Title: CPSC 583 Final Project Report: Visualizing the Consumption and CO_2 Emissions Relating to Different Foods Around the World

Name: Sara Hall

Course: CPSC 583

Due Date: April 14, 2021

Table of Contents

1. Introduction

This project was about visualizing a dataset on the per capita rates of both the consumption and the CO_2 emissions associated with different food sources. The dataset has 1431 rows and four columns which include the country names, food source names, rate of per capita CO_2 emissions, and rate of per capita consumption. Overall, it has information about 130 countries and 11 different food sources including beef, dairy, wheat, lamb and goats, poultry, fish, pork, rice, eggs, nuts, and soybeans.

My main goal with visualizing this dataset was to find a way in which to display the large amount of data such that it could be quickly and easily understood. When I first encountered the dataset, I thought it could be valuable to show how rates of consumption and CO_2 emissions compare for different food sources to give people the information and data needed to inform decisions about sustainable food choices. In this report, I will outline the design process, starting from the selection of the dataset and sketching, prototyping designs, finalizing a static design, and then adding interactions. I will then discuss areas in which I think the visualization could be improved moving forward.

Overall, my design process began with sketching, which led me to want to be creative. As a result, I first tried to make bubble charts shaped like footprints but ended up deciding this was not an effective approach. Instead, for my final visualization, I combined maps showing consumption and emissions by scaling the colour of countries in orange and green, respectively. I also added a bar chart to allow for easy and accurate comparisons between the food sources. Finally, I added interactions to allow people to explore the dataset if interested, allowing them to zoom in on the maps, and show the data for different countries in the bar chart. This final visualization is effective in terms of many of the principles that were discussed in class. It uses length, position, and carefully selected colour scales to show the different variables. It also has minimal chart junk and maximizes the data to ink ratio. While it may not be the most creative of designs, it should be intuitive and easy to follow. In the future, it could be improved with more interactions, possibly allowing multiple food sources and countries to be selected at a time, and by the addition of smooth animated transitions between views of the data to improve aesthetic and increase user engagement.

2. Data Description

2.1 Data Descriptions

This dataset contains data about the consumption and amount of carbon dioxide (CO_2) emissions of different foods around the world. There are four columns and 1431 rows. The first two columns are both nominal variables; the first is the country being considered, and the second is the category of food. There are 130 distinct countries and 11 distinct food categories. The remaining two columns are both quantitative. The third column contains food consumption in kg/person/year, which is a measure of the amount of food supplied by a country for consumption and may include food waste(*Food Carbon Footprint Index 2018 | Nu3*, n.d.). The fourth column has the associated CO_2 emissions in Kg/person/year, which was calculated by multiplying the supply for human consumption by the median of emissions in the world for that food source(*Food Carbon Footprint Index 2018 | Nu3*, n.d.). This dataset came from one of the Tidy Tuesday projects on GitHub, which was pulled from nu3(*Food Carbon Footprint Index 2018 | Nu3*, n.d.; Kulma, n.d.). Nu3 collected the data from the Food and Agriculture Association of the United Nations Balance Sheets(*Food Carbon Footprint Index 2018 | Nu3*, n.d.). I have chosen this dataset because we are living in

a time where being environmentally conscientious is crucial and being able to see the emissions levels of different food sources might help people to make more sustainable choices as consumers.

2.2 Pros and Cons of Data Set

There are many pros to working with this dataset. Since it has both nominal and quantitative variables, it will give me a chance to work with both. Additionally, the information in the dataset is relevant to today's modern society, and visualizing it well has the potential to be impactful. There are also no empty cells, so I will not have to worry about that. The relatively small number of attributes (four) should also make it a manageable dataset to work with. One of the cons of working with this dataset is that there are a lot of rows. However, since they can be grouped into countries, it should be manageable. If I struggle with the large number of rows, then I can also group the countries into larger regions like continents. While the Food and Agriculture of the United Nations Balance Sheet seems like a reliable original source, it is worth noting that the data was pulled from them by nu3, a company that sells health foods and advocates for plant-free diets(*Food Carbon Footprint Index 2018 | Nu3*, n.d.). As a result, there may have been some bias. Additionally, the Food and Agricultural Association of the United Nations did not have information on CO_2 emissions for all food types, so a couple of other reports were also used(*Food Carbon Footprint Index 2018 | Nu3*, n.d.). This could lead to inconsistencies in the measurements. While having the CO_2 emission for each food category in each country makes it seem like local emissions are being referenced, these values are median emissions worldwide relative to the amount of consumption in that country. While this does make the emissions more of an estimation, in some ways it is also a good thing because it prevents countries with high export rates from having artificially high emissions(*Food Carbon Footprint Index 2018 | Nu3*, n.d.). The measure used is also only sensitive to production on the farms; any upstream or downstream emissions are not included(*Food Carbon Footprint Index 2018 | Nu3*, n.d.). Finally, while this dataset was published in 2019, the actual data were collected in 2013/2014. As such, they may be a bit outdated.

2.3 Data Set Decision

Overall, while the dataset is a bit old and needs to be interpreted with care since it only represents a subset of emissions (only CO_2, and only during farm production), it is still potentially useful. Having a good visualization to show the differences in CO_2 emissions for different food sources has the potential to help people evaluate their nutritional choices in terms of environmental sustainability. Lastly, visualizing the combination of nominal and quantitative data will provide a good learning opportunity.

3. Design Process

This section shows the progression of design, starting from sketches all the way to the final static visualization and interaction prototypes. In my designs, I wanted to highlight how the rates of consumption and CO_2 emissions change around the world in different countries. I also wanted the user to be able to get an idea of which foods are more environmentally friendly, emitting less CO_2, and which are less environmentally friendly. I explored many different representations in my sketches but ultimately decided that maps would be the best way to show global trends and that a bar graph would be best for allowing easy comparisons.

3.1 Sketchable Data Subsets

The first subset I chose looks at the three largest North American countries – Canada, the USA, and Mexico[1]. I thought that this would give an interesting view of emissions and consumption of different

food types in North America. While they are certainly not the entire continent, I figured they had the potential to be representative. Additionally, by choosing a small number of countries, it allowed me to focus on different ways of representing the eleven different food types and their associated rates of consumption and CO_2 emissions. While this may not represent the full global data, it was a good starting point for playing with different ways to represent the different attributes. By looking at all the different food types, I also started getting a sense of how challenging it is going to be to show the large difference between the emissions related to beef, but also scale things so that differences between other food types are visible.

The next subset I chose looks at the most populated country in each continent, with Europe split into east and west, so I looked at the USA, Brazil, Germany, Nigeria, Russia, China, and Australia[2]. I then looked at all eleven food types and their related consumption and CO_2 emissions. I chose this subset because I wanted to get a more global view of the data, and I thought the most populated countries would provide a good starting point. One thing I did struggle with for this subset was hand sketching a large number of data points.

Finally, because I started before the assignment changed, I also sketched a third subset[3]. Because I wanted the opportunity to play more with geographical representations, I decided to focus on the most western European countries – Portugal, Spain, France, Belgium, Luxembourg, Switzerland, Italy, Netherlands, Germany, Ireland, and the UK. Since I already played with showing all eleven different food types in the previous two subsets, I decided to limit the subset to the two grains on the list of food types – rice and wheat. This allowed me a bit more freedom to play with geographical representations.

3.2 Design Direction

Due to the categorical nature of the food types and countries, I did do a fairly large number of bar charts. By nesting variables along the axes, they allow position to be used to represent many different categories. One thing I played with a lot was different ways of grouping these variables. For instance, I tried grouping food categories under country (Figure 4.III), and then I switched them so that country was grouped under food type (Figure 4.IV). Since emissions and consumption are both measured in units of kg/person/year, they can be plotted on the same axis for easy comparison which is rather handy. I also played a lot with using area and colour to represent the emissions and consumption variables. While I know that area does have the potential to be misleading and colour is not very quantitative in nature, I do still find myself drawn to their ability to quickly show differences briefly. I think with a computer, they could also be made more precise by adding something like a mouseover function so that exact values can be seen. In many of my sketches, I used "heat map" colours, with warmer colours indicating higher values.

I tried to use nested areas to layer information in a lot of my sketches. For instance, in some of my sketches, I used an outer rectangle to show overall emissions or consumption and then nested different food groups or countries within them (Figures 1.X, 2.VIII, 2.IX). In my early sketches, I used a lot of different icons to visually represent the food types, but after learning about chart junk in class I started trying to be more conscious about using less unnecessary ink and encoded more attributes using position and text labels.

In many of my sketches, I also tried to incorporate geographical position. In some cases, this involved relative placement on the page (Figure 6.III, Figure 6.IV), and in others I tried to overlay the data on top of an actual map (Figure 6.V).

3.2.1 First Sketches

For my first subset, I liked the visual comparisons using size and colour on a treemap (Figure 1.X), which is why I decided to focus on that to elaborate on and make variations of for the second set of ten sketches.

3.2.2 Variation

For the second subset, I liked the easy comparisons that were enabled by the bar graph idea (Figure 3.VIII, 3.X), so I decided to focus on these for the subsequent ten sketches. Finally, for the third subset, I really liked the idea of using geographical encodings for the different countries (Figure 5.IV), which is why I focus on that for the last ten sketches.

3.3 Process

I learned a lot about my dataset working on these sketches. There are three important takeaways for me. The first is that in my final visualization I will likely want to either use something like a logarithmic scale for rates of emissions and consumption or include some sort of zoom function so that difference between groups with smaller rates of consumption and emissions are still visible. The second is that I will likely want to use multiple charts in my visualization so that I can highlight different views of the data. Finally, I still have work ahead of me in terms of finding a way so that my visualization will scale to show data for all eleven food types in all one hundred and thirty-one countries. A large challenge with this will be maintaining clarity and not letting my final visualization get too messy.

3.4 General Design Direction

Based on the feedback I received for my initial sketches, I chose to elaborate on the footprint design (Figure 7). The main advantage to this design is that it is emotive and underlines the massive differences between the amounts of foods we consume, and the amount of CO_2 that is released into the atmosphere by consuming those foods. By scaling the area of circles to kg/person/year, it is possible to show lots of data on one page and emphasizes the large culprits for both consumption and CO_2 emissions. The area scaling and attempts to maximize data density were inspired by Tuft's Integrity Principles, as discussed in several lectures. While the symbols I used do add ink, they also increase legibility, are very intuitive, and help viewers gain insight into the data at a glance. A large trade-off with choosing this design is that it can be challenging to judge differences in area perceptually, as has been discussed in lectures. The footprint design while emotive may also be unintuitive for the viewer as it is an unconventional design and is a highly unusual shape.

To show how the rate of CO_2 emissions changes with the rate of consumption, I also decided to try keeping all of the bubbles the same size and then plot them on axes to show consumption and emissions. This was motivated by position being one of the best visual variables to encode data for all data types, according to Mackinlay's Ranking, as discussed in Lecture 5.

For the third variation, I pulled in some concepts from the sketches I made using maps. I coordinated the view of the bubble map "footprints" with colour intensity encoded maps to show global changes in food consumption and CO_2 emissions. In this way, I was able to highlight both global trends and differences in food categories. This is my favourite visualization and is likely what I will focus on for the final hand-in.

For my first two variations, I used an online colour picker tool to generate my palette as discussed in the lecture on colours. However, it occurred to me that the colours I used might not be great for someone with red-green colour blindness so I will need to experiment more for my final implementation. The pale yellow and green I used to represent Europe and Africa, respectively, may also be challenging for some people to distinguish.

Finally, while interaction was not supposed to be the focus of this hand-in, after seeing my first visualization with 1430 nodes for each of rates of consumption and rates of CO_2 emissions, I decided that the best way to make the visualizations more legible in the second and third implementations was through the use of interaction.

All codes are available at https://github.com/SaraAnnHall/SaraAnnHall.github.io. As a footnote, something I need to figure out is how to scale the visualization to the browser that is used. For me, they work in Google Chrome, but if I open them with Firefox the layout is wonky. The Emoji symbols also do not scale properly in Firefox, please open the links in chrome if possible.

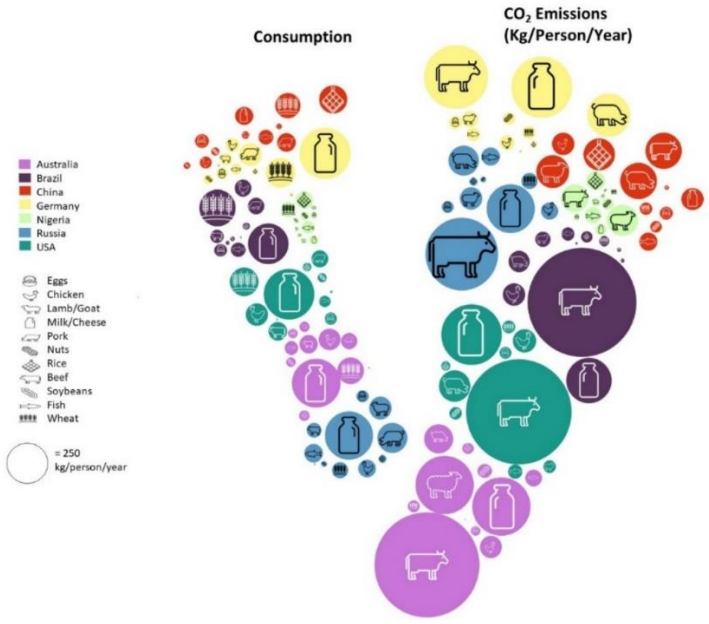


Figure 7. A cleaned-up version of the bubble map "footprint" concept I decided to elaborate on for this hand-in. Note that the symbols are not properly scaled. I also should have included a better size legend but note that the areas are proportional to rates.

3.5 Prototyping Variations

All three visualizations can be accessed from the links on my GitHub page (https://saraannhall.github.io). Individual links will also be provided in the descriptions.

3.5.1 Variation A

For my first variation, I decided to show the footprints as bubbles grouped into larger circles to represent the footprint of both consumption and emissions around the world (Figure 8, https://saraannhall.github.io/forced_fixes_2.html). Within each of the two larger circles showing consumption (on the left) and emissions (on the right), there are 1430 bubbles, one for each food category in each country. The areas of the bubbles are scaled to the rates of consumption and CO_2 emissions. I also implemented a tooltip to show country name, continent name, food type, rate of consumption, and rate of CO_2 emissions on mouseover.

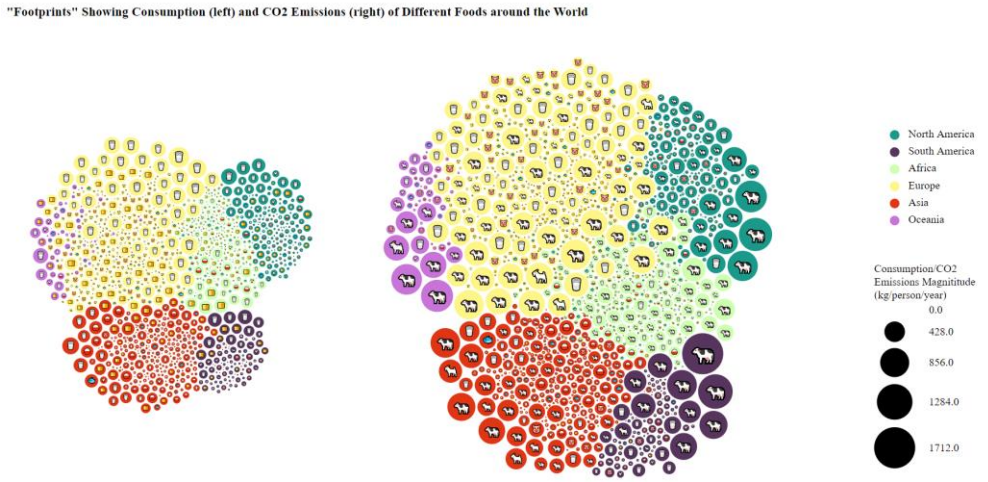


Figure 8. "Footprints" are shown as large circle-like shapes, one on the left for consumption, and one on the right for CO_2 emissions. The bubbles are clustered and colour-coded by continent, and their area is scaled to rates of consumption and emissions, in units of kg/person/year. In the running version, a tooltip appears when a bubble is moused over and shows the country name, continent name, food type, rate of consumption, and rate of CO_2 emissions. Running version available at https://saraannhall.github.io/forced_fixes_2.html.

