Skylar Wurster

CSE 5544 Data Visualization

Professor Jian Chen

February 26, 2019

Midterm exam part 2

1.

Good visualization - <https://www.nytimes.com/2014/04/23/upshot/the-american-middle-class-is-no-longer-the-worlds-richest.html>

Bad visualization - <https://www.washingtonpost.com/graphics/politics/trump-promise-tracker/?utm_term=.3ad4b7ad6e35>

1)

Good visualization

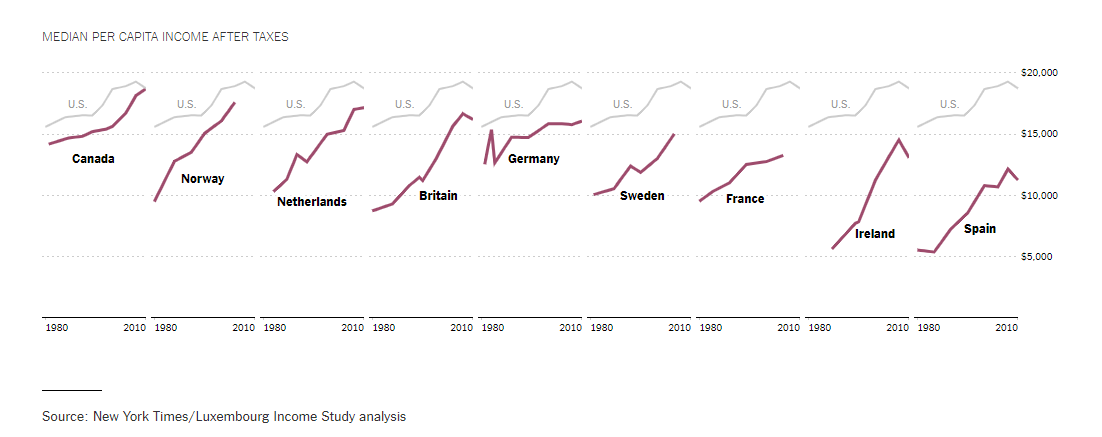
This piece from the New York Times is mainly aimed at the task of identifying the middle class’s income in the United States compared to other countries. Another task it has (as a consequence of the comparison task) is to show a trend in the United States middle-class income.

Bad visualization

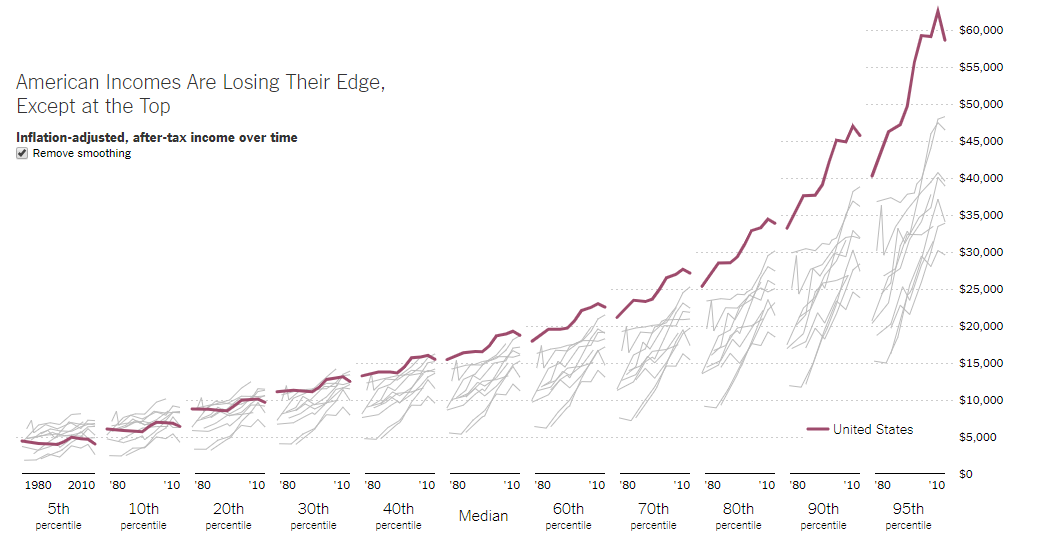
The Washington Post addresses the task of showing how President Trump has followed through with promises he listed in “Contract with the American Voter” that he created on Oct 22. Though this is ongoing and update, the task is not to show temporal change. Another task within the main task is to show more than just binary classification of whether a promise was kept or not – the visualization shows more information about whether the promise is being worked on, if it’s stuck, or if a compromise has been made.

