

# What makes people go to movies

Data Visualizations by Emma van den Broek, Inèz Hemme, Jonathan Jeuring, Door Kolkman, Mette Laros and Jakub Stachurski

## Ratings

Rotten Tomatoes dataset:

<https://www.kaggle.com/datasets/stefanoleone992/rotten-tomatoes-movies-and-critic-reviews-dataset>

Rotten Tomatoes ratings over the years

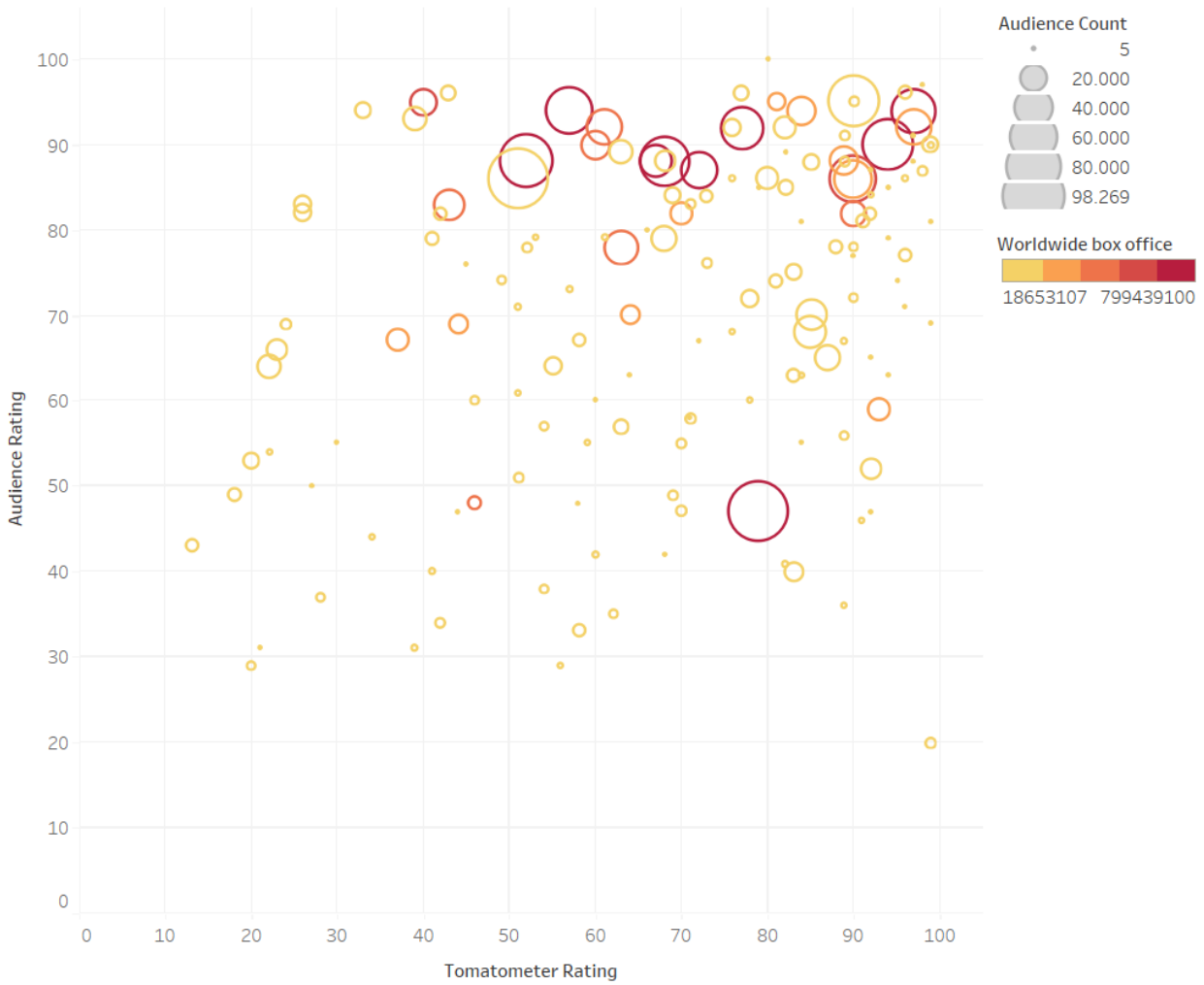


The trends of Avg. Audience Rating, Avg. Tomatometer Rating, Avg. Audience Count and Avg. Audience Count for Original Release Date Year. Color shows details about Avg. Audience Rating, Avg. Tomatometer Rating and Avg. Audience Count. The data is filtered on Original Release Date, Audience Count and Original Release Date Year. The Original Release Date filter keeps non-Null values only. The Audience Count filter ranges from 5 to 6732586. The Original Release Date Year filter has multiple members selected.

To this the box office dataset was added:

<https://www.kaggle.com/datasets/jonbown/worldwide-box-office-rankings-1977>

## Rotten Tomatos movie ratings and popularity in 2019



Sum of Tomatometer Rating vs. sum of Audience Rating. Color shows details about Worldwide. Size shows sum of Audience Count. Details are shown for Title and Movie Title. The data is filtered on Original Release Date Year and Tomatometer Count. The Original Release Date Year filter keeps 2019. The Tomatometer Count filter ranges from 100 to 574. The view is filtered on sum of Audience Rating and sum of Tomatometer Rating. The sum of Audience Rating filter keeps non-Null values only. The sum of Tomatometer Rating filter keeps non-Null values only.

## Actors Datasets

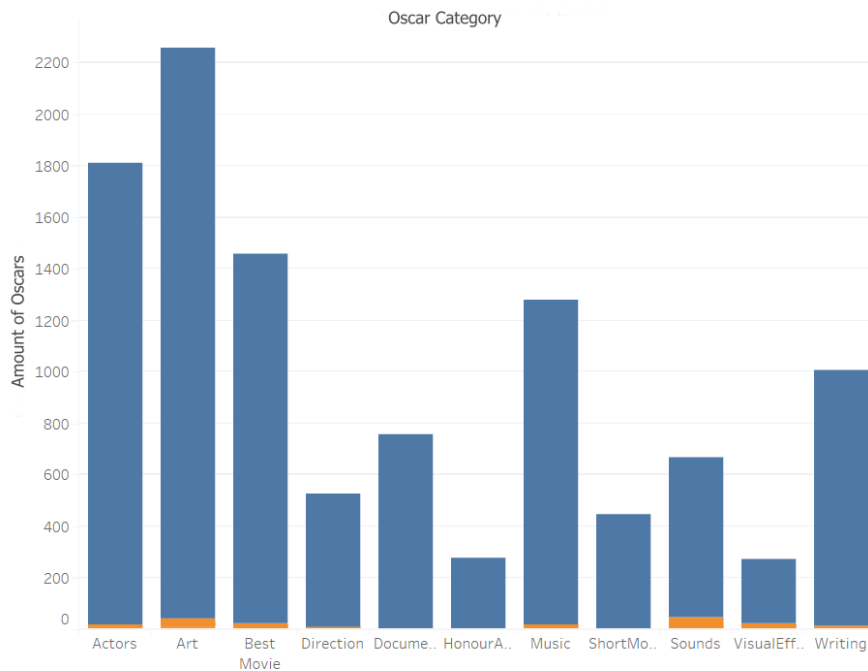
Rotten Tomatoes dataset

<https://www.kaggle.com/datasets/stefanoleone992/rotten-tomatoes-movies-and-critic-reviews-dataset>