3.5.2 Variation B

For my second variation, I decided to implement a scatter plot (Figure 9, https://saraannhall.github.io/scatter2.html). The rate of CO_2 emissions are represented on the y-axis, plotted against the rate of consumption on the x-axis. Both scales are linear, to prevent any data distortions. Again, the circles are colour coded by continent and the food-types are labelled using symbols. The drop-down menus can be used to filter by both food category and continent.

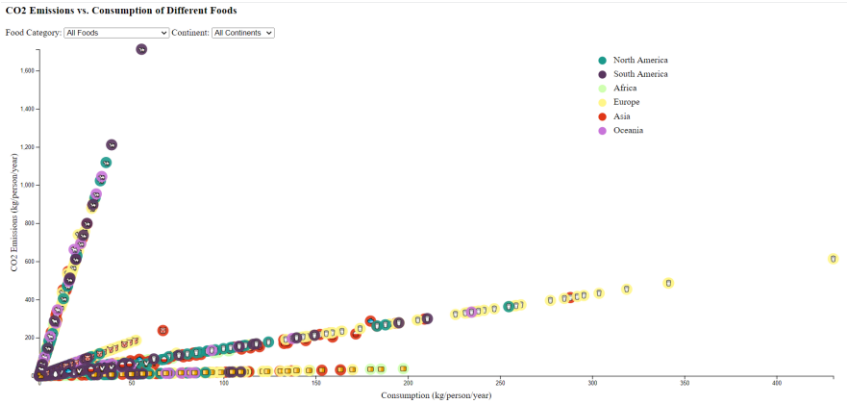


Figure 9. A scatter plot with the rate of CO_2 emissions plotted against the rate of consumption. Circles are colour-coded based on continent and symbols show the different food categories. A tooltip shows more information, and by using the drop-down menus, the data which is displayed can be filtered by food category and continent. Running version available at https://saraannhall.github.io/scatter2.html.

3.5.3 Variation C

For my final variation, I decided I wanted to combine a better way of showing geographical information along with the bubble "footprints" (Figure 10, https://saraannhall.github.io/map_with_bubbles.html). A dropdown menu changes the map on the left from showing rates of emissions scaled in orange and rates of consumption, scaled in green. Initially, these show averages for all food categories, but the next drop-down menu can be used to select particular food categories. On the right, bubbles show the magnitudes of consumption and CO_2 emissions by scaling with area and using the orange and green colour scales. They are also filtered by the food category drop-down menu.

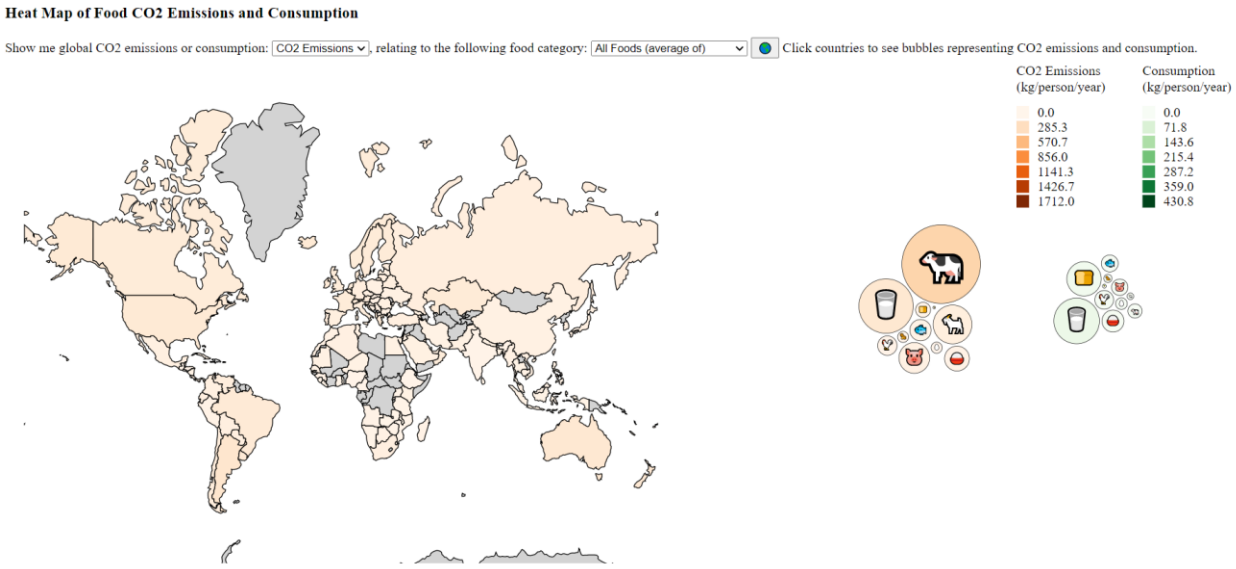


Figure 10. Third variation. Shows data on a map to highlight geographical trends, and as bubbles to highlight the differences between rates of consumption and CO_2 emissions for different food categories. Running version at https://saraannhall.github.io/map_with_bubbles.html.

3.6 Implementation Process

While starting to try and implement a footprint shape, I decided to start with arranging the nodes into two larger circle-like shapes, using D3's force layout functionality. Upon doing this, however, I came to realize that this design might not be optimal for the data. One large problem with clustering the nodes based on continents is that some continents have a lot more countries, making it look like they both consume and emit more. Additionally, as discussed in class, it can be hard for us to judge circle areas and comparing the sizes of two footprints is likely to be more challenging. Therefore, I decided next to keep the circles a constant size and plot them with the rate of consumption on the x-axis and the rate of CO_2 emissions on the y-axis. Due to a large number of nodes, I then decided to add drop-down menus to filter by continent and food-type. With the scatterplot, occlusion is still a problem. Finally, because I was drawn to showing the global trends, for my third iteration I wanted to combine a map with bubbles. This final visualization is my favourite so far and is likely what I will elaborate on in the future. I would like to scale the colours differently so that smaller differences can be seen and are not washed out by the high magnitude of beef and lamb emissions, and milk and wheat consumption. I would also like to allow the user to select multiple countries to compare using the bubbles and play around with different button types and zoom functionality.

3.7 Final Static Design

For the final static design, I decided to elaborate on the third variation from prototyping, with the map next to a bubble chart (Figure 10). However, because this implementation was supposed to be static, rather than only show a single map changing colours based on the selected consumption/emissions and the food types in the drop-down menus, I decided to employ small multiples to show as much on the one page as possible and maximize the data density by making use of parallels (Tufte's Guidelines, Lecture 21). Therefore, I made a grid with eleven columns, one for each food category, and two rows: one for CO_2 emissions and one for consumption, and a small map in each cell (Figure 11). Consumption is shown on the maps in the top row in greens, and emissions are shown in the maps in the bottom row in oranges. After noticing that a lot of the values were hard to see using linear sales mapped to D3's interpolateGreens and interpolateOranges colour scales, I decided to use a logarithmic scale instead. As a result, it is now easier to see the differences between some of the lower values. Finally, because the colour scales are challenging to make exact comparisons between, I decided to use an additional view, in the form of a bar graph. This was motivated by length and position being some of the best visual encodings (Mackinley's Rankings, Lecture 5). For the static version, the bar chart shows the global averages for both emissions and consumption to allow for the user to get a quick idea of how the different foods compare overall. As secondary information, I included colour legends, axes and labels for the bar chart, and symbols used as labels for the food types under the maps.

The code can be found at https://github.com/SaraAnnHall/SaraAnnHall.github.io/blob/main/static2.html and the live visualization can be found at https://saraannhall.github.io/static2.html.

For interactions, I wanted to allow the user to filter the data to see the data pertaining to different countries and food types. I also wanted to enable the user to zoom in on the maps since they are quite

small, and on the bars since categories like nuts and soybeans are small and hard to see. In general, I was inspired by the "Visual Information Seeking Mantra" from Lecture 19 – "Overview First, Zoom and Filter, then Details on Demand". I wanted to show the big static overview first, allow the user to select a map of interest so that it zooms in, and then from there, further allow the user to select a country on the map and have the bar graph change to display that country's data. In this way, I wanted to connect the two views of the data. I also wanted to enable the user to change which food category was displayed on the maps by clicking the bars in the bar chart. Then for details on demand, I decided to implement a tooltip to show precise values. Overall, I wanted to implement selection tools to allow the user to explore with zooming and filter the data.

In terms of interaction alternatives, another thing I considered was allowing the user to draw a lasso around several countries and then show the data for the selection either in one bar chart as averages, or in several bar charts, one for each country. However, with 130 countries, this could get overwhelming. I also could have kept the drop-down menus from my third prototype instead of clicking on elements, however, I decided that added too much extra junk to the page unnecessarily.

There were not a lot of interactions that seemed entirely incompatible with my final static visualization, however, there were some things that I decided against since I thought they would make less sense. Given the abnormal shapes of countries, rectangular brushing for selection did not make seem the most suitable. I also did not think it was necessary to rearrange the data. I could have made it so that the user could move positions of the bars on the axes, but I decided with only 22 bars, this needed. Finally, I did not think it was important to allow the user to change encodings as I think the encodings that I have chosen already show the data well, and I explored a lot of alternatives in my sketches.

Figure 11. Final static visualization. Small map multiples on the top show rates of consumption and CO_2 emissions for all 130 countries and 11 food types. The bar graph on the bottom shows the global averages for the different food categories. The order of food categories from left to right is the same on the maps and the bars, and actual names and values are displayed on mouseover.

3.8 Prototyping Interactions(s)

For the first interaction, I wanted the user to be able to select one of the small multiples and zoom in on it. To do this, I decided that clicking on the small map would open up a zoomed-in view as shown in Figure 12.

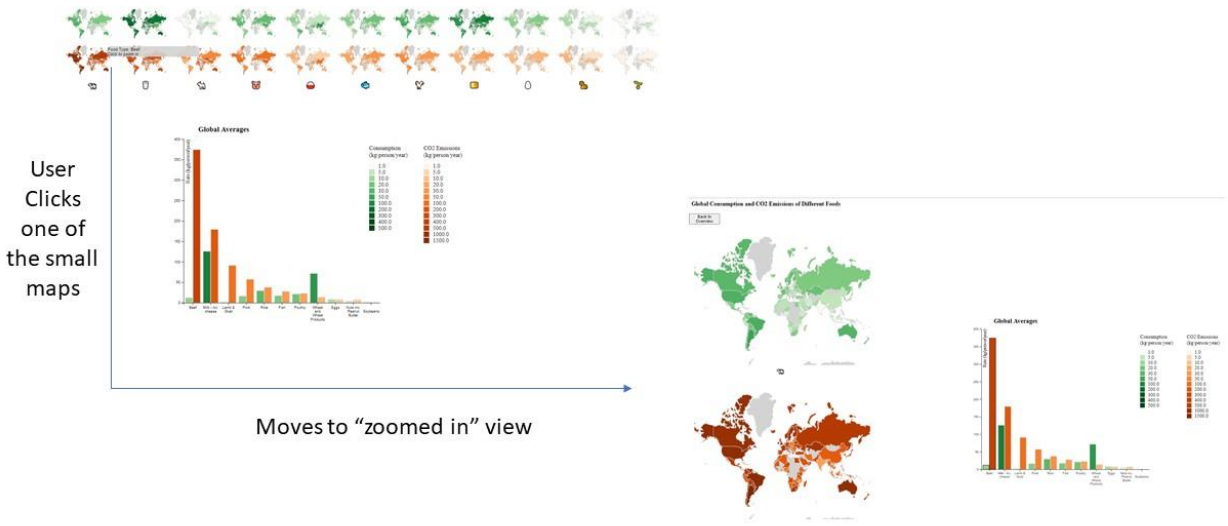


Figure 12. The first interaction allows the user to click on one of the small maps to zoom in on the maps relating to that food type. The bar graph moves to the left, and the bars for the selected food type are emphasized by a black outline.

For an additional selection mechanism moving into the zoomed-in view, I also made it so that the user can click on one of the bars, as outlined in Figure 13.

Figure 13. Clicking on a bar also move the visualization into the zoomed-in view.

Next once in the zoomed-in view, I connected the bar chart with the maps. When a country on one of the maps is clicked, it becomes selected – emphasized with a black outline on both maps – and the bars change to show that country's data (Figure 14).

Figure 14. Shows what happens if a country in a zoomed-in map is clicked.

Following this, to enable the user to switch between different food types shown on the maps, bars can be clicked on the map (Figure 15). Only one country and food type can be selected at a time.

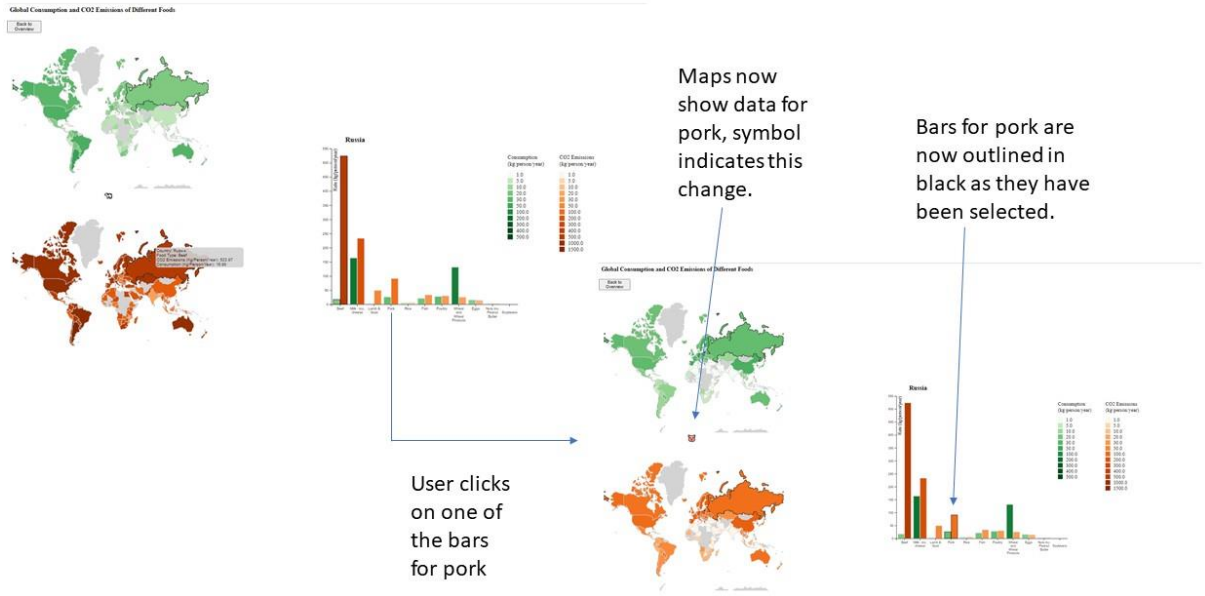
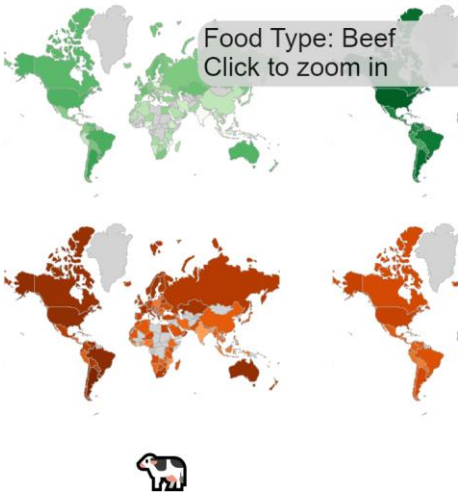


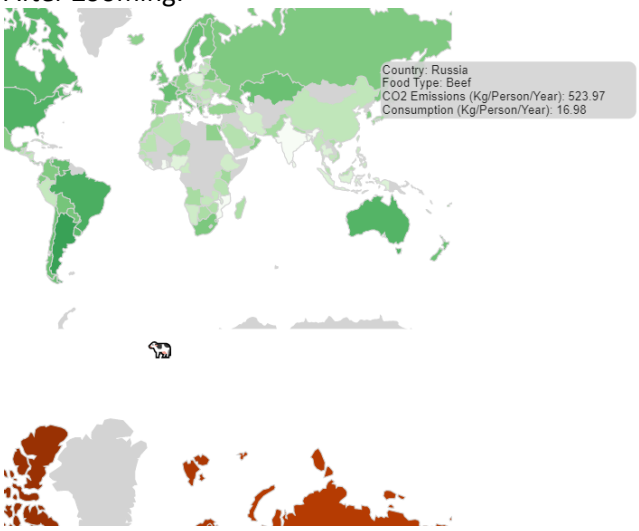
Figure 15. Shows what happens when a bar is clicked in the zoomed-in view.

