

## Design critique 7

### Patrick Jonk, 10001336

This is an evaluation of the data visualization "Track National Unemployment, Job Gains and Job Losses", published by Andrew Van Dam and Renee Lightner of the Wall Street Journal on December 4th, 2015.

The visualization consists of two pairs of interactive graphs.

#### Domain and task

The domain of this visualization is economics [1]. I think the readers of the Wall Street Journal can be expected to have background or special interest in economics or management. Nevertheless, I think this visualization is also understandable for the layperson.

The task that this visualization achieves is to inform the public on the change in unemployment rate. The unemployment rate can be explored throughout time. It is possible to zoom in on the different sectors that make up the market. Data can be filtered for different demographics.

#### Tufte's principles

The graphical integrity of a visualization is given by three principles [2]:

- clear labelling and appropriate scales
- lie factor
- data variation, not design variation (lecture 2 slides)

I do not think all these principles have been met.

In all the graphs of the visualisation, all scales are present and appropriate.

The lie factor is the (size of the effects shown in graphic) / (size of effect in data). I would like to argue that the lie factor in the visualization is constant for all datapoints. In the first graph, the height of the graph and the number of sectors that have risen or fallen is proportional. This also seems to be the case in the second graph. In the third graph, size does not encode any data and so we can not apply the lie factor here. In the fourth graph the unemployment rate and height are proportional so the lie factor is one. In other words, there is no bias in the size of datapoints in the graphs for certain points

In the first two graphs there is a distinct difference in graph design, which I think is justified by the differences in data displayed there.

In the last two graphs however, I think that the same data is shown twice with only a graph design difference. One could argue that the first graph shows the unemployment rate more detailed per month, whereas the second gives a better overview. Nevertheless, I think the use of two graphs for the same data is not justified here.

The visualization design principles are [2]:

- Maximize data-ink ratio
- Avoid chart junk
- Increase data density
- layer information

In all four graphs the data-ink ratio is maximized. All visual elements take up no more ink than they have to.

Also, there is no chart junk in the graphs. The graphs look clean and contain no redundant design elements.

It can be argued that graphs two and three have a low data density. Graph two is essentially an interactive table with a bar chart in it. Tables have a low data density as compared to other types of graphs as they get large. Graph three contains the same data as graph four, only the former takes up more than double the area than the latter. Because data density is defined as (number of data items) / (area of data in graphic), this would mean that the data density in the fourth graph is half the data density of the third graph. This obviously does not comply with the data density principle.

The layering of information is non-existent in all graphs. I don't think this is a bad thing. Layering of information is especially useful if there's an extremely large amount of data. This is not the case here.

## **Graphic design principles**

### **contrast**

In the first three there's an important contrast in color. In the first two graphs this contrast emphasizes the difference between sectors that have gained and sectors that have lost jobs. The rising sectors are represented by cool colors such as green and blue, while the falling sectors are represented by warm colors such as yellow and red.

In graph three a red - green contrast emphasizes the difference between months with respectively high and low unemployment.

Graph four does not use contrast at all.

### **Repetition**

Repetition plays an important role in the visualization. The colors used are repeated in graphs one, two and three. In graphs one and two, the same colors represent the same data values. In these graphs, the order is also an element that is repeated. By changing the order in graph two, the order in graph one will also change.

### **Alignment**

Between graphs three and four, alignment plays an important role. All data points in the same time

frame are aligned vertically.

### **Proximity**

In graph one all positive and all negative are grouped together. In the other graphs proximity doesn't really play an important role.

### **Visual encodings**

The visual encodings used in this visualization are color, position and size [3].

In graph one, datapoints of rising and falling sectors are close together. Also, the variable on the horizontal axis is time. This means that data points close together in time are also close number of markets. The color of the circles encode the percentage of growth in the sectors.

In graph two, the size of the graphs is encoded by the size of the bars in the chart. Colors encode the monthly change in size. The order of data points can be changed to sector name, sector size etcetera. Therefore, in this graph position can have different meanings.

In graph three data is encoded by position and color. Position encodes the data point's moment in time and color encodes the unemployment rate.

Graph four is a basic linegraph. Here, the only visual encoding is position.

I think these visual encodings are all appropriate. All encodings add information without being excessive.

### **Subjective dimensions**

What I think is good about this visualisation are the aesthetics, vividness and playfulness. All graphs look clean, neat and inviting. Because especially graph one and three are very colorful they really catch the viewers attention.

Due to the various the interactive features implemented in the graphs, the visualization has a high degree playfulness. Especially between graphs one and two, where selecting some data in one graph can completely change the other graph. This really invites the viewer to explore these features, while in the meanwhile to viewer absorbs the data.

### **Intended goal**

I think the goal of this visualization is to inform the public that the rise in unemployment rate caused by the recent recession has almost been recovered from. In addition tot that the unemploment rate is put in an historic perspective. The first set of graphs show how during the recent recession the number of jobs per sector dropped, but that this number has been rising again from about 2010. The second set of graphs shows that although recessions happen every so many years, the recent one has been a very large one.

What I would do differently, is first of all, is that I would use classical bar chart for graph two, instead of the one that is used. The table plus barchart that is used is too much of a mess in my opinion. Although less information can be shown in a standard bar chart, this information can be added to the

graph through interactivity.

Second, I would combine graph three and four into one interactive linegraph. In my opinion the use of two graphs is excessive and not aesthetically pleasing.

Although the use of the rainbow color is discouraged in the literature because it can confusing [4], I don't think it's a problem here. Because in graph one all positive and negative data points viewed separately, the use of color is not confusing.

## **Sources**

1. T. Munzner, "A Nested Process Model for Visualization Design and Validation," IEEE Transactions on Visualization and Computer Graphics (InfoVis '09), vol. 15, no. 6, pp. 921–928, 2009.
2. CS 171: Visualization Design Principles Hanspeter Pfister, slides lecture 2
3. S. Carpendale, "Considering Visual Variables as a Basis for Information Visualisation".
4. D. Borland and R. M. Taylor II, "Rainbow Color Map (Still) Considered Harmful", IEEE Computer Graphics and Applications, vol. 27, no. 2, pp. 14–17, 2007.