

Charles Perin^{1,2,3}
1-INRIA-AVIZ
2-Univ. paris-Sud
3-CNRS-LMSI
charles.perin@inria.fr

DATABASE FROM IMDB:
All MOVIES SINCE 1990

People

first_name
last_name
people_id
movies_id: [*]

Movie

title
release_date
people_ids: [*]
movie_id
budget
genres: [*]
Opening Box Office*
Viewers Rating*

MATRIX VIEW OF A MOVIE: to Explore the Opening Box Office (dollars) and Viewers Rating. Each Cell Shows the Movies with both the Column and the Row Dimensions

The selected dimensions of the movie, both horizontally and vertically

Loading a movie using its Imdb unique ID

Dimension selection among the dimensions of the movie

Manually adding dimensions of the movie

Cell Visualizations

Several ways of observing the data distribution

In blue the Opening Box Office
In orange the viewers rating
The y axis encodes the number of movies, the x axis the opening box office or the viewers rating

165 Stripped Charts
Opacity encodes the frequency

165 Dual Line Charts
Height encodes the frequency

165 Bar Charts
Height encodes the frequency

165 Line Charts
Height encodes the frequency

165 ScatterPlot
Opacity encodes the frequencies

The actor Will Yun, the writer Marc Bomback and the producers Hugh Jackman and Hutch Parker are not reliable:

The distribution of their Opening Box Office and/or Viewers Rating is spread or binomial, or their number of data points is low

INITIAL ESTIMATION VIEWS

Each selected dimension is associated a weighting slider and a Line Chart in both the Opening Box Office view and the Viewers Rating view

6 additional nD sliders weight the movies for which n dimensions are in common with the explored movie: the more dimensions in common, the more similar the movies

in this initial setting, interesting information is not highlighted

SCATTERPLOT MATRIX VIEW TO DISCOVER SIMILAR MOVIES

The actor Hugh Jackman and the producer Stan Lee were involved in all the X-Men movies

Also, when exploring the different X-Men movies, we realize that their Opening Box Office (average 80.64M) and Viewers Rating (average 7.26) are high

We can expect high values for both indicators

EXPLORING INDIVIDUAL MOVIES

Showing movie details by hovering its associated dot
The darkest a dot, the more the associated movie is similar to the explored movie

Unsurprisingly, the series of X-Men movies is the most similar to The Wolverine

OPENING BOX OFFICE ESTIMATION

Setting the dimensions sliders:

- Main actor and Director have strong weights
- Producer Stan Lee has a strong weight
- Budget is important for Opening Box Office
- Genres (brown Line Charts) are visually weighted so that they have reasonable impact

The explored movie belongs to a series: Highly similar movies (6D+) have strong weight
The less similar, the lower the weight

The Opening Box Office mean value is 55.8M (Vertical red line)
Using the Mean view, we identify a close spike to this value
Hovering this spike displays the associated value in the top-left corner

Finally, the estimated value for the Opening Box Office is 54M and is quite lower than the 80.64M average of the X-Men series

VIEWERS RATING ESTIMATION

Sliders weights are set differently than for the Opening Box Office:

- Higher weights for the director and the writer
- Lower weight for the Budget: advertised blockbusters make money, but the quality of the movie does not depend on its cost and special effects only
- Very high weight for similar movies and very low weights for less similar movies

We brush movies (lightgray area) involving mainly Hugh Jackman, James Mangold, Scott Frank and Stan Lee

The mean value for the brushed movies is the vertical orange bar, indicating a Viewers rating of 7.2
This value is close to the 7.26 average value of the X-men series