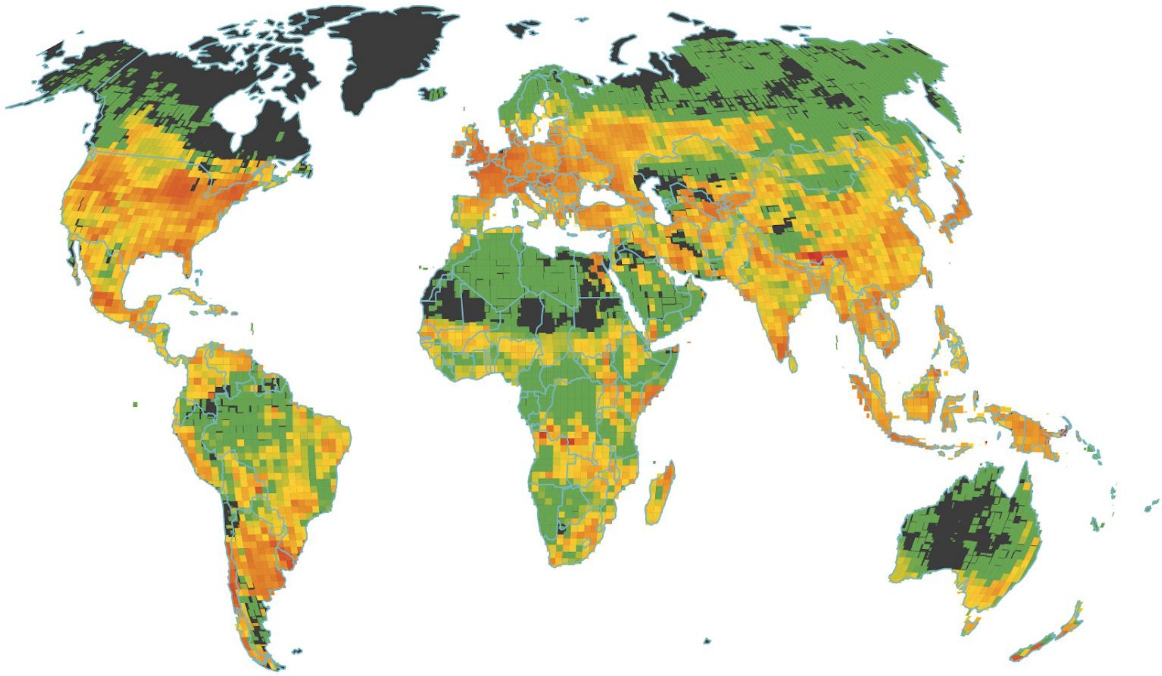
2050 Food Production

A study between different climate scenarios



Process Book

Romain Caristan Kyle Gerard Frank Schmutz

Introduction

This project was proposed in the context of the Data Visualization course (COM-480) during the fall semester of 2018. The goal of this class is to learn good ways to visualize data by using methods of design and tools that were taught to us during the semester like D3.js for example.

This process book intends to show the reader the process and the different ideas that led us throughout the semester to the final visualization, hosted on the website http://foodproduction.me.

Overview

To face the challenge of global warming, scientists try to predict different climate scenarios depending on how the society will behave during the 21st century. These predictions include parameters in a wide range that can go from a political point of view, such as environmental policies led by governments to face the problem, to models of consumptions that can lead to different levels of energy consumption.

Our goal, in this project, is to analyze and provide a clear and user friendly visualization of how the food production will be impacted by the different climate scenarios compared to the one in year 2000.

Motivation

How food production will be impacted in the future is an important problem, as we want to know how a more sustainable economy would impact the different world regions or on the other hand if high-energy consumption models of society would leave some regions behind in terms of food production, leading them to be dependent on other countries. We are also going to study the impacts on wider world regions rather than only countries, such as differences between continents or differences between high and low per capita income countries.

Target Audience

The target audience is not be limited to only one type of person, it can be a wide range of individuals going from researchers that want to have a clear visualization of the predictions and the data without having to look and analyze all the CSV files, to anyone interested in the subject of global warming that want to be informed of how the different climate scenarios can impact our future economy and model of consumption.

Inspiration

We noticed that future food production was one of the topics related to climate change that many people were not aware of. After a bit of research we realised we didn't know a lot on that subject ourselves. Therefore we saw this project as an opportunity to educate people as well as ourselves on the future of food production with the impacts of different climate change scenarios. Also, we wanted to program an interactive map, which is something we'd never done before, so naturally this was a very good project on that front. Lastly, we saw the fact that we would be working with real data collected by Stanford University as a really good opportunity to gain experience on real world problems.

