

Design Critique 7 - Data Processing - Mounir Hader

Introduction

This design critique concerns a visualization by Andrew Van Dam and Renee Lightner. The purpose of the visualization is to enable readers of The Wallstreet Journal to track data about unemployment, job gains and job losses in the. The complete visualization consists of four charts and can be divided in two main parts which allow the user to achieve several tasks. These possible user tasks are explained in the next two paragraphs.

Job gains and losses

The first part of the visualization contains data about job gains and losses in 20 different sectors and consists of two interactive charts. The upper chart sorts the sectors based on monthly change in sector size. The sectors are represented by rounded, colored boxes. On the positive y-axis the sectors with a monthly gain in jobs are aligned, whereas on the negative y-axis we find the sectors with a monthly loss in jobs. Sectors with no change in job totals are aligned at the bottom of the figure. The color of a box represents the percentage change from the previous month in each sector's payrolls. The bar chart located directly below the upper chart shows details about job gains and losses per sector in the month that is selected from the upper chart.

National unemployment

The second part of the visualization contains data about unemployment in the US from January 1948 until November 2015. It again consists of two interactive charts. The upper chart shows all months represented by blocks and colored according to nine different unemployment rate intervals. The second chart shows the same data represented in an interactive line graph. Both charts can be manipulated by setting filters for specific population segments.

Analysis

In this paragraph the visualization is analyzed and discussed, making use of design theory like Tufte's design principles, which are considered an excellent standard for graphic design.

The first chart is very abundant in data and maybe even a bit overwhelming at first glance. It becomes clear very soon that we are looking at a time scale, representing every sector at a positive or negative monthly growth. There is no scale distortion in the x-axis, which represents continuous months from February 2006. Although the scale of the y-axis is not distorted either, the choice of units can be questioned. According to the axis labels the axis should an absolute amount of sectors either rising or falling. This would mean the minus sign is not necessary here and factually incorrect. Since omitting the minus signs would lead to a chart that looks unnatural and thus probably more confusion for the user, the current design choice seems a reasonable one.

Integrity principles of Tufte (Tufte, 1986)

One of Tufte's integrity principles concerns the use of clear, detailed and thorough labeling and appropriate scaling. As just described, the uppermost chart succeeds in that for most part, except for the negative y-axis labeling. Furthermore, the bar chart with detailed information about the sectors should specify which unit was used under "Sector size" (which most likely is the number of jobs). Tufte's second graphical integrity principle concerns a chart's *Lie factor*. None of the four charts shows a scale distortion or a difference in graphic effect and data effect sizes. Thus, the *Lie factor* equals exactly one.

Tufte's third and last integrity principle says a chart should show data variation and not design variation. This visualization fully succeeds to measure up to this principle by the consistent use of axes and scaling: all charts show data counting up from zero, not from some arbitrary value. Also, the ranges of the color codes are evenly distributed (the intervals at which the color codes change are all evenly big).

Design principles of Tufte (Tufte, 1986)

Data-ink ratio maximization

An assessment of the design aspect of the visualization can be made using Tufte's visualization design principles. Of these principles, one in particular seems to be taken into account in this visualization's design process: maximization of the data-ink ratio. The main idea behind this principle is that the data is the visualization's priority. Only elements that will help the data to be understood more effectively should be included. On the other hand, elements should not be just removed for the sake of simplicity if they contain important information.

This principle is the driving force behind the merely white background throughout the entire visualization. It is the reason the axes in the upper chart do not consist of continuous lines, but only contain ticks, specifying the year and the number of sectors. The spaces between contiguous boxes also help to maximize the data-ink ratio. More or less the same is going on in the visualization's third chart: the colored block chart showing unemployment rate. White lines are used to separate blocks from each other, maximizing the data-ink ratio. One might even argue this principle is the reason for putting the total number of jobs *inside* the bars of the second chart, in white font of course, thus using less ink.