Oscars

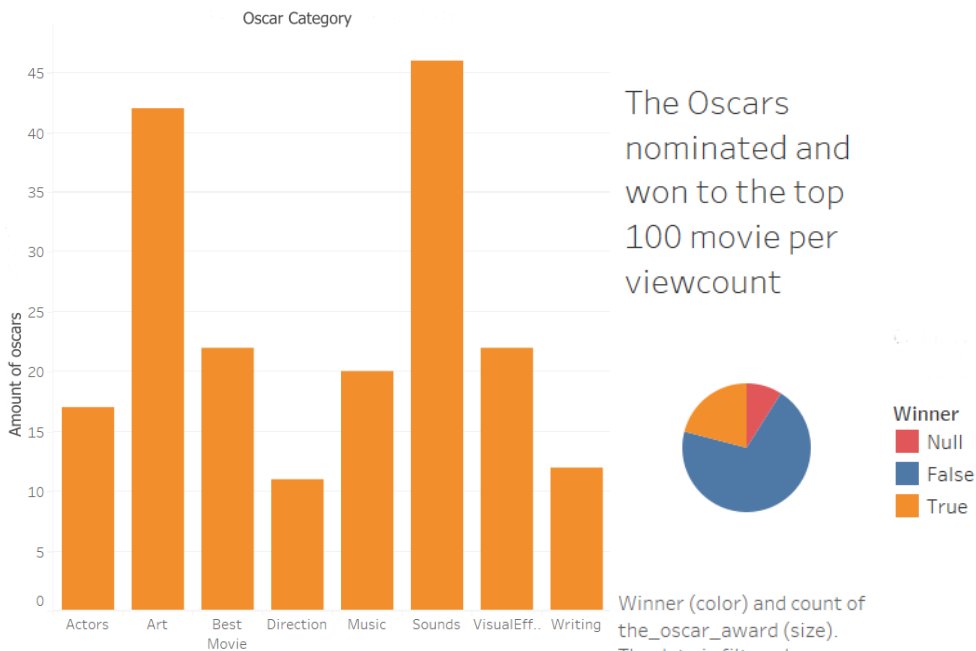
<https://www.kaggle.com/datasets/thedevastator/oscar-award-nominees-and-winners-1929-2020-movie>

### Amount of oscar nominations per category



Count of the\_oscars for each Category (The!Oscar!Award1) (group). Color shows details about Movie Title.

### The oscars of the top 100 movies per viewcount

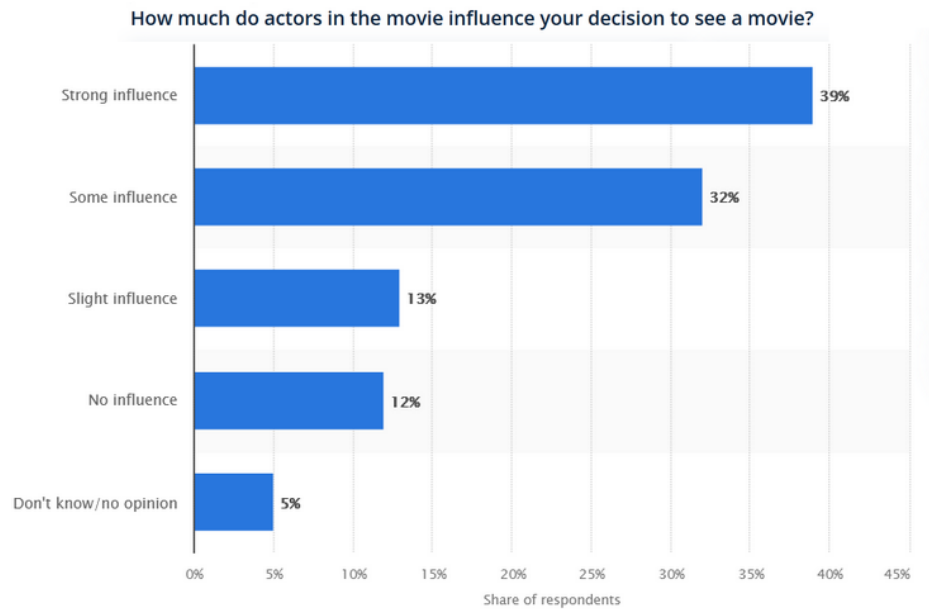


The Oscars nominated and won to the top 100 movie per viewcount

Count of the\_oscars for each Category (The!Oscar!Award1) (group). The data is filtered on Movie Title, which excludes Null. The view is filtered on count of the\_oscars, which includes everything.

Winner (color) and count of the\_oscars (size). The data is filtered on Movie Title, which keeps 101 of 101 members.

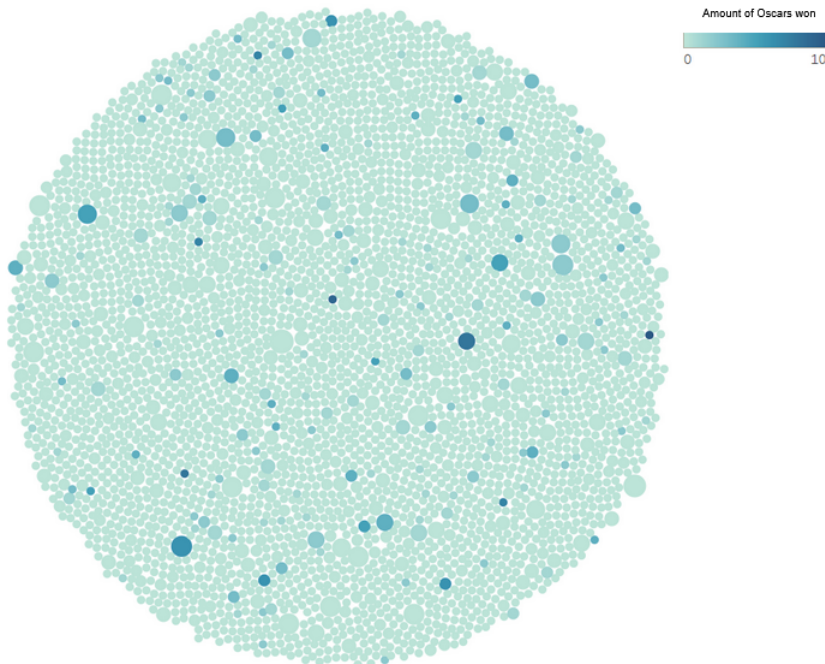
Count of the\_oscars for each Category (The!Oscar!Award1) (group). The data is filtered on Movie Title, which excludes Null. The view is filtered on count of the\_oscars, which includes everything.



Source: <https://www.statista.com/statistics/682973/actors-in-the-movie-influence/>

**Visualisation from Statista used to show the influence of actors on decisions to use movies.** <https://www.statista.com/statistics/682973/actors-in-the-movie-influence/>

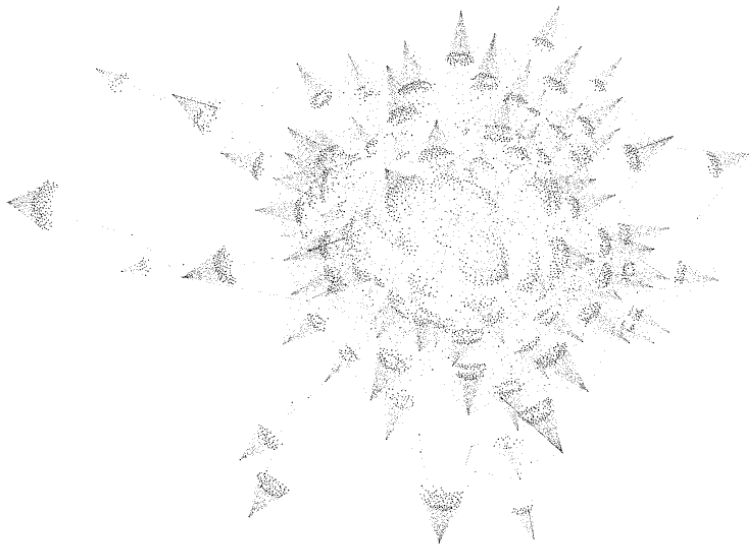
Actors in the top 100 showing the amount of oscars won