Tooltips are also used to show some details on demand in both the overview and the zoomed view (Figure 16).

Before Zooming:



After Zooming:



Bars:

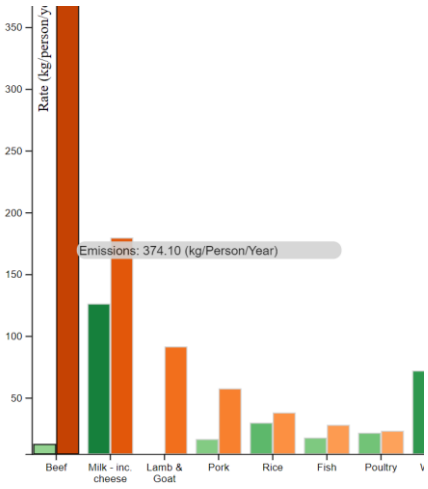


Figure 16. Shows the available tooltips that are displayed on mouseover.

To enable the smaller bars to be seen, the bar chart can also be zoomed and panned. In doing this, the scales of both the x and y axes change (Figure 17).

Figure 17. Outlines zooming and panning interaction in the bar chart.

To go back to the original overview, a button in the top left corner can also be pressed (Figure 18).

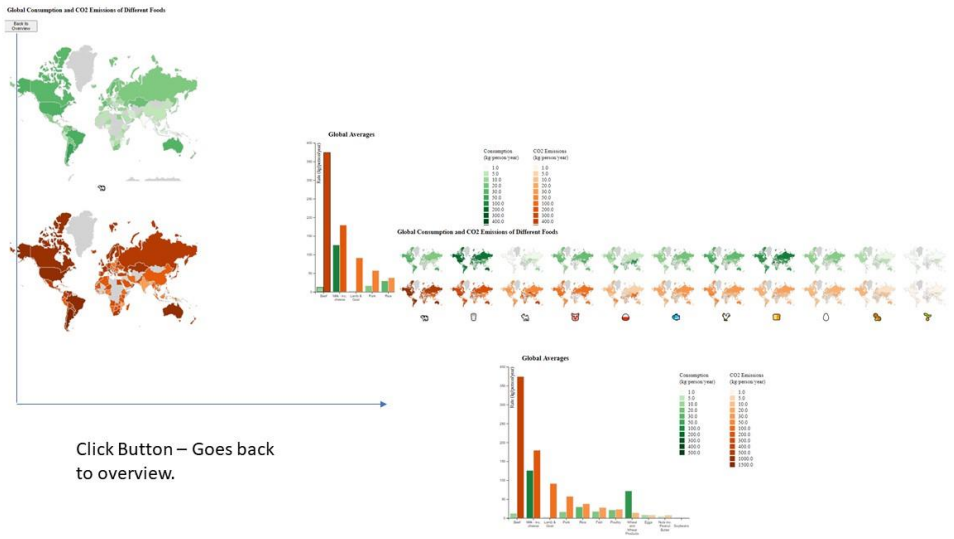
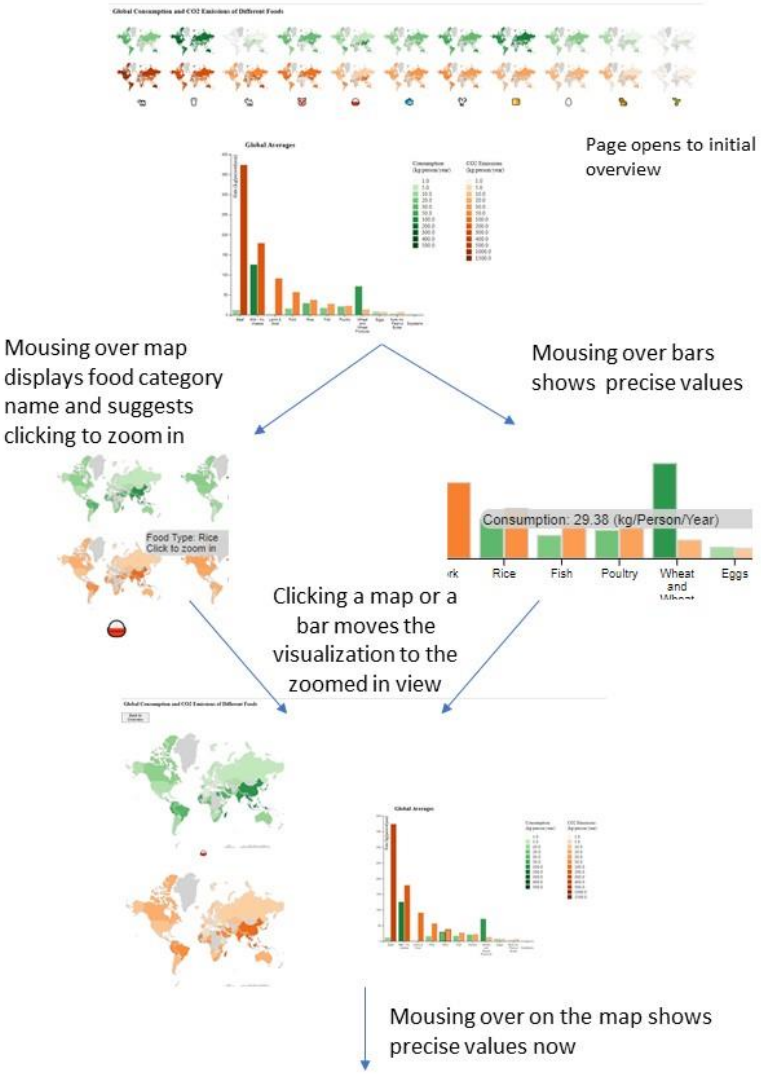


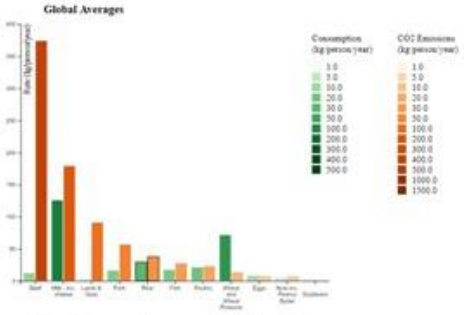
Figure 18. Clicking the button goes back to the overview.

4. Final Implemented Visualization

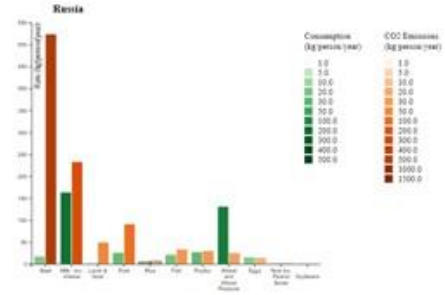
The final visualization can be found at https://saraannhall.github.io/interactive3.html, and the code is available at https://github.com/SaraAnnHall/SaraAnnHall.github.io/blob/main/interactive3.html.

The visualization opens to a view of the final static version as described above, which makes use of small multiples to display the entire dataset. From there, all of the interactions described in the previous section are available. To enter the zoomed-in view, the user can select a food category of interest by either clicking on one of the maps or one of the bars. Then to see the data for a particular country rather than the global averages in the bar graph, it can be clicked on the map. To change the maps to display data for a different food category, the bars on the bar chart can be clicked. Current selections are emphasized with a black outline in both the map and the bar chart. Because some of the categories have really small values, the bar chart can also be zoomed and panned to see these. Finally, when the user is done exploring the zoomed-in view, by clicking a button in the top left, they can return to the overview. A step-by-step guide for how the visualization looks and works is shown in Figure 19, due to the number of interactions, this figure spans multiple pages.

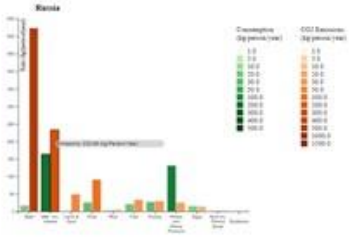

Page opens to initial overview
Mousing over map displays food category name and suggests clicking to zoom in
Mousing over bars shows precise values
Food Type: Rice
Click to zoom in
Consumption: 29.38 (kg/Person/Year)
Rice
Fish
Poultry
Eggs
Clicking a map or a bar moves the visualization to the zoomed in view
Mousing over on the map shows precise values now



Clicking changes the data displayed in the bar chart



Clicking a bar changes the data displayed on the maps to a different food type



Bar chart can be zoomed and panned

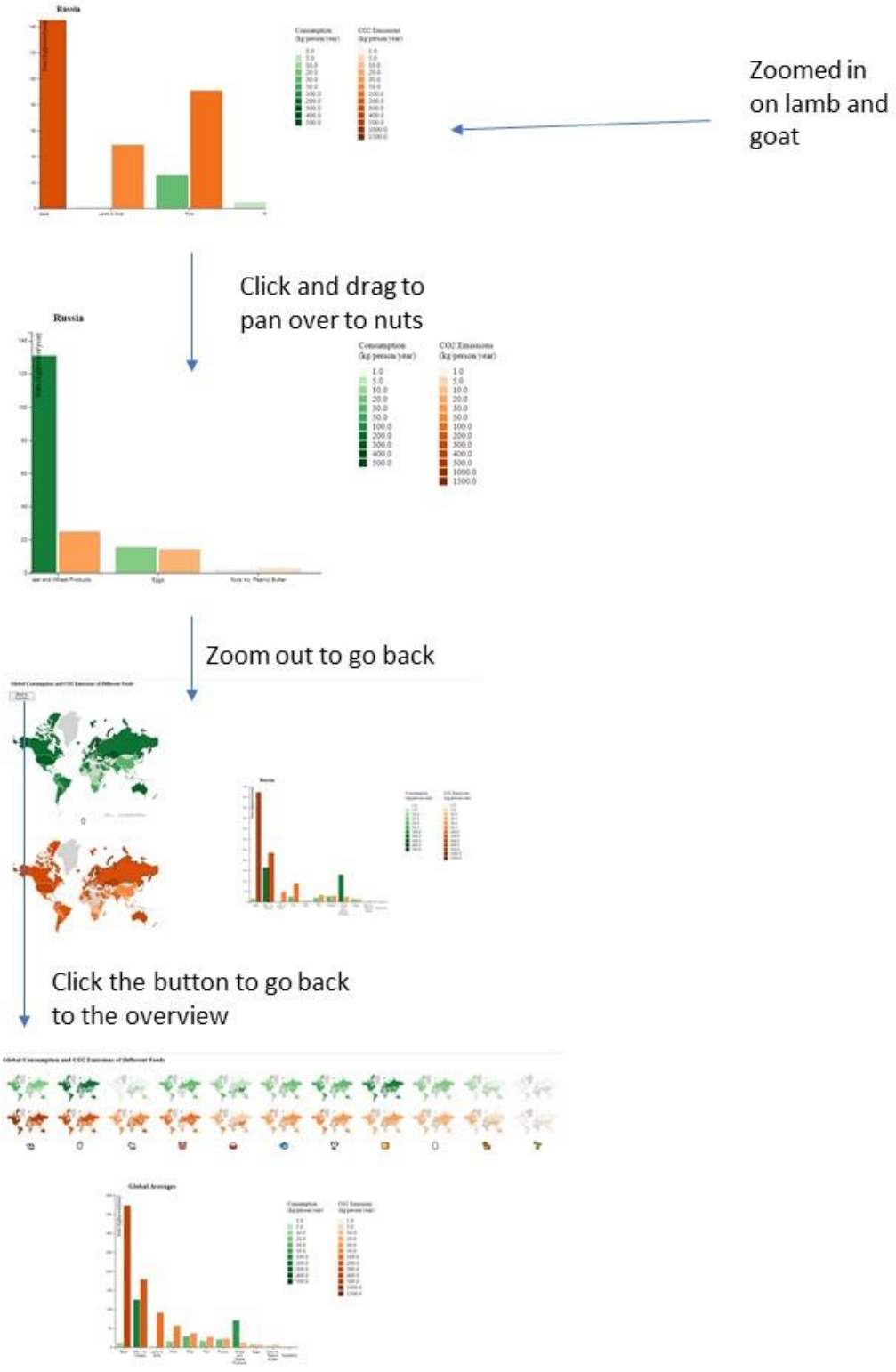


Figure 19. Outlines how the final interactive visualization looks and works.

By using the small multiple maps initially, the user can get a quick overview of how the consumption and CO_2 emission rates relating to different foods vary across the world. The bar chart further allows a quick way to see globally how the rates of consumption and emissions compare for different food sources. This allows the user to see how high the CO_2 emissions for beef are relative to the other food categories, and relative to consumption. By looking at the maps, it is also clear that globally, dairy and wheat are the two food categories with the highest yearly per capita rates of consumption. Through the use of interactions, the user can then explore the dataset, zooming in on maps of interest and clicking on countries to change the bar graph so that the rates of consumption and CO_2 emissions can easily be compared. Overall, the visualization allows the user to get a quick overview initially, and then allows them to explore through zooming, selection, and details on demand by mousing over the objects. This means that if someone is busy and not paying attention, they can get the important ideas from the overview, but if they are interested then the interactions allow them to dig deeper into the dataset.

4.1 Process Reflection

Starting out, I was focused on the idea of representing the data as footprints, which is what inspired my first three visualization prototypes. However, as I made these prototypes, I started to realize some of the problems with using circles. While we had discussed several times in class how people find it hard to discern differences between circle areas, I wanted a creative and emotive visualization. However, after the first prototype, I realized how messy this design was (Figure 8), and I decided it was not the most effective way of representing the data. From conversations with my peers, I also decided that one of the most effective ways to show geographical information was through maps, as these are highly intuitive. This led me to my third prototype (Figure 10), which is what I elaborated on for my final static visualization. In this one, I held on to the footprint idea and continued to use circles to show differences between food categories. However, moving into the final static design, I was inspired by class discussions to use a bar graph instead because lengths allow for much more accurate visual comparisons than areas. Additionally, after talking about interactions in lecture 19, I realized I needed to start with a broad overview and then zoom in, which I had not done in my third prototype. As a result, for the final static visualization, which I also used as a starting point for the final interactive visualization, I made a grid at the top with small map multiples showing the entire dataset – rates of CO_2 emissions and consumption for 11 food types in 130 countries. Then because using colour scales from light to dark allows approximate comparisons but not exact, I also displayed a bar chart with the global averages for each food category.