Avoid chart junk, increase data density and layer information

Tufte's other design principles are in accordance with the previously mentioned data-ink ratio maximization, especially the avoidance of chart junk. Increasing data density and layering information are also somewhat similar to the first design principle, but aren't used very much in the visualization (due to the fact that only one variable at a time needs to be visualized). The line graph at the bottom of the visualization is an exception: here multiple lines are drawn in the same chart when setting specific filters. This is in accordance with Tufte's principles of increasing data density and layering.

Graphic design principles of Robin Williams (Pfister, 2014)

Contrast

What we usually notice when viewing a visualization, is contrast. It drives the cognition and thus interpretation of the data represented. The concept of contrast can also be found back in this visualization. The colored boxes, bars and blocks throughout the entire visualization are in big contrast with the merely white backgrounds. Specifically the more dark colored show big contrast and therefore are more eye-catching. This is the reason that when looking at the upper chart, the clump of dark red colored boxes is the first thing that pops out. By doing so, it directly supports the statement of a recession in that period. At the bottom of the upper chart, the sectors that remain unchanged are colored grey, which is in much less contrast with the white background and thus is less outstanding.

A similar argumentation holds for the bar chart and box chart. Contrast is not really used in the line graph. When setting a filter a blue line appears which is not very distinct because of the lack of contrast with the black "overall" line. This makes it harder to read the lines separately.

Repetition, alignment and proximity

The visualization is very consistent in using colors, fonts and headers. Also the alignment of data elements is consistent throughout the visualization. The same sequential color pallets are used for job gains/losses and unemployment rate (although the scale is inverted). Nothing seems to be placed on the page arbitrarily. The symmetric, organized and sophisticated look of the page is acquired by strong alignment and makes the whole more visually appealing to the user.

Other points

One might question the choice of colors in this design: the many green and red elements don't seem to be very colorblind friendly. However, color blindness is essentially the inability to distinguish different hues of the same intensity from each other. Distinguishing different intensities from each other is usually not a problem for colorblind people (Pastukhov, 2013). Since this visualization uses green and red in different intensities the different blocks and boxes should be distinguishable, also for the colorblind.

It is good practice to enable users to seek information according to the mantra: "Overview first, zoom and filter, then details-on-demand" (J. Heer and B. Shneiderman, 2012). This mantra can be interpreted and applied to the visualization by noticing the first chart is an overview of all months. By selecting a specific month the bar chart below it will "zoom in" on the data and present data details to the user. In the unemployment visualizations this isn't really the case. Instead of selecting a month there and getting all the detailed information of that month, the user needs to select which details to see (by setting filters), which are then shown in the charts.

Opinion and discussion

The overall visualization looks very tidy and well-designed. The color choices and white background make the whole attracting to read. The design is simplistic and direct: no three-dimensional elements, yet very vivid and not at all boring to read. From the beginning it is clear what information the authors wish to display and what story they want to tell (how recessions, job losses and unemployment co-evolve). The interactive components of the visualization are user friendly and project a certain playfulness.

What is probably confusing in the beginning though, is that the header of the second chart doesn't change when zooming in on different months. An other point of criticism is that the black highlights of the boxes in the upper chart disappear too fast. A user might be interested in tracking a specific sector over several months, but it is pretty hard to follow the black dots without losing track of them (the user should move his/her mouse at a very high pace or another selection of black dots appears).

Altogether, we can conclude the visualization achieves its intended goal in presenting data in an understandable and comprehensive way. The design invites users to play with the interactive elements and enables them to search for specific data. For me, the visualization was clear and interesting, and would pass with a more than average score.

References

1. Tufte, E. (1982). "The Visual Display of Quantitative Information"
2. Pfister, H. (2014). "CS 171: Visualization" (powerpoint). Retrieved from <https://cdn.mprog.nl/dataviz/slides/02-Design.pdf>
3. Pastukhov, S. (2013). "Visualizing for the Color Blind". Retrieved from <http://www.dundas.com/blog-post/visualizing-for-the-color-blind/>
4. Heer, J. and Shneiderman, B. (2012). "Interactive Dynamics for Visual Analysis"