Critique on the Ebb and Flow of Movies visualization

The visualization depicts how movies performed over time at the box office visualized by the amount of tickets sold (in US dollars, adjusted for inflation) throughout the years 1986 until 2008. The movies that are presented on the visualization are those which have been reviewed by the organization behind the creation of the visualization, the New York Times.

By hovering over a movie element in the visualization (a title pops up), one can see how much approximately a movie has earned in total at the box office (by looking at the color and the area of the movie element), as well as the life-cycle of the movie (by looking at how long the line draws on the x-axis and the corresponding height). Furthermore, by clicking on one of the movie elements, one gets access to a pop-up window with more information about the movie. The visualization gives the possibility to move through time by means of a horizontal scrollbar or by dragging. The box office revenues are stacked for each period. This allows for a total box office comparison between periods, and the fact that the visualization spans a wide period also allows see yearly box office trends.

The x-axis is clearly labeled. However, there is no label for the y-axis. You can see on the side of the visualization that the height of the colored area shows the box office revenue per week, but this does not allow an accurate interpretation of the data.

On first glance, it looks like the movies on the outsides of the visualization are the most profitable ones, although this is not always the case (as can be seen by the colours). Moreover, the areas representing the profits can both go in a positive and a negative direction on the y-axis. It may seem that the movies which have their areas in a positive direction are more profitable than the ones in a negative direction, but this is not the case. It could even be interpreted as the movies on the bottom of the visualization are actually losing money rather than turning a profit. So, one could note a lie factor.

Most of the variation in the visualization comes from the data. The representation of one movie does, however, depend on other movies that were running during the same time. This changes the design at the same time, so it could be concluded that part of the variation does come from design variation. Almost all ink is used to represent the data, so the data-ink ratio is maximized.

A lot of data is represented in the visualization, so the data density is indeed high. The movies with different box office profits differ by color, so separate layers are used. Other attributes of different movies, like the shapes of the areas, the tone or the weight are very similar, so different layers are not used to great extent.

Color contrast and outlines makes it possible to distinguish separate movies. The white border that defines the border of a movie is in contrast with the colors used to indicate the domestic gross. As you hover over a movie the outline becomes black. In terms of color and style the data visualization is simple as there are only four different colors and the colors corresponding to the two lowest groups in terms of total box office are very similar. The style is consistent throughout the whole visualization. The shape of the separate data areas, however, depends on not only the individual data but also the surrounding data which makes the shapes unique.

All data points are aligned in case of the time component (x axis). Also alignment can be found in the height of the separate data points, however since the data does not start at the same height which makes it hard to compare data that looks alike. Although it does show relative differences between the movies, the encodings used do not leave room for showing details about the given information.

Things we would do differently are for example to provide the total box office revenue for each movie in the message box, since movie data is stacked and the y-axis is not labeled this is difficult to interpret from the visualization. Secondly, we would prevent the visualization from going out of bounds. Thirdly flexible coloring of the movies, where for example as total box office revenues increases, the darkness or amount of red increases. The simple grouping of wide revenue groups makes it hard to compare movies in terms of total box office revenue and could imply large differences between movies in different groups while they are small and small differences between movies within the same group while they are large. Another consideration would be the make the y-axis only increasing upward from zero instead of increasing in both directions. This would allow interpreting the y-axis better.

Problem 2:

* Two of Bertins visual variables and discussion in relation to our visualization:
  + Position: the position is on a quantitative scale and allows us to see how much the average life expectancy and the average income per capita are for each variable (country).
  + Color: each color represents a different continent. In this visualization, Africa is depicted in blue for example. The colors are thus extremely selective.
* What is the designer trying to convey and how does the visualization achieve this task:
  + Some of the tasks we think the designer tried to allow us to accomplish:
    - Differentiate between countries and continents
    - Rapidly differentiate the amount of the average life expectancy
    - See over time how average life expectancy and income per person evolve
    - Rapidly identify the correlation between average life expectancy and income per person
  + The visualization does achieve these tasks by using different colours and allowing us to scroll over time. Furthermore, the y- and x-axises are well scaled so we can easily see the amount of average income and/or the life expectancy. Finally, the fact that the year count is depicted in the background allows us to keep track of time.

Visualization:

*http://www.gapminder.org/tools/#\_locale\_id=en;&state\_time\_value=1983;;&chart-type=bubbles* 