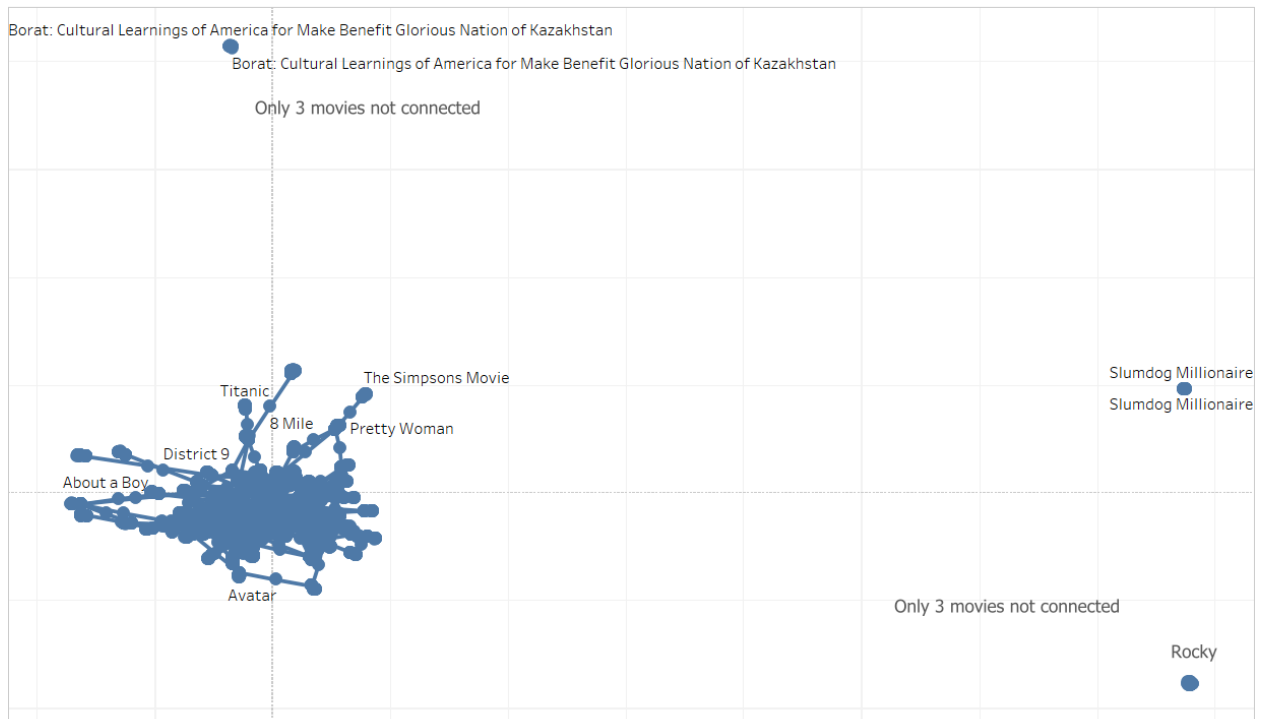
Actors. Color shows count of the\_oscars\_award. Size shows count of TabularDataset.csv. The marks are labeled by Actors. The view is filtered on Actors, which has multiple members selected.

The following program was used: <https://gephi.org/users/>  
The setting used is called Force Atlas

# Actors who act in the top 100 movies per view count on rotten tomatoes connected by role in the same movie

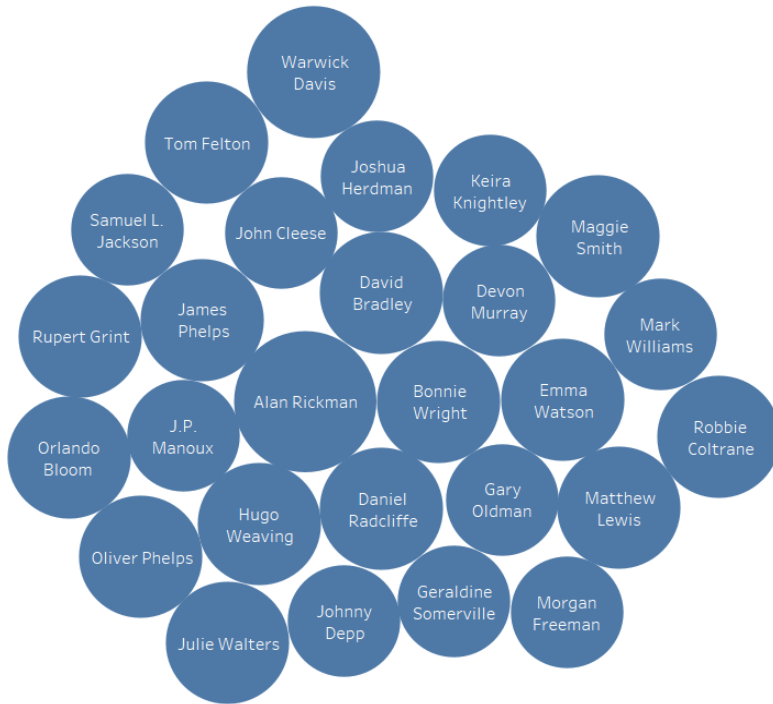


## Connection between actors in top 100 movies per viewcount



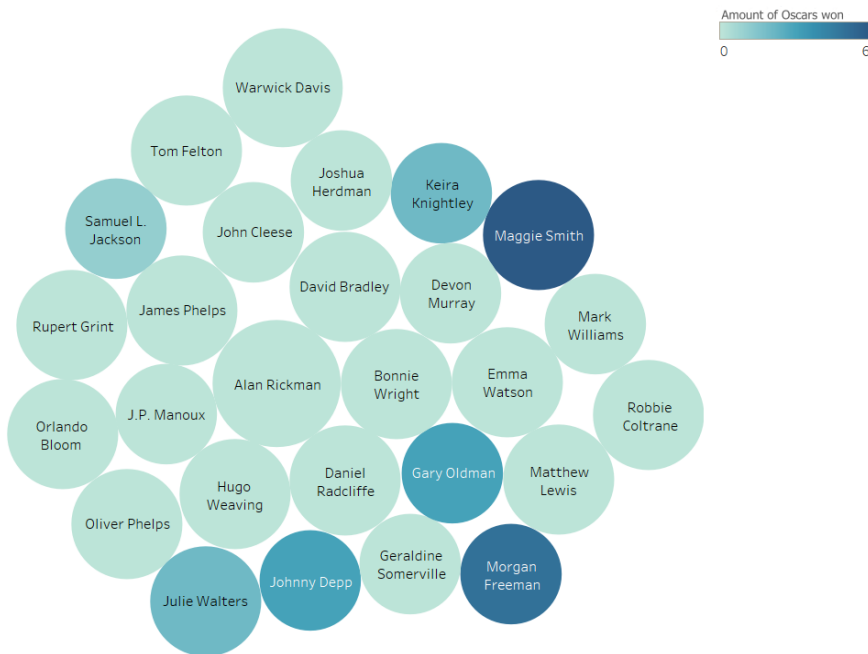
X vs. Y and Y. Details are shown for Source, Target and Movie Title. For pane Y: The marks are labeled by Movie Title. The data is filtered on Actors, which keeps 4.114 of 4.114 members. The view is filtered on Exclusions (Movie Title,Source,Target,X,Y) and Y. The Exclusions (Movie Title,Source,Target,X,Y) filter keeps 9.502 members. The Y filter keeps non-Null values only.

