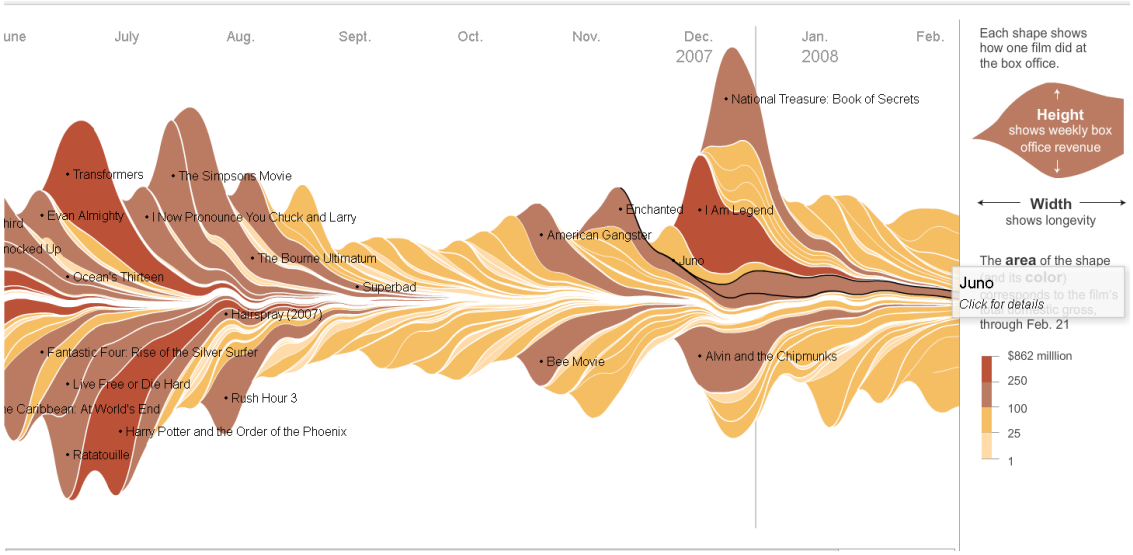
Problem 1: Design Critique



The problem domain of the visualisation above is to visualize the box office receipts from movies that were shown in the cinema from 1986 to 2008. These box office receipts are adjusted over inflation.

This visualization achieves four clear tasks. The first task achieved is to give insight in the box office revenue of each movie throughout he time it was shown in the cinema. A second task s to give insight in the total revenue spend at box offices weekly in the observed period. A third task achieved is to show the time that a movie was shown in the cinema, which is shown by the width of each ‘moviesurface’. The last task achieved is providing extra information about each movie when clicking on it.

We will now discuss some topics related to Tufte’s principles of graphical integrity. When considering the labelling of the scales we have concluded that the scales are not properly labelled. The y-axis does not contain any scale at all. The x-axis, however, is properly scaled: it shows the time span of the data as it is supposed to do. Secondly: The lie factor of this visualization is high. This lie factor is supposed to measure the effect in the visualization compared to the effect in the data. There are three parts of this visualization that contribute to this high lie factor. First the way the movies are sorted does not make much sense: our first interpretation was a positive and negative scaling in the y-axis, but there is no actual scaled y-axis. The movies are agglomerated in the visualization. This ensures people will think that a movie on top, has a high box office receipt, which isn’t necessarily true. The size and the colours of ‘the moviesurfaces’ are hard to link to the actual height of the box office receipt. All these parts contribute to the high lie factor. In the third place: the visualization shows data variation by the different size of the movie surfaces. It does also show design variation because of the different shapes the surfaces have.

We will now discuss Tufte’s visualization design principles. The ink-ratio of this

visualization is maximized, because the ink is mostly used to represent the data. There is no chart junk here, but more scaling (and thus more chart junk) would have been better. The data density of this visualization is high because there are many (maybe even maximally many) data items stored within the area of data in the graph. However, on the left side of the graph (1986…) the data density is quite low, because there are not so much observations which a high box office receipt. There are different layers of information in this visualization. For example when you click on a coloured surface, information about that movie is shown.

When talking about visualization the graphic design principles cannot be refrained from discussion. Here contrast is implemented by using different colours. This contrast ensures that the distinction between different movies is clear. The meaning of each colour is not clear: it is not intuitively logical that red means a high box office revenue and yellow means a low revenue. Contrast is also used in the different sizes (height and width) of each movie surface. Repetition is used here because movies are visualized during different time intervals. The colours are repeating in this visualization. This ensures that you directly see in which category a movie belongs and which movies are in the same category. Alignment is not so good in this visualization: there is no numeric scale. This means that you can compare movies to each other, but you cannot link a movie to a specified scale. Proximity is good here since related items (movies shown in the same period) are together.

