PROCESS BOOK

May 5, 2021

CS573 FINAL PROJECT - EXPLORING TIME SERIES MACRO NUTRIENT DATA ACROSS THE GLOBE

Junying Li,
Jason King,
Ryan LaMarche,
Michael Osei

Contents

Overview and Motivation

Macronutrients, in modernity, are becoming a leading topic of discussion within our society. Due to several biomedical technological advancements, the introduction of non-organic foods, GMOs, and genetically engineered food products has become more and more widely spread throughout our communities, as well as globally. Global energy imbalances and related obesity levels are rapidly increasing [1].

A big concern moving forward is finding different avenues in which we can further educate, inform, and inspire our communities to live a more nutritious lifestyle. With that being said, the consequences of not understanding the nutritious lifestyle can not only be harmful on a micro level but can cause a severe impact on the macro level. Global obesity rates are steadily rising [2].

Globally, one of the biggest technological problems we will face comes in the form of providing nutritious foods to the world. As we move forward in achieving this goal we must communicate, inform, and showcase how nutrition is impacting our world today.

Related Work

Visualizations About Obesity

Website: https://ourworldindata.org/obesity

This website provides several visualizations on how obesity infects the death rate and the share of adults that are obese all over the world across time. The spatial visualization here inspired us to design an interactive world map.

Macronutrients And Obesity: Revisiting the Calories In, Calories Out Framework

According to this paper, recent clinical research has studied weight responses to varying diet composition, but the contribution of changes in macronutrient intake and physical activity to rising population weight remains controversial. To estimate the contributions of changes in macronutrient intake and physical activity to changes in population weight, they conducted dynamic time series and structural VAR analyses of U.S. data between 1974 and 2006 and a panel analysis of 164 countries between 2001 and 2010. Their structural VAR results suggest that, on the margin, a 1% increase in carbohydrates intake yields a 1.01-point increase in obesity prevalence over 5 years while an equal percentage increase in fat intake decreases obesity prevalence by 0.24 points [3], as we can see in figure-1.

