



MoVis Movie Society Visualizer

Hanna Huber (00925230)

VU Information Visualization, June 2017

About MoVis

MoVis is a tool to analyse the cast of movies produced in the USA between 1950 and 2000 with respect to the following case attributes:

- Age (in the year, the movie was produced)
- Gender
- Origin (country of birth)

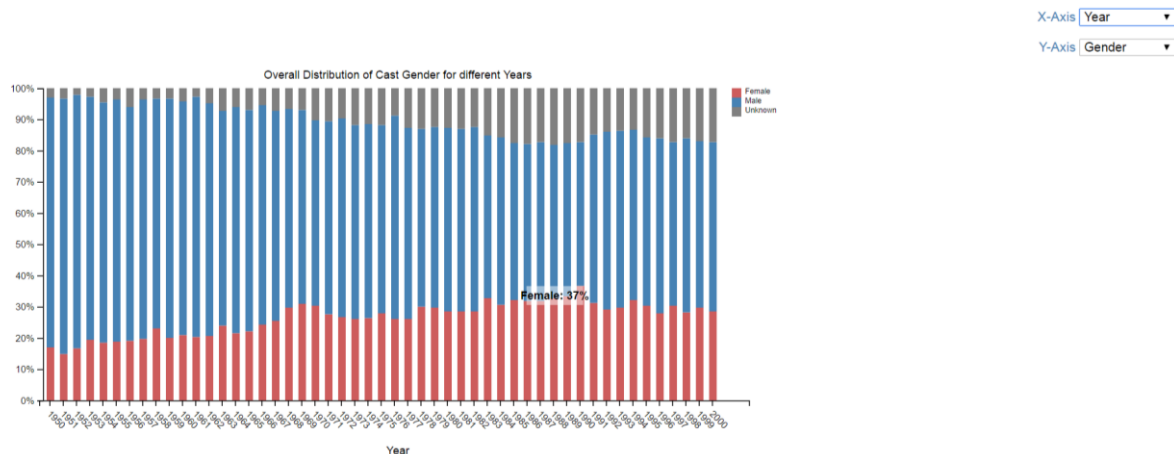
The distributions of the respective attribute among the movie cast can be compared for different movie attributes:

- Year (when the movie was produced)
- Genre

For each movie and attribute, the movie is represented by a list of attribute values, where each entry corresponds to one member of the cast. Each value is a normalized density value (e.g. if 4 out of 40 cast members have the same age, the density of this value is 0.1), so that movies with different cast sizes can be compared. This list represents the cast attribute distribution of the respective movie.

Distribution Overview

Select the cast attribute whose distribution you want to examine on the x-axis.
Select the movie attribute for which you want to compare the distributions on the y-axis.



1 Gender vs. Year

Distribution Overview

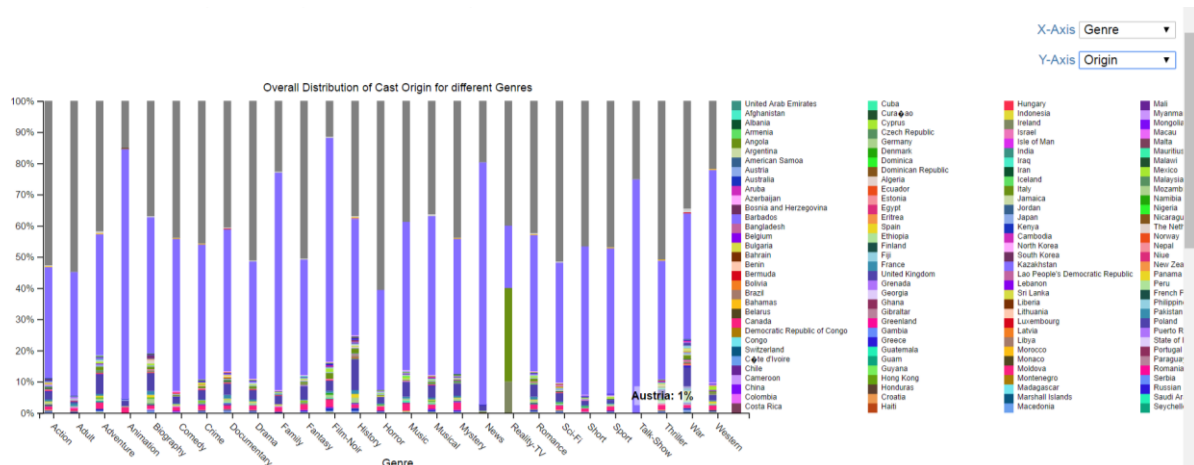
In this view, the user can choose the cast attribute on the y- and the movie attribute on the x-axis. For each x value (e.g. all movies produced in 1988), the corresponding cast attribute distribution are averaged, creating a mean cast attribute distribution. These distributions can then be compared vor different x values.

Interactive Methods

- The attributes of both axes can be changed.

- Additionally, a tooltip helps to identify individual values of interest and provides details on demand.

These methods can be used to solve specific tasks. For example, they reveal that the overall share of female cast members increased since 1950, reached a peak of 37% in 1990 and has slightly decreased since then (see Figure 1). They also show that while most cast members were born in the United States, movies of the genres War, History and Biography have a larger share of cast members born in Germany or Austria (see Figure 2).



2 Origin vs. Genre

Filtering Individual Distributions

Additionally, the attribute selected in the overview, can be analysed further in this view. Each cast attribute distribution is represented by a single piecewise linear function graph.