2)

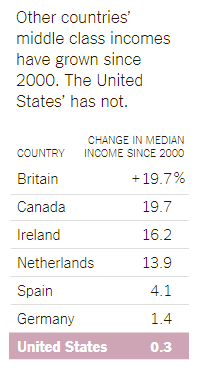
Good visualization



The dataset here must be quantitative number representing median per capita income (in US dollars) after taxes collected each year, most likely. This attribute is captured in US countries besides the US. Since other countries don’t use the US dollar, it is assumed that a currency conversion was done for the data with a conversion rate during some point in the year the data was taken. So, we have 10 countries with 30 data points each. This data is mapped to a line graph, which is excellent for showing trends. The US’s line graph is duplicated for multiple small views, where each view compares the US’s line graph to another countries line graph.

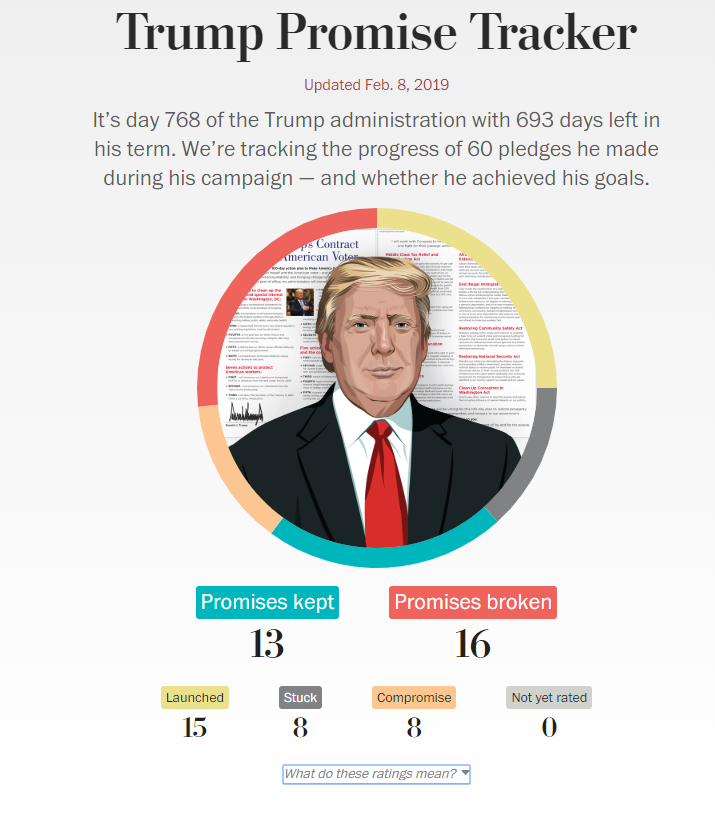


Further down in the article, we have another graph. This doesn’t use just the median per capita income, but income from each 10th percentile, along with the 5th and 95th percentiles. Again, we see a multiple small views approach where each graph is segmented by the percentile. The line graph is used very similarly to the line graph before, except now all countries are depicted in gray and combined on one graph, causing overlap.



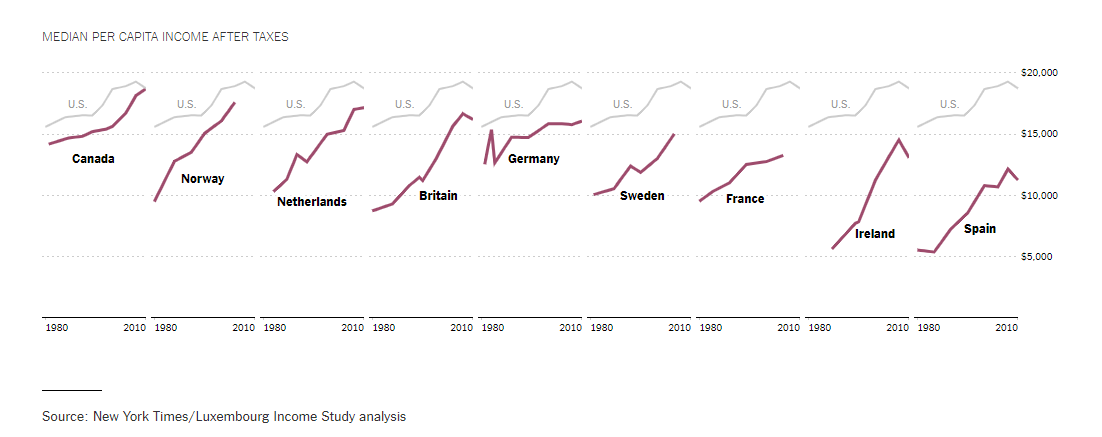
Though not really a visualization, this last chart is simply a single value put next to the country. The visual mark here is the exact text driving the visual mark.