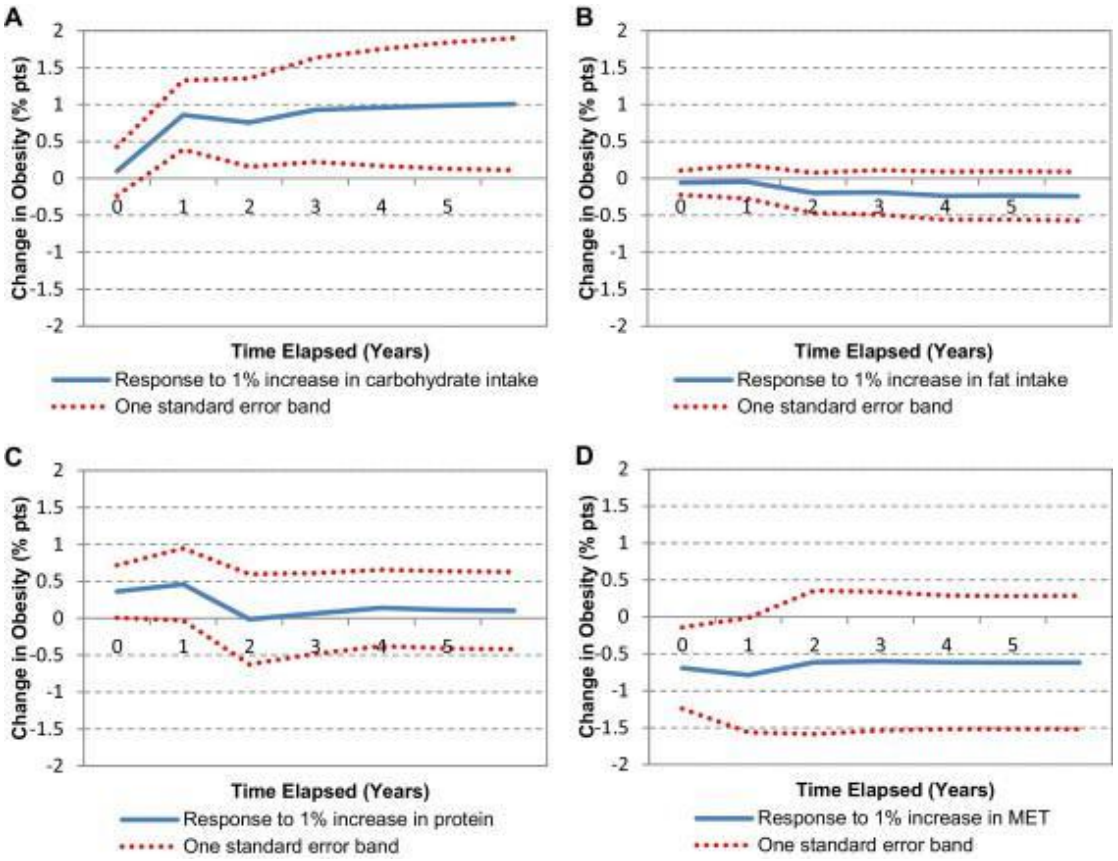

Figure – 1

This paper inspired us to design a time series visualization. In addition, the research result confirms that micronutrients play an important role in the change in obesity, thus we think this topic is very meaningful and worth to be discussed.

A New Approach for Visualizing Quantified Self Data Using Avatars

In this paper, the researchers described a new approach of visualizing Quantified Self data. The visualization of the information is proposed as an avatar that maps the different activities the user is engaged with, along with each such activity level, as graphical features. An initial prototype is shown in figure-2 [7].

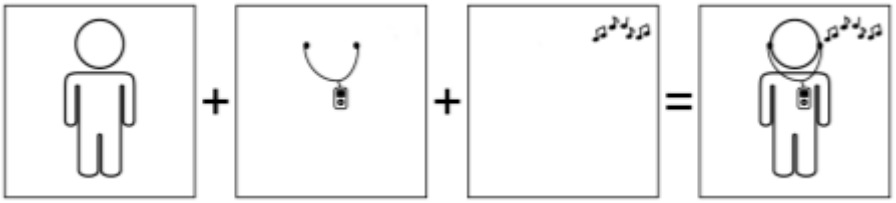

Figure-2

This paper inspired us to implement an SVG avatar that reacts to changes in the selected data points. For instance, when a user selects a country that eats a lot of fat and has a high obesity rate, the avatar's belly will grow in size proportional to the data. If the user selects a country that has a high amount of protein in its diet, the avatar will grow more muscular and healthier to reflect that. We believe this personification of visual elements will provide a more pleasant experience to users who are less interested in the numbers and statistics but would just like to explore the data at a very high level.

Questions

For most visualizations, there are limitations on how the users themselves can explore or interact with data, especially high dimension data. In addition, how to visualize high dimension datasets without overwhelming users always worth to be discussed. Considering obesity is becoming one of the leading risk factors for early death these years, we also very curious about how people's heath influenced by macronutrients.

Thus, in this project, we are trying to answer the following questions:

1. How do macronutrients influence people's overweight rate?

2. How to make new users, who do not have any level of expertise in nutrition or global health studies, are able to consistently explore the data using our visualization and draw meaningful conclusions?

Data

Macronutrients Data

Dietary compositions calculated by the OWID author (Hannah Ritchie) based on food supply statistics from the UN Food and Agricultural Organization database: FAOstat [4].

The FAO provides annual figures from 1961 by country on daily caloric supply, fat supply (in grams), and protein supply (in grams). To calculate the daily per capita supply of carbohydrates, we assumed an energy density by macronutrient of 4 kcal per gram of both protein and carbohydrate and 9 kcal per gram of fat (based on established nutritional guidelines reported by the FAO). The daily supply of carbohydrates was therefore calculated [5] as:

```
((Daily supply of kcal) - (Daily supply of protein * 4 + Daily supply of fat * 9)) / 4
```

The quantity of calories from each macronutrient was then calculated based on the energy density figures given above (e.g., calories from protein were calculated by multiplying the daily supply of protein in grams by 4).

The share of calories derived from each macronutrient could then be calculated by dividing the number of calories derived from a given macronutrient by the total daily caloric supply.

Protein of animal origin includes protein supplied in the form of all meat commodities, eggs, and dairy products, and fish & seafood. Protein of plant origin was derived as the difference between total protein supply and that of animal origin.

The original data set from FAO is a CSV file. In order to query and showcase datasets easier, we made use of MongoDB. Besides ingesting data from CSV to database, we also made a backend API for querying the data in different ways.

Overweight Rate Data

Data source:
https://ourworldindata.org/grapher/share-of-adults-who-are-overweight

The dataset is a CSV file. We made use of d3.csv() to read the data.

Exploratory Data Analysis

We use an interactive map with tooltips to initially look at our data. It's perfect to search information of one individual country, but it had limitations on presenting data across time, also it could not showcase the data of multiple countries. So, we decided to design a time series chart and link it to the map, then it would be perfect to show our data across time. Moreover, a stacked bar chart shall be helpful to present data of selected multiple countries. Thus, we decided to add them to our design too.

