Climate Dashboard

Adrian Andersson

Index Terms—Climate, Precipitation, Carbon dioxide, Temperature, Choropleth Map, Dashboard, Prediction, Visualization.

1 Introduction

Computers are becoming more and more powerful and storing data is cheaper than ever before. This has paved the way for information visualization as subject matter. Information visualization is a fast growing field of study where often large data is treated and processed in order to visualize patterns or trends for human interpretation. We humans are generally bad at interpreting numbers but good at interpreting images. Information visualization together with data mining methods can be used to visualize hidden patterns and complex relations that are otherwise hard to detect. An important feature of information visualizations is the ability to interact with the data for exploratory purposes.

This paper presents the making of a web-based dashboard (multiple grouped visualizations) for climate data for the course TNM048 at Linköping University. To implement the dashboard the JavaScript library *data-driven-documents*, more commonly referred to as *d3*, was used.

2 Background and Related Work

Climate impact has become a dominant topic in politics and lifestyle since the late 20^{th} century and more recently with the rise of the movement *Fridays for Future* led by climate activist Greta Thunberg.

The main purpose of the dashboard is to visualize changes in climate over time as well as attempting to showcase correlations between various data sets. The available data sets contain measures of temperature, precipitation, natural disasters, and carbon dioxide. Carbon dioxide is one of so called greenhouse gases which contribute to global warming but there are several well known others, for example Methane, and Ozone. Data and visualizations for these other greenhouse gases are not available in the dashboard as the workload would exceed that of the scope of this course. If however they were to be included, the same visualization technique would be used on them as with the carbon dioxide.

A similar dashboard provided by *Climate.gov* has been used as a guideline. Although their dashboard is more extensive it lacks several useful features.

3 Data

The data used to in the visualization come from various online sources. The data is static and therefore the most recent dates are missing. Because different data sets are used the interval of dates of which the data is available varies. The data sets used are presented in table 1.

Table 1. The data sets used in the dashboard.

Link to data set	Available dates	Periodicity
Carbon dioxide	1958-2020	Monthly
Temperature	1743-2013	Monthly
Precipitation	1901-2015	Yearly
Natural disasters	1900-2015	Yearly

The data set of temperatures has several missing values for most countries at the early time period. The precipitation and carbon dioxide data sets are complete and has no missing values within their respective time frame while the natural disasters data set has some missing values. All data sets exists as *csv* files. The natural disasters data set contains several categorized disasters such as flood, wildfire, etc. All data sets contains global measures with the exception of the temperature data set which contain measure for each individual country.

4 Method

The method describes how the data presented in section 3 is visualized and how it is motivated.

4.1 Map

The temperatures are displayed on a choropleth map over the entire earth. The displayed temperature is the average over a chosen period of time, more on this in section 4.3. The countries are filled by completely opaque colors. This means that roads and landscape are not visible on the map. Furthermore there is a layer of labels containing the names of continents, countries, and cities, depending on the level of zoom.

The color scheme used is diverging with blue in one end of the spectrum and red on the other, with a yellowish tone in between. The diverging color scale is chosen to represent a deviation of a neutral midpoint [3], in this case 0°C. The color blue represents a colder temperature while the red color represents a warmer temperature. The yellow color is considered to be neutral. The colors are divided into eleven classes and a twelfth class is added to display missing data with the color black. A legend is added to the map to let the user know what the colors represent. There is also the ability to hover over a country with the cursor to display a tooltip with the exact average temperature.

The colors are divided into classes and then assigned to each country. By using classes similar countries are clustered with the same color by temperature and geospatial information. An alternative to classes would be to linearly interpolate colors using the same color scheme. This could potentially make minor changes to a country's temperature more visible. The main issue however is that the temperatures ranges from -20°C to 30°C. The span is too large. An alternative solution to this is to display deviations from an average temperature for each country. This would result in a smaller span of temperatures and therefore increase the visibility of minor changes as the number of classes remains the same.

4.2 Charts

The rest of the visualizations on the dashboard are different kinds of charts. When the horizontal axis consists of dates in a strictly increasing fashion it is common to use line charts [3]. An exception is made for the precipitation data is it represents deviations from an average. All charts can be seen in figure 1.

The precipitation is visualized as a bar chart. As the precipitation data is a deviation or anomaly from the average precipitation the bar chart contains negative values. This is implemented as bars originating from the center of the vertical axis. When hovering a bar a tooltip appears with the exact date and value.