Actors credited in 5 movies or more in the top 10



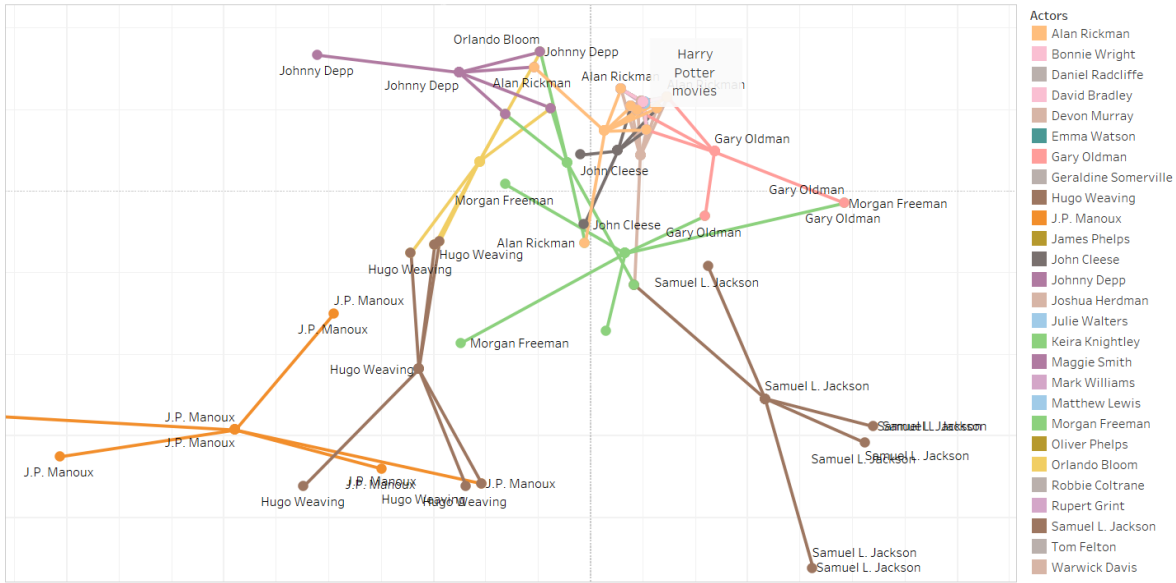
Actors. Size shows count of TablueDataset.csv. The marks are labeled by Actors. The view is filtered on Actors, which keeps 27 of 4.114 members.

Actors credited in 5 movies or more in the top 10



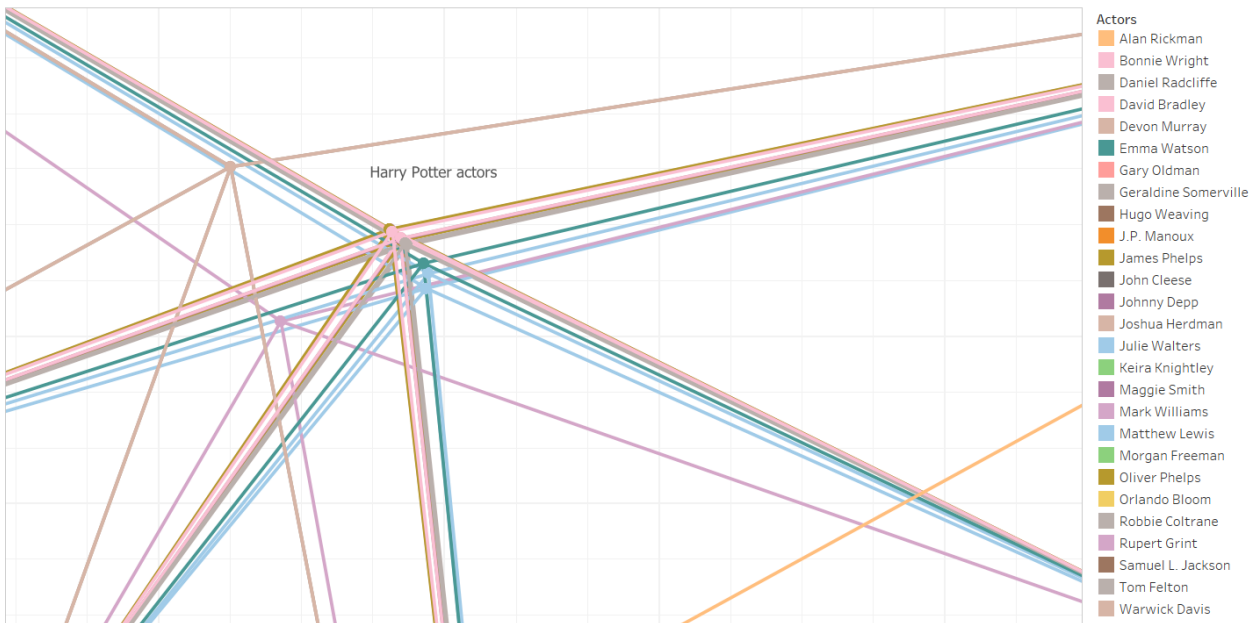
Actors. Color shows count of the\_oscar\_award. Size shows count of TablueDataset.csv. The marks are labeled by Actors. The view is filtered on Actors, which has multiple members selected.

Connection between actors with more than 5 roles in top 100 movies per viewcount



X vs. Y and Y. Color shows details about Actors. The marks are labeled by Actors. Details are shown for Source, Target and Movie Title. The view is filtered on Exclusions (Movie Title,Source,Target,X,Y), Actors and Y. The Exclusions (Movie Title,Source,Target,X,Y) filter keeps 9.502 members. The Actors filter keeps 27 of 4.114 members. The Y filter keeps non-Null values only.

Connection between actors with more than 5 roles in top 100 movies per viewcount



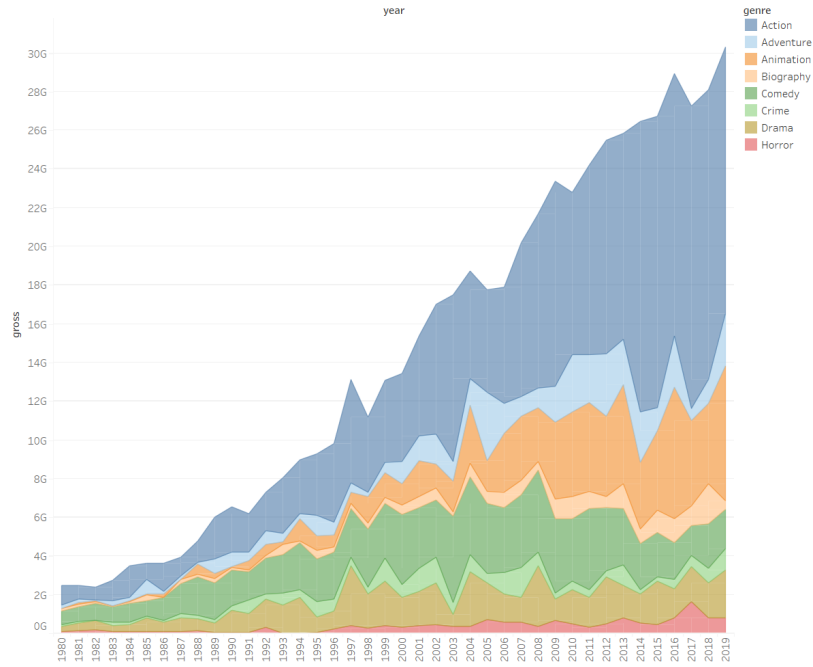
X vs. Y and Y. Color shows details about Actors. The marks are labeled by Actors. Details are shown for Source, Target and Movie Title. The view is filtered on Exclusions (Movie Title,Source,Target,X,Y), Actors and Y. The Exclusions (Movie Title,Source,Target,X,Y) filter keeps 9.502 members. The Actors filter keeps 27 of 4.114 members. The Y filter keeps non-Null values only.