Design Evolution

In the beginning, we considered 5 different visualizations in our project, which are map, time series chart, stacked bar chart, and SVG avatar. The design evolutions of each visualization are described as follows.

Map

Since our dataset covers spatial data, according to the textbook [6] of our course, for spatial data, understanding and comparing the geometric shape is the common target of users' actions. Thus, we think designing a map would be a great showcase of the spatial data.

We created a wireframe sketch of the world map, as you can see in figure-3.

Figure-3

After that, we decided to add a title above and a timeline sketch below. The sketch of this design is in figure-4. It seems that we were getting closer to what we want.

Figure-4

Then, it occurred to us that tooltips shall be a great choice to show more information about the selected country. So we decided to add tooltips for the map. When we mouse over a country, the tooltip will show the country name and other information we would like to present, as figure-5 presented.

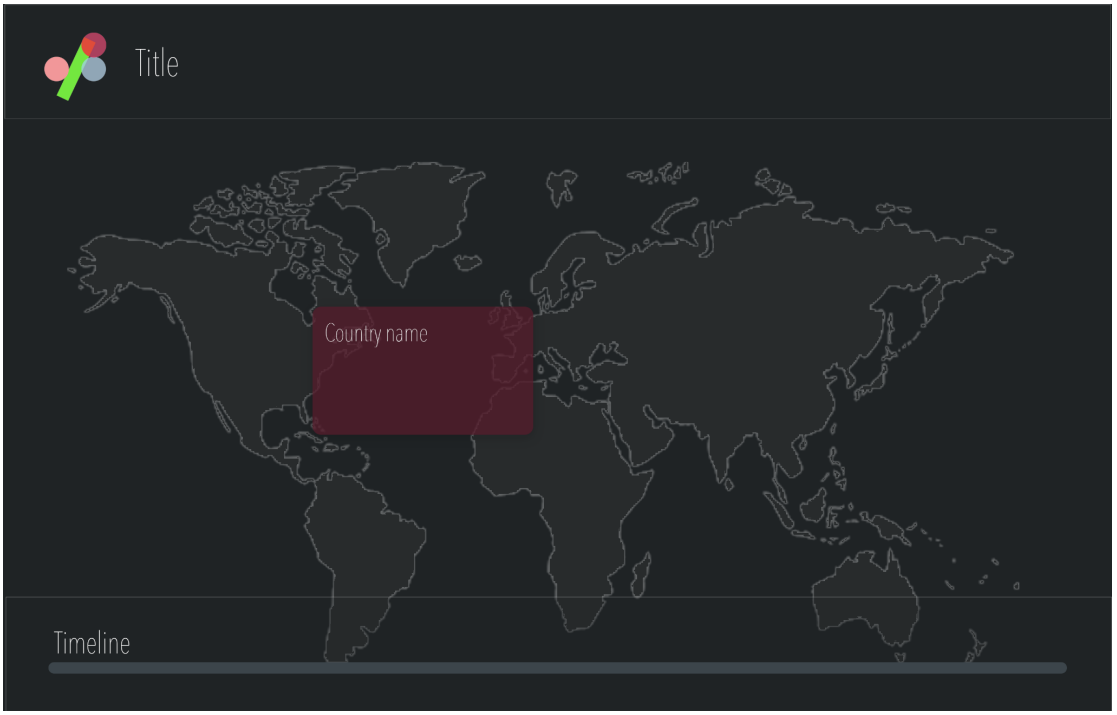

Figure-5

Time Series Chart

We would like to design a visualization that could show data points across time, and our dataset has time attributes inside. In this case, a time series chart shall be a good choice. In addition, we considered linking the time series chart with the map by making use of a cross filter.

Stacked Bar Chart

When individual countries are selected, a good way to showcase the share of macronutrients of selected countries is to make use of a stacked bar chart.

SVG Avatar

We planned to implement an SVG avatar that reacts to changes in the selected data points. For instance, when a user selects a country that eats a

lot of fat and has a high overweight rate, the avatar's belly will grow in size proportional to the data. If the user selects a country that has a high amount of protein in its diet, the avatar will grow more muscular and healthier to reflect that. This visualization is being used to analyze, either to consume existing data, according to the three levels of actions that define user goals in chapter 2 of our textbook [6]. Furthermore, the avatar visualization is mostly designed for enjoy goal. It can show the change of human body shape under different macronutrient data in a straight and interactive way, which shall be impressive for users. Our initial design of SVG avatar was as figure-6.