Moving into the interaction design, because the maps are so small and just allow for a broad overview, I wanted the user to be able to zoom in on the maps. To do this, I decided to implement a click interaction. One thing I would like to improve in the future is to somehow make the available interactions more obvious so that the user knows what can be clicked on without adding chart junk. After the discussion in Lecture 24, a nice way to do this might be to add a small animation to interactive elements on mouseover to signal that they can be interacted with. I also added another way to get the zoomed-in view – clicking on a food category on the bar chart. This addition may be redundant, and if I had more time, it is something I would want to test with more people. Once zoomed in on the maps relating to a selected food type, I wanted the user to be able to see the data for each country in more detail. This is why I decided to implement a click function on the map that changes the data in the bar chart. I also wanted the user to be able to change the data displayed in the map without going back to the overview which is why I implemented an additional click function on the bars to change the map. Finally, I was not initially

planning on implementing a zoom and pan function on the bar chart, but after playing with my visualization, I was disappointed by how hard it was to see the food categories with smaller values like nuts and soybeans. This function is something else I would want feedback on if I had more time, as I am not entirely sure zooming in on the y-axis of the bar chart makes sense. Overall, for my interactions I implemented a lot of click functionality as it is something that I find intuitive but could probably use more testing with other people. After the class discussion in Lecture 23, in the future, I would also really like to add some small animations when transitioning between views to increase user engagement.

Altogether, I changed the interactions from what I had in my third prototype after being inspired by the "Interaction Mantra" introduced in Lecture 21. I started with a broad overview showing all of the data points, then added a zoom function to show maps of interest. Then I further added functionality to select which country is displayed in the bar graph and which food category is displayed on the maps in the zoomed-in view. Because I found it was hard to see the smaller food categories on the bar chart, I further implemented zooming and panning, which was not something I had initially intended.

5. Discussion

When I began sketching, I tried to explore as many different representations as I could. When it came to selecting one to focus on for the visualization prototypes, I was drawn to the more creative and emotive sketches that I had done. As a result, I wanted to design bubble charts shaped into footprints to represent the differences between consumption and CO_2 emissions. As I learned more about maintaining data integrity, which visual variables make for the best encodings, and the need to make sure visualizations are effective above all else, I came to realize that I had not made the optimal choice. In my first prototype, I had these bubble charts shaped into large circles, but with a large number of nodes and the challenge of judging differences in circle sizes, the only really clear thing was that beef emits a lot of CO_2, and generally a higher mass of CO_2 is emitted than the mass of food which is consumed. It also did not display the geographical information, which is one of the interesting aspects of the dataset. As a result, for the final static visualization, I chose to focus on my third prototype which combined a map with a bubble chart. However, when I initially designed this prototype, I did so before learning about the optimal ways to use interaction. Instead of showing a good overview of the entire dataset, I started with a view of the world averages. This meant that to get a good sense of the data, the user would have to spend time using all of the interactions rather than getting a strong initial overview. Because of this, for my final static visualization, I was inspired by the class discussions about small multiples, so I decided to use 22 small maps to display all of the data points, with two maps for each food source, one showing emissions and one showing consumption. Then because colour intensity allows more for relative comparisons than exact, I added a bar chart with the global averages of consumption and emissions for the different food sources. In this way, I used length to allow viewers to quickly compare the rates of CO_2 emissions and consumption. For interactions, I then implemented zooming in on a map of interest, and then further selection of countries to change the data displayed on the bar chart, and details on demand using a tooltip on mouseover. I also enabled zooming on the bar chart to allow the smaller categories to be seen.

A large advantage of the final static visualization that I created is that by using small map multiples, all of the data is displayed at a glance. The bar chart then further emphasizes the differences between the rates of consumption and CO_2 emissions for different food sources and allows comparison using length, one of the better encodings for quantitative variables (Mackinlay's Rankings, Lecture 10). By using the small multiples, data density is also increased, both important factors according to Tufte's Graphical Theory

(Lecture 21). They show a lot of data in a small amount of space, and it does not take long to get an idea of how rates of consumption and CO_2 emissions of different food sources change around the world. This leads to the advantages of the final interactive visualization. It starts with a view of the final static visualization to allow a quick overview of the dataset. I then tried to make use of the Visual Information Seeking Mantra from Lecture 19. From the overview, I enabled zooming in on a map of interest, then selection of a particular country and/or food source, then detail on-demand using the tooltip. A huge advantage with this design is that if someone is not interested in spending time interacting, they can get a valuable general idea of the data from a glance at the overview. Then if they are interested, and have questions, it is possible to use the interactions to further explore.

One disadvantage of this visualization is that the combination of the bar graph with the maps may be redundant. However, I do think that the length of the bars allows for more accurate comparisons between values, while the map is best for conveying the relative spatial information using colour intensity. Another disadvantage is that I am not certain how intuitive it will be to other people that interaction is available. In the future, this could potentially be fixed by emphasizing the interactive elements somehow on mouseover. Another possibility is adding text instructions, although this may be construed as chart junk. Finally, one last disadvantage is that on the average computer screen, the small map multiples are quite small. As a result, this visualization might benefit from a larger screen or a different layout for the maps. Another disadvantage of my final visualizations is that they are not overly creative and memorable. However, on the flip side, they make use of traditional plots that many people are likely to be able to interpret.

If I had more time and resources, I really would have liked to be able to iterate more in the design process as discussed at the beginning of the semester. I initially got caught up in the idea of doing bubble chart footprints, but upon finding them ineffective, doing some more sketching might have been beneficial. I would also really like to get far more people to look at my final visualizations to get feedback on what others find effective and not effective. Particularly in terms of the interactions, it would be really helpful to get people to play around with them and then say how easy or challenging it was to navigate the visualization. As a smaller side note, with more resources, I would also design better symbols to represent the food sources than the emojis that I used. With more time, I also would have liked to experiment more with zooming in on the maps. Instead of having the one bar chart on the side, it might be useful to have encodings change as the user zooms in. For instance, initially, a map could show global averages for emissions and consumption, but then as the user zooms in on a country, the bar graph for that country is overlayed.

One main insight from this project is to not overlook the value of doing lots of different sketches. While it felt a bit tedious at the time, those sketches helped me as I moved forward and tried to figure out how to translate my ideas into code. Using subsets of actual data while I was sketching also gave me a really good feel for the dataset and how different designs might work. I have also really come to understand the value of an iterative design process, with lots of work at the beginning of the project figuring out how best to represent the data. Without all of this work at the beginning it would have been challenging to start coding effectively. While I was coding, I also developed an appreciation for how much online support there is for making visualizations in D3. However, I was often also frustrated by how incompatible different versions of the library can be. Moving forward, something I want to work on is finding the balance between creativity and efficacy. Figuring out where to make trade-offs is something I struggled with while doing this project and is something I will need to be mindful of in the future. Finally, while I learned a lot about

D3, I did not explore a lot of the animation functionality, which is something I would like to make use of moving forward, given that when used well, they can increase both efficacy and engagement (Lecture 23).

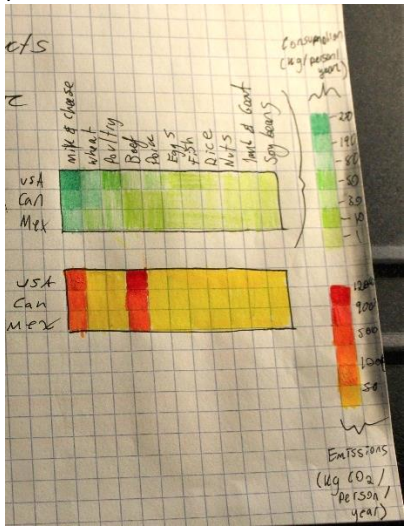
6. Conclusion

Over the course of this project, I have followed the design principles discussed in class to create both a static and an interactive visualization of a dataset on per capita rates of consumption and CO_2 emissions of different food sources around the world. My final visualization makes use of maps with colour scales to show the geographical differences, and a bar chart to allow for easy and accurate visual comparisons. I then implemented interaction to allow for zooming, filtering, and details on demand. Overall, I am very pleased with what I have created, although I know with a few more iterations of the design process, it could be improved. The final visualizations show all of the data, and the main trends are easy to see – beef produces the highest CO_2 emissions relative to the other food sources, and globally, wheat and dairy are two of the most consumed food sources. By completing this project, I have learned how to apply the design principles that were discussed in class and have gained an appreciation for the importance of all the steps along the design process.

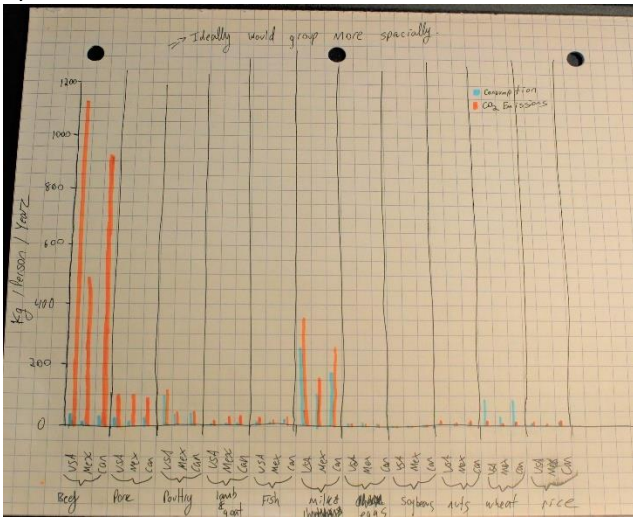
7. Sketching Figures

Subset 1 Sketches (Set 1)

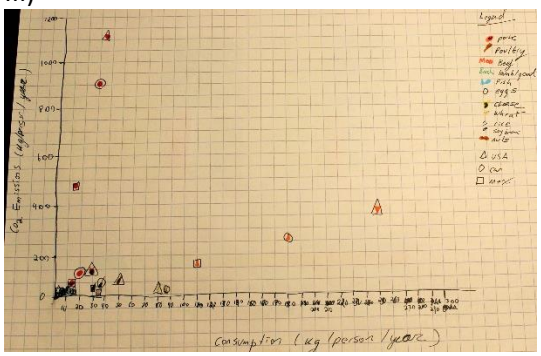
I)



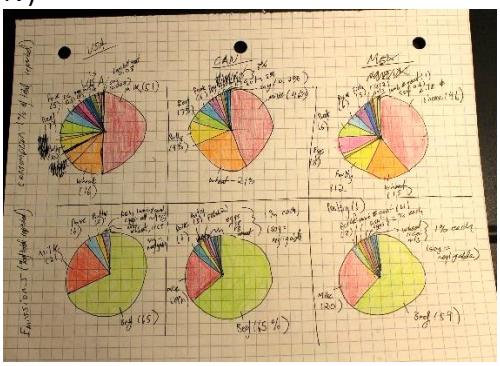
II)



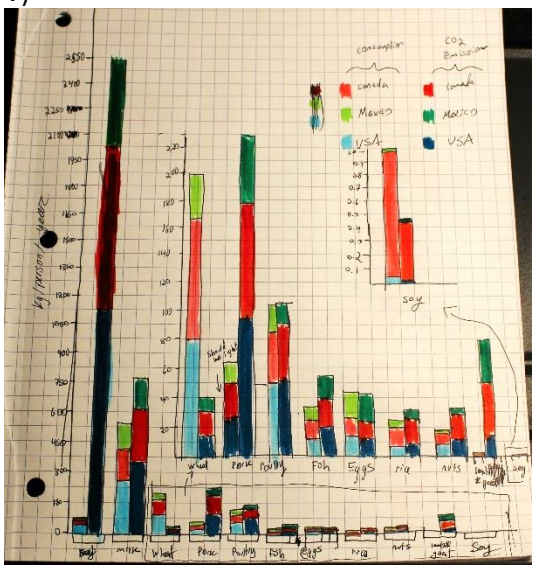
III)



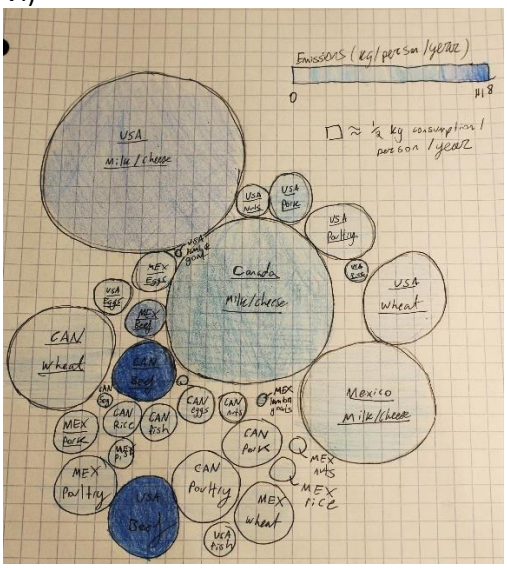
IV)



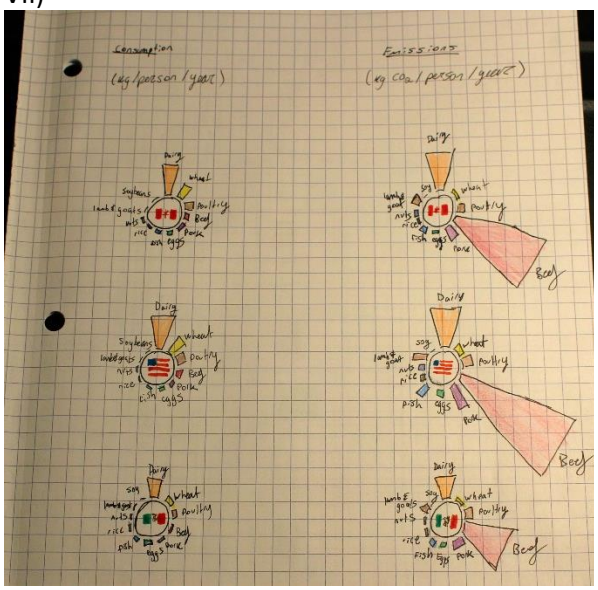
V)



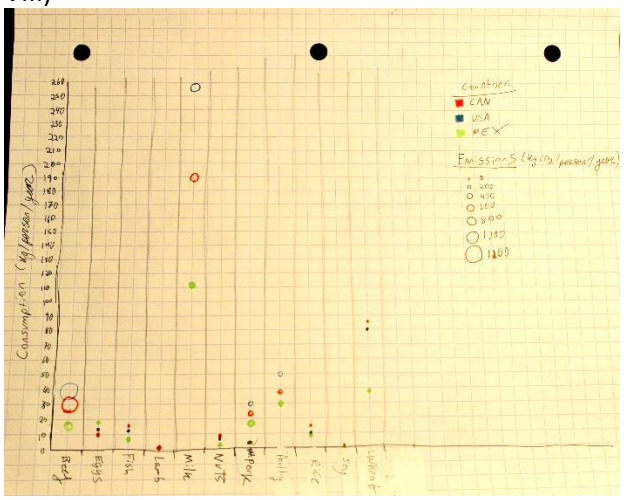
VI)



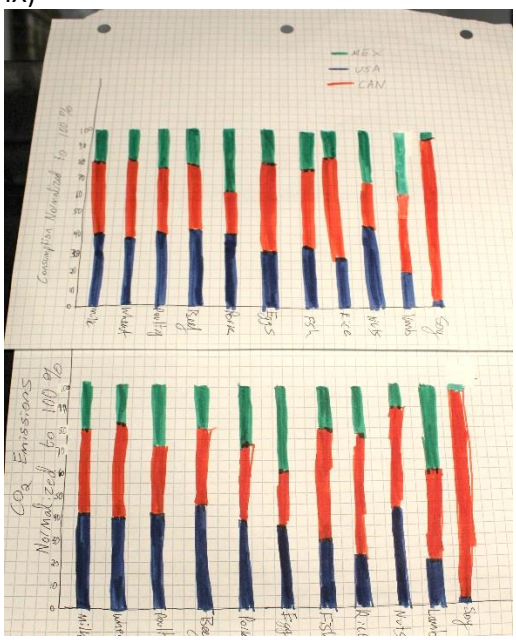
VII)



VIII)



IX)



X)

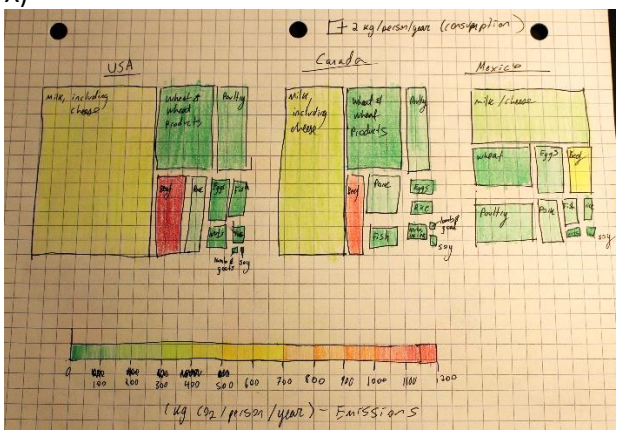


Figure 1. I-X. This subset was comprised of the CO_2 emissions and consumption for all eleven food groups, for the three largest North American countries – Canada, USA, and Mexico. These are the initial 10 sketches. In case the PDF images are unclear, here is a link to the jpegs: https://uofc-my.sharepoint.com/:f:/g/personal/sara_hall_ucalgary_ca/EhetpoyTORVFvVpxfmVF_mIBzPbSkYo5_KumJekDQbl42Q?e=MntsWN.