# Genre

## Movie trends by genre popularity and profit

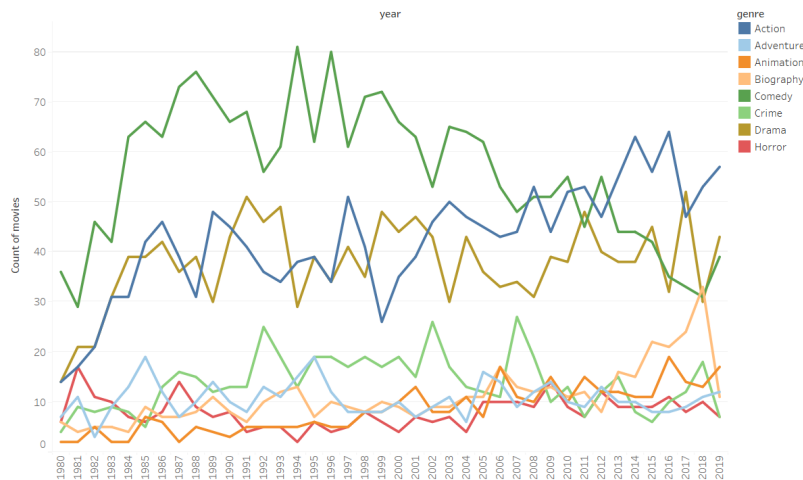
<https://www.kaggle.com/code/kbevinayers/movie-trends-by-genre-popularity-and-profit/input?select=movies.csv>

Genre Grosses Over The Years



Sum of gross for each year. Colour shows details about genre. The view is filtered on genre, sum of grossandyear. The genre filter keeps 8 of 19 members. The sum of gross filter keeps non-Null values only. The year filter excludes 2020.

Movie Releases by Genre Over Time



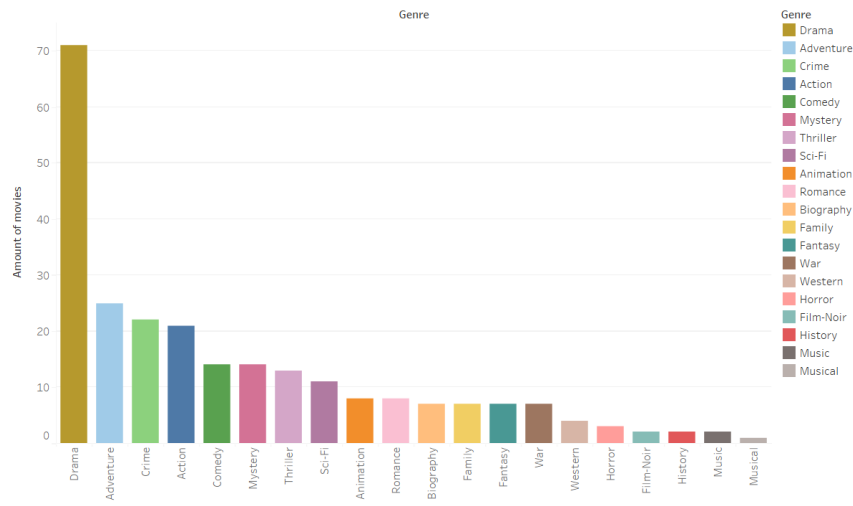
The trend of count of movies csv for year. Colour shows details about genre. The view is filtered on yearandgenre. The year filter excludes 2020. The genre filter keeps 8 of 19 members.



# IMDB Dataset for Top 100 Movies

<https://gist.github.com/stungeye/a3af50385215b758637e73eaacac93a3>

Top 100 IMDb Movies Genres



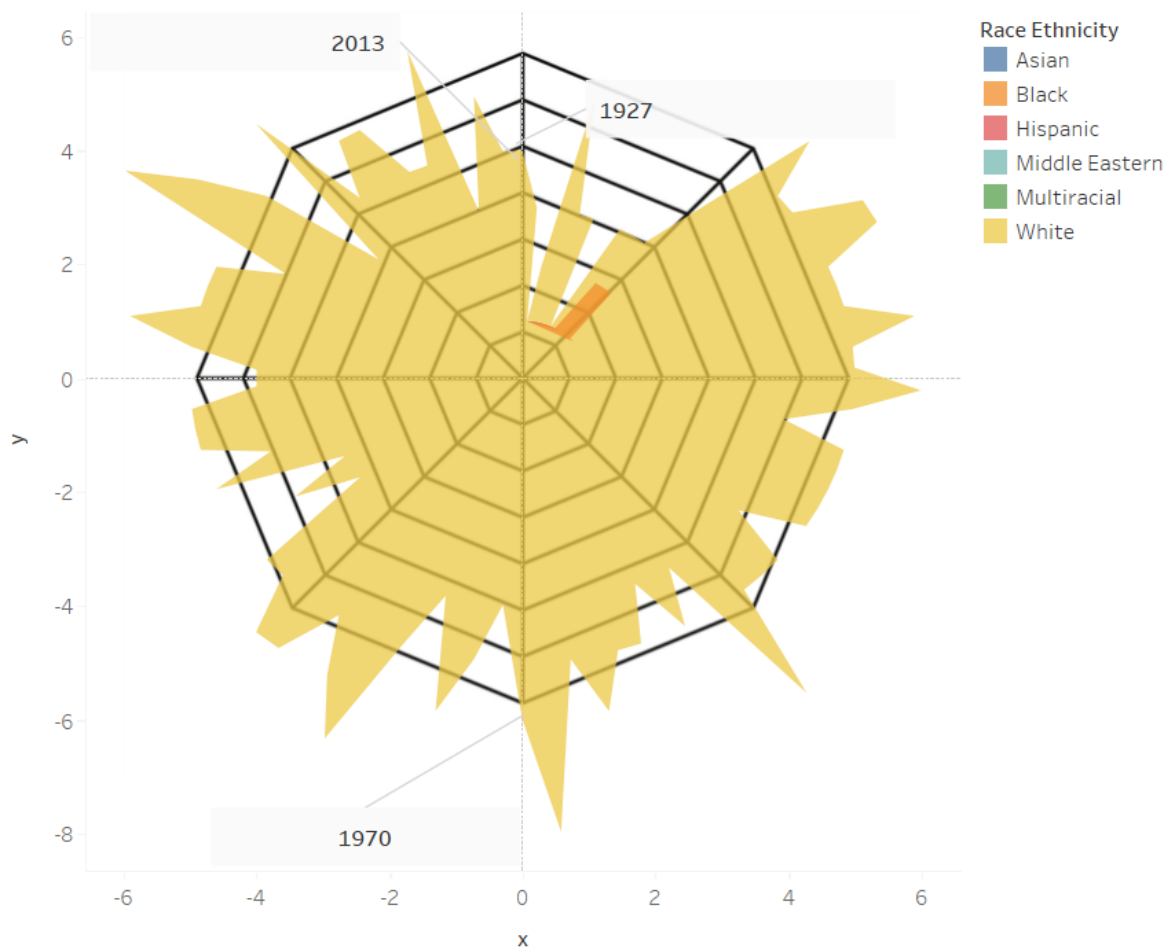
Count of Genre Column for each Genre. Colour shows details about Genre.