But we deviated from our proposal a little by removing the SVG avatar from our design in the end. Because we found that, design an interactive SVG avatar linked with other charts and also be able to react with changes in selected data points, is technically very creative, which means it will take lots of time to design, implement, test, redesign, etc. And we obviously don't have enough time for that in our final project. In addition, it might make users overwhelmed by placing so many different kinds of visualizations together on one web page.

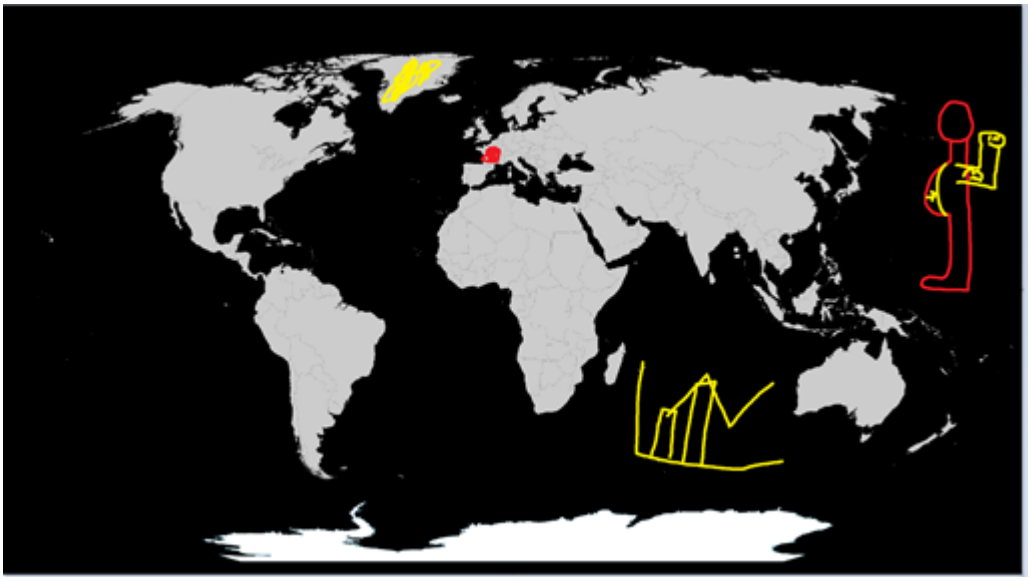

Figure-6

According to the four nested levels of vis design, there are two major concerns at play with idiom design. One set of design choices covers how to create a single picture of the data: the visual encoding idiom controls exactly what users see. Another set of questions involves how to manipulate that representation dynamically [6].

We justified the design decisions we made using the perceptual and design principles above and found that our visualizations can not only show users a picture of our dataset in an efficient way but also let users be able to manipulate data dynamically by interactive techniques. In conclusion, the visualizations we designed are validated.

Implementation

Overview

When users visit our project, a welcome page will show up at first, as figure-7.