I. Shows a heatmap for the consumption of the different food types (top) and the related CO_2 emissions (bottom). Both are measured in kg/person/year, and darker colours indicate higher numbers.

II. Bar chart with countries grouped by food type. Blue lines show the consumption (kg/person/year) for the food type in that country and orange lines show CO_2 emissions.

III. Scatter plot showing consumption versus emissions, with icons indicated food type, and surrounding shapes indicating country.

IV. Pie charts for emissions and consumption in all three countries.

V. Radial charts for emissions and consumption in all three countries.

VI. Bar chart with light colours showing consumption, and dark colours showing emissions. The lower end of the chart is also shown on additional axes so that smaller differences are visible.

VII. Bubbles with size indicating consumption, colour darkness indicating emissions, and text labels showing the country and food type.

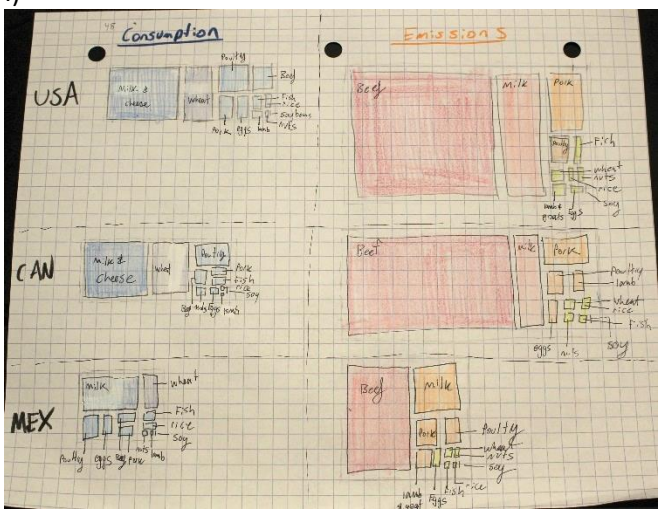
VIII. Uses position to show food type and consumption, colour to indicate the country, and size to indicate emissions.

IX. Stacked bar charts for both emissions and consumption of all food types, each normalized to 100% to allow for relative comparisons. Colours are used to differentiate between countries.

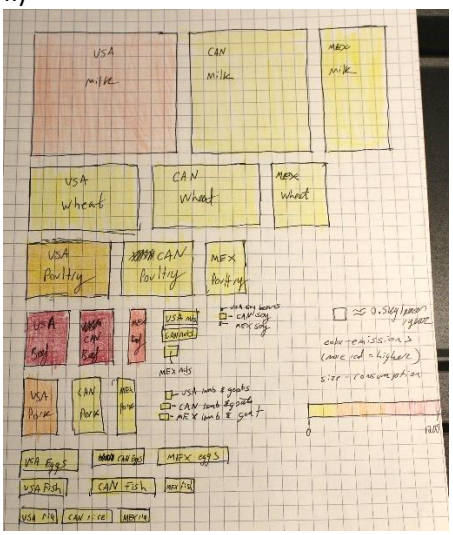
X. Uses rectangle size to show the magnitude of consumption, and colour to show the magnitude of CO_2 emissions. There is a grouping for each country.

Subset 1 Sketches (Set 2)

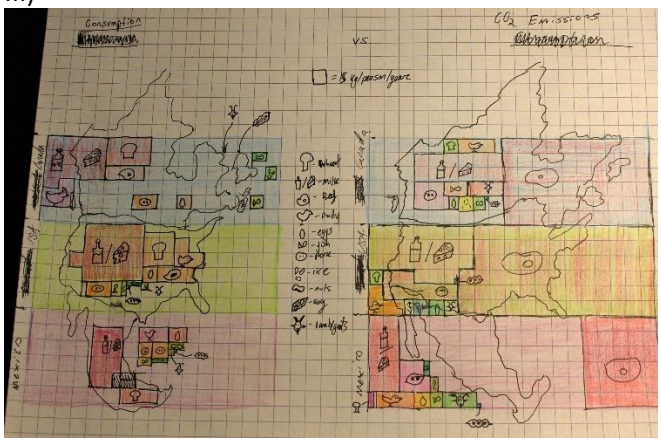
I)



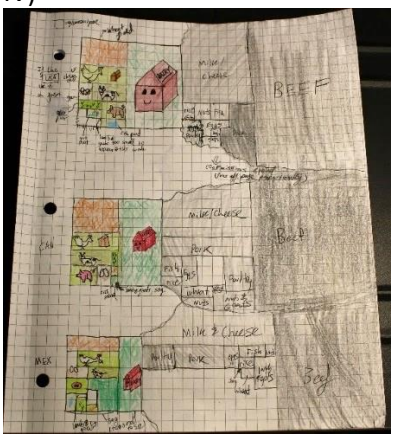
II)



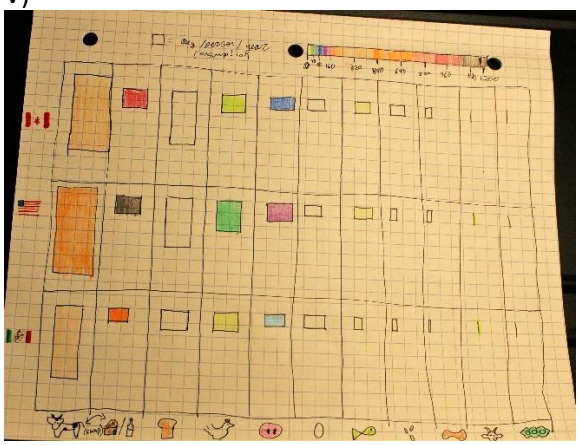
III)



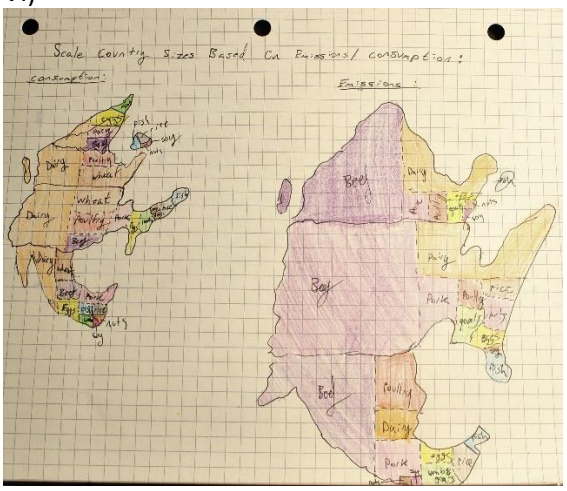
IV)



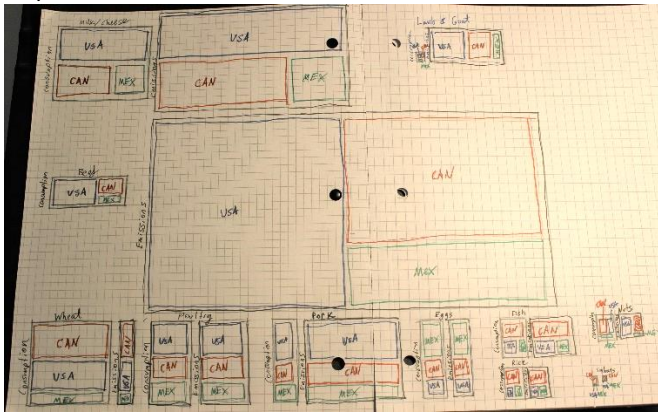
V)



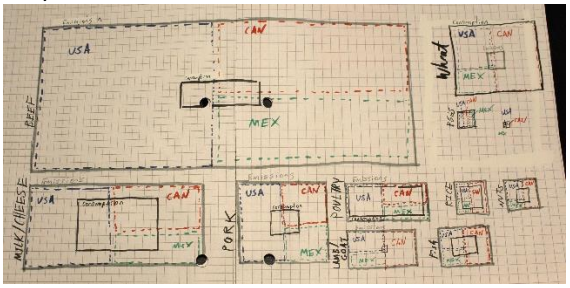
VI)



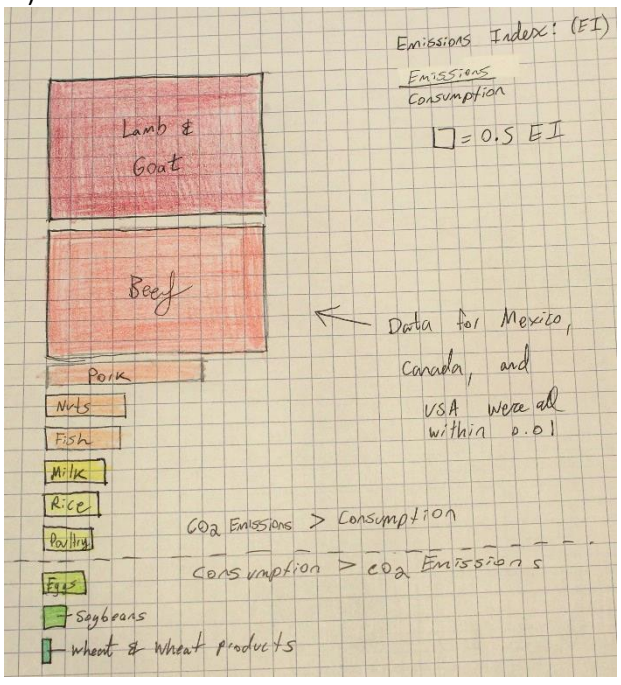
VII)



VIII)



IX)



X)

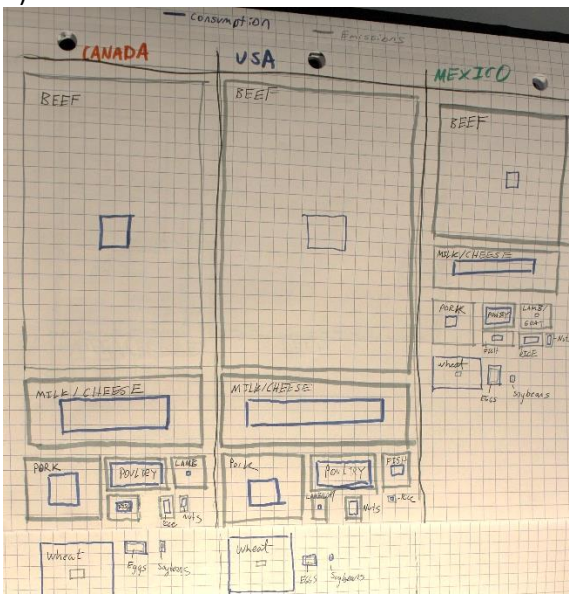


Figure 2. I-X. These are the subsequent 10 sketches for the USA/CAN/MEX subset. I decided to focus on elaborating Figure 1.X, using the size of rectangles.

- **I.** Rectangles are separated into groups for emissions and consumption for each country. The areas are proportional to the consumption (right) and emissions (left). Colour intensity is also used to emphasize differences.
- **II.** Rectangles are sorted from the highest rate of consumption (top) to lowest (bottom). The areas are proportional to consumption and colour indicates CO_2 emissions.
- **III.** Rectangles of equal size for each country are overlayed on top of the countries outline. Smaller rectangles inside them then have area proportion to consumption (left) and emissions (right). Both colours and icons are used to differentiate the different food types.
- **IV.** Squares on the left show the relative consumption of different foods for each country, with "CO^2 clouds" pluming out to the right, divided into sections with areas proportional to emissions.
- **V.** Rectangles with size relating to consumption and colour relating to emissions are placed on a grid with rows corresponding to countries and columns corresponding to food types.
- **VI.** Consumption is shown on the left and emissions on the right. Each countries size is proportional to consumption/emissions, and sections show the relative contribution of the different food types.
- **VII.** Rectangles for consumption and emissions of different food types are place side by side for comparison. Within each larger rectangle, smaller rectangles indicate the contributions of the different countries.
- **VIII.** Rectangles showing consumption (black) and emissions (grey) are overlayed, with the relative contribution of each country shown by smaller coloured rectangles.
- **IX.** Emissions were divided by consumption to get an index. This is then shown by both sizes of rectangles and colours for each different food type. Since these numbers were stable across the three countries, it was only drawn once to represent all of them.
- **X.** Three columns separate the data for the three countries. Within each column, the sizes of grey rectangles show emissions, while the sizes of blue rectangles show consumption. Food types are arranged generally from higher emissions at the top to lower at the bottom.

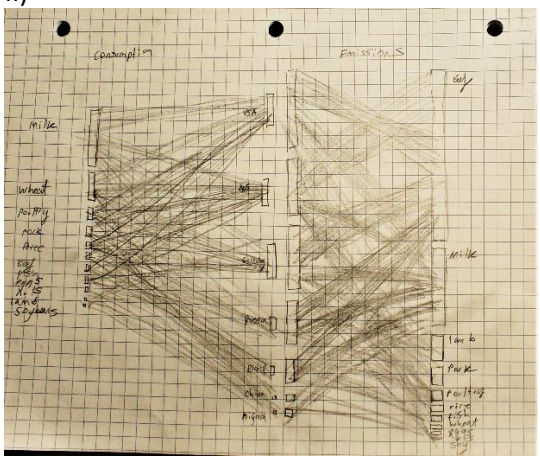
Subset 2 Sketches (Set 1)

I)

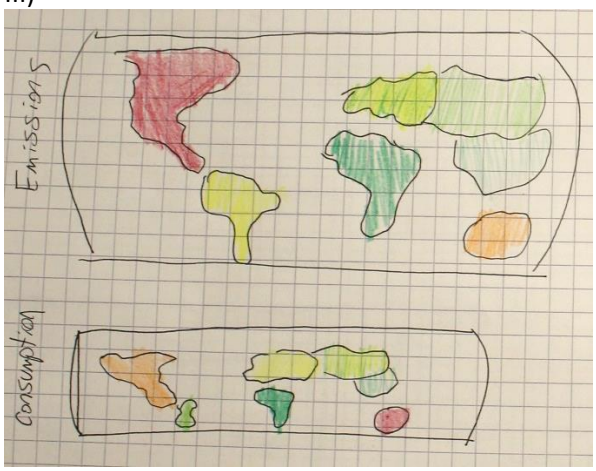
Consumption	Milk - inc. cheese	Wheat and Wheat Products	Poultry	Pork	Rice	Beef	Fish	Eggs	Nuts inc. Peanut Butter	Lamb & Goat	Soybeans	Sum
USA	254.69	80.43	50.01	27.64	6.88	36.24	12.35	14.58	7.86	0.43	0.04	491.15
Australia	234.49	70.46	46.12	24.14	11.03	33.86	17.69	8.51	8.73	9.87	0.19	465.09
Germany	258.7	83.41	17.75	51.81	3.34	13.16	11.16	12.2	7.96	0.73	0.88	461.1
Russia	163.57	130.99	27.5	25.66	4.71	16.98	20.82	15.43	1.62	1.4	0.05	408.73
Brazil	149.28	53	45	12.6	32.13	39.25	10.01	8.98	0.67	0.62	3.62	355.16
China	32.66	63.36	13.2	38.43	78.18	5.12	21.01	18.76	6.39	3.13	3.66	283.9
Nigeria	7.91	20.82	1.73	1.47	28.23	2.33	16.03	3.46	6.93	2.72	2.6	94.23
Sum	1101.3	502.47	201.31	181.75	164.5	146.94	109.07	81.92	40.16	18.9	11.04	