Visualization

We are now going to describe with more precision the data that we used to do our analysis and visualization and explain the process that led us to the final solution.

Dataset

The data was provided to us by the Stanford University Natural Capital Project (https://naturalcapitalproject.stanford.edu). It consists of 5 CSV files representing different socioeconomic scenarios coupled with future possible climate scenarios. These files contain one row per pixel of a map with the associated production of calories, predicted population for years 2000 and 2050 with multiple other variables such as calorie per hectare.

The future climate scenarios we are going to study are described in the data as being Shared Socioeconomic Pathways (SSP), an assumption of a future consumption model based on multiple parameters such as which domains the investments focus on and socioeconomic government policies. These SSPs are associated with a future climate scenario under the form of Representative Concentration Pathways (RCP). They represent the radiative forcing in the future (how global temperature will increase due to greenhouse gases emissions. RCP 2.6 represents a radiative forcing of +2.6W/m2. The higher the forcing value is, the higher the greenhouse gases concentration will be.

Processing of the data

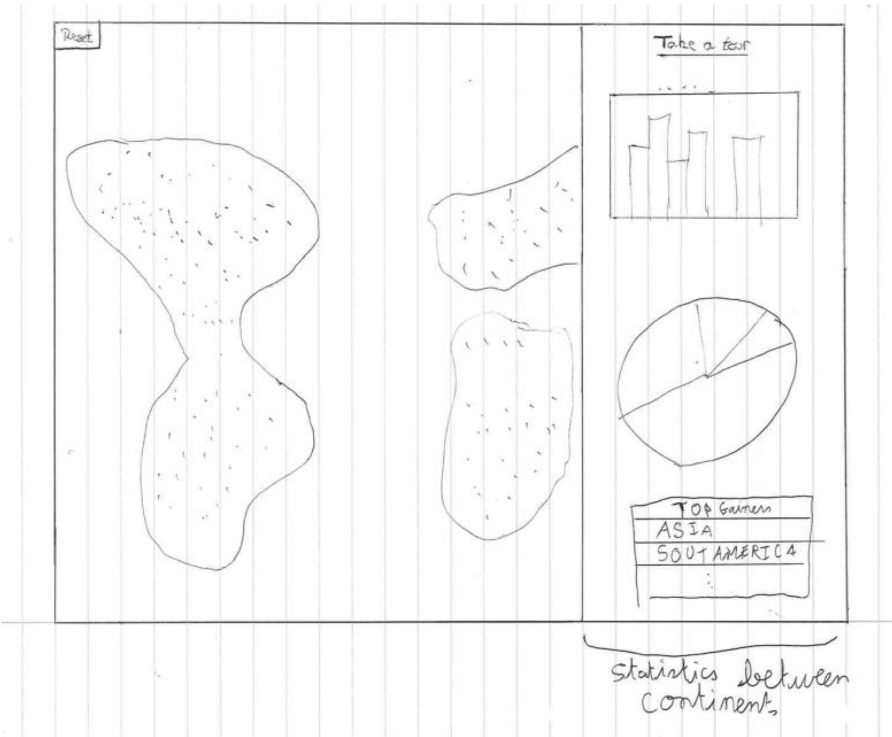
We needed to perform a lot of preprocessing on the data as it only contained informations about the pixels with the associated latitude and longitude.

Our first step was to assign each pixel to a country or a continent as it is a useful information for or visualization (latitude and longitude can not give the user a precise idea of the location of the pixel). The goal was then to aggregate the data by country and continent to extract user friendly information. This step was done using a Jupyter Notebook in Python taking the polygons of all countries and continents and implementing a function that for each pixel outputs us the country and continent it is into.