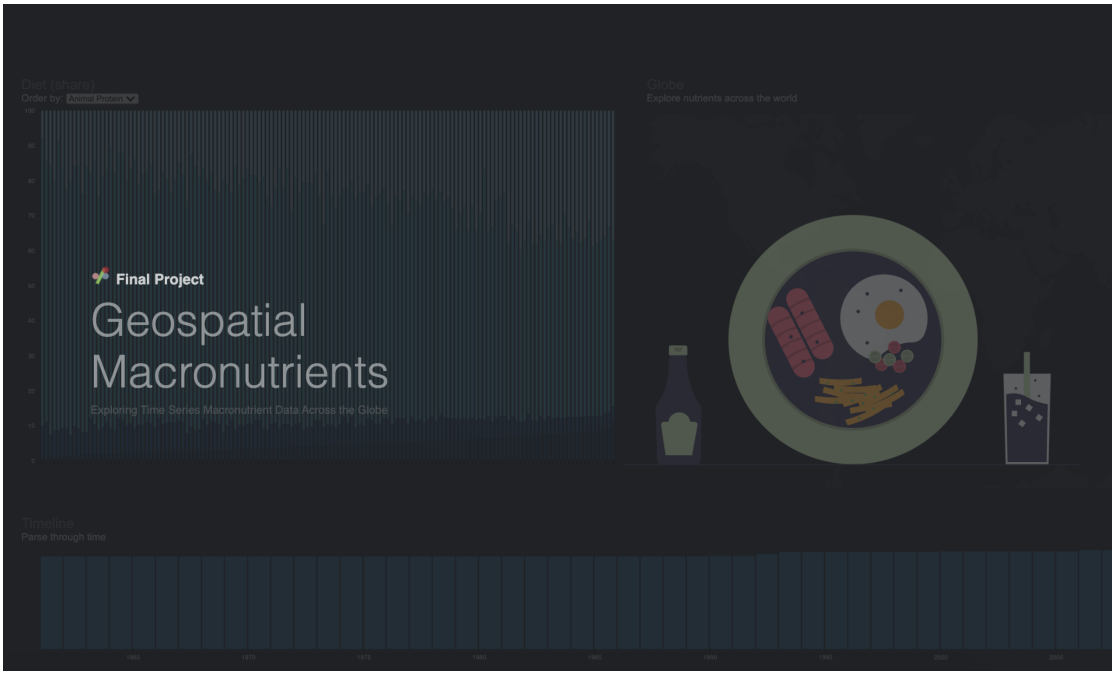

Figure-7

Then the page will show the visualizations after a smooth transition. The overview of our page is as figure-8 shows. The stacked bar chart is placed on the left side, the world map is placed on the right side, and the time series chart below.

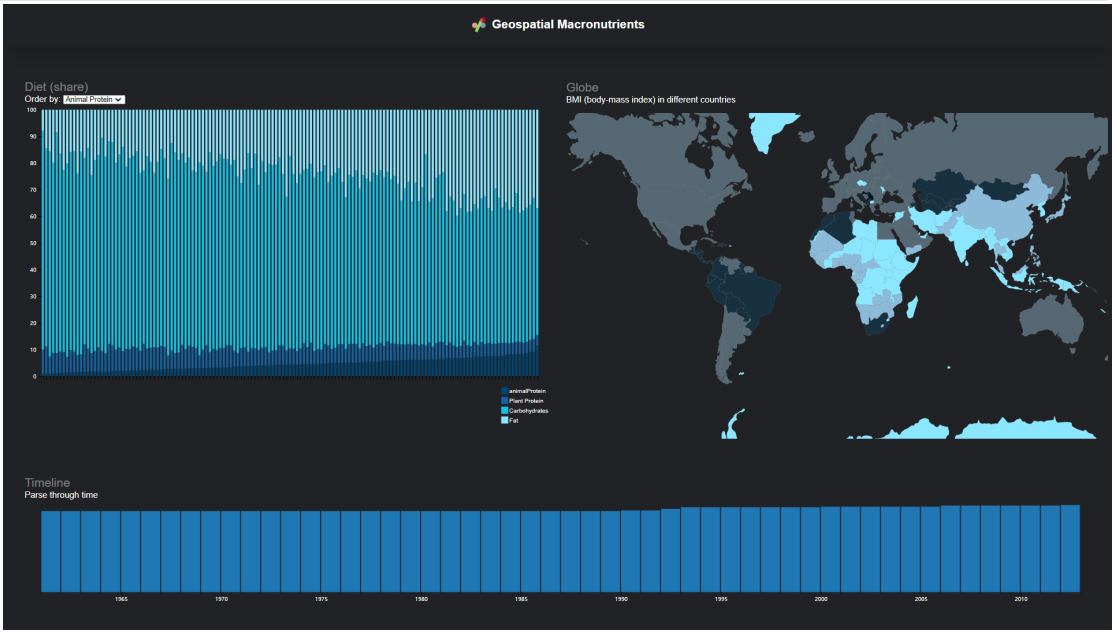
Geospatial Macronutrients
Diet (share)
Order by: Animal Protein
Globe
BMI (body-mass index) in different countries
animalProtein
Plant Protein
Carbohydrates
Fat
Timeline
Parse through time

