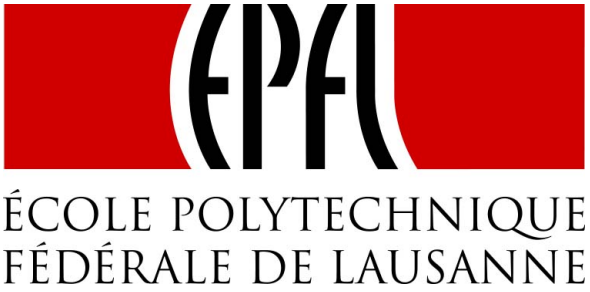


ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Data visualisation

Process book

Charles GALLAY
Thomas BATSCHELET
Valentine SANTARELLI

December 22, 2017

Charles GALLAY
Thomas BATSCHELET
Valentine SANTARELLI

1 Introduction

Environmental changes have certainly been a major topic for the last decades but we could argue that due to the massive amount of information that is regularly given by news, it can be hard to filter out the actual relevant facts that influence climate evolution.

A lot of raw data on the matter are available but hard to interpret to get the full picture. That's the seed of our motivation behind this project, getting a full picture of environmental changes happening for last decades and awaken (the force) awareness. Therefore, this project implements a visualization of this raw data to take a better grasp at the different aspects of the actual changes.

The resulting tool will be an intuitive visualization that present actual knowledge by giving an objective insight about environmental changes without arguing global warming. This way, the curious visitors can develop their own opinion on environmental changes based on objective informations.

In more depth, we gathered temperature, CO2, and ecological footprint data at country level for the last five decades to show plausible correlations.

Questions

Through our visualization, we asked ourselves the following questions.

- Does the ecological footprint of a country directly impact its climate changes ?
- Can we visualize the correlation between CO2 emissions and temperature changes at country level ?

2 Datasets

The several datasets used come from the Worldbank group website and Kaggle;
Average CO2 emissions [kT] per country and per year from 1960 to 2013

Population per country from 1960 to 2016

Average Temperature [C°] per country and per year from 1743 to 2013

Ecological footprint [gha] per country and per year from 1961 to 2013

The temperature dataset gave the monthly temperature average of each country. Since we decided to focus on the years we aggregated the data accordingly. To make use of the data, we decided to merge all the datasets into a csv file depicting the country ISO codes and all corresponding data through all the years. To do so, we used pandas library and had to clean the data as many values were missing. We decided to keep only values with years starting 1960 to 2013 as they were the most reliable and complete. We also added a feature representing the difference of temperature for each country with respect to the year 1950. This will ease our will to show the climate change through the recent years. For the footprint datasets we needed to map the countries name and ISO code. Unfortunately, we realized that there are missing data for certain countries in the ecological footprint dataset.

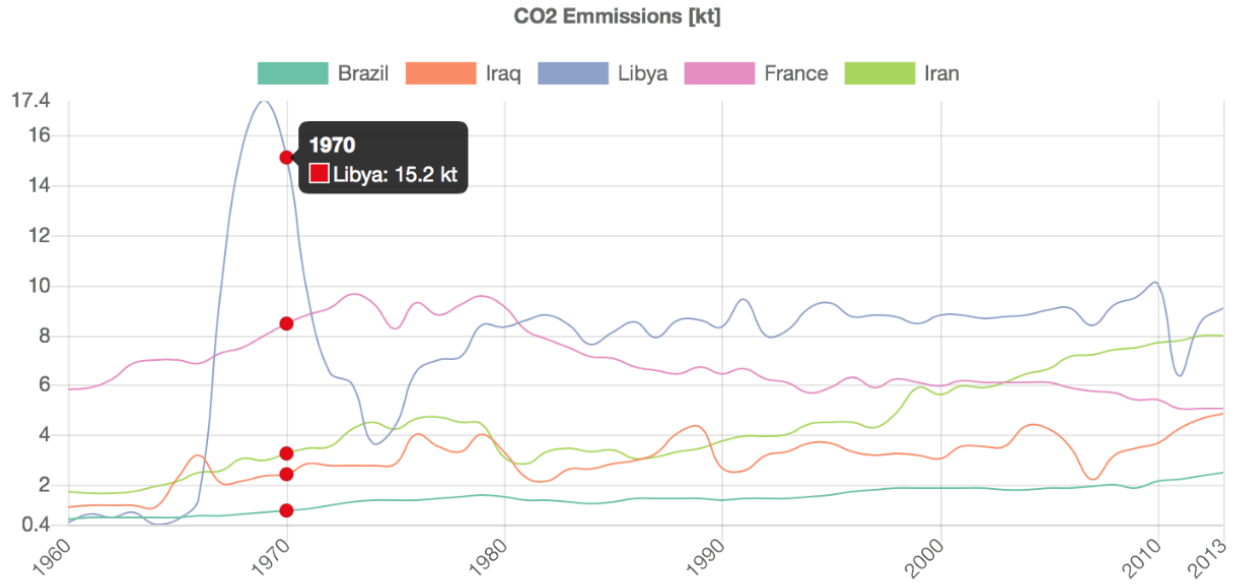
Furthermore, we aggregated all countries to compute the mean values in order to show the overall evolution of the earth.