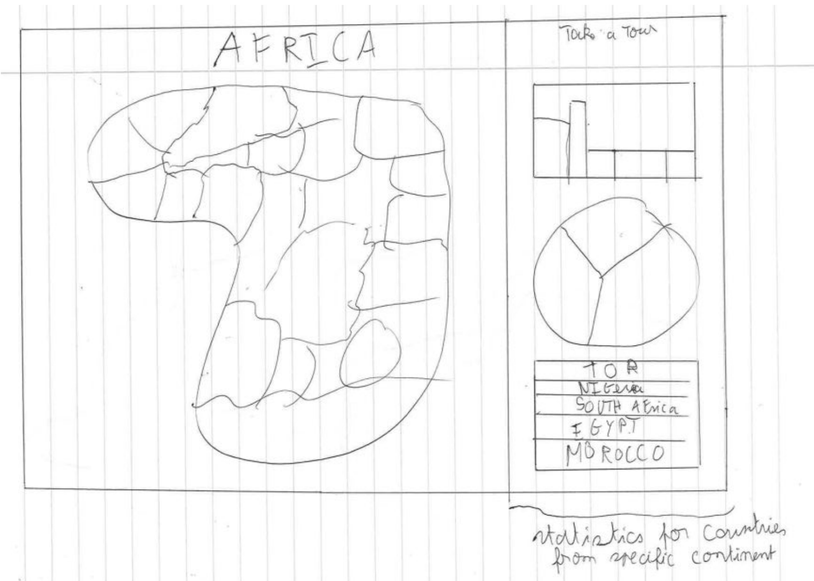
The second step of our preprocessing was to aggregate the data by combining multiple pixels that were close on a geographical point of view. That step was necessary as to reduce the loading time of the data in our website, in fact this parameter is really important as it is the first impression of the user on the website and a homepage taking a long time to load can leave the visitor with a bad feeling on the continuity of their journey through the website or can even cause them to leave it. We only kept columns we needed, i.e., the calories produced in 2050, the population in 2050 and the difference of calories produced in a year between 2000 and 2050. For each SSP scenario we created one csv file for the entire world, one for each continent and one for each country. As the surface area of the regions shrink, i.e., when looking at continents or countries compared to the world, the area size on which we do a single aggregation also diminishes, which leads to a higher precision when looking at a single small region. This is very useful as we can show this more detailed view when we zoom in on the region.

First ideas of design

Our first idea was to do a globe with statistics printed when a country is selected. We quickly realized that this was not a good solution as we could only see one side of the globe so we had no idea of what was the result on the other side of the world. Hence this idea was quickly dropped even if the idea of showing results on a map associated with statistics stayed.

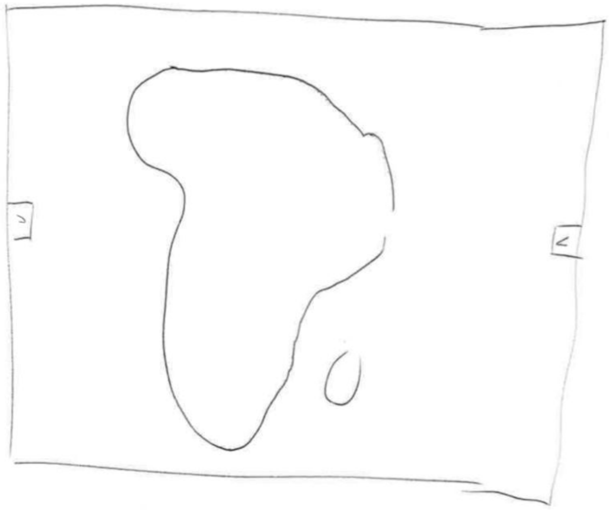


We then moved on to a next idea of a flat map with two sliding panels containing statistics about the selected region using a mercator projection. The TA's correctly pointed out that mercator was not an appropriate projection for the data we were trying to display and suggested using Winkel Triple instead. We also realized that the sliding panels were clunky and that having them able to slide did not help the visual aspects of our project as well as masking a part of the map. So we decided to put a single panel containing all the relevant information on the right side of the screen, which clearly indicate where is shown which type of information (visual on the left and analytics on the right).



The problem with this design is that it could be very difficult to extract meaningful information from the visualization without some assistance on where to look: there simply are to many combinations of regions / climate scenarios to remember. So we decided to take a storytelling approach to give the most meaningful insights to the user from the beginning and only then let them freely look around on the map to answer any further questions they might have.

