Project 2 - Final Submission

INFO 3300 (Spring 2019) | April 11, 2019

Srishti Belwariar (sb2355), Nancy Engel (ne96), Frank Rodriguez (fsr32)

I. Data Description

The intended purpose of this visualization is to highlight the changing proportions of available cropland and crop wastage for individual countries in a given year for a given specific crop. Through this visualization, we intend to show both (1) how certain countries may experience changes in how much land they can allocate towards crop cultivation and (2) how much crop wastage occurs in a given year.

Multiple data sets were analyzed in this project, all coming from the Food and Agricultural Organization of the United Nations. We used numerous separate datasets for this visualization (for cereal crop production, cereal crop wastage, sugar crop production, sugar crop wastage, starchy roots production, starchy roots wastage, and global lands available for cropland) that spanned from approximately from 1961 to 2017 (*Source: http://www.fao.org/faostat/en/?#data*). We noticed that some datasets do not have the same year intervals so we decided to only show data from 1980 to 2010. We used the different crop wastage and production data sets depending on what the user selects as the crop type and displayed this information in the bar graph. We then used the cropland data set to shade the globe. In order to properly report values, we needed to filter values and process the values in order to make sure if there were null or unreported values, we could detect them. The datasets were organized such that if a country did not have data for a specific crop and year, then it would not appear on the dataset at all. To handle this situation, our script searches for current countries (found on a tsv file) and if no data is found, then the visualization will process it appropriately such as showing no bars in the bar graph or shading the country gray on the globe. We noticed that the sugar data set had many missing information from many countries, which may be due to how many countries do not have the environment to cultivate sugarcane crops. This explains why for many countries, there is no data for sugar crops. However we decided to include this data set in the visualization in order to allow the user to make comparisons between countries that can cultivate this product, which is an important crop for consumption.

All the data sets contained more information than we needed so we used only the following variables:

- world-country-names.tsv
    - Current Country names
- landForAgri.csv
    - Share in Land (representing % of land used for crops)

    - Year
- All Loss_[…] csv files
    - Losses (in 1000 tonnes)
    - Year
- All Production_[...] csv files
    - Production (in 1000 tonnes)
    - Year

In the globe, the Share in Land attribute was scaled sequentially and used with a color scale. The slider used the years provided from these datasets to retrieve appropriate data. In the bar graph, the bar lengths were determined by taking the ratio of a country's Losses and Production (in 1000 tonnes). If, for any of these attributes, data was missing, we set in place some visual feedback to notify the user of this.

In addition to these datasets, we also used the resources listed Appendix B. We needed to work with D3 Topojson and world-110m.v1.json for the globe and national boundaries.

II. Visual Design Rationale

Overview

In regards to the front end design of the visualization, the color scheme, element attributes, and layout were designed keeping in mind the theme of global agriculture. The colors thus were selected to be more towards green and blue, with gold as an accent color for highlighting specific elements such as user selections. The design has been made on a dark blue background to make the visualization markings (colors and shapes) stand out more for the user to directly focus on them.

Crop Type and Country Selection Sections

There were two sections where the user is able to interact and select specific variables to interact with the visualization. The first section allowed the user to select from three different crop types. The team decided to not add more crop options as more options would make the layout more cluttered and not add much to the message the visualization is meant to convey. Each item has an icon image (from Flaticons), the aggregated crop category, and what are some crops common to that crop category. The icon presents a visual for the user to immediately determine what the option is referring to using a familiar object. The category label and list of crops allows the user to see what types of crops the option will include in the data presented by the bar graph below.

In the second section, the user is able to select at most 8 countries from the globe and add them to the selected country bank. When the page first loads, there will be a few selected countries already in the box to show the user how it looks when countries are added to the bank. The bank

has a fixed height to implicitly show that the number of countries a user can select is limited to just 8. This section also has annotations that explicitly tell the user how to add countries to the bank to avoid any confusion on how the interactivity components work. For example, there are annotations that state the user can click to add countries into the bank and that at most 8 countries can be added to the bank. The steps with which one can remove a country from the bank is not explicitly stated. This is because users typically will move their cursor over an item they want to remove, which will trigger the cell to turn gray and be crossed out to show the user can click and thus delete it. This is similar to how certain websites will require users to hover over an element for an "X" to appear rather than explicitly stating how to delete an object. We know that not having information more clearly presented on how the user can start interacting will be a challenge for users to use the visualization. However, for design, we decided that our small annotations would be sufficient to at least encourage the user to move their mouse around and then discover interactive components in that process.

