Analysis based on Tamara's framework about Colombian Cinema

Introduction:

The movies produced in Colombia are a part of what is known as independent cinema. In general, it is understood as independent cinema all cinematic product that wasn't produced by one of the 6 big distributors, these being Warner Bros. Pictures, 20th Century Fox, Paramount Pictures, Universal Pictures, Sony Pictures Entertainment and Walt Disney Studios.

Movies produced on Colombia, compared to those in Hollywood, present considerably less popularity with general audiences, i.e. those who see movies in the cinema. In comparison, in spaces las international contests Colombian movies tend to win awards and fame. This is a tendency among independent movies around the world, not exclusive to Colombia. Still, there are some rare success cases in which an independent movie that, at first glance, doesn't seem to have that much in different with other independent movies. "Belier Family" and "El abrazo de la serpiente" are examples.

The Ministry of Culture have a data base with info from all Colombian movies from 2006 to the present. Based on this info, each year the Ministry of Culture makes a report when the year is ending, to which they refer as a "Anuario Estadistico", this is a summary about the status and events in Colombian Cinema in the year that just passed. Among the information in these summaries there is which movies earned the most, which cinemas received more people, which award have been won and by who, which cooperation with other countries happened, etc. Mainly based on these documents is how the film industry in Colombia have analyzed its status until this day.

For our analysis we have selected as success indicators the amount of money earned and the amount of spectators a movie had.

The WHAT:

As part of the information given to us by the Ministry of Culture we received in PDF format the "Anuarios Estadisticos". From these archives we deduced the most relevant information for this analysis and extract it to the formats we needed.

We saw a most data had a direct and strong relation with time, reason why most data is sequential.

Most relevant data was quantitative like the number of spectators, money earned, number of days in the cinema, money invested in publicity, etc.

There is also categorical data like the movie genre, the different cities that present it, among other.

All this data is organized in tables.

The WHY:

After talking with our client and checking the data we came to the conclusion that to create a successful visualization we need to:

- Compare different dimensions to discover which pattern are follow by the uncommon successful independent movies, if there is any pattern.
- Allow the visualization of multiple dimensions from a centralized point, giving a general panorama of the behavior of Colombian films.

We believe that with an interactive visualization, our client can solve its doubts, which can be summarized as ¿Why Colombians don't go to the cinema and watch local movies?

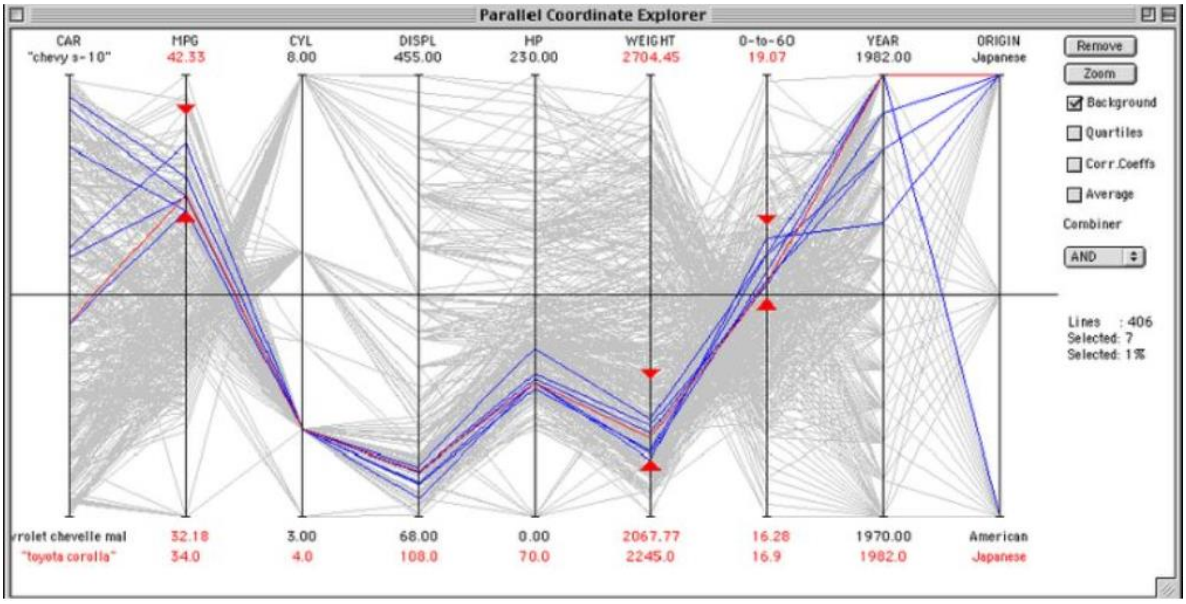
The HOW:

We are going to juxtapose and filter the different data in a way that hundreds of movies can be compared at the same time. Also we want to allow the user to select the data they desire, allow navigation, show addition information being in parallel graphs or in embedded information.

Parallel Coordinates:

We chose to use parallel coordinates as this graph allows the analysis and comparison of multiple data, through many dimensions.

It is important to manage the scale of the data because different dimensions have their own range for its information.



It is typical to use the technique known as brushing along with parallel coordinates because this allows the user to manage the range in the dimensions of the data the are interested in.

It is recommended to not use more than 12 dimensions because it becomes hard for the user to analyze more dimensions at the same time.

We understand that at the beginning this visualization may need a little bit of training and explanation, but in the long run it allows to identify patterns in a very fast and easy way.

The Solution:

Originally we planned to realize the analysis using multiple graphs like bar charts, circle charts, maps and probably a tree map. At that moment we still hadn't access the information, even more we weren't sure we will be able to access it.

After some reunions with the Ministry of Culture we gain access to de database, this allow us to understand with what we were working.

After the session in class in which our peers check our proposal we chose to stop using many simple visualizations and start using parallel coordinates.

We believe it is of most importance to allow an easy way for the user to interact with the visualization and we believe it is possible with or proposal.