Bad Visualization



This visualization maps the proportion of promises outcomes to the edge of a pie chart. That is, that if a certain promise outcome has a proportion of 0.45, then a slice taking up 45% of the area in a pie chart is made corresponding to it. One more visual mapping that doesn’t follow a pattern is the legend’s font size. We see that “13” and “16” under “Promises kept” and “Promises broken” have a larger font size (2.8em via the inspector) compared to “15” and “8” under “Launched” and “Stuck”, respectively (which have 2em font size via the inspector).

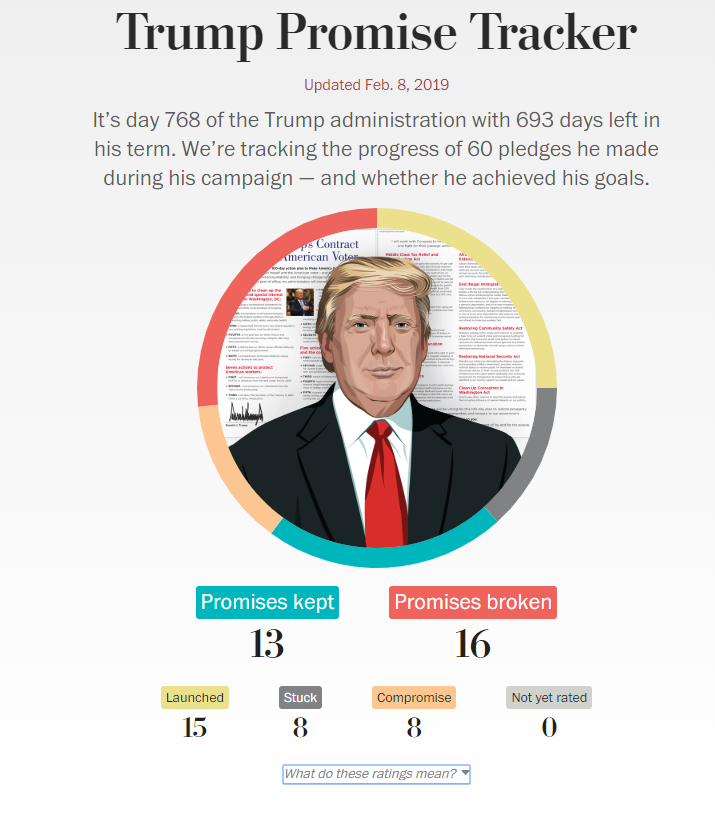
3)



This visualization is good. To begin with, we have clearly labeled axes and title. There is no broken y-axis that might be used for tricking viewers, and the x-axis are consistent across each multiple view. The task of comparing the United States median per capita income to other countries over time is clearly addressed with the line graph. Since that is the only comparison we care about (i.e., we don’t care about Ireland vs. France), it makes sense to use multiple small views instead of a single line graph with 10 lines for each country. One thing to note is that this graph may hide data. We only have a view of 10 countries here. Of these 10, most are very well off. Additionally, only going back to 1980 could be hiding important trends seen before that date. The paper would potentially make a stronger argument if the date started earlier and if more countries were compared against.

The chart itself is very expressive – facts are encoded and only those facts. No chart junk is present, except for perhaps the choice to use purple to highlight the compared country, and gray for the US. Therefore, due to the potential hiding of data brought up earlier, the chart could be more effective.