# Diversity

## Diversity of Oscar winners

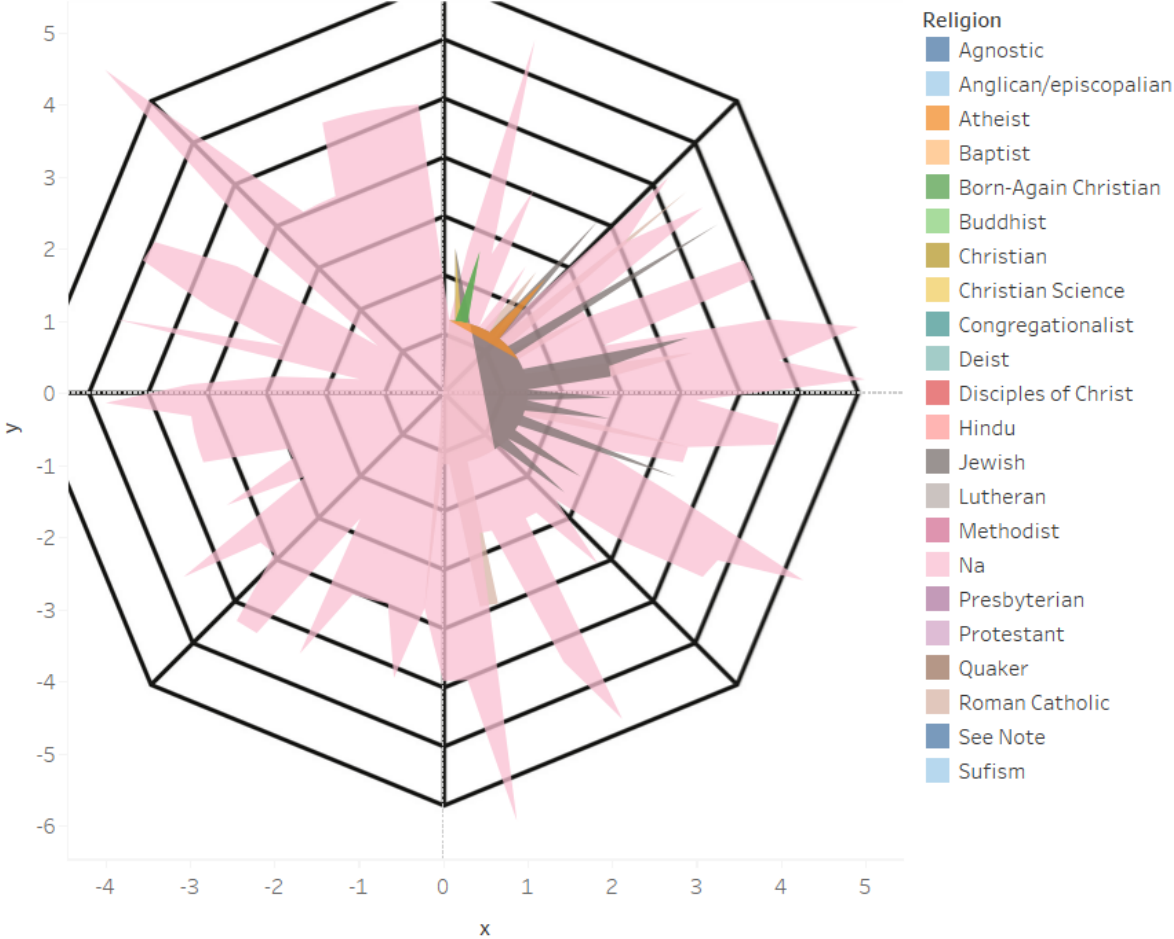
For the diversity of oscar winners, the dataset Academy Awards Demographics was used from <https://data.world/crowdfunder/academy-awards-demographics>. The tutorial for how to make this graph is from <https://www.youtube.com/watch?v=wWUXTQZZW2w&t=61s> and <https://www.youtube.com/watch?v=826y-AJVHR8&t=868s>.

### Etnicity of Oscar winners



X vs. y. Color shows details about Race Ethnicity. Details are shown for Year Of Award.

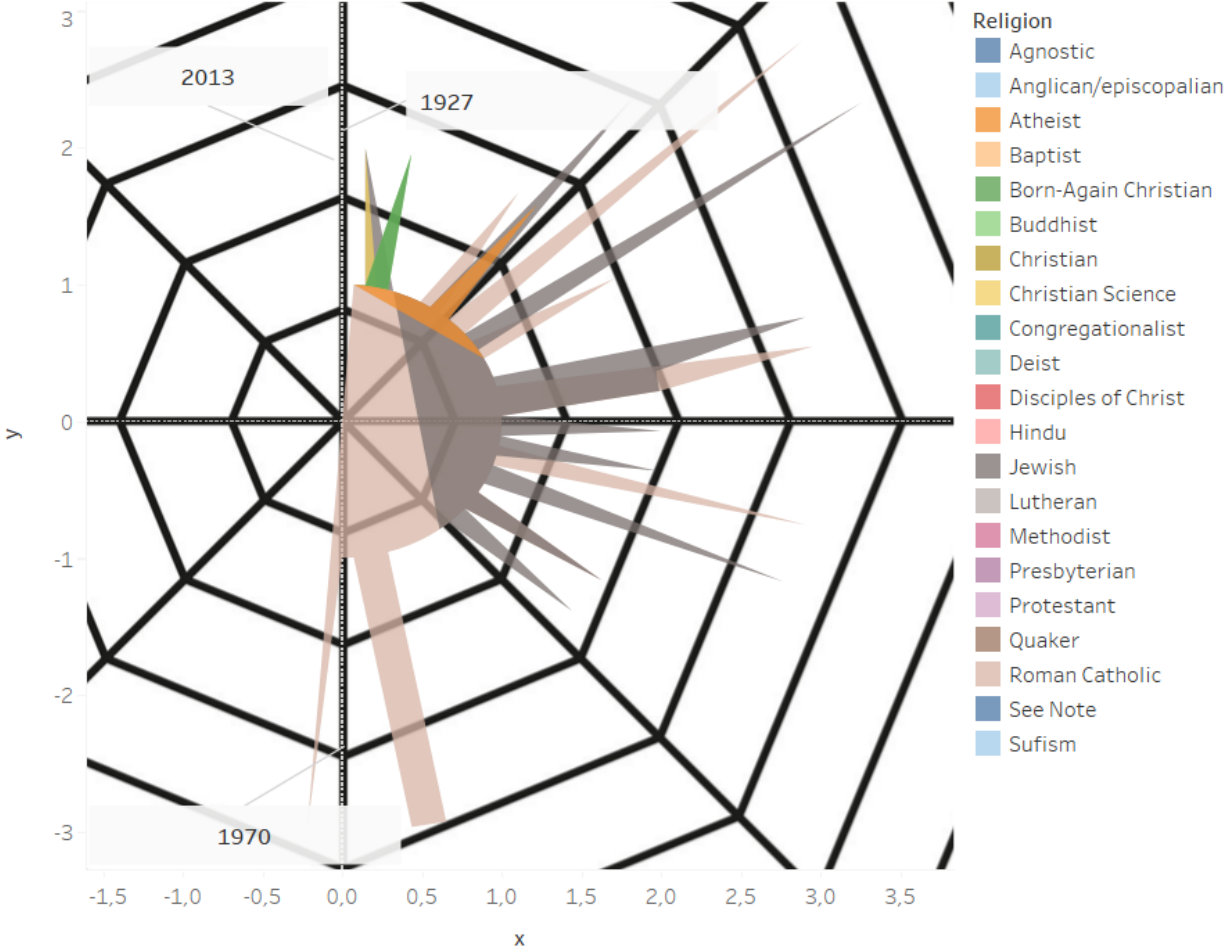
# Religion of Oscar winners



X vs. y. Color shows details about Religion. Details are shown for Year Of Award.

