The Proposal

* With your group come up with a 1 - 2 page proposal about the questions that you intend to ask of the data.
* This should include:
  + Initial plots, figures or tables.
  + References to column names and the analysis that they may provide.
  + Additional datasets that you plan on including in your analysis like the weather data. This means links, columns that you'll join on, etc.
  + What you plan to cover in the final report and how you plan on organizing it.
* Datasets:
  + Movielens.org
  + TMDB
  + IMDB
* Questions:
  + Movie genres
    - Changes in trends in the past few decades
  + Ratings
    - Ratings correlation with release date
    - Ratings correlation with genre
    - Avg rating per user
    - Distribution of ratings (i.e. are most ratings either very positive or negative?)
    - User rating correlation with tagging
    - Is there a correlation between number of ratings for a movie with the average rating for that movie? More ratings means higher rating?
    - Ratings on different platforms (movielens, imdb, etc.). Is there a correlation? Look at rotten tomatoes
      * Are different platforms biased towards specific genres, directors, etc.
* Plots:
  + Statistics
    - **Movies per year**
    - **Genre per year**
    - **Highest rating**
    - Top tags
    - Graph to show number of movies by genre