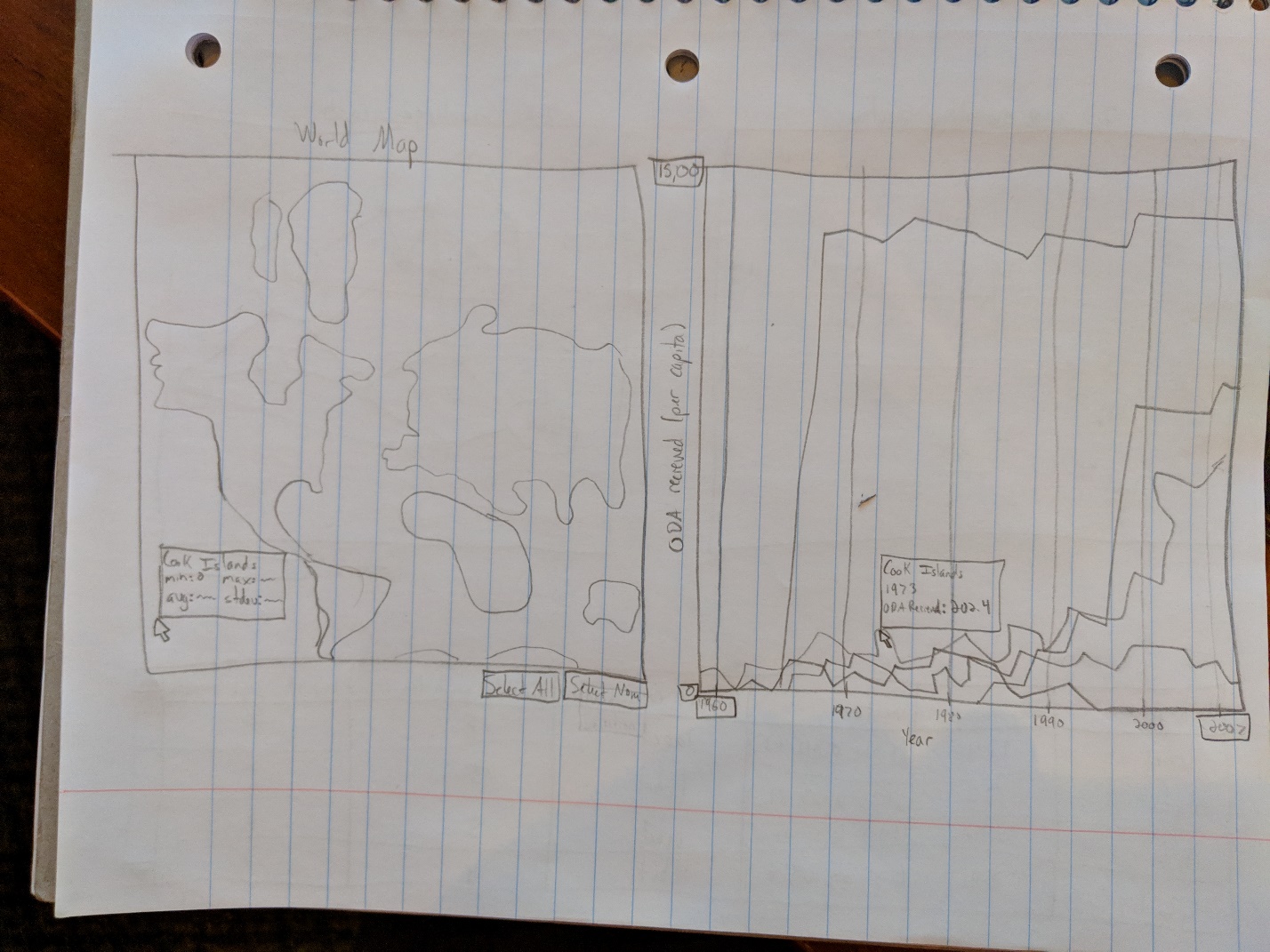
4)



This visualization’s task is to simply show the current outcome of promise outcomes. There is chart junk added into the center of the pie chart – Donald Trump’s face here does not contribute to the data whatsoever. This is a very ineffective. Another detail added that was not needed were the change in font sizes for the legend. This creates an information scent that will lead the user to think that promises kept and broken are in fact higher than all other promise outcomes, when we see that “Launched” has 15 but is in smaller font and generally overshadowed. This distracts from the data meant to be shown. Fix these issues by removing the image of Trump and making a smaller, more representative legend where each entry has the same font size.

A different chart type would make this data clearer; a simple bar graph with ordered entries from smallest to largest would effectively map the data attributes to visual markings while allowing comparisons between each of the entries.

2.



Describe your design:

This visualization uses multiple views to allow a user to see geographical positions of the countries selected and use the map as a filtering tool.

On the left, we have a world map. On the right, we have a line graph including each country in the dataset. When the visualization begins, all countries will be present on the line graph, assigned random but unique colors. Similarly, each country on the dataset will be highlighted on the world map on the left with the same color as their line. Since not all countries in the world are included in this dataset, those countries are grayed out on the world map.