3 Exploratory data analysis and Design

Evolution

A first insight to present is the evolution of our variables (temperature, CO2, EF) through time in order to clearly highlight the excepted changes in temperature and other data over the last decades. To get an idea on the entire world's environmental changes, we started with an average of all countries, then we refined the visualization to the country level.
Our choice for the visualization of this insight is a ChartJS line chart with an adaptable Yaxis scale for each variable to better highlight the variations among each countries. To select the country whose data the user want to visualize, we introduce a world map where the user can click on a desired country and details will then be depicted. We can select many countries at once.

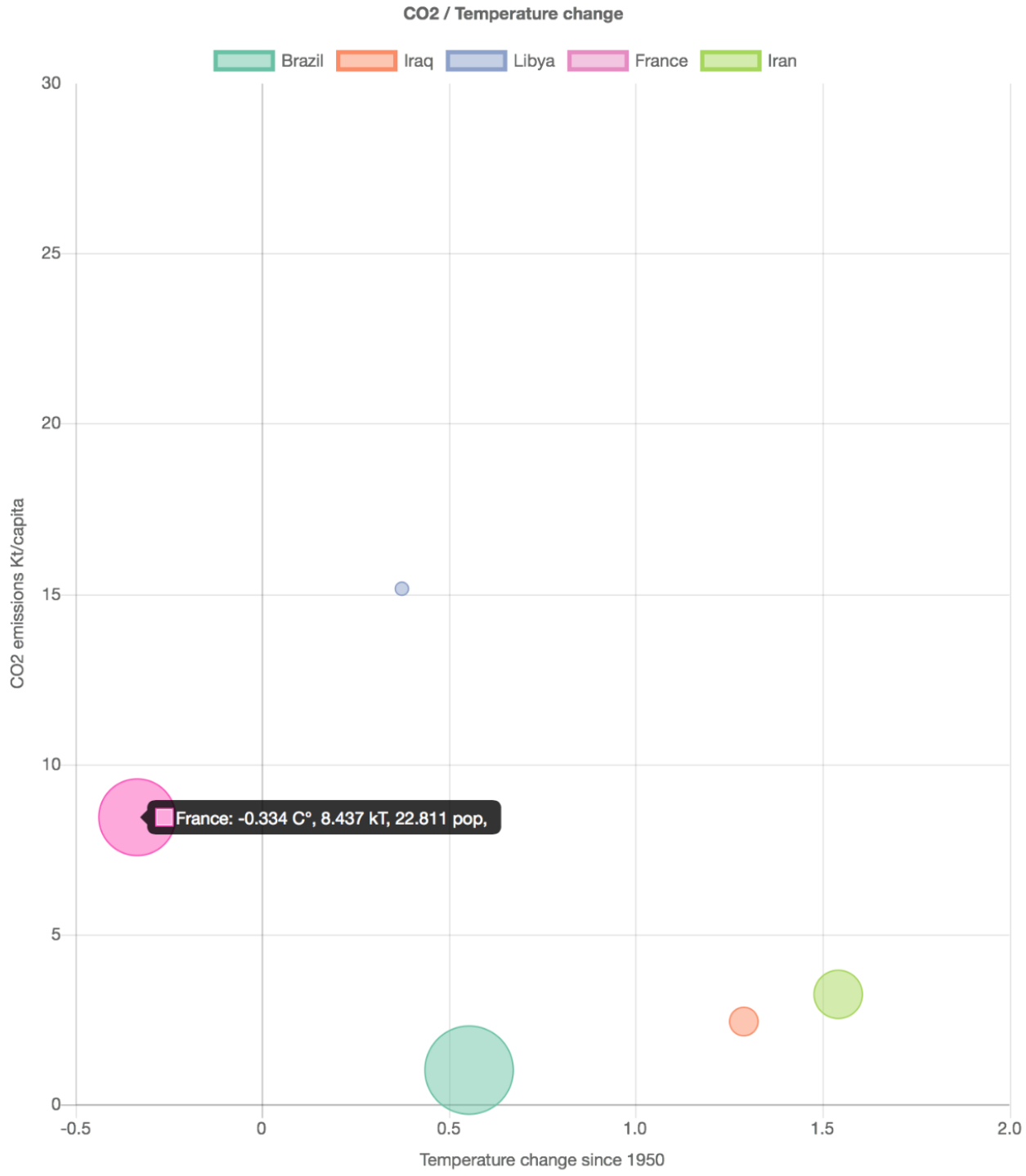
Figure 1: Line Chart and tooltip



Correlation

A second insight is to show the correlation between the temperature, CO2 and EF.
For this we have two types of visualization:
First, each line chart (temp/CO2/EF) is put on under the other to ease their comparison and highlight possible correlations in their evolution.
Second, we implemented a bubble chart of the CO2 emissions in function of the temperature variance with respect to the data collected from year 1950. In addition, the radius of each bubble represent the population of the country. This shows the impact that a bigger country has even if its CO2 emissions are lower.

Figure 2: Bubble chart and tooltip



Comparison

A third insight consist in showing that the global warming doesn't affect every country equally.
For this, we wanted to able able to compare environmental changes between two countries. This was done by simply adding lines/bubbles for each country in the lines/bubble charts. Still in order to get a better grasp of environmental changes at a country level, we included another visualization: a world

choropleth map whose colormap depends on the variable to visualize.

Figure 3: World map

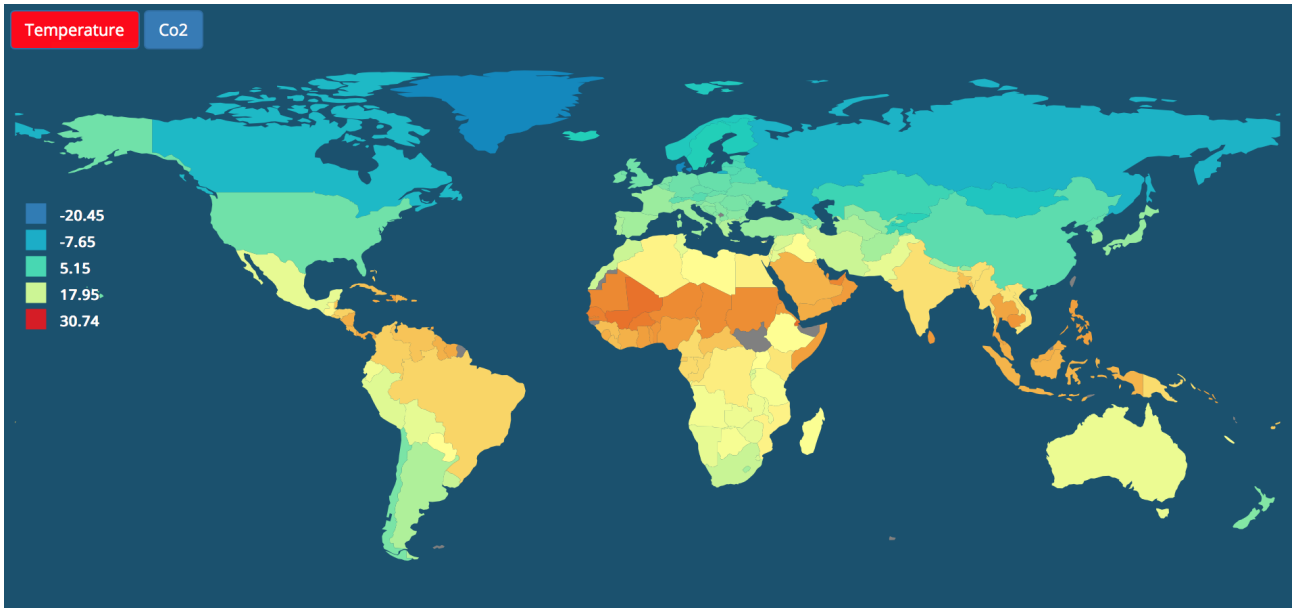
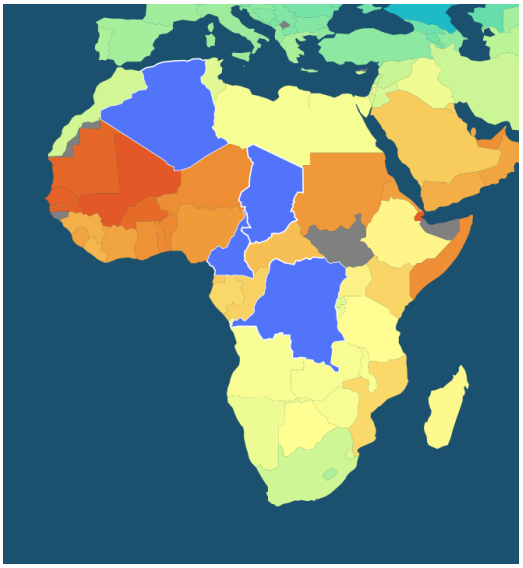