Figure-8

Map

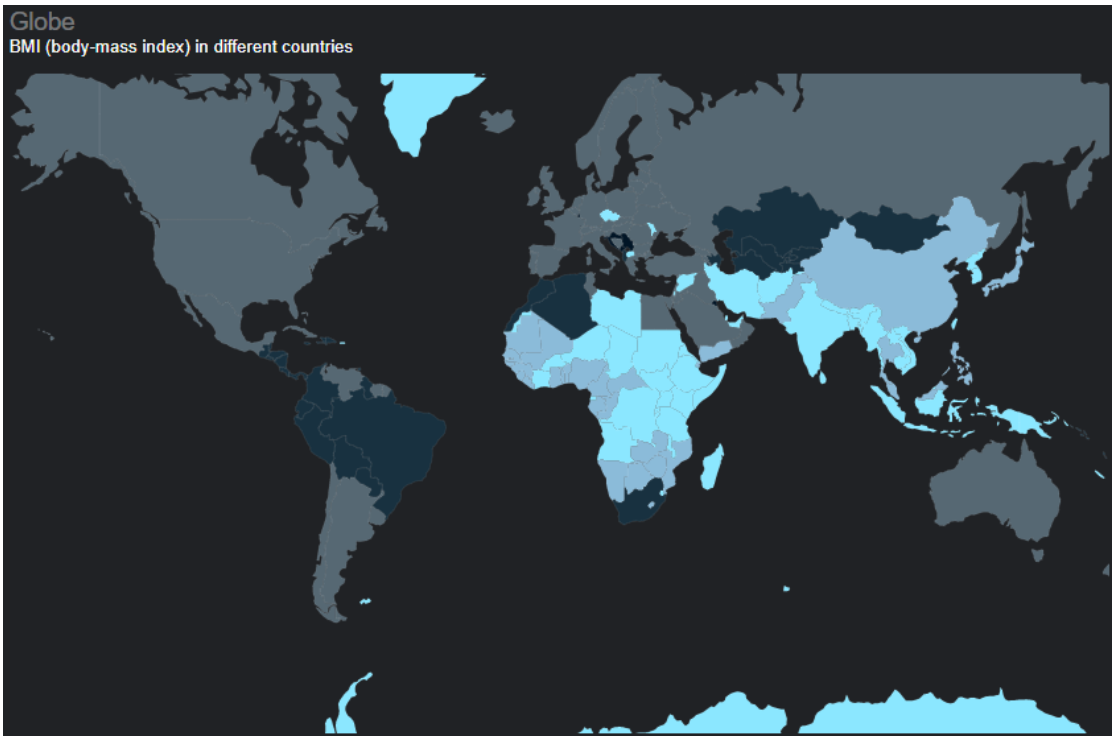
Globe
BMI (body-mass index) in different countries

Figure-9

We made use of a world map to visualize geospatial data. In this map, color represents BMI in different countries. The color encoding is saturation. The darker the blue is, the higher the overweight rate does the country has. For instance, India is filled in light blue and Brazil is filled in dark blue, which means the overweight rate of India is low and the overweight rate of Brazil is high.

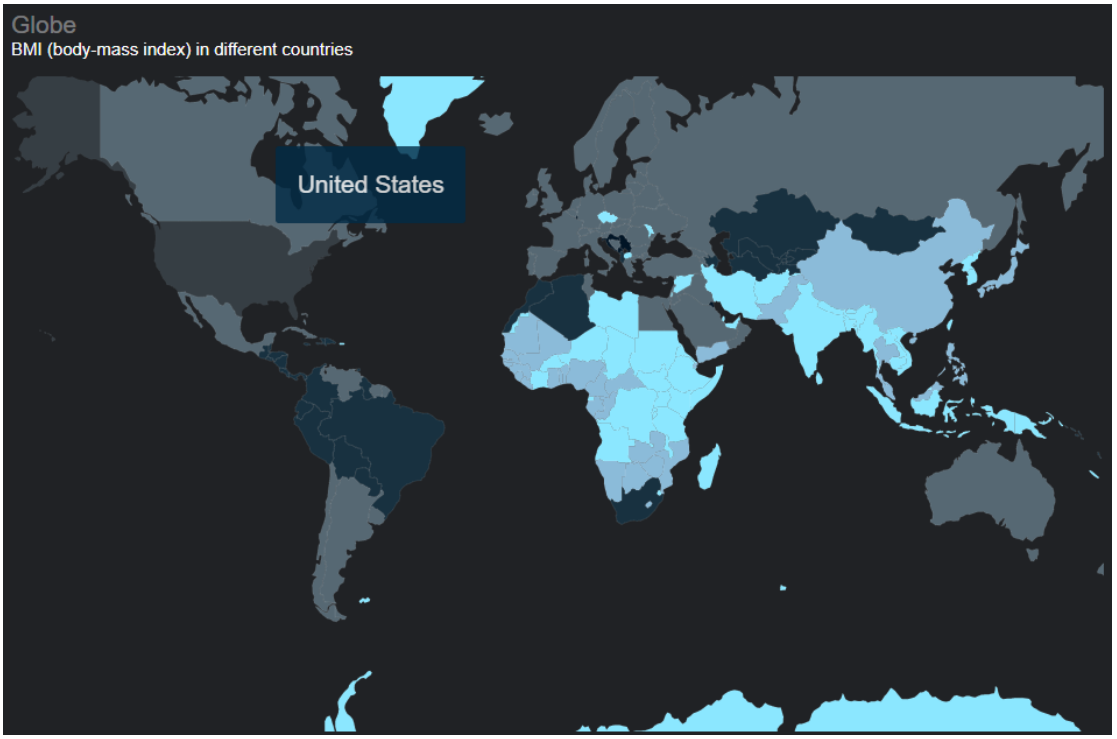

Figure-10

When the user's mouse over a country, the fill color of this country turns darker to emphasize its area. Meanwhile, a tooltip will show up beside the mouse, and show the country name.