Emissions	Beef	Milk - inc. cheese	Lamb & Goat	Pork	Poultry	Rice	Fish	Wheat and Wheat Products	Eggs	Nuts inc. Peanut Butter	Soybeans	Sum
Australia	1044.85	334.01	345.65	85.44	49.54	14.12	28.25	13.44	7.82	15.45	0.09	1938.66
USA	1118.29	362.78	15.06	97.83	53.72	8.8	19.72	15.34	13.39	13.91	0.02	1718.86
Brazil	1216.19	212.63	21.71	44.6	48.34	41.12	15.98	10.11	8.25	1.19	1.63	1616.73
Germany	406.09	368.49	25.56	183.38	19.07	4.27	17.82	15.91	11.21	14.09	0.4	1066.29
Russia	523.97	232.99	49.03	90.82	29.54	6.03	33.24	24.98	14.17	2.87	0.02	1007.66
China	157.99	46.52	109.61	136.02	14.18	100.05	33.55	12.08	17.23	11.31	1.65	640.19
Nigeria	71.9	11.27	95.25	5.2	1.86	36.13	25.59	3.97	3.18	12.27	1.17	267.79
Sum	4534.26	1568.69	661.87	643.29	216.25	210.52	174.15	95.83	75.25	71.09	4.98	

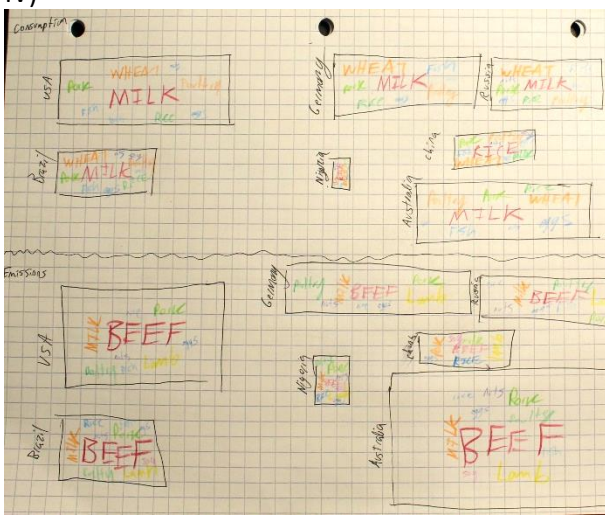
II)



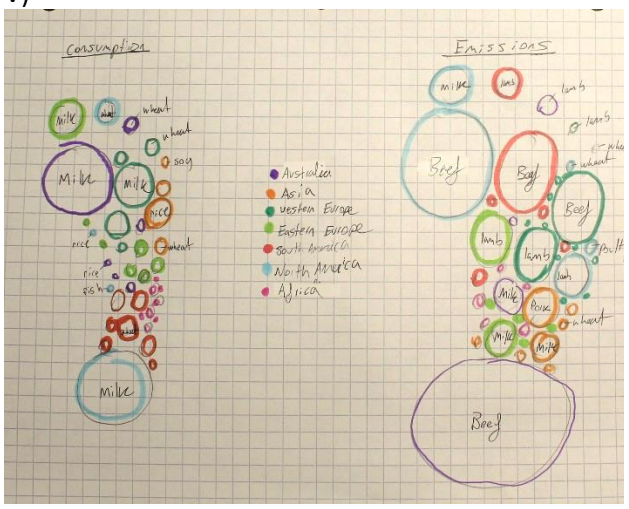
III)



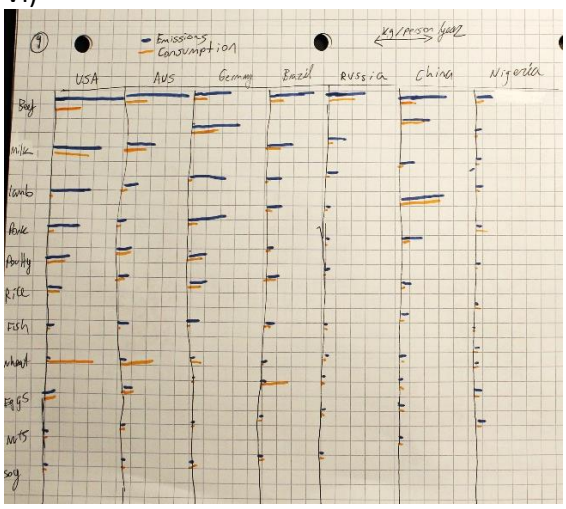
IV)



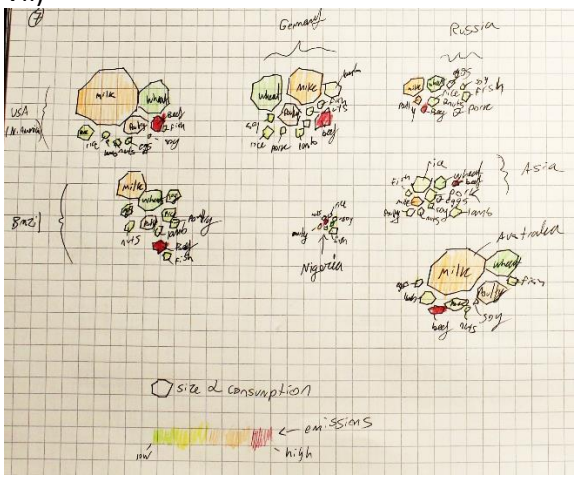
V)



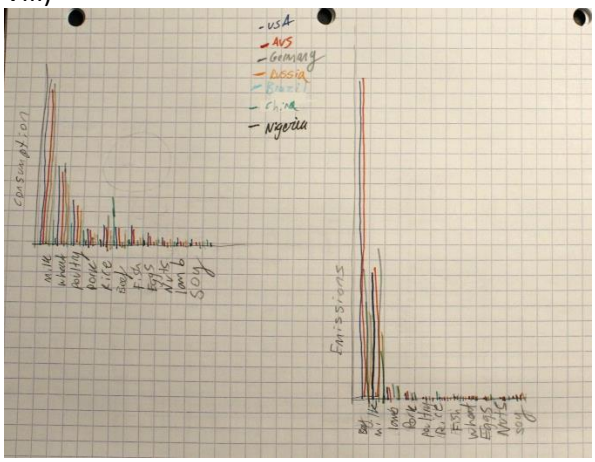
VI)



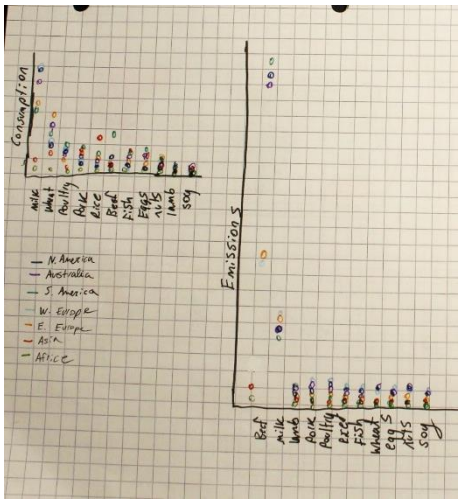
VII)



VIII)



IX)



X)

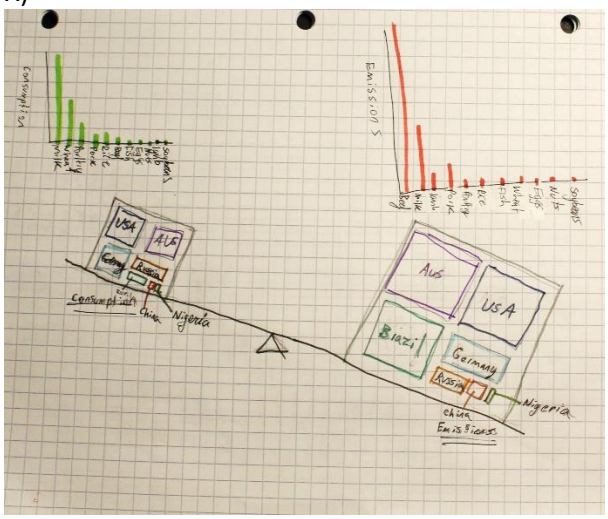


Figure 3. I-X. This subset looks at both consumption and CO_2 emissions for all eleven food types in the most populated country from each continent.

- **I.** Heat maps for both consumption (top) and emissions (bottom). Warmer colours indicate higher rates, while cooler colours indicate lower.
- **II.** Flow chart showing how each food type is consumed by the different countries, and then how much CO_2 emissions they produce.
- **III.** A very rough global heatmap for both consumption and emissions. Each country is considered "representative" of its continent, and the colour of the continent shows low rates as being greener, and high rates as being redder.
- **IV.** Word clouds for both consumption and emissions in each country. The size of the word is proportional to the rate of consumption/emissions. Colour is also a rough guide, with warmer colours indicate higher rates.

V. Footprints comparing consumption (left) vs. emissions (right). Circle size is proportional to rates, and colour outlines indicate the country. In scaling up, continents could each have a country, and mousing over could show the country name the bubble refers to.

VI. Columns show food type and rows show country. Each cell has a bar chart with orange showing emissions and blue showing consumption.

VII. Hexagons for each food type grouped into the shape of each country, and placed spatially on the page, have size proportional to consumption and colour showing emissions, with warmer colours indicating higher CO_2 emissions.

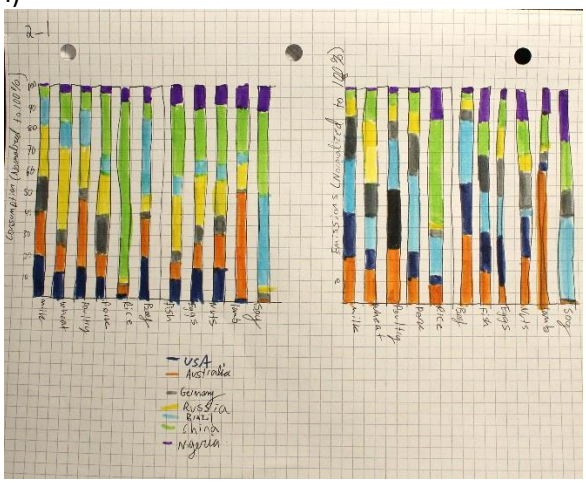
VIII. Separate bar charts for consumption (left) and emissions (right), countries are grouped by food type, and colours are used to differentiate the countries.

IX. Boxes on a scale are proportional to consumption (left) and emissions (right). The contribution of each country is shown within the box. Separate bar charts also show how different food types contribute to global emissions and consumption.

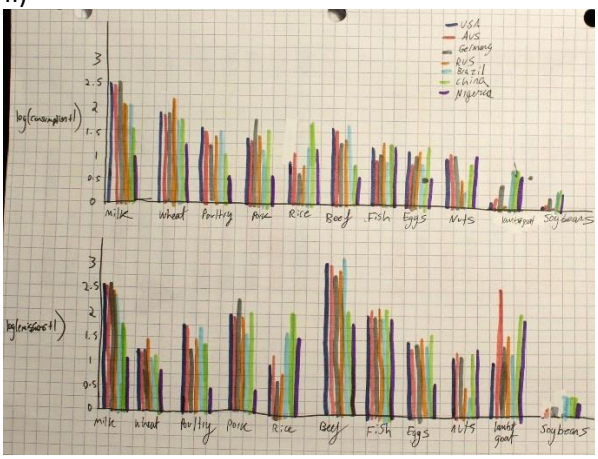
X. Scatter plot with horizontal position indicating food type and vertical position indicating the rate of consumption/emissions. Colours are used to differentiate between continents (with the largest countries being considered representative, here). Again, would scale possibly by allowing the user to mouse over and see the country names.

Subset 2 Sketches (Set 2)

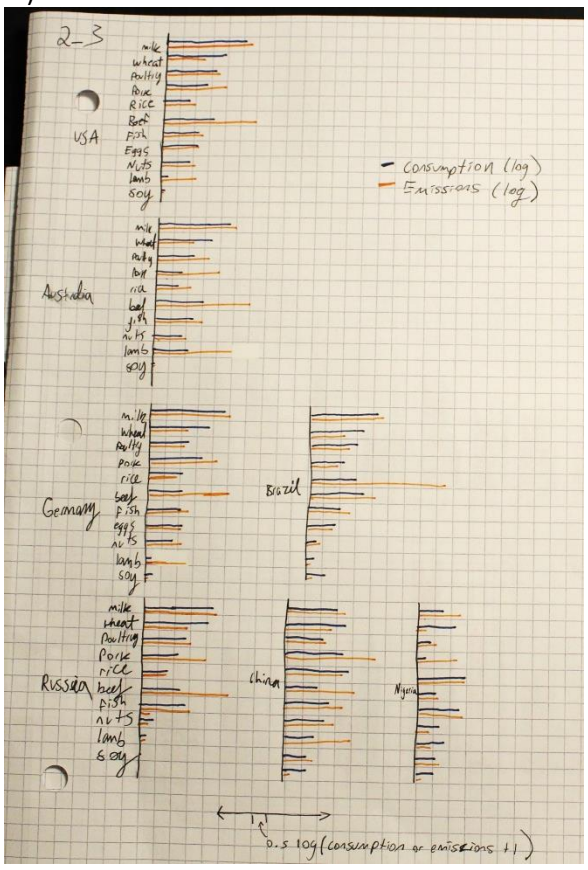
I)



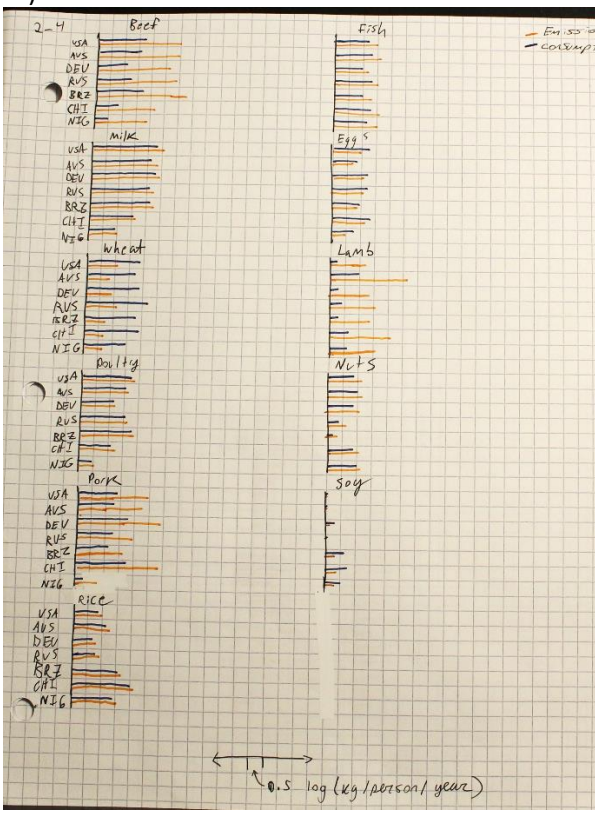
II)



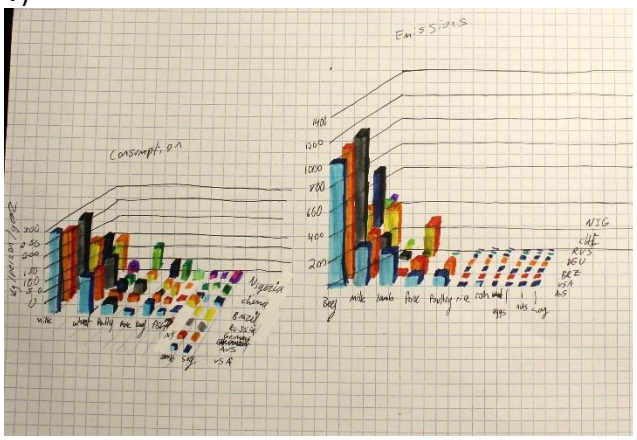
III)