Both the carbon dioxide levels and the quantities of natural disasters are represented as line charts. By hovering the charts a tooltip appears with additional information on the exact date, based on where the cursor is located. The tooltip displays date and carbon dioxide levels or

• *Adrian Andersson, student at Linköping University, Sweden. -mail: adran117@student.liu.se.*

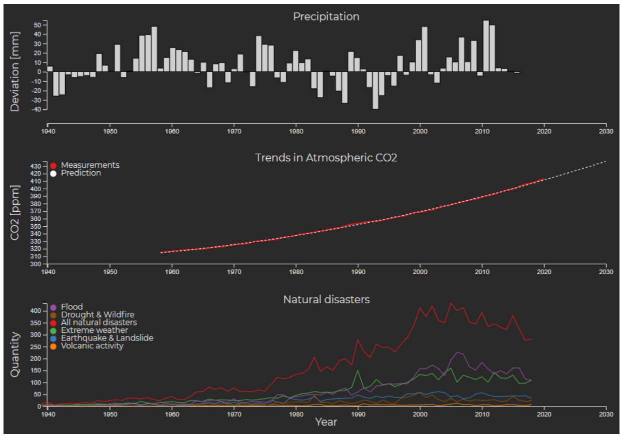

Fig. 1. The charts of the dashboard.

all of the quantities for each category of natural disasters depending on which chart is hovered. Both charts contain a legend with color and title of each line in the chart. Having multiple lines in the same line chart however presents a problem as it gets messy. To combat this the legend is removed from the natural disasters chart when the cursor hovers the chart, this creates space for the tooltip and the two will not overlap each other. Furthermore the disasters are arranged in decreasing order inside the tooltip and uses the same colors as in the legend and chart. The tooltip for the natural disasters chart is visible in figure 2.

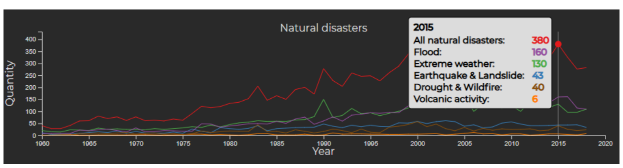

Fig. 2. A tooltip for the natural disasters chart.

The color schemes for all charts are qualitative, meaning that the colors are independent of each other and do not represent differences in magnitude [1].

4.3 Context Bar

To enable interaction with the dashboard a context bar of all existing dates are added at the bottom of the map and charts. The brush in the context bar allows an interval of dates to be chosen and the charts and maps are updated according to the chosen interval. The initial dates are chosen to be between year 1960 to 2030. This time span is deemed as the most interesting period across all charts and map. Moreover the brush can only be moved and stretched in one dimension.

4.4 Preprocessing and Merging of Data

In order to use the data sets some preprocessing was needed. The precipitation is measured in inches as it is provided by an American source. The inches were converted to millimeters as it better fits our purpose with the dashboard.

Furthermore the data set of temperatures needed to be merged with a *GeoJSON* object to use with the map. A GeoJSON object is an encoding format of various geographical data structures based on JavaScript's *JSON* object [2]. The chosen GeoJSON object has, along with the names of countries, a feature called *MultiPolygon* which is the coordinates of each country's borders.

The merge maps the dates and temperatures of the data set to the countries available in the GeoJSON object. This presented a problem as the names of countries did not always match. For example the data set contained outdated names as "Burma" while the GeoJSON counterpart contains the name "Myanmar". To combat this the data set names were replaced to the updated version available in the GeoJSON object before being merged with a small *Python* script. The *Python* script produced a new file with the *.geojson* extension containing the merged data.

The categories of natural disasters has been revised and some categories were removed or combined with other categories. This decision is based on the content of those categories (very little to no useful information) or that it could naturally be combined with another category. Such a combination is the two categories "extreme weather" and "extreme temperature" which can be grouped under the label "extreme weather" as extreme temperature is a subcategory of extreme weather.

4.5 Data Mining

To present some information to the user that is otherwise not available through the data alone a prediction was made. The available data for carbon dioxide is used with a polynomial regression technique to fit a curve to the data. To minimize the error the least squared error method is used. The curve stretches to year 2030 in an attempt to display the general behaviour of the graph for the coming ten years. The prediction curve is plotted in the same chart as the carbon dioxide levels with a dashed line. The decision to use a polynomial regression curve was based on the appearance of the carbon dioxide graph. The resulting regression curve is visible in the middle chart of figure 1.

5 Implementation

This section accounts for the implementation details of the dashboard as well as limitations. As stated in section the JavaScript library *d3* has been used to create the various visualizations. The map is created with the open source JavaScript library *leaflet*. To create the prediction model the JavaScript module *regression-js* is used.