If the user clicks on the country's area, our page will turn to be like figure-11. The map will highlight the selected country. And since the map and stacked bar chart are linked together, when we select one country on the map, the stacked bar chart will display the bar of this country on the right side of it, close to the map, as you can see in figure-11.

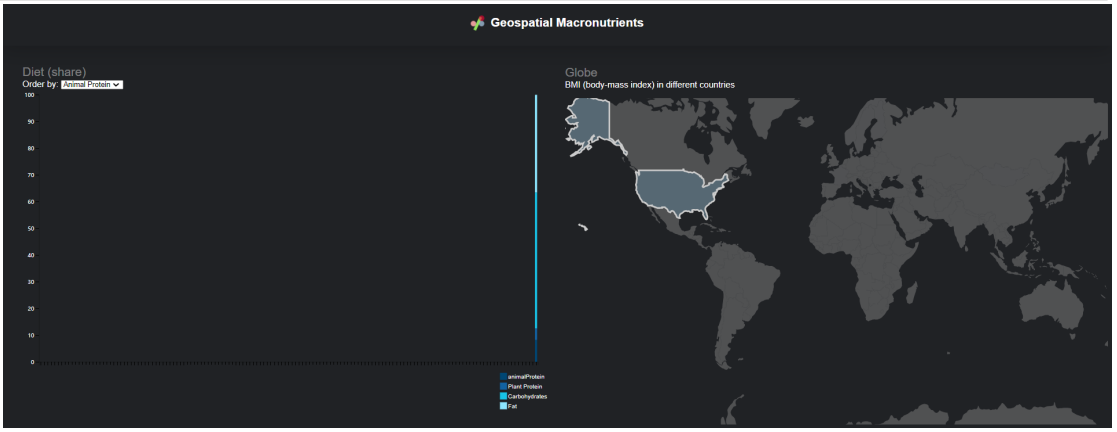

Figure-11

Users can also compare multiple countries' data by selecting (clicking) them on the map. For instance, when I select the United States, Canada, Mexico, and Russia, the map will highlight these countries and the stacked bar chart will put the bars of them together so that I can compare the macronutrients share of these countries. As we can see in figure-12.

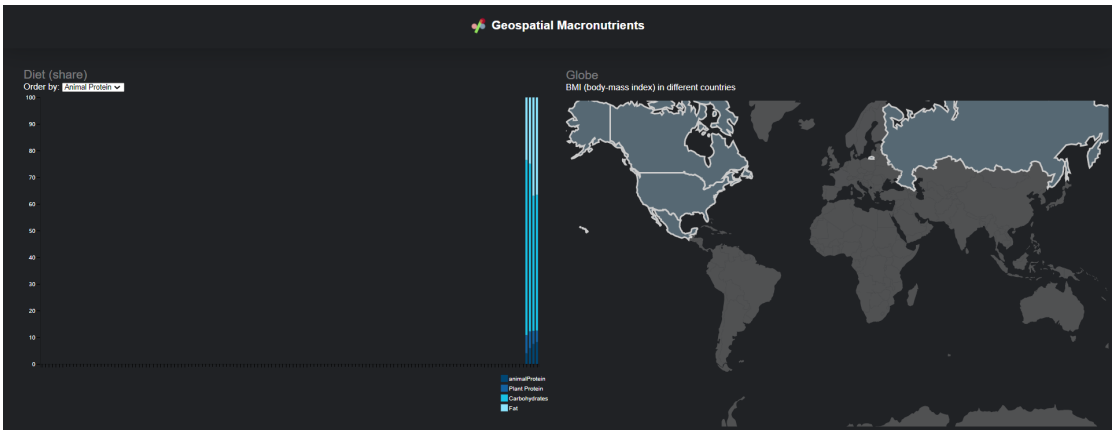

Figure-12

Time Series Chart

The time series chart's x-axis represents the number of years, and the y-axis represents the number of countries that have data for the given year. The initial look of our time series chart is as figure-13.