The visual encodings used here do not make you understand the story told by the visualization very fast. The visual encodings used are: size, colour, width and height. These are appropriate visual encodings, but the way in which they are combined does not enable you to interpret the data fast. All information is shown somehow, but it is not easy or intuitively understandable.

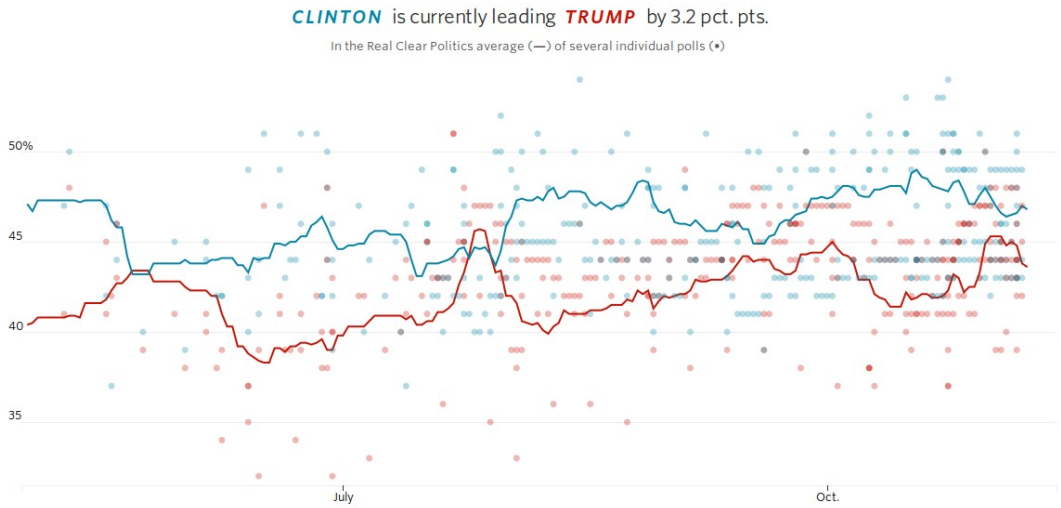
This visualization is intriguing because of its interactive elements and beautiful shapes and because of its unique style. The level of interactivity provides much playfulness.

The goal of this visualization, to show the box office receipts from different movies in a specified time, is achieved. However this is not shown in a easy understandable way and the information is not scalable at all.

We would like to change a few things about this visualization:

- Add a scalable y-axis and ‘flip’ all observations that are not pointing down.
- We would choose a different visualization form, maybe a bar graph.
- We would like to add a gradient of colours: from light to dark, where dark stands for a high box office revenue. This is intuitively more clear than the colours red, yellow and brown.
- Add a interactivity element: when you press a movie a pop-up window will show the total box office revenue of that movie.

Problem 2: Questions corresponding to the readings



Part one

Position

The position of elements in a graph give a clear vision on the values that data points represent. With predetermined ratios, it is easy to see which data points outperform others, which are further in time or any other measurement variable.

Answer

In this graph, position is used very well to first of all show what percentage of votes the candidate has on a particular point in time and second, which candidate is leading the elections. The viewer is able to see very fast whom of the candidates has more votes and how these percentages change over time.

Color

Color is the visual perceptual property corresponding in humans to the categories called red, green, blue, and others. The HSV color system contains hue, saturation, and value. Hue refers to the color names such as red, blue, green, and yellow. Saturation refers to the intensity of the color, or how bright it is. Value refers to the lightness or darkness of a color. Choropleth maps often use color value to differentiate between characteristics that are being mapped.

Answer

This data visualization makes use of two colors which fit very well with their purpose.

The visualization shows the polls of the American elections in which there are

two parties each carrying a distinct color, red and blue.
Saturation is used to show the difference between the average poll (which is shown very bright) and individual polls which are a bit vague.
In this visualization they also use different values of the color but it’s not very clear what these different values mean.

Part two

The designer is trying to show the broad diversity in polls. It clearly shows the average of the expectations, but also shows the upper and lower bounds and how the individual polls are scattered.

This visualization should help you to make a prediction about the upcoming elections. With showing the average expectation, this graph really achieves this goal. At the same time, with showing the individual polls as said before, it also shows the uncertainty of this prediction.

This and last, this visualization also shows some trends in the expectations of the elections. It is clear to see that in the two months shown, the percentages of votes for both candidates have grown. Also the graph shows clearly some peaks and drops in the average expectations, however these are not explained.