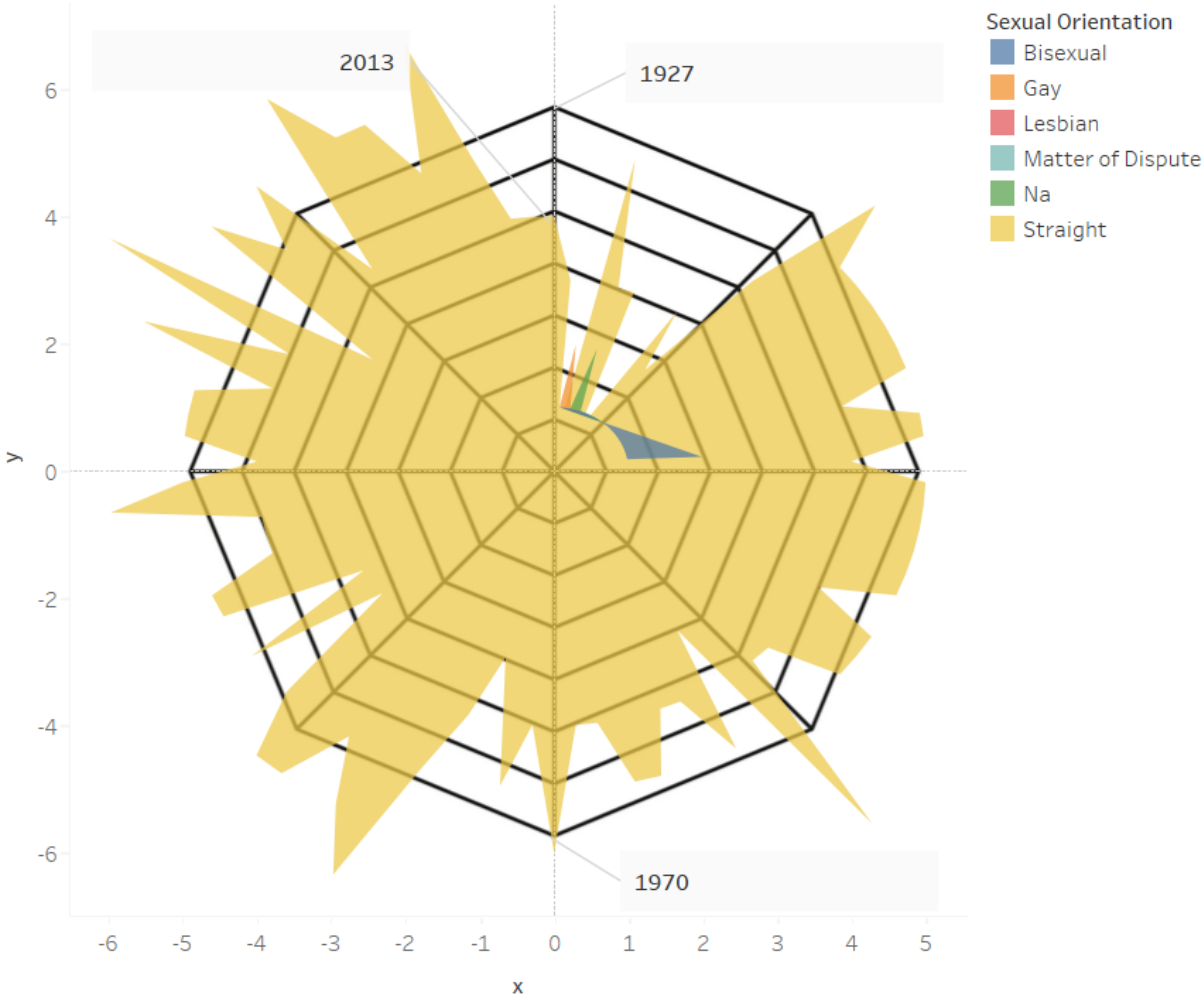
Because there were so many Oscar winners of whom their religion was not available, another graph was made on the religion excluding these values.

# Religion of Oscar winners



X vs. y. Color shows details about Religion. Details are shown for Year Of Award. The view is filtered on Religion, which excludes Na.

# Sexual orientation of Oscar winners

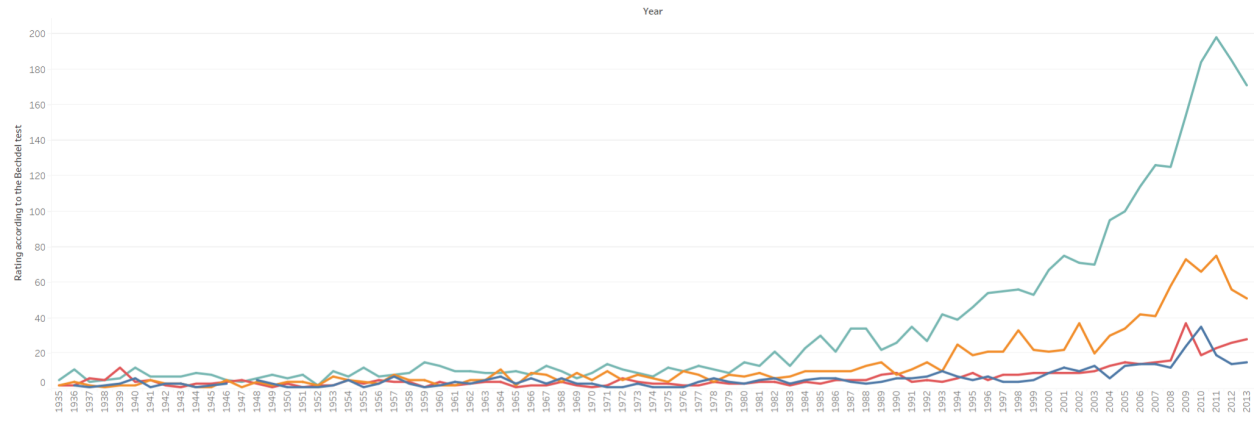


X vs. y. Color shows details about Sexual Orientation. Details are shown for Year Of Award.

# Bechdel tests

The visualisations for the Bechdel tests were made with the data from the dataset Bechdel Test from <https://data.world/sharon/bechdel-test>.

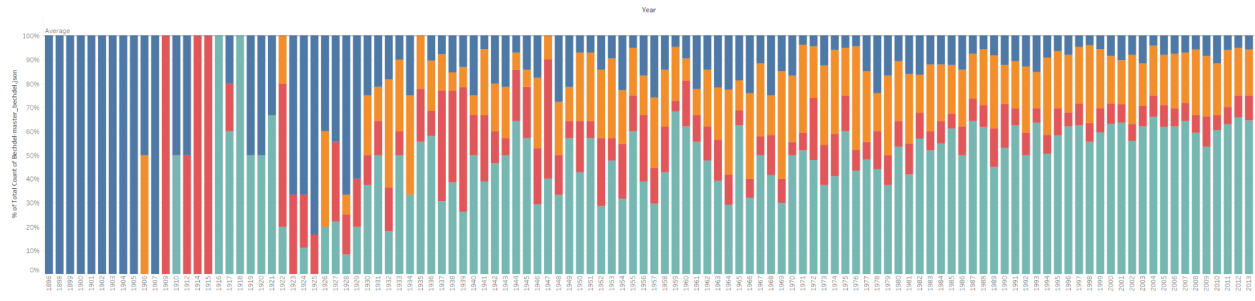
Bechdel scores over the years



The trend of count of Bechdel-master\_bechdel.json for Year. Color shows details about Rating. The data is filtered on Title, which keeps multiple members. The view is filtered on Year and Exclusions (Rating,Year). The Year filter keeps 79 of 116 members. The Exclusions (Rating,Year) filter keeps 390 members.



