Visual Analytics Assignment 4

- Akash Malhotra

Part 1:

The above stacked area chart is meant to show the overall trend of the greenhouse gas emission over the years. Area encoding was chosen because it is an effective channel to visualize and compare the quantitative fields. The area chart is combination of line chart and area, and thus it is also useful for seeing trends over time.
Filter transform (as a drop-down menu) is used to choose the desired greenhouse gas, whose variations we want to visualize over time. Countries can be identified by the user using a tooltip. Legend for country is not included, as there are ~30 countries in the dataset, and that would exceed the reasonable number of colors that should be used in a visualization as a sole identifier. Also, big legends are not aesthetic.

This Visualization not only shows the quantity of a particular (user selected) greenhouse gas emitted, but also it is effective in inter-country comparisons and to see the over-all trend. For example, we see that for most gases the quantity of emission is has decreased over time, except for HFCs. We also notice that there are some peaks, for example in 2007 due to emissions by China. This makes sense, as it is right before the 2008 Financial Crises.

The next bar chart uses length encoding to encapsulate Total Emission value per capita, as length encoding is often the most effective channel to get comprehend quantitative fields. It gives us a good idea of quantity of greenhouse gases emitted by different countries. Due to unavailability of data of specific gases, we can only visualize the total greenhouse emission in this chart. Per-capita chart was chosen over total emission by a country as the latter can be misleading, as it doesn't encapsulate the size of the country. For example, USA is the leading overall pollutant, however, Australia is the leading pollutant per capita. Transform filter for year was applied. With the help of the slider, we can see the variation in emission of total greenhouse gas emission over time. Since the values are sorted, we can also see, how rankings change over time.

Next we see specifically, the total emission of each pollutant by each country using a heatmap. It uses position encoding to identify a country and pollutant and color to identify quantity. To monitor the change over time, it is a very effective tool. Thus, again, using filter transformation we use slider to see change in the heatmap over the years. We use log scale to clearly see the difference in different tones of colors, which also "normalizes" the quantities of different gases. Overall, we see that the heat map becomes less yellow from 1990 to 2018. But also, we see that some countries like Russia and Turkey have started to emit more Nitrogen trifluoride and HFC respectively.

Part 2:

We use bar-chart for plotting emission by top countries. Transformation window and transform filter was chosen to select these 10 countries automatically and they changed over time. The slider filters the year. From the first sight, it looks like there is not much difference in the emissions of top 10 countries, and that it doesn't change over time.

This chart is misleading because

1. It is ambiguous. y axis is total emission of greenhouse gases. It is the sum of greenhouse gases emitted by each country in that year. Different gases are harmful in different quantities. There is no normalization between gases.
2. The y axis is on the log scale. So, it appears that different countries emit almost the same amount of greenhouse gases. Thus, ignoring the fact that United States emit 10 times more than United Kingdom. To suppress the difference in emissions:

    - The y axis scale starts from 1 (and not from the higher baseline).
    - The bars are not sorted
    - Countries are not color-coded

3. Also, the emissions are not normalized by per capita, which completely ignores the size of the country.
4. The title subtly suggests that being top 10 is positive

Next, we use transparent area charts to visualize the emission of different gases over time. Since Altair doesn't have support for pie chart yet, area charts are the next best thing to compare quantities in a time series setting. At first sight, due to large quantity of CO2 being released, the viewer thinks that CO2 is the most significant pollutant and compared to it rest of the pollutants are insignificant. The above chart is misleading:

1. The title introduces the bias. It would indicate that more major the pollutant is, the more area it will have on chart
2. Plots the literal number of gas emissions, without taking into account the effects of different gases in different quantities. For example, Hydrofluorocarbons deplete ozone layer and is effective in even small quantities, and CO2 can be tolerated in much larger quantities relative to HFC. But this chart ignores this fact and draws the attention of the viewer to CO2, more than the other gases which are potentially much more harmful.