Figure 4: Country selection



Zooming on time

As we have a large time scale, it would be interesting to focus our attention on a specific year range. For example, for the data story we want to show the evolution of CO2 emissions after the Kyoto protocol which is an international treaty aiming to reduce greenhouse gas emissions adopted in 1997.

To enhance details in the time scale, we implemented a time slider, whose handles enable to select a year range. This year range is applied to the x scale of our line charts of the variable's evolution and

thus enables a zoom on the selected period.

But what about getting the data from a particular year? We included a third handle in our time slider: the time selector. This time selector sends the selected year to the map and the bubble chart to display the corresponding data. The time selector is also shown on the line charts by highlighting the corresponding data point in red. By hovering that point a tooltip appears with the country's name, the value of the variable, and the year selected.

Figure 5: Time slider

Awareness and motion

How to catch one's interest ? Even with the data in plain view, how to highlight the urgency of the global warming situation ? How to raise awareness when changes are not visible at our scale on a daily level ? Introducing motion, a third dimension in our 2D visualization, highlights the environmental changes and transport the user in another dimension where time goes by so quickly.
For this we implemented a play/pause button animating the time selector, and thus the country's color in our map, the time selector in our line charts, and the bubble chart.

Figure 6: Play Pause button

4 Visualisation structure

The interface

The user will arrive on the web page with a map of the world and a panel. He/she will be able to choose between following the storytelling or immediately dive in the discovery of the data. SCREENSHOT

The interface is divided into three parts: the world map and the panel and the time handler.

World map

The world map will have different utilities;

- it will be used to display the choropleth of the evolution of temperature's variation, the CO2 emissions, and the Ecological footprint per country by selecting the corresponding button widget.

- it will also be used to select severals countries to be compared on the panel. If no country is selected then the visualization will display statistics about the entire world. If one or more countries are selected the statistics will take into account the data corresponding to this country. The user can zoom in and drag the map to select the desired country that will be highlighted.

Panel

The panel will display different graphs and statistics on temperature/CO2 emissions/EF and its variations using the visualizations described below. It is placed on top of the map, and it is scrollable to look at all the charts at the map at the same time. But it is also can be expendable by clicking on it to have a better look at the chart.

- A line plot of variables evolution in time . The variables are the temperature, CO2 emissions and EF and each of them has its own graph since their scale and unit differ. Each chart represents a variable (grouping), and each line a country (continuity).
- The next graphs will be similar except they will represent the evolution of the variation of the variables in time.
- A bubble chart representing the CO2 emissions in function of the temperature with the size of the bubble representing the population of the country to see its impact on the CO2 emissions.

Time handler

The time handler is composed of a time range/selector slider on the years and a play button. The time handler, at the bottom of the page, has different utilities:

- it can select a range of years to take into account in the displayed statistics in the panel and choropleth map .
- it also comes with a play button what will animate the time selector and update the choropleth map and the bubble chart and the line charts.

5 Deviations

We preferred visualizing the temperature variations instead of the absolute one because it has more sense to look at the gain of Celsius degrees in the last decades. We choose a reference deltaT = 0C° for the year 1950 because the data are more complete starting from this year.

We decided to cancel the visualisation of the ecological foot print using 3D planets, the monthly data visualisation, GDP visualisation because of lack of time.

We wanted to put the radius of the bubble as the ecological footprint but since we have missing data, we opted for the population of the country instead.