

# Assignment 2

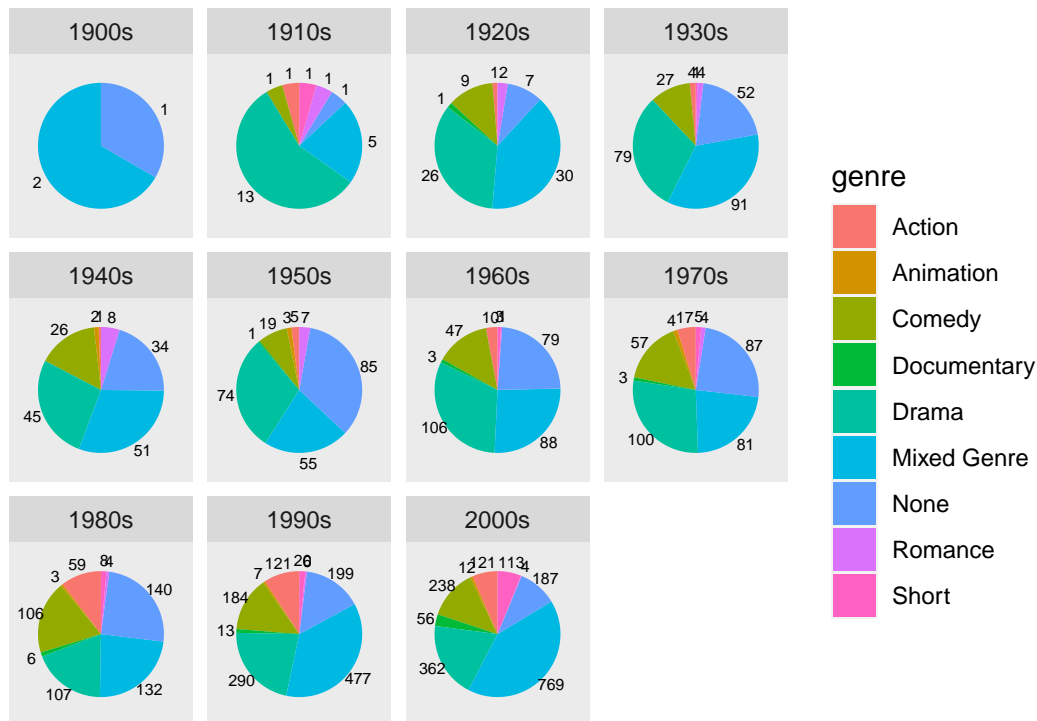
Amy Ly

2022-04-22

## Data Wrangling

## Original Plot Critique

Number of Films made per Genre



There were very few genres and films made in the early 1900s  
Many films by the 2000s are of mixed genres.

For my original plot, I used pie charts to show which movie genre became popular as time went on. The pie chart shows the relative counts of each genre out of the total for that year. The primary perceptual task is the perception of angle. It is also important to judge the data by the areas and arc lengths of the pie slices.

The original plot did not employ the easiest of perceptual tasks. It relied a lot on the area of pie charts (to convey information about frequency) and colors (which conveyed information about what movie genre was being identified). These are some of the hardest perceptual tasks.

A possible improvement would be to compare the genres on a common scale and length. For example, utilizing a side by side bar chart so that the magnitude of frequency of genres can be showcased and then have the

position of the bars be on a common scale to allow for easy comparison. Another possible visualization would be to do a line chart showing the changes in proportion over the decades.

The pie charts are faceted by decades and the only grouping that is utilized are different colors that represent genres. There was no particular meaning in the colors chosen. In hindsight, I should have chosen a color palette that was more colorblind friendly.

To help colorblind people see my visualizations, I will improve the plot by using a palette that incorporates sequential gradient of hue and luminescence. To make it easier to differentiate trend lines, I will also include different geom\_point symbols.

To abide by Gestalt's Law of Similarity, a possible improvement to color selection may be to use different shades to identify each slice within a group of particular genres. This may encourage people to compare slices of 's also encouraged to compare a group of slices to another group.

If I were to visualize the data as a series of connected data points, then the trend of proportion for each genre over the decades would abide to Gestalt's Law of Connectedness and show that the data points are related to each other. However, I know there is a higher proportion of mixed genre films and it may be better to separate the genres and count them individually. Since there are 7 different genres, we will not be overwhelming anyone with too many groupings.

In regards to the polishing of the original plot, it is not sufficient. The proportion text labels are difficult to read, a subtitle could have been added, and the caption did not include the source of the data. Since I plan on improving the visualization by using bar or line charts, I will also need to add on additional axis labels.

## Improved Plot