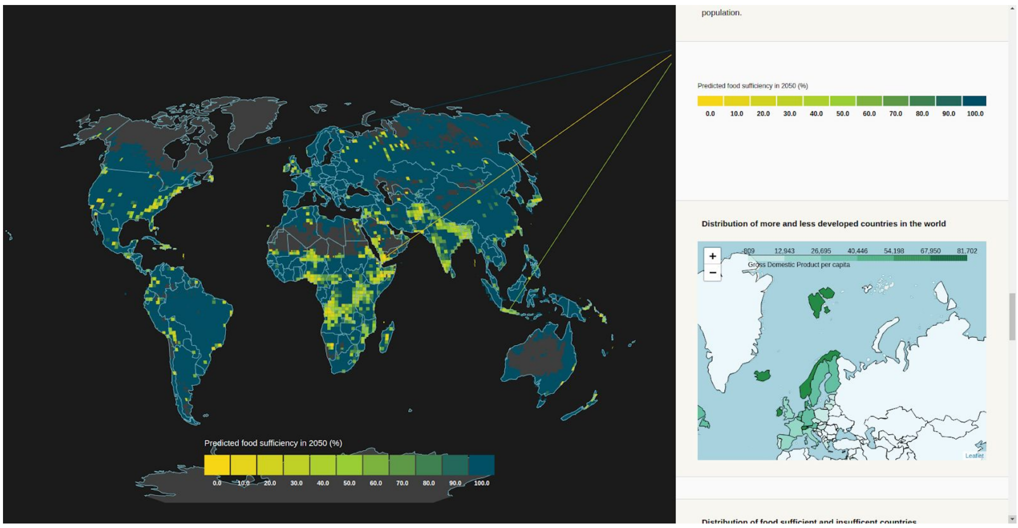
Finally we decided to separate the storytelling from the free exploration mode in two separate pages so we could optimize each page to its fullest potential.

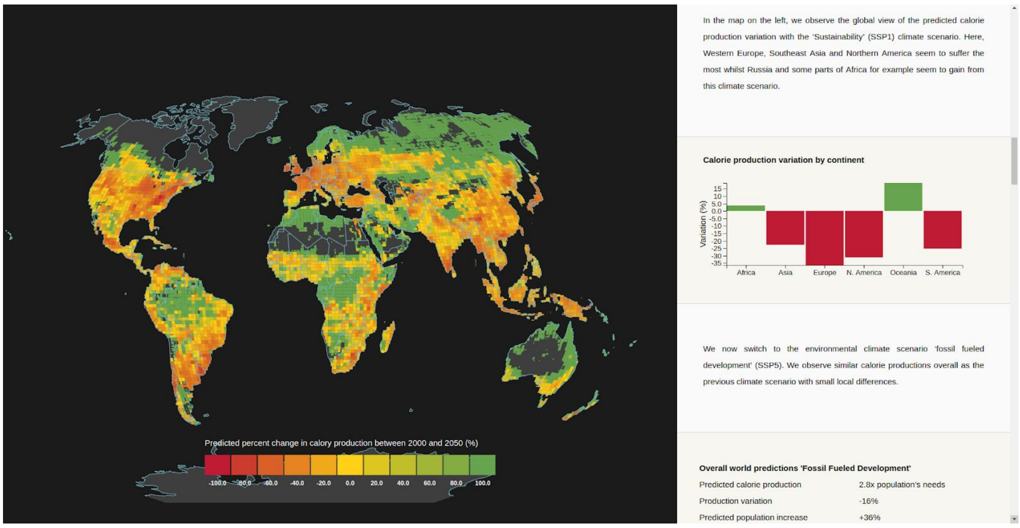


Final implementation

Our final implementation has diverged from our first ideas of designs after getting feedback from the assistants and asking a few external people what they were thinking about our visualization. Our main goal was to make it understandable for everybody what the visualization was about and what was intended in the analysis.