Interactive Methods

- The movies can be filtered by their movie attributes, keeping only the movies of interest in the chart.
- Mean distributions are displayed if the “average” box is checked.
- Additionally, for attributes origin and age, gender-specific distributions can be displayed, allowing direct comparison.
- The line chart is linked to the bar chart, so if the cast attribute is changed above, the chart below is updated accordingly.

This way, it is easy to switch between different combinations of attributes. The gender-specific option combined with the filtering option reveals, for instance, that in Action movies, the age actresses is more concentrated between 20 and 38, while the male distribution is spread more widely (see Figure 3).

Implementation

This project was implemented in HTML, CSS and JavaScript using the library D3.

Web Page Structure

File: [index.html](#)

The webpage is structured into a header and two chart views. In both of them, an animated loader provides visual feedback about the loading process to the user. At first, only the bar chart is loaded,

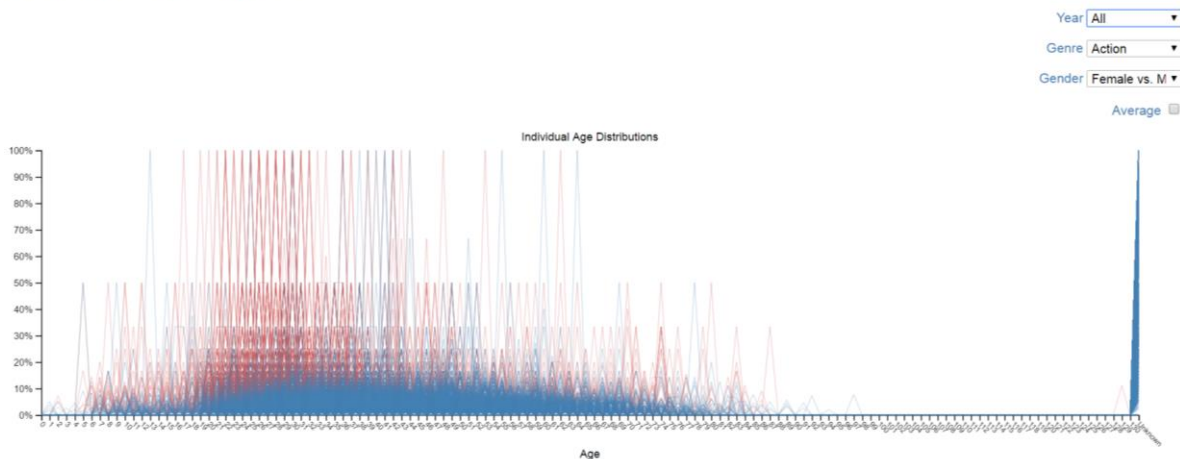
which allows the user to start exploring the data more quickly. A dropdown list allows the selection of the x- and y-axis attributes.

In the second view, only the heading is visible, informing the user about the additional view. By clicking on it, the view opens and the line chart is loaded according to the data which is visible in the bar chart at this moment. By clicking again, the line view can be closed again.

Global variables which are used by all scripts are declared at the end.

Filter Individual Distributions

Select one or more movie attributes to filter the results.



3 Female vs. Male Age in Action Movies

Loading data & Updating Views

File: main.js

This file serves as an interface for the two charts and controls the data loading. Once the page is loaded, the cast attribute files are loaded using `d3.csv()` and processed to a suitable structure (e.g. densities are normalized) to be passed on to the chart. Then, the bar chart is initialized.

Once the line view is toggled, it is either opened or closed based on the current state. In the former case, an additional check is performed to see whether line view has already been initialized before. If this is NOT the case, additional gender-specific files which are only used in the line chart are also loaded using `d3.csv()` and processed. Above that, the cast data from the bar chart is restructured for the use in the line chart. All variables are passed on and the line chart is initialized.

In case the line chart has been toggled before, main only calls for an update.

Main also contains a callback method which is called in case the y-axis of the bar chart is changed. In this case, both the bar as well as the line chart are updated.

Bar Chart

File: movis-bar.js

A stacked bar chart is created using `d3.stack()`. It uses local variables which are assigned according to the user selection using different getter methods. This way, different scales, domains, ranges, etc. are created for different attributes. For the large number of categorical color classes needed for the origin data, color generator `Colorgorical` [7] was used.

A tooltip is added to display the attribute key and value of a specific bar in the stacked bar chart. This way, also small entries can be analysed.

A legend shows the color range and domain, describing the fill values of the bars.

In the update method, current bars, axes and the legend are removed and the chart is re-initialized.

Line Chart

File: movis-line.js

This file is structured similarly to the file above. Additionally, unique movie attribute values are extracted and added to the filter dropdown list. Line paths are created using `d3.line()`. If only the filter values change, an additional update method only redraws the paths while keeping axes, title and labels unchanged.

References

Data & Processing

1. IMDb: <http://www.imdb.com/>
2. IMDbPY: <http://imdbpy.sourceforge.net/>
3. Tableau Desktop: <https://www.tableau.com/de-de/products/desktop>
4. Matlab: <https://de.mathworks.com/products/matlab.html>

Implementation

5. D3: <https://d3js.org/>

Icon:

6. freepik: <http://www.freepik.com/>

Color palette

7. Colorgorical: <http://vrl.cs.brown.edu/color>