Team Members:

Mu-Ruei Tseng, GTID: 903550800

Justin Higgins, GTID: 902983119

Dataset Chosen: movies.csv

List of analytic tasks our visualization supports:

1. This visualization allows users to see the relationship between gross and budget, IMDB score, director popularity and movie popularity respectively.
2. It allows user to select a region in the graph and show points located in that area and also its corresponding points in the rest of the graphs. This help users to gain the knowledge of what other attributes points in the selected area tend to have.
3. The visualization will show the genre composition in that selected area with a pie chart.
4. Below the pie chart, the visualization will show a word cloud with the thirty most common listed keywords.
5. Each item in each scatterplot can be clicked on to show a tooltip with data about the individual data item.
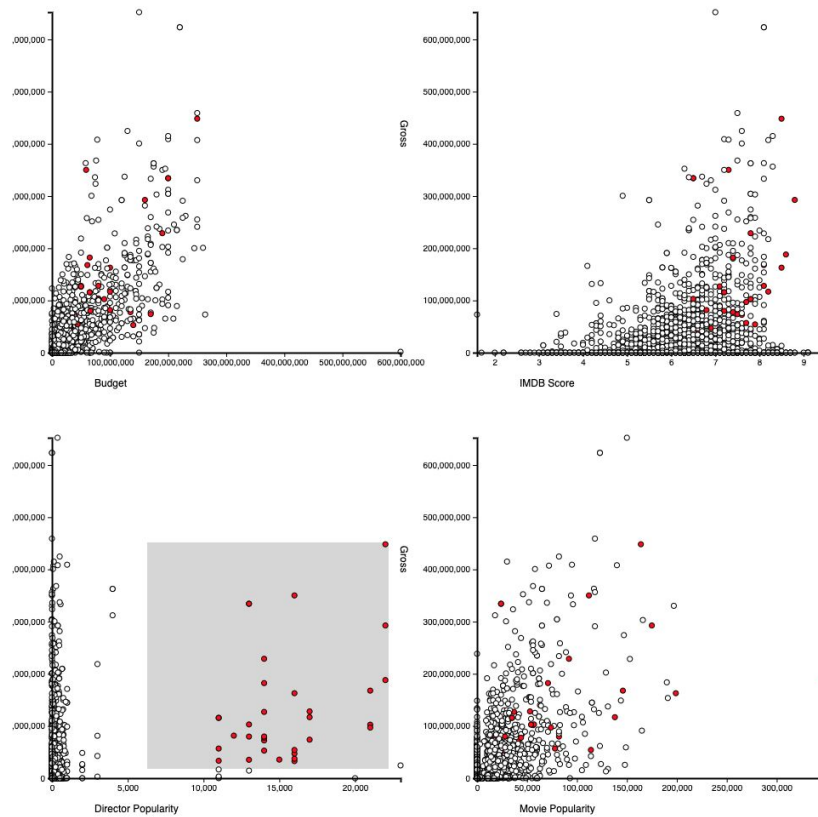
Design Overview:

We started by thinking about what fields interested us from the data set. We decided on wanting to learn how certain fields in the data set relate to how much money the movie made at the box office. We included the movie's budget, IMDB score, and the popularity of the director and the movie itself, these two measured by Facebook likes. We accomplished this first level of analysis using a brushed and linked scatterplot matrix where each plot's Y axis is the movies' box office gross. Each X axis is one of the four aforementioned attributes. Beyond that, we created two supplemental charts to show more data about a selection of items. The first chart displays a pie chart, showing what genres compose the selection. The second shows a word cloud of most common listed keywords relevant to the selection. The goal behind the supplemental charts is to show even more data that may relate to box office gains or losses. Below are a list of analytical questions that can be answered using this visualization.

● Do the most highly rated movies on IMDB make the most at the box office?
● Of the highest budget movies, are there more action movies or comedies?
● Are the highest grossing movies actually popular on Facebook?
● Is there a relationship between the most popular directors on Facebook and the genres of movies they typically produce?
● Are there common keywords between the highest grossing movies?
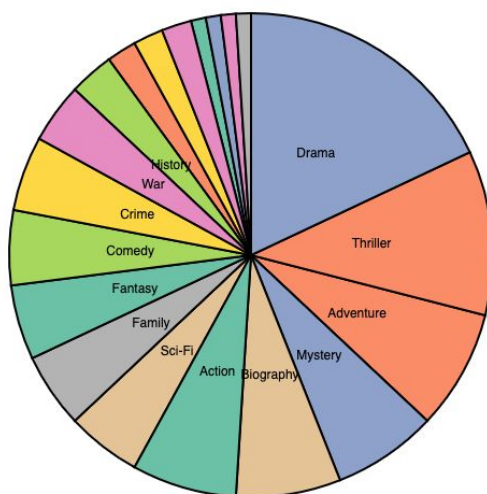
Screenshots:

1. These are the relationship scatter plots that show the corresponding location of the points in the selected area in other attributes.

2. This is a pie chart of genres for the selected area in the scatterplot matrix. Genres with a small percentage are not given a label.

Genre Chart

3. This word cloud shows the 30 most popular keywords listed in the selection of movies.

Keywords

4. When a data item is clicked, a tooltip appears with information about the item.