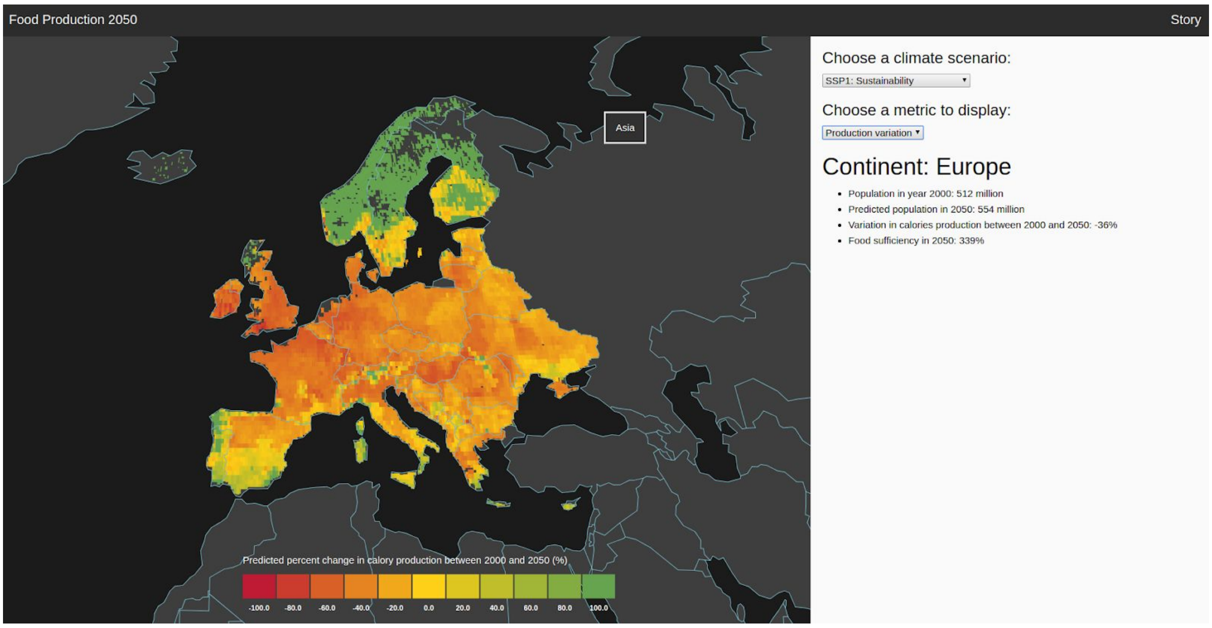
The final website consists of two HTML pages, the first one is presented as follows: an introduction to the problem where the context is explained followed by a guided analysis (storytelling) through the map comparing different world regions and different scenarios, accompanying each region with textual and graphical explanations. The user can see precise values by placing his mouse on the bars of interest on the charts. We selected interesting regions to look at by looking at our processed data. We did not include regions that didn't give any meaningful insight towards a better understanding of the problem in our storytelling. At the end of the page it is proposed to the user to explore the whole map by himself by clicking on a button.





The user then arrives on the second page where he can see the full map and zoom on it, the projection we used was the Winkel-Tripel projection as it was the best fit for what we were trying to show as it respects the ratio of the size of the countries and continents. He can also select countries or continents to see statistics about them on the right panel. If nothing is selected, the world map statistics are shown. The user can also select the different study scenarios or different metrics (either the variation in calorie production between 2000 and 2050 or the food sufficiency) to explore them. The choice of not implementing a comparison between the scenarios was on purpose as our goal in the map visualization was to explore each scenario and compare different regions within each one rather than comparing multiple scenarios on one only region or country. Also we think that observing the transition between the two scenarios gives a better idea of where the changes are happening than having to look at two different maps where it could be hard to track the small variations.

To conclude, our design uses a storytelling approach in order to show the mechanics of the interactive map to the user and give him some of the important insights to be gained from our visualization, before letting him answer his own remaining questions by exploring the interactive map by himself.



Evaluation

We learned a lot about the data using our visualisation. Indeed, plotting the the data points themselves instead of averaging by country was very useful because it allowed us to find disparities in food production inside countries (in large countries like Russia as well as relatively smaller countries like Spain).

We gained insight on which regions of the world are predicted to face hardships related to food self sustainability. For example, the African continent is predicted to produce 10% of the world's food whilst housing 20% of its population. Even worse, for countries like Nigeria, whose population is predicted (Sustainable scenario) to increase by 165% and will only be able to produce 66% of its population food demand.

Finally we also learnt that the region where we live, Western Europe and more particularly Switzerland, is going to suffer. If the predictions are correct, no matter which climate scenario is chosen, Europe is the continent whose calorie production variation between 2000 and 2050 decreases the most. Chocolate alone may not be able to feed us!

Contributions

Preparation

The group meetings were done as follows:

- Each member presented his ideas for the development of the project
- We discussed each presented idea until we all agreed whether or not we should pursue that idea or do it in a different way
- For each problem nobody knew how to solve we brainstormed some ideas and asked the TA's for advice if we couldn't find a solution
- We implemented complex ideas requiring skills from multiple team members by doing peer programming
- We assessed our progress and plannified tasks for each team member to be completed by the next team meeting

Peer assessment

There was no difficulty working with each other. Every member was motivated to work on the project and improve it. Everyone proposed a lot of ideas and helped determining which ones were good and should be implemented and which ones were not worth spending time on.

Everyone was punctual and attended all group meetings having done their assigned tasks beforehand and was ready to help if a member struggle with a specific task.

We were all able to advance our skills in data visualization by helping each other and giving useful references to learn more on specific subjects.