Bechdel rating deviation

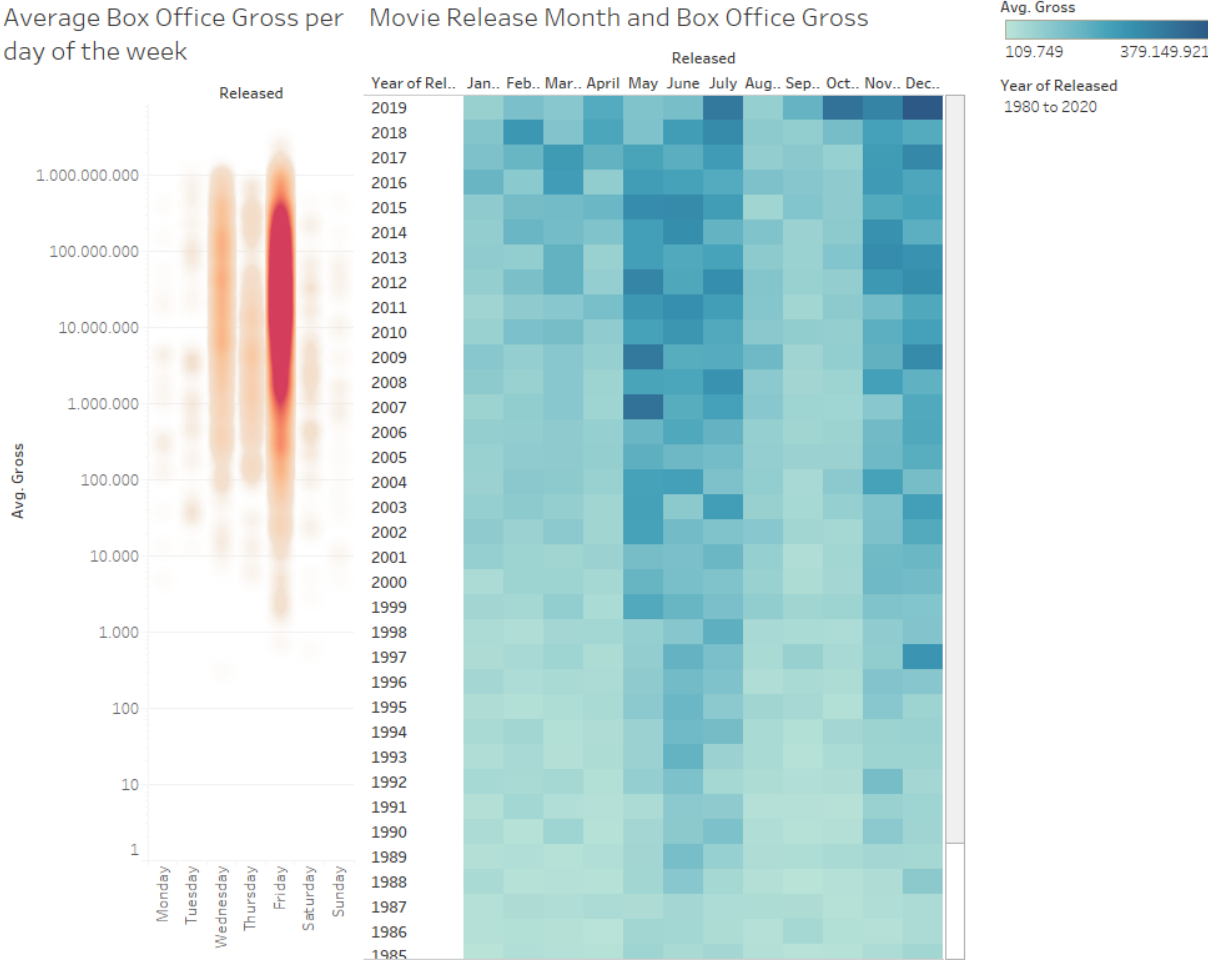


% of Total Count of Bechdel-master\_bechdel.json for each Year. Color shows details about Rating. The view is filtered on Year, which excludes Null and 2014.



# Release circumstances (Interactive dashboard)

The interactive part of the visualization is the common filter for the release years.



Movie.csv  
<https://www.kaggle.com/danielgrijalvas/movies?resource=download>

# Movie direction aspects

IMDb dataset:

<https://www.kaggle.com/datasets/carolzhngdc/imdb-5000-movie-dataset>

This dataset was adapted to ensure that there were no more duplicates and null values.

The colour summaries:

<https://www.thecolorsofmotion.com/>

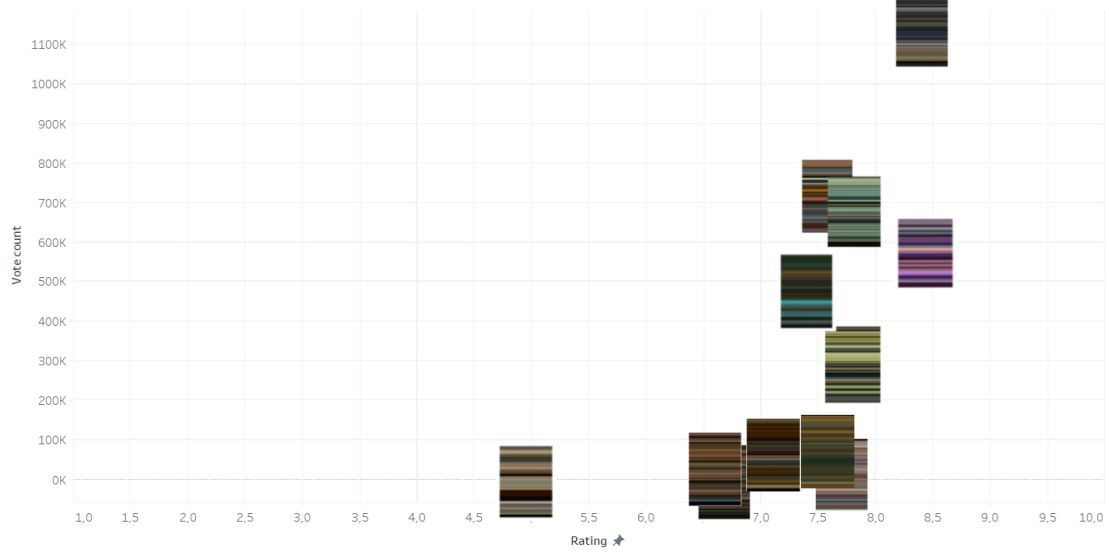
World map by amount of movies released in country of top 4000 IMDb



Map based on Longitude (generated) and Latitude (generated). Size shows count of Country. Details are shown for Country.

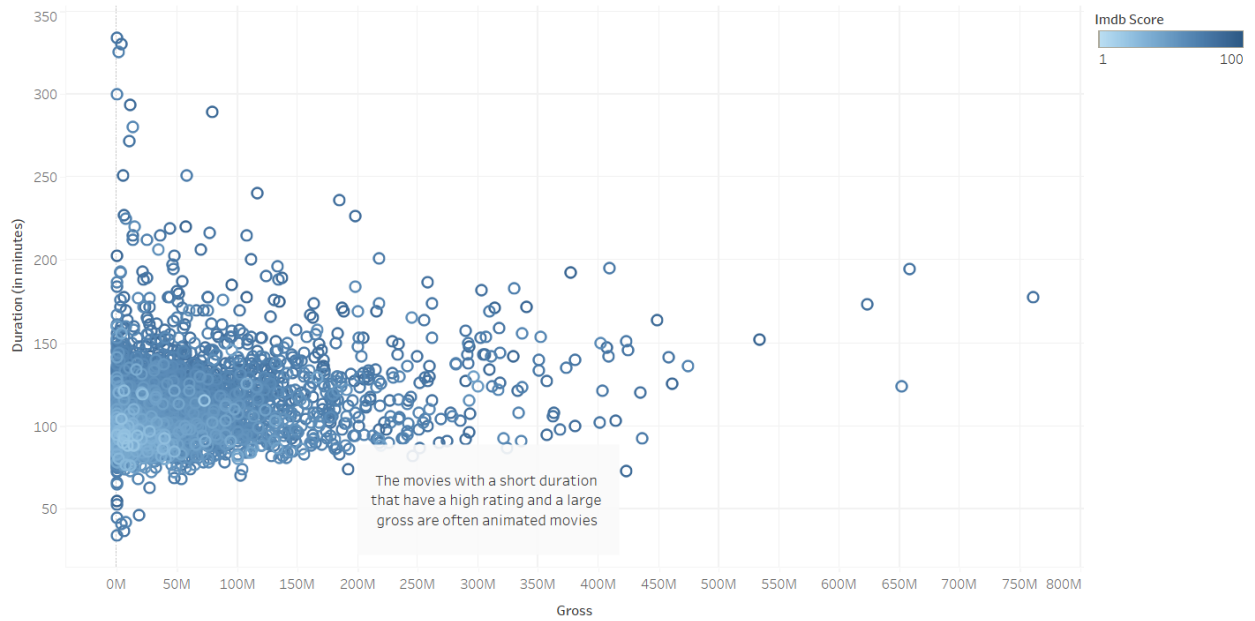


Vote Count vs Rating with colours of movies in 2017 & 2018



The trend of sum of Vote count for Rating. Details are shown for Title.

Duration vs Gross 1981-2016



Gross vs. Duration. Color shows details about Imdb Score.