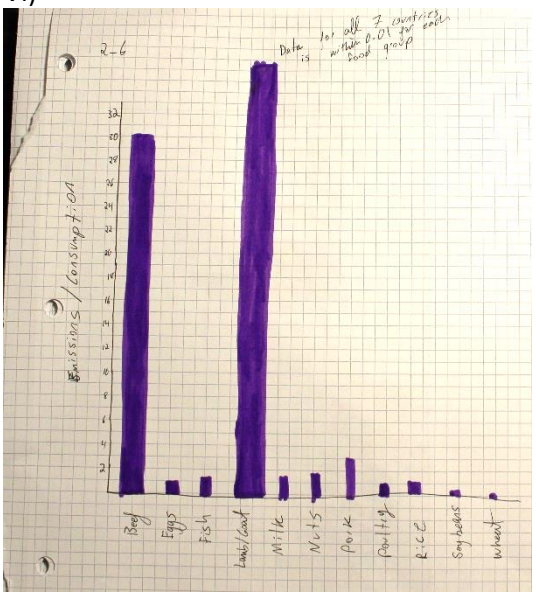
IV)



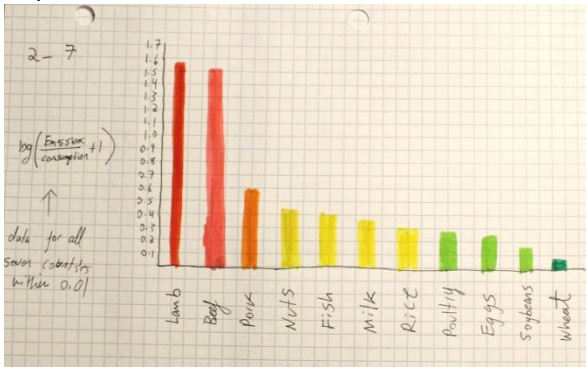
V)



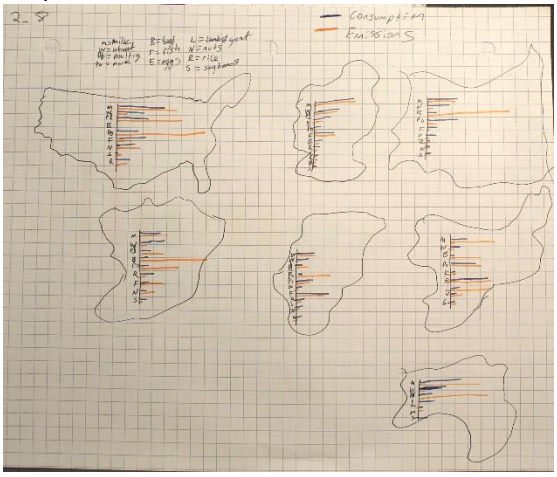
VI)



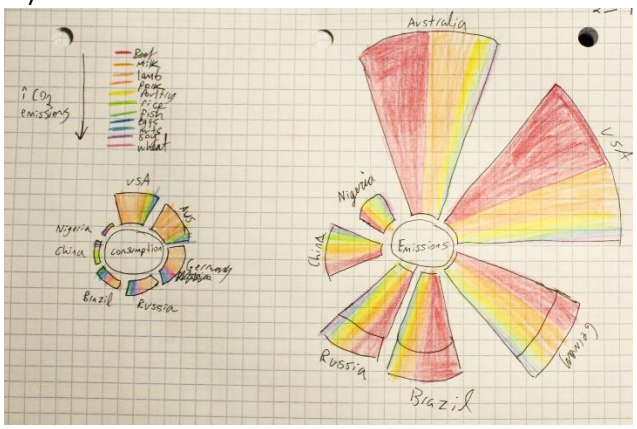
VII)



VIII)



IX)



X)

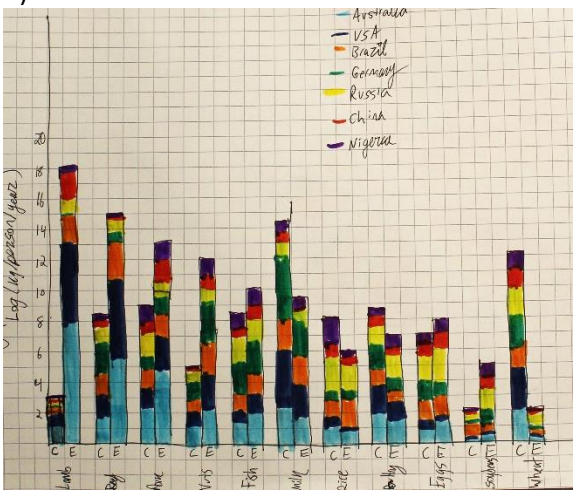
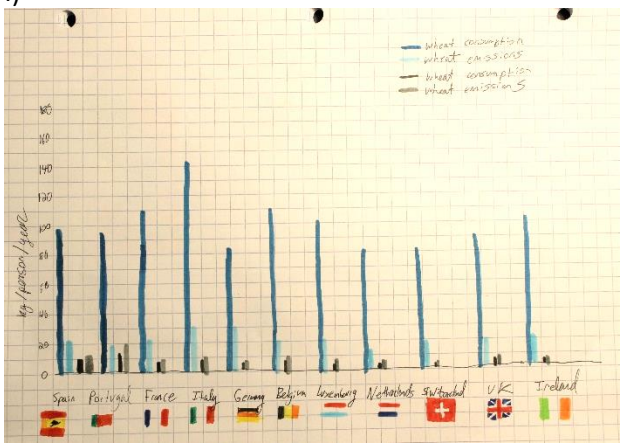


Figure 4. I-X. For the next 10 sketches, I decided to elaborate on the bar chart idea shown in Figure 3.VIII.

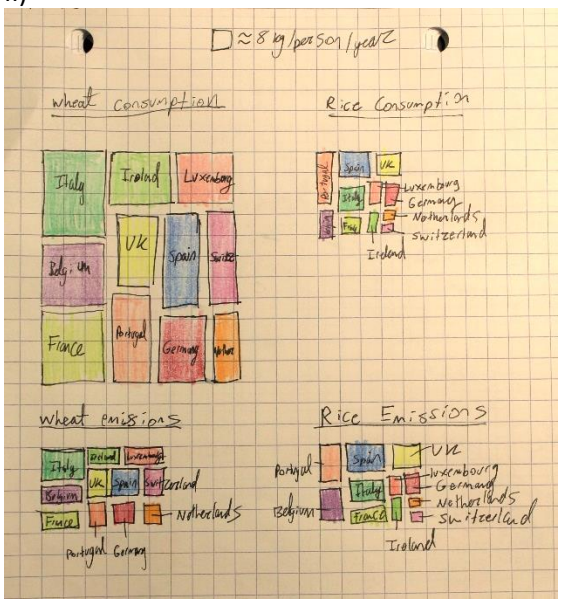
- **I.** Uses separate axes for consumption and emissions and normalizes each to 100%. The position on the horizontal axis shows food type and stacked bars use colours to differentiate between countries.
- **II.** Shows emissions (top) and consumption(bottom). Both have a y-axis that is log(rate+1). One was added to avoid negative numbers. The horizontal position shows food type and the colour of bars show countries.
- **III.** The log of consumption and emissions are plotted on separate charts for each country, with vertical position showing food type and horizontal position show log(rate+1). Consumption is shown in blue and emissions are shown in orange.
- **IV.** Similar to 4.III, but instead the food types are each shown on different charts with vertical position showing food type.
- **V.** Three-dimensional bar chart using the additional dimension so that emissions and consumption can be compared between country and food type. Consumption and emissions are shown on different charts.
- **VI.** Bar chart showing Consumption/Emissions for each food type. These values were within 0.01 for all countries, so one chart is used to represent all seven.
- **VII.** Bar chart showing log(consumption/emissions + 1) for each food group. These values were within 0.01 for all countries, so one chart is used to represent all seven. Food groups are sorted from highest to lowest, and warmer colours show a higher index.
- **VIII.** Horizontal bar charts showing the emissions (orange) and consumption (blue) of each food type (horizontal axis), place on top of the country outlines, geographically arranged on the page.
- **IX.** Radial bar chart showing global consumption (left) vs emissions (right). Areas are proportional to rates for each country, and colours show the contributions of each food type.
- **X.** The stacked bar chart on a log scale vertical axis. Separate bars show emissions and consumption for each food type along the horizontal axis, and colours show the different countries.

Subset 3 Sketches (Set 1)

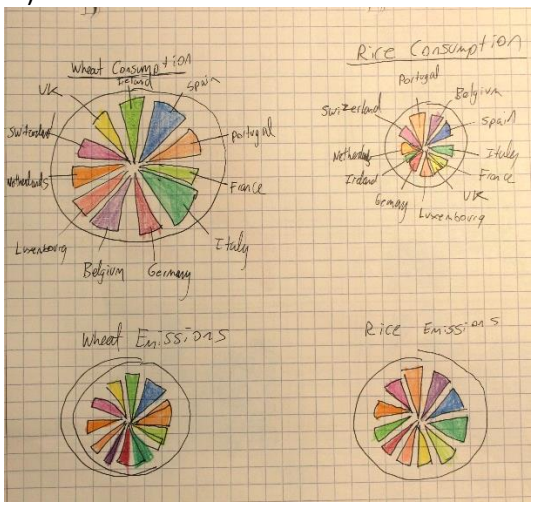
I)



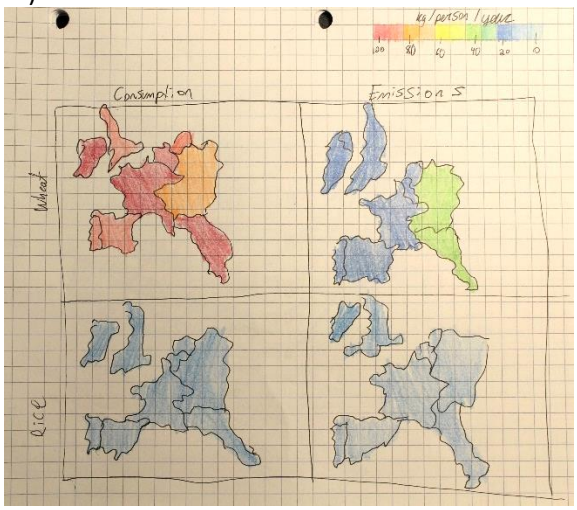
II)



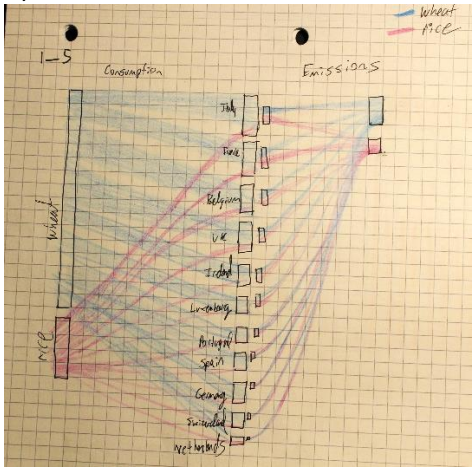
III)



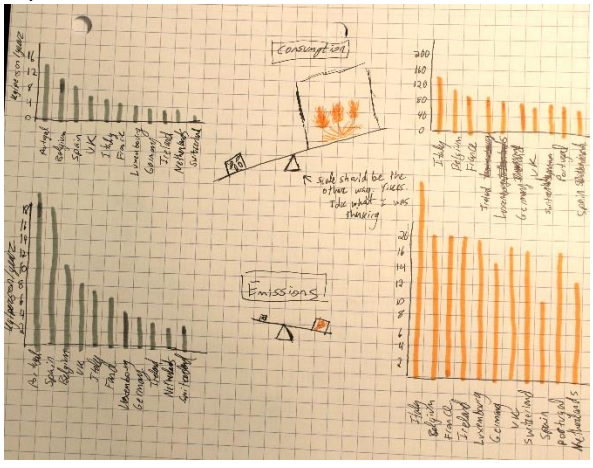
IV)



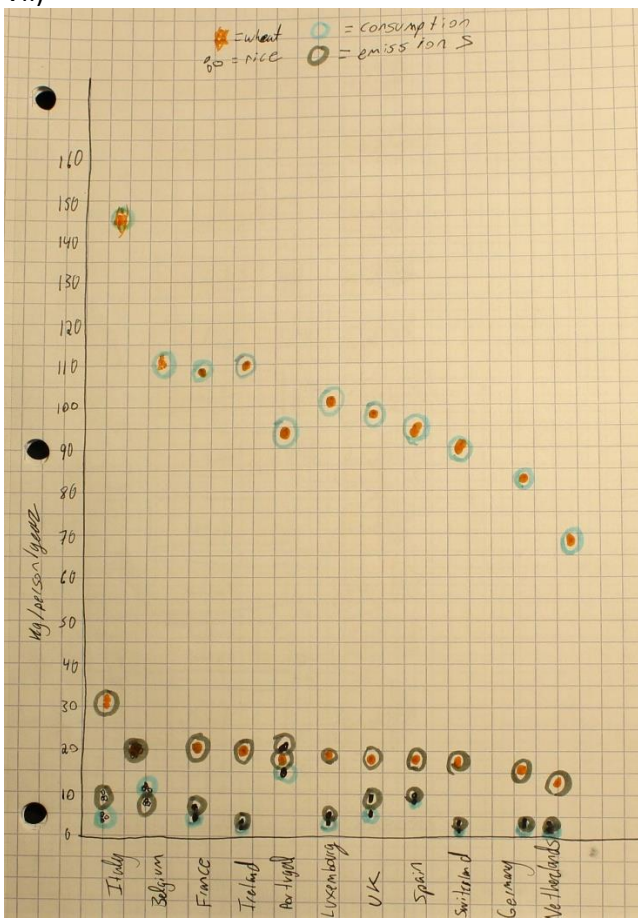
V)



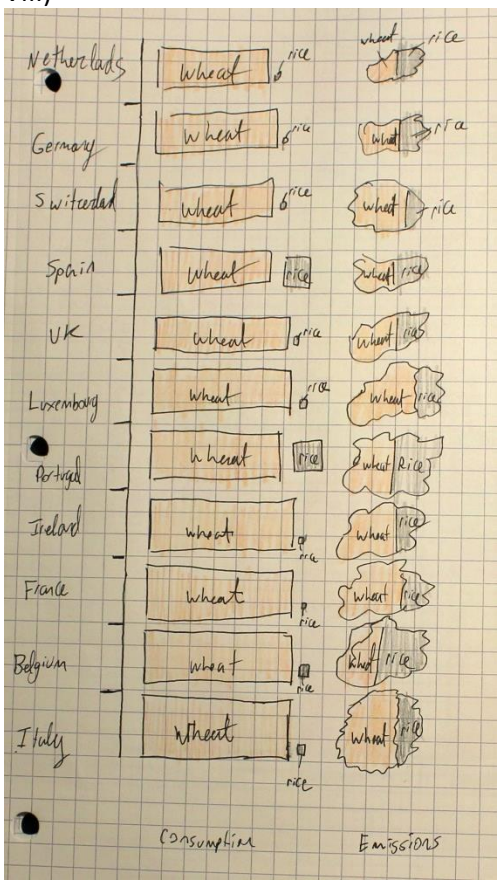
VI)



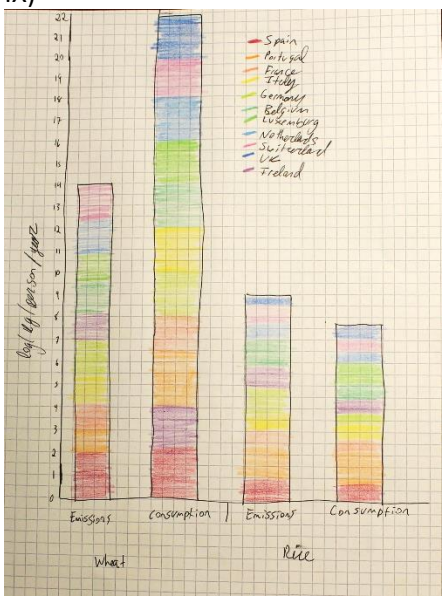
VII)



VIII)



IX)



X)

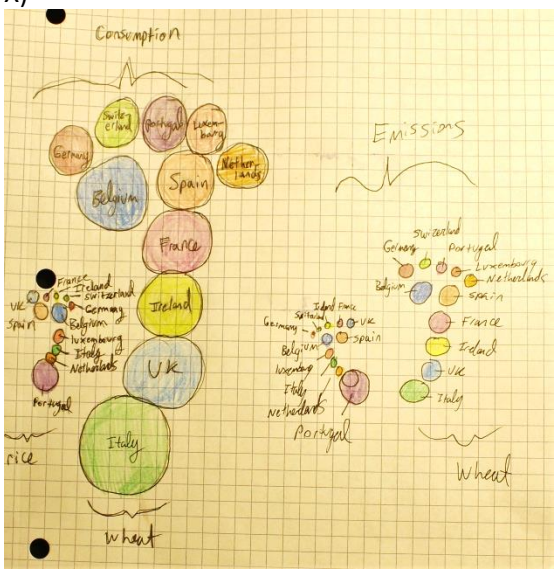


Figure 5. I-X. This subset looks at wheat vs. rice (the two grains in the dataset), for the eleven westernmost European countries.