Animated Globe

For our first visualization, we created a rotating interactable globe. The purpose of the globe was to present data about land usage and act as a visual list of countries from which users can click to present data on crop production and wastage. We decided to use a globe over typical earth projections because a globe most accurately depicts the countries of the earth, and given that users can compare different countries by looking at their land usage, we felt that this was important. Additionally, a globe allows for more interactivity than a static map. We allow the user to rotate the globe in the direction of their choosing, pause, and play the graph. This facilitates interactivity and attention of the user.

We went through several phases in the design of the globe, and we discussed the benefits and tradeoffs of many options. For example, an initial thought of ours was to outline the countries on the map so that users can continuously see the land usage for that country. Since land usage is mapped as the color of a particular country, we believed this would be a good idea. However, we decided to simply color the entire country "blue" to show that a country was clicked because a simple outline is hard to see on the graph, especially with some of the smaller countries. We realize when a country is selected, the user loses the land usage information for that country, but coloring the entire country blue is very clear to the user and after clicking another country or changing the year, the color of the country is recovered again. Another visual element that we discussed about the globe was the color scale. As a group, we debated the most appropriate color scale. One of the options was a sequential color scale. For example, we could use a scale that maps low land usage to light green and high land usage to dark green. After testing this approach, we realized that the majority of countries have low land usage with a few outliers such as India. Thus, the majority of the globe looked white or a very light green. After testing a few other color scales, we decided to use a

red-green color scale that appropriately demonstrated the differences between the countries of low land usage and high land usage.

Bar Graph

The bar graph highlights the proportion of crop of a specific type a country produces that goes to waste for a specific year. This graph is interactive to accommodate for user inputs such as changes to the country selection bank, crop type, and the year. The x axis is limited to a maximum of 50% due to how, overall, no country was wasting more than this threshold. The y axis shows the countries the user has selected in their countries selected bank. The overall bar graph width on the page is less than 50% of the web page to ensure the user is able to see the information close to the globe, which also provides relevant information. We decided that having a large or wider bar graph would be detrimental for users to easily refer to information provided both in this chart and on the globe.

The horizontal bars are markings that represent the proportion of crop production that is lost or wasted. All bars are of the same height and color to ensure that the representation of the information does not appear skewed. The only channel we use to show differences in percentages of lost crops is bar length. If there is no data available for a specific country, a "No data is available" label will appear where its bar would have appeared.

This bar chart was designed to be simple in terms of its styling to ensure that the information was easily apparent and interpretable. We see that there are trade offs with this decision as it may not entirely fit well with the overall web page's style. However, the main priority of this visualization is to ensure that the information conveyed is clearly present for the user. Initially, we pitched having the bar widths be proportional to how many countries have been selected (i.e. the width of the bars when 3 countries have been selected would be greater than if 7 countries have been selected). However, this complexity in styling would also skew perception of the data if the user adds/removes countries after the bar graph has been generated. Thus, we accepted that a simpler bar graph would not fit as well with the overall web site theme but it would convey the information clearly for the user to interpret with less bias.

III. Interactivity Design Rationale

Overview

This visualization used interactive components to allow the user to select what types of data they wanted to focus on in order to update the bar graph and globe. There are some interactive elements that help provide feedback to the user, but the bulk of the interactivity in this project is towards allowing users to filter and focus in on specific information dynamically.

Crop Type and Country Selection Sections

The sections where the users can select the crop type and specific countries are made not using D3 features but rather pure Javascript. When deciding how the layout and features should be designed, it was important to choose layouts that fit well with the theme of the website while also providing user affordances in a clean manner. It was decided to use button selection options (rather than checkboxes or radio buttons) to highlight selections made as the user because this form would fit better with the step-by-step process the user takes when navigating the visualization. Additionally, buttons allow more flexibility with design to better convey affordances for users.

