as we head into the summer blockbuster season. The type of regression models we consider using include: Linear Regression, Lasso, and Ridge Regression. We will also have a dashboard with a user interface that allows users to enter details regarding a pending release and the machine learning model will predict the movie's gross revenue. Now to talk more about our data exploration phase, here is Kylie Hicks.

1. **Data Exploration, Database, & Cleaning- Kylie (3 mins)**

We obtained our main movie dataset from [Kaggle](https://www.kaggle.com/datasets/danielgrijalvas/movies). The dataset was originally scraped from the IMDB website. IMBD is an online database that contains information related to movies, TV series, video games, as well as streaming content. This particular datset caught our eyes because it is robust and we could seemingly draw direct conclusions from linear regression models.

We obtained additional datasets containing the birthdates, number of nominations, and number of awards won by the lead actor/actress listed for each movie in the movie dataset. These secondary datasets were obtained by webscraping Wikipedia pages. Wikipedia is an online encyclopedia allowing free public access. Information provided by volunteers and contributors through open collaboration. The dataset containing the birthdates was obtained by a script that visited each actors' individual Wikipedia page. The second dataset, containing the Academy Award nominations and awards won for leading and supporting roles since 1927, was obtained by scraping [Academy Award Nominations](https://en.wikipedia.org/wiki/List_of_actors_with_Academy_Award_nominations#List_of_actors).

1. **Machine Learning Model- David (4 mins)**
   1. **Describe linear regression**
   2. **Explain Ridge model**
   3. **Talk about limitations of the model**
   4. **Share results/accuracy score**

1. **Conclusion & Demo- Kaiya (5 mins)**
2. Describe Dash Plotly
3. Model demo:
   1. Avatar 2
   2. Top Gun