I. Bar chart with countries on the horizontal axis and rates (kg/person/year) on vertical, blue shows wheat, grey shows rice, with lighter indicating emissions, and darker indicating consumption.

II. Four large rectangles show consumption and emissions for wheat and rice. These are subdivided into contributions by countries, which are colour-coded and labelled.

III. Pie charts showing wheat and rice consumption and emissions across Western European countries.

IV.

V. Flow chart showing the consumption of wheat and rice in different European countries, and then how they contribute to CO_2 emissions.

VI. Central scales show contributions of wheat and rice to consumption and emissions in Europe. Separate bar charts show the breakdowns in the different countries.

VII. Scatter plot with countries represented by position on the horizontal axis and rate (kg/person/year) on the vertical axis. Rice and wheat are represented by different icons and consumption vs emissions are represented with different colour circles around the icons.

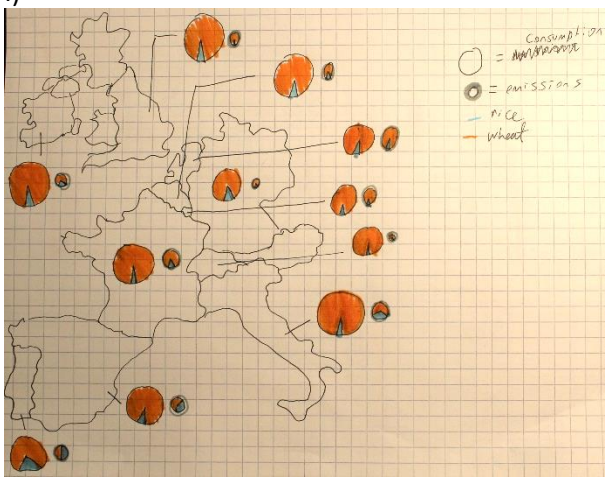
VIII. Rows are used to separate the different countries. Rectangles on the left show consumption, with orange being wheat and grey being rice. Clouds on the right represent emissions. Sizes are proportional to consumption and emissions, respectively.

IX. Stacked bars show different countries' (shown by different colours) contribution to emissions and consumption, groups under wheat and rice on the horizontal axis.

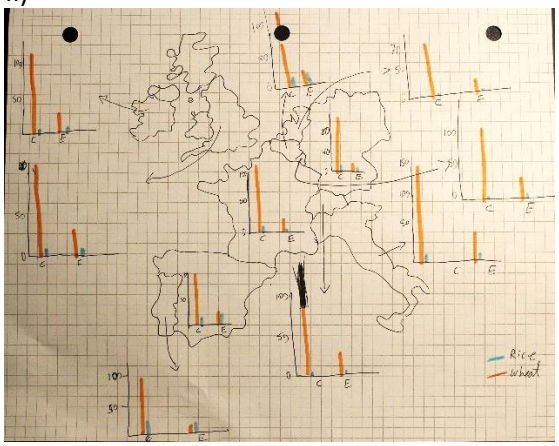
X. A set of footprints for consumption (left) and emissions (right). The left footprints of each pair are for rice and the right footprints are for wheat. Countries are both colour-coded and labelled with text. The sizes of the circles are proportional to rate (kg/person/year)

Subset 3 Sketches (Set 2)

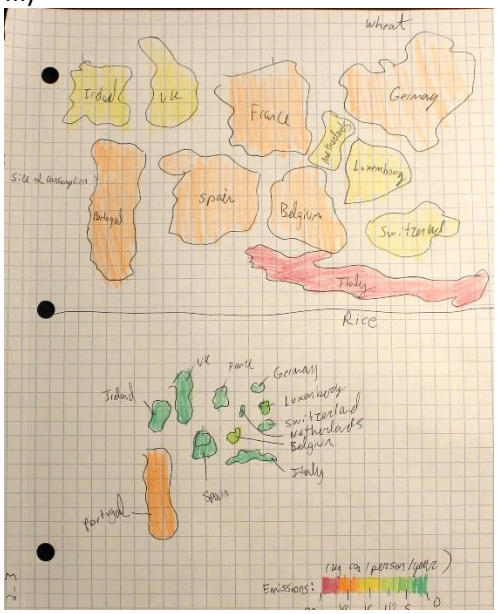
I)



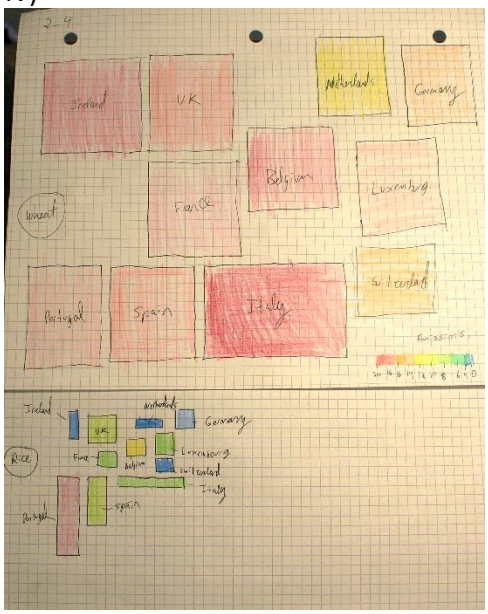
II)



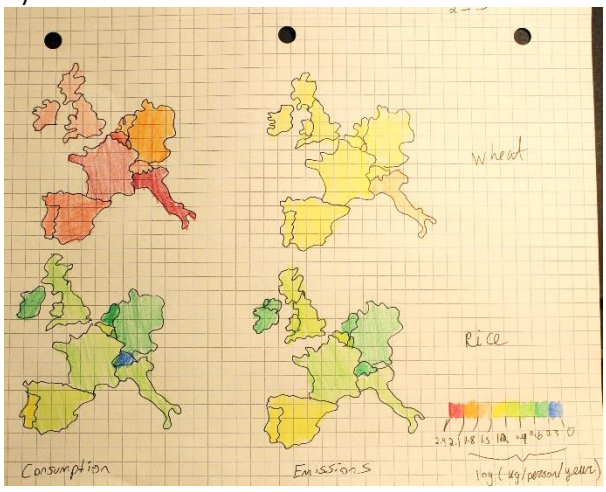
III)



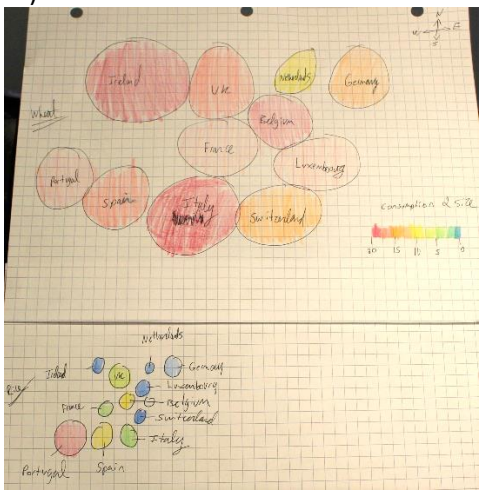
IV)



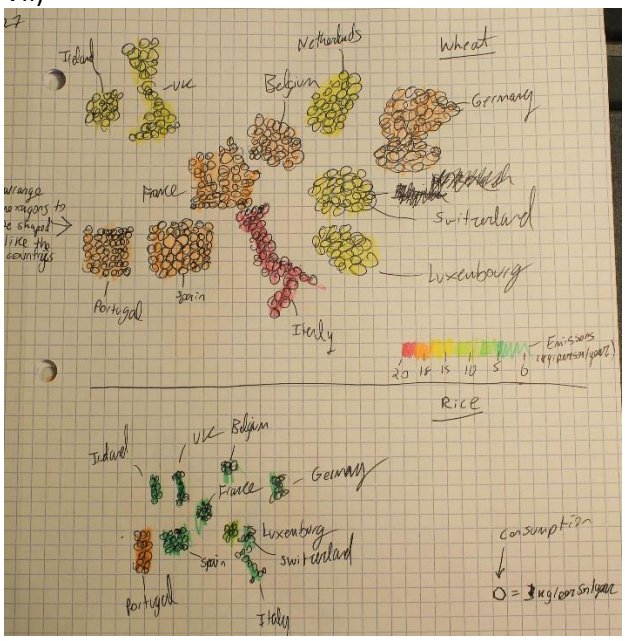
V)



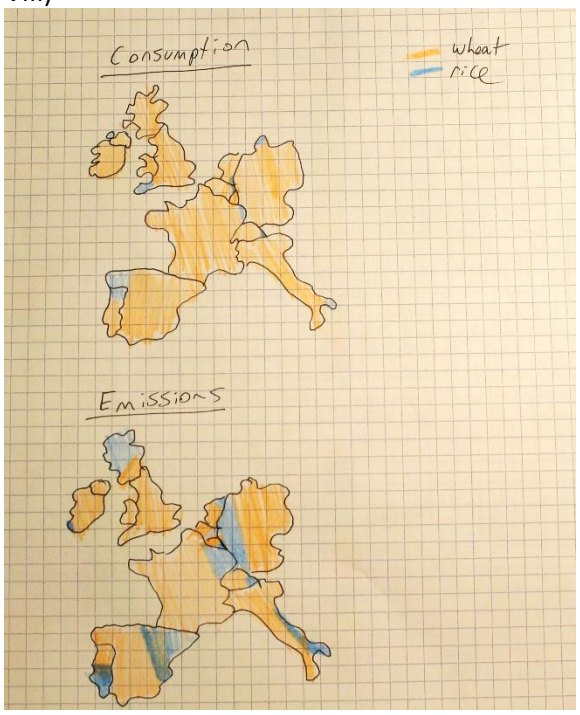
VI)



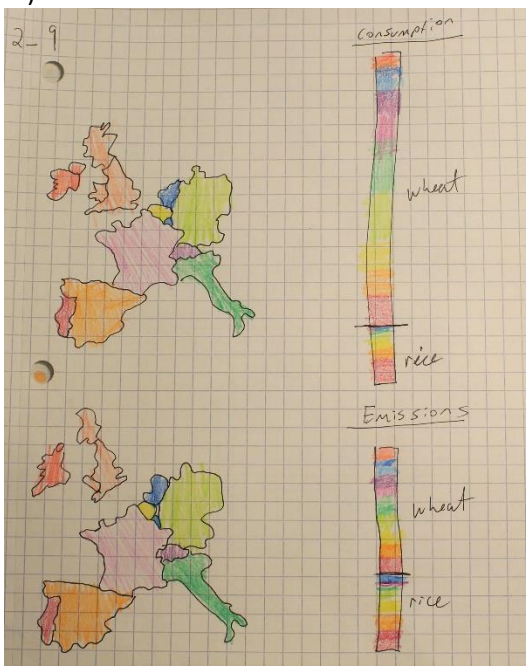
VII)



VIII)



IX)



X)

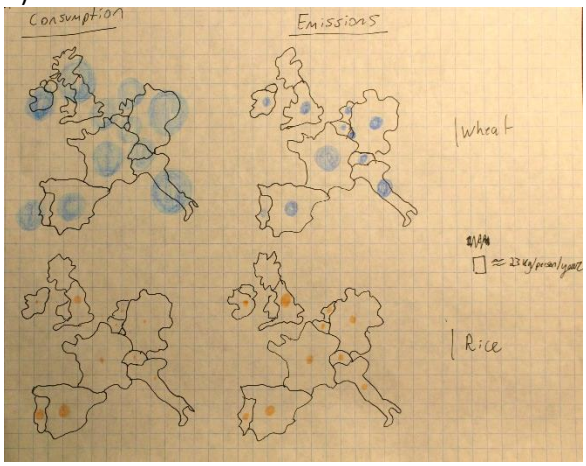


Figure 6. For the second set of ten sketches looking at Western Europe, I decided to focus on the geographical style representation (Figure 5.IV).

I. Each country on the map is associated with two pie charts, the areas of which are proportional to emissions (outlined in grey) and consumption (outlined in black). The percentage related to wheat is shown in orange and the percentage related to rice is shown in blue.

II. Each country on the map is associated with a bar graph. Consumption and emissions are on the horizontal axis, with the rate on the vertical axis. Wheat is shown as orange bars and rice is shown as blue bars.

III. The size of each country is scaled to show the consumption of wheat (top) and rice (bottom). The related emissions are then shown using colour, with warmer colours indicating higher emissions.

IV. Rectangles are used to represent countries instead of their actual shapes but are still placed approximately geographically. Wheat is shown on the top and rice is shown on the bottom. Areas of the rectangles are proportional to consumption and the related emissions are then shown using colour, with warmer colours indicating higher emissions.

V. Maps are placed on a grid with columns showing consumption and emissions, and rows showing wheat vs. rice. The magnitude of rates (kg/person/year) are using colour, with warmer colours indicating higher rates of emissions/consumptions.

VI. Circles are used to represent countries instead of their actual shapes but are still placed approximately geographically. Wheat is shown on the top and rice is shown on the bottom. Areas of the rectangles are proportional to consumption and the related emissions are then shown using colour, with warmer colours indicating higher emissions.

VII. Small consistently sized hexagons are used to roughly build the shapes of the countries, with the number of hexagons proportional to the rate of consumption. Related emissions are then shown using colour, with warmer colours indicating higher emissions. Wheat is shown on the top of the page, and rice is shown on the bottom.

VIII. The idea is sort of like a pie chart but using the shape of countries instead of circles. Actual magnitudes are not shown, but the relative contributions of wheat and rice to consumption (top) and emissions (bottom) are shown by colouring in the percentage of the country in orange (wheat) and blue (rice).

IX. Countries on the map are colour codes, and overall emissions and consumption are represented by proportional bars. The bars are then coloured proportionally based on how much each country contributes.

X. Translucent circles that are proportional to consumption (left), and emissions (right) are overlayed on top of the countries with which they are associated. Wheat is shown in orange on the top row, and wheat is shown in blue on the bottom row.

8. References

Food Carbon Footprint Index 2018 | nu3. (n.d.). Retrieved January 24, 2021, from https://www.nu3.de/blogs/nutrition/food-carbon-footprint-index-2018

Kulma, K. (n.d.). *tidytuesday/readme.md at master · rfordatascience/tidytuesday · GitHub*. Retrieved January 24, 2021, from https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-02-18/readme.md