

Information Visualisation *Channelling Hans!*

Michael McNulty
Romain Ducarrouge

Task 1

Objective

The objective of the visualisation is to provide the end user with a clean representation of how countries key metrics have evolved in the past 115 years. The visualisation presents the following information to the user:

- **Countries**, represented as a circle
- **Region**, encoded to a specific colour for each circle.
- **Life Expectancy**, encoded to the y Axis of the graph.
- **GDP** per capita, encoded to the x Axis for our graph.
- **Population**, represented as the area of the circles.
- **Years** are also represented on the visualisation in the background.

The user can also see the evolution of country's situation throughout the years by selecting the country simply by hovering over the data circle with their mouse. When selecting a specific country, additional information about the country is provided such as the flag, value for GDP and Life Expectancy.

Considerations

With regards to the dataset, we initially identified several potential issues with the appearance and disappearances of countries throughout the year. This means that during the visualisation some data points would appear later (e.g. Taiwan in 1950). We addressed this using *D3 transition()* to smoothly introduce and fade out countries if required.

While getting to know the data we also identified the minimum and maximum values for the Life Expectancy, GDP per Capita and Population in order to understand the range of these values to generate the proper scale for axis and circumference for our bubbles.

We also made sure to include the code populating the Year in the background of the SVG element before populating the bubbles to place the text in the background.

Regarding the display of the selecting country (bubble), the approach chosen was to reduce the opacity of all the other bubbles to 'highlight' the selected element on mouseover events, then reverting the process when the mouseover event is over.

Another element that required some attention and discussion was regarding the size and colour of the bubble to have a palette of colour that would be aesthetically pleasing while not being too aggressive. With the same objective in mind we also analysed various combinations for the transition timer in order to modify the speed of changes between years. We determined that the current settings (and combinations of *transition()*, *ease()* and *duration()*) allows to clearly identify the situation of a country at any given year but also visualise the transition to the next year.

Challenges

Identifying the change in data format for the year's values.

The years start in 1900 and increase by 10 until 1950 and then we get yearly data.

Choosing between using the version 3 or 4 of D3.

We decided to use the version 3 based on the examples viewed in the practical sessions and started implementing our code for the project.

As we also decided to implement a display of the country flag when selected by the user, we decided to implement a third-party API call to populate the flag. Unfortunately, our data does include the new country South Sudan (code: SS) which is not recognized by the API. This was handled by having an SVG copy of the South Sudan flag in our folder and populating the flag from the local folder when the country was selected rather than querying the API for the .png image.

Designing a clear legend.

Here we had to use the same colour scheme as for the bubbles. This meant we had to find a font colour that would be visible on all background colours for the rectangles in the legend. After some trials and errors using the following attributes: `font-size`, `fill`, `paint-order`, `stroke`, `stroke-width`, we realized that using these only would not allow us to have a clear legend. To fix this, we decided to use a black colour for the text and modify the opacity of the background for the one purple rectangle to deliver a clearer visual.

Additional Features

In our visualisation, we decided to implement the following additional functionalities:

The **change of animation speed** which we believe is valuable as some people may want a quick look at the whole picture while some people may prefer a slower speed. We believe the change of speed offers the user an alternative the repetitive playing and pausing of the animations.

We have also implemented an additional **country id card** which allows to display information regarding the country selected by the user. The information presented included the country name, flag, year, population and GDP.

United Kingdom, 1900

Population: 37995759

Income: 8013



Task 2

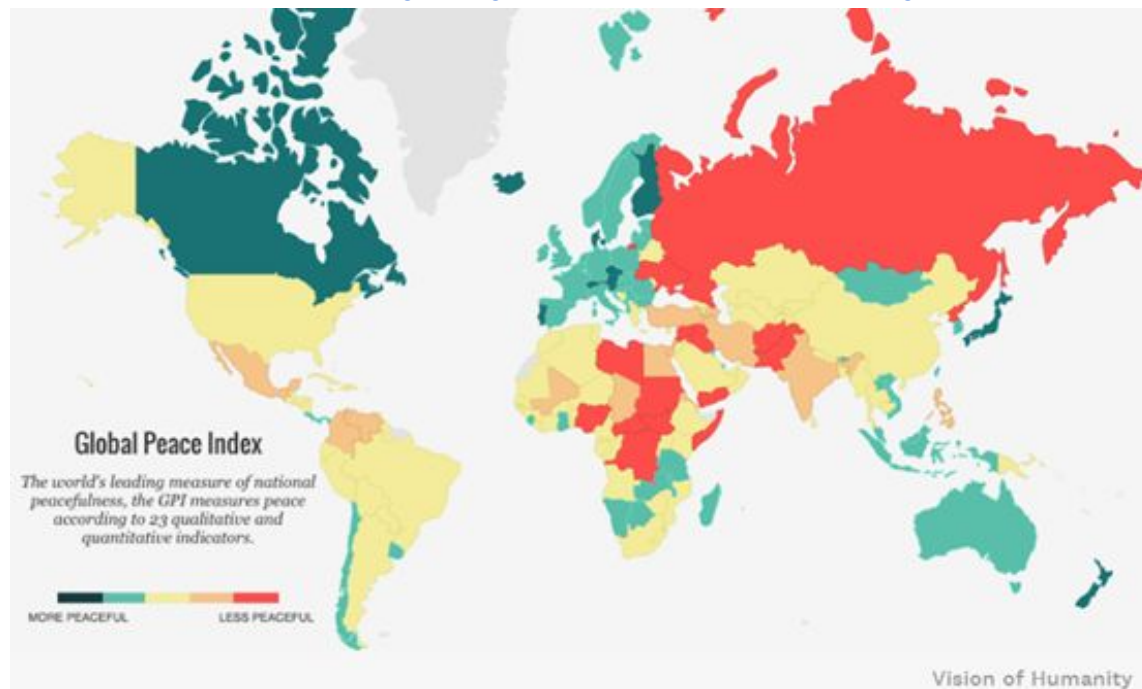
1 What Is the Purpose of Your Visualisation? (150 words)

The objective for this visualisation is to extend the insight gained from the provided dataset. First, we initially decided to present the data grouped by regions as well as by government types. This allows to get a more granular view of the initial dataset for the user looking to investigate the data at a region level or to study the data based on the government type. On this first dashboard we can see the Life Expectancy shown on a world choropleth map. We also display two vertical bar charts that show the LifeExp, GDP and Population for the various regions and governments.