The first section allows the user to select a crop category. For all options, the cursor is transformed into a pointer to make the user aware that those options are clickable. When selecting or changing one's selection of a crop type, the option that is selected is outlined with a gold border. This feedback allows the user to recognize which crop they have selected and that the visualization has registered that preference for the other interactive components to use. The gold border contrasts well with the darker web page background, which helps with identifying what option is selected at a quick glance.

The second section allows the user to select at most 8 countries from the globe. There are some countries preselected which appear in the selected countries bank, which are shown as golden divs/buttons. The selected countries in the bank are golden buttons similar to how selected crop options have a golden border to highlight selections collectively. When the user hovers over a country button in the selected country bank, the div/button containing that country turns gray, the name is crossed out, and the cursor is turned into a pointer. With this feedback, the user is made aware that they can remove any selected option from the bank at a click. To add a country, the user clicks on a country on the globe and automatically the country is added to the country selection bank. Once the user selects 8 countries, the user will not be able to add another country. If they attempt to, their selections will not be stored in the country selection bank. The visualization limits the user to 8 selections to ensure the interactive bar graph does not become over populated and cluttered, which can impact how the data is interpreted by the user. There is no feedback shown to emphasize that the maximum number of country selections has been reached. Because this interactive section has many interactive components, it was best to just explicitly state the limit rather than have pop up messages once the limit has been reached. Having too many feedback elements can make the design cluttered and even frustrating to work with the visualization.

Animated Globe

The marks of the animated globe are the interactable country elements, and the channels for these marks is color which corresponds to the land usage of that specific country for a selected year. A large purpose of the globe, as outlined in the previous section, was to serve as a sort of visual list

that users can select countries from. Thus, we wanted the user to be able to click on all the countries of the globe. Since we are using a globe and not a static map, we had to implement another interactive element of rotation. Specifically, the user can manually control the rotation of the earth, pause its natural rotation, or start its natural rotation.

An interactable component of the globe that we discussed in depth, was the method of rotation. Initially, we wanted a globe that rotated, had countries that could be selected, and could be dragged. After implementing this, we saw that dragging actually decreased the usability of the globe. There are two main problems with dragging that we found: the clicking of countries sometimes interferes with the dragging action causing the earth to behave in an unexpected fashion, and secondly, the performance of the system was impacted by dragging. Wanting as smooth as an experience as possible, we opted to use buttons for the earth instead. Although buttons do not have the novelty of dragging, buttons clearly show the user the interactable actions that can be done which facilitates usability and discoverability.

The method of manually moving the globe was an interactable component we discussed, but we also discussed the implementation of the buttons causing the rotation. The implementation of the buttons that allow for manual rotation is not clearly explicit to the user; nonetheless, there were several design decisions made about this implementation. Rotation of the globe is typically achieved using a time-based system. D3 timers are used as well as the current real time. This worked great for our primary implementation, but not so much for moving the globe in the opposite direction. Doing this would effectively mean moving backwards in time. A great deal of time was spent trying to achieve a system that would smoothly switch directions on a time based system, but the math to achieve this was not working out. Thus, we opted for a non-time based system, relying on updating the angle of rotation at every iteration of the *rotate* function. A trade off of this approach is that we achieved our goal of smoothly changing the direction of rotation, but the speed of rotation can vary from computer to computer based on the computer's processor.

Bar Graph

The bar graph itself does not use interactive components where the user clicks on the graph elements to see transformations; rather it is the output of selections the user makes using the other interactive features of the visualization. The bar graph does update automatically from changes in selections. We decided that the updates would happen instantaneously rather than giving the users a refresh or update button because this would lessen the number of actions the user would have to go through to generate the visualizations they need. Additionally, when using the slider to change years, it is much easier for the user to see periodic changes without having to interpret their use with the slider. For example, the user may be interested in seeing how a country's percentage of wastage changes over a twenty year period. Rather than them having to shift a year and then refresh the chart, the bar graph automatically is modified so the user can see trends without having to ever leave

the slider itself. Additionally, the bar graph updates when a new crop is selected; the bar graph automatically updates to reflect the new crop, but with the countries that were already selected at that point.