As the context bar was implemented there was a considerable delay when moving the brush. Unfortunately the computations of the new average temperatures for each country was too many when a new time frame was selected. By instead updating the charts and map only when the brush is released the delays are avoided.

6 Results

The results presented in this section has been improved after the evaluation stage and is thus a slightly altered version. The dashboard is presented in figure 3 below.

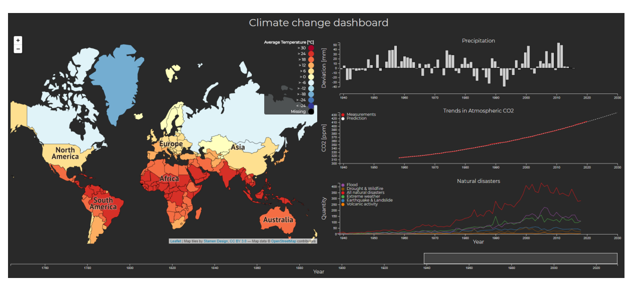

Fig. 3. The resulting dashboard.

The map is considered to be equally important as the charts combined and is thus the same size as the three charts combined. At the bottom of the page the context bar is located. The entire dashboard is visible at the same time with a standard 16:9 ratio monitor. Tooltips and legends are available for every visualization and when brushing over the context bar the entire dashboard updates according to the chosen time frame.

7 Evaluation

In order to evaluate the dashboard a user test was set up. Evaluating information visualization techniques are recognized as an important step during development. The test consisted of three parts where the user would interact freely, perform tasks and answer questions, and lastly be interviewed. The goal of the evaluation was to gain insight to what was missing and what could have been implemented differently. The participants were mostly fellow students with the same level of education and in the same field. At the time of evaluation the prediction for the carbon dioxide levels used a linear regression model, this has since been changed to a polynomial regression model.

The tasks consisted of using the dashboard to answer a given question. A typical question could be "Would you say the average temperature during the 21^{th} century is higher than during the period 1880-1900?" or "What was the precipitation anomaly for the year 2001?". All of the participants had the exact same answers with the exception of one student with impaired sight. The poor sight led to some errors as the text was too small to read.

During the short interview afterwards the participants were asked to give some feedback and some suggestions to what might improve the dashboard. Some participants thought the sizes or layout of the charts and map should be altered while others disagreed. The suggestions of improvement given by the participants were all valid and well motivated. Furthermore all of the participants would have preferred if the brush updated the charts and map as it moved instead of having to release it. Several participants had difficulties choosing a specific time frame for the brush as there is no indication, except the axis ticks, what time period is chosen. Overall the participants thought that the dashboard was useful and liked the appearance of it.

8 Conclusions and Future Work

The following section is dedicated to conclusions regarding the progress of the dashboard and the future work one might expect after the evaluation stage as well as important features that should be included but are not.

The evaluation stage presented much needed feedback and suggestions to improve the dashboard. The temperatures of the map should be changed to display a deviation from an average temperature instead of the average temperature. This would in turn increase the visibility of minor changes as the span of possible deviated temperatures is smaller. Another feature that would improve the usability of the dashboard is the ability to filter the data by country by choosing one or several countries in the map. This would update the charts to display the data for those specific countries instead of the global data. However the chosen data sets are global with no country specific data available which means that the data sets would have to be replaced. Other greenhouse gases as previously mentioned in section 2 should also be visualized in the dashboard as they contribute to climate change and would make for a more complete dashboard.

With respect to people of poor eyesight there should also exist a feature to enlarge the charts and maps. As the evaluation revealed some people had trouble reading the smaller texts and thus misinterpreted the data. Future work might also include adding a snapping function to the brush and to display the chosen time frame.

The evaluation resulted in some much needed feedback and gave some interesting suggestions on what the dashboard was missing. Moreover it would have been desirable if the participants of the user study had a more varied background and different levels of cunning in the field of information visualization as the dashboard is meant to be used by anyone.

Using this dashboard one can conclude that there is indeed an ongoing climate change. The carbon dioxide levels and temperatures, as well as reported number of disasters are rising.

References

[1] C. Brewer and M. Harrower. Colorbrewer.org: An online tool for selecting colour schemes for maps. "https://pdfs.semanticscholar.org/8c9a/f5bd12b36e450ba564f644008b871dba5cdf.pdf", 2013.

[2] H. Butler, M. Daly, A. Doyle, S. Gillies, S. Hagen, and T. Schaub. The geojson format. "https://tools.ietf.org/pdf/rfc7946.pdf", 2016.

[3] C. O. Wilke. Fundamentals of data visualization. "https://serialmentor.com/dataviz/".