We also wanted to tie in the original data with an external one in order to identify to which extend the country's population were affected by conflicts and war. To show these we decided to implement two choropleth maps that first highlight the countries involved in conflicts and the magnitude of the conflicts on the map. It also displays the parties involved in the wars through the tooltip.

2 What Similar Visualisations Exist? (150 words)

First, <http://all-that-is-interesting.com/global-peace-index-2016-rankings>:



This visualisation is similar to ours although it also accounts for non-conflict related event such as terrorism attacks. It is interesting to see that while it represents a continuous Index, it was encoded with hue rather than gradient. It is also encoded on the basis that green is commonly associated with positive and red with negative elements.

Second, [http://www.cfr.org/global/global-conflict-tracker/p32137#/:](http://www.cfr.org/global/global-conflict-tracker/p32137#/)



This map is a beautiful, clean, live representation of current conflicts and events in the world. One thing worth noting is that the map reflects events that impact US interest and as such may not include all currently ongoing conflicts in the world.

Both sources decided to use geo-representation to show the location of the conflicts. One decided to make a choropleth map which enables the viewer to see which country it is relating to. The second one used a symbol map (usually better than choropleth maps for showing raw data totals) and uses a fixed sized circle to show the precise location of the event rather than encoding the importance of the event to its size, this makes sense due to the fact that a lot of conflicts happen in same areas. Using circle size here would just have made the visualisation complicated to interpret for the user.

3 Why Is Your Visualisation A Good Solution? (150 words)

Learning from what other sources did well on their visualisation we believe that our solution is good as it enables access to clear and relevant information showed directly thanks to appropriate encoding as well as additional information available through the Tooltip functionality. We have generated a choropleth map for the following reasons: we have the dataset encoded per country rather than battle location, we do not have a count of the victims for each conflict, we only have access to a binary feature representing less than 1000 deaths or more. Had we access to more specific data about conflict casualties, we could have represented this attribute using colour density, rather than a categorical encoding such as red/yellow. This would have

allowed for a more informative visualisation, with the possibility of comparing two countries that were both classified as *high* intensity.

Putting these datasets together also provides the possibility to see the effect of the war on the country's population.

4 What Data Manipulation Was Required To Create Your Solution? (150 words – it is acceptable not to have required any data manipulation)

In order to integrate the new dataset we have had to make a lot of modifications to the location of the conflicts. They were often using overly descriptive terms such as “Madagascar (Malagasy)” which lead to issues with Tableau’s Geographic Role. To address this we decided to use the *Replace All* feature present in Excel to normalize all the names of the conflict locations. We also used the built-in functionality of Tableau that highlights unknown value and fixed it to get the maps working correctly.

We have also had to remove the years that were only part of one dataset, for example in the GapMinder dataset the years start in 1900 while the UCDP dataset only starts in 1946.

We also attempted to merge the two datasets using the Tableau feature *union* but that would not allow us to see the countries at war in a certain year and see the impact on the population between that year and the following one in order to visualise its impact.

There were multiple categorical values for the government types which were overcrowding the visuals. To clean this we decided to remove categories that were not deemed significant such as the Netherlands state which disappeared after a few years and was leaving an empty column in our chart. By having our columns closer together as a result of this, a comparison based on bar height was easier to make.

5 What d3 Resources Did You Use To Create Your Visualisation? (150 words)

Initially we looked into using geomap packages from d3 in order to generate the choropleth map. We ran into several issues when trying to implement the code using his library and thought about using Google Charts (file “*task2_initial_idea_GoogleCharts*” included in the .zip folder) but we realized that for this project we should use D3 or Tableau.

We decided to use Tableau to generate two dashboards. The first one only displays additional information about the original dataset. The second one focuses on the world conflict dataset that we obtained from <https://www.prio.org/Data/Armed-Conflict/>.

Tableau is a powerful visualisation tool which allows extremely fast turn around and generation of clean and interpretable visual representation of datasets.

Considerations

Considering ordering the bars of the bar chart by decreasing order to make it easier to compare the values across all the government types. Unfortunately, because we also animate based on the year the values changes and so would the order. We believe it would be less clear to the user if the order of the categorical values changed on our axis.

We also researched the possibilities to implement a break in our y axis in order to make the visualisation better by reducing the disparity and not having very small and very big bars on the same graph. This would make a clearer graph. Unfortunately, this is a feature that is not implemented in Tableau. The only workaround that was suggested online is to superimpose two bar charts where the bottom one contains the x axis details and the top one does not. Hence this would mimic a break in the y axis but would not be a clean way of doing so.

We also investigated other graph types that would allow a clearer view on the data such as bubble graphs but we soon realized that if we encoded the population to the size of the bubble this would make it even harder to read as the y axis value is encoded to the centre of the circle.

We also removed categorical data features relating to government types which were not significant and only led to less clear visualisation.