IV. Insights and Trends

The purpose of this visualization is to provide context as to what proportion of crops produced are wasted for each country and how this connects to the amount of land that's available for crop cultivation. We also wanted to ensure that comparisons were made between countries and over time. Before developing this visualization, our goal was to try to find some trend between how much available cropland affects crop wastage. The initial assumption was that countries with higher percentages of cropland available could tend to have higher wastage levels due to the resources they had available. In order to explore this, we decided to choose three major crop groups and large datasets that spanned numerous decades. We noticed that this trend is not necessarily apparent as there are other factors (i.e. crops imported, population size, GDP, etc.) that can influence whether countries are efficiently managing their crop wastage level given cropland constraints. However, there is still a trend where certain countries with higher cropland allocations tend to have higher crop waste proportions. Overall, we intend to show that countries that have an overall limited cropland allocation tend to have lower wastage percentages due to resource constraints.

Appendix A: Task Allocation

1. Initial dataset and design discussion and wireframing (All members)
    a. Time taken on design: ~2 hours
    b. This task required us to meet a few times to decide on the datasets we wanted to utilize in the visualization and what the general layout ought to be. Because FAO provides numerous datasets on various crops and attributes, we needed to spend time on finalizing what information was necessary to show a trend. This part of the project was mostly creating wireframes.
2. User selection (Srishti)
    a. Time taken on design planning: 30 min
    b. Time taken on development and implementation of design: 5+ hours
    c. This task area includes the styling and implementation behind the features that allow users to select a crop type, add/remove at most 8 countries from the globe to the countries selection bank, and ensuring all user selections were properly refreshed and stored in order for the other parts of the visualization to work. The work in this area was specifically using CSS and pure Javascript (not necessarily the D3 library). Additional time was spent on reevaluating the styling to ensure possible actions were easily inferrable by the user.
3. Interactive Globe (Frank)

    a. Time taken on design planning: 1 hour
    b. Time taken on development and implementation of design: 6+ hours
    c. This task includes creating the processing land use data, globe, implementing rotation, intractability, buttons, slider, and color scales. The main reason that this section took a while was because there was a lack of documentation on globes and even less on adding interactivity and rotation to them. Showing the globe itself was not too time consuming as it is based on a homework we did in class, but interactivity was slightly more difficult because how we wanted to click on each country, but also show the changes of each country's color when the user moved the year slider. The slider was also slightly challenging to implement because the slider documentation provides details on how to set up a basic slider, but nothing beyond that. However, the most difficult part was implementing the globe rotation. I spent a large amount of time attempting to calculate the correct angles of rotation on a time based system. Eventually, I decided to simply increment and decrement the angle on each iteration with the cost that different people may see slightly different speeds based on processor.
4. Production and Losses Bar Graph (Nancy)
    a. Time taken on design planning: 30 minutes
    b. Time taken on development and implementation of design: 5+ hours
    c. This task involved creating the bar graph and scales associated with it, as well as making the bar graph update dynamically based on which crop is selected and based on which country is selected. This task took a long time because I ran into a lot of issues that required debugging (like when I got the graph to update for each country, I could no longer select a crop so I needed to restructure a bunch of the code and come up with a solution)! There were a lot of moving parts that needed to fit together; the datasets needed filtering, I needed to make sure the data was matching the correct year, and that existing data was being removed and rewritten at the correct times, and that the graph looked nice and was easy to read.
5. Final web design and styling (Srishti)
    a. Time taken on design planning: ~1 hour
    b. Time taken on development and implementation of design: 2+ hour
    c. Implementation (i.e. preparing style sheets, setting final color scheme and making final modifications based on how cohesive it was to the overall theme) was done by Srishti. General styling attributes were set but all members were free to modify, align, and reformat their respective visualizations as long as these changes were in accordance with the overall website's theme.
6. Documentation (All Members)
    a. Time taken on preparing document: ~4 hours

Appendix B: Resources and References

- Food and Agriculture Organization of the United Nations (http://www.fao.org/faostat/en/?#data)
- Google Fonts: https://fonts.google.com
- D3 Documentation: https://github.com/d3/d3-shape/blob/master/README.md
- D3 Topojson
- Flaticon icons (by Freepik): https://www.freepik.com/
- D3 Simple Slider: https://unpkg.com/d3-simple-slider,
- D3 Simple Slider Documentation: https://github.com/johnwalley/d3-simple-slider