The world map is interactable. First, note the “Select all” and “Select none” buttons. Clicking this will either highlight all countries in the dataset or highlight no countries. When this is done, the line graph on the right will gray out any lines of countries not selected and make them more transparent, but still visible. When a country is not selected, the user may click on that country on the world map to select that country and only that country. They may also shift click that country to add it to the list of currently select countries. Doing so will highlight only the selected countries on the line graph. Likewise, users may select lines on the line graph to add them to the list of selected countries, and the corresponding country on the world map will be highlighted. Finally, control clicking will either add the country if it’s not added or remove it if it has already been added. For precise details, a pop-up text box will appear when hovering over the line graph showing the exact country and value for the closest data point to the mouse. Similarly, hovering over the world map will create a text box with the country’s name and the min, max, average, and standard deviation for that country.

Both views allow for zooming and panning. In the world map, users may scroll wheel to zoom in or out on selected areas, then use the middle mouse button to pan around the scene when clicking and dragging. On the line graph, since there is such a large range in the y-axis, users may enter in a custom y-axis min and max, as well as x-axis start and end year. This allows them to brush through different domains and ranges. When the visualization starts, all years (1960-2007) are shown and the y-axis max is set above the maximum value in the dataset, which was about 15,000. Note that inputted y-values cannot “cross”, meaning the max is less than the min. Input processing will prevent that, and require that the max is always greater, and there is a difference of at least 1. Similarly, the x-axis cannot “cross”, so the right side must be dates closer to 2007, and the left side must be older. There must also be a difference of at least one year.

To summarize, there is one list of selected countries. Users may select a single country by left clicking the country on the map or the countries line on the line graph. Users may add a country to the selected countries list by shift clicking that country. Users may add or remove a country by control clicking on the country.

Questions answered through my design:

Question: How does country X ODA count per capita compare to country Y?

Solution: Select only those two countries and observe their differences on the line graph.

Q: Is there a trend between geographical placement and ODA count per capita?

S: Select a region on the globe by selecting countries that are geographically close to each other and view their line graphs on the right.

Q: What countries exhibit ODA received per capita in the range between Y1 and Y2?

A: Using the Y-axis on the line graph, type in Y1 to the minimum and Y2 to the maximum. Select the resulting countries shown in the graph.

Q: What countries exhibit ODA received per capita in the range between Y1 and Y2 sometime during the years between X1 and X2?

S: Using the Y-axis on the line graph, type in Y1 to the minimum and Y2 to the maximum. Using the X-axis on the line graph, type X1 into the minimum and X2 into the maximum. Select the resulting countries shown in the graph.

Q: What is the exact number of ODA received per capita in year X for country Y?

S: Select country C from the world map. On the highlighted line on the line graph, hover the mouse over the line. Observe the pop-up window. Slide left and right over the line until year Y is found, and observe the ODA received attribute in the window.

Q: What are basic statistics about country C?

S: Hover the mouse over country C on the world map and observe the pop-up window with the min, max, average, and standard deviation of ODA received over time.

Rationale:

This visualization provides a nice overview first and allowing for zooming and filtering for details on demand. Since all countries are highlighted to begin with, the user has a general overview of the kind of data seen, and perhaps a few clusters or trends. The user may then “Select none”, and selectively add countries one at a time they’d like to examine. Perhaps they highlight 4 adjacent countries on the world map and see on the line graph that they have similar spikes around the same times. Or they may select 5 countries on the line graph that seem to be clustered to see which geographic region they reside from. This is very useful filtering and zooming. Details on demand happens when a user hovers their mouse over data points on the line graph or over a country on the world map, as explained in the 3rd paragraph to my description of the design.

The visual encodings from this dataset is a simple mapping to a line graph where all countries have lines. There is no difference in size, scale, or other visual elements between the countries, only the raw data. What has been added is a random but unique color assignment to each country. It would be too time consuming to issue a specific color to each country, and since there are more countries than colors we can remember at once (around 19), it would be useless to use color to classify the countries. However, this is still useful for allowing our eyes to focus on one color/country at a time and see it’s corresponding spot in the other view (world map or line graph).

Color is used once more in this graphic. That is to show what is selected. Any country that is not selected has its transparency set to around 90%, so that it is nearly blended with the background, but still available for the user to make a comparison the visible transparent bars around the selected lines.

No sorting is used, but preprocessing must be done in order to create the min, max, average, and standard deviation for each country, visible when hovering over the country on the world map.