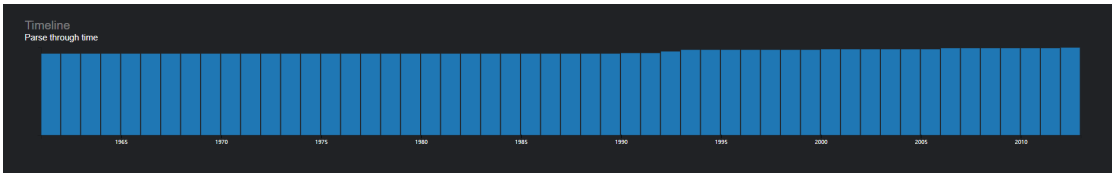

Figure-13

Users are able to create a window that can slide horizontally to choose a time zone. And it's easy to change the left and right bound of the sliding window. As you can see in figure-14.

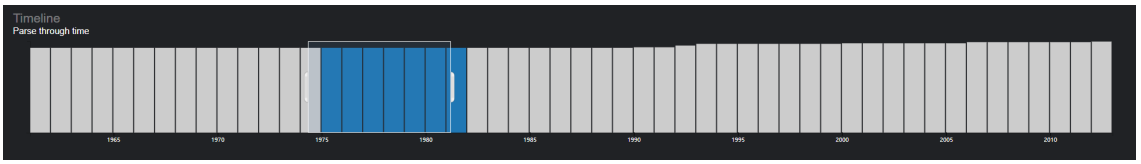

Figure-14

The time series chart is linked to map and stacked bar chart. When we slide the window, map and stacked bar chart will display the data points of selected years. Thus, we are able to figure out the change of the share of macronutrients and overweight rate of countries across time. We can also use this feature to observe the change of selected countries across time.

Stacked Bar Chart

The stacked bar chart is used to visualize macronutrient data of all the countries. We use 4 different colors to represent 4 kinds of macronutrients. There is a legend below to explain what does each color represents. It also has a select box above, which can change which attribute does the chart order by. As figure-15 presents.

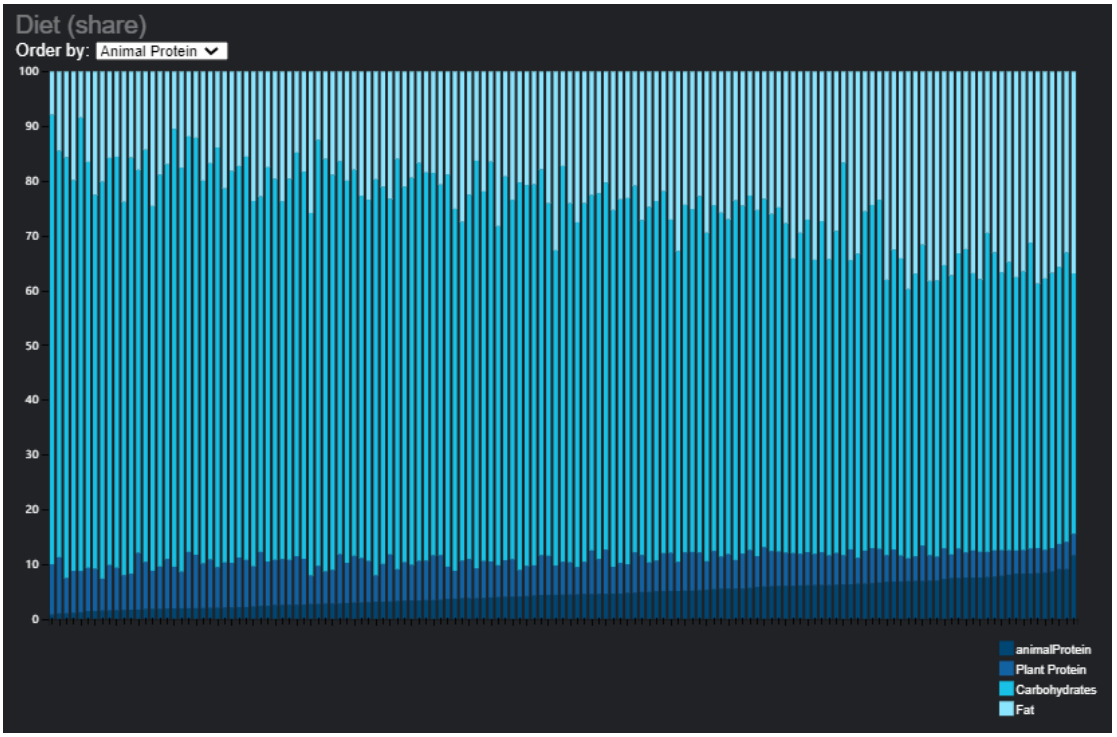

Figure-15

The stacked bar chart is interactive. When user's mouse over bars, a tooltip will show up beside the mouse. The tooltip shows the selected country's name and the percentage of selected nutrient of this country. As we can see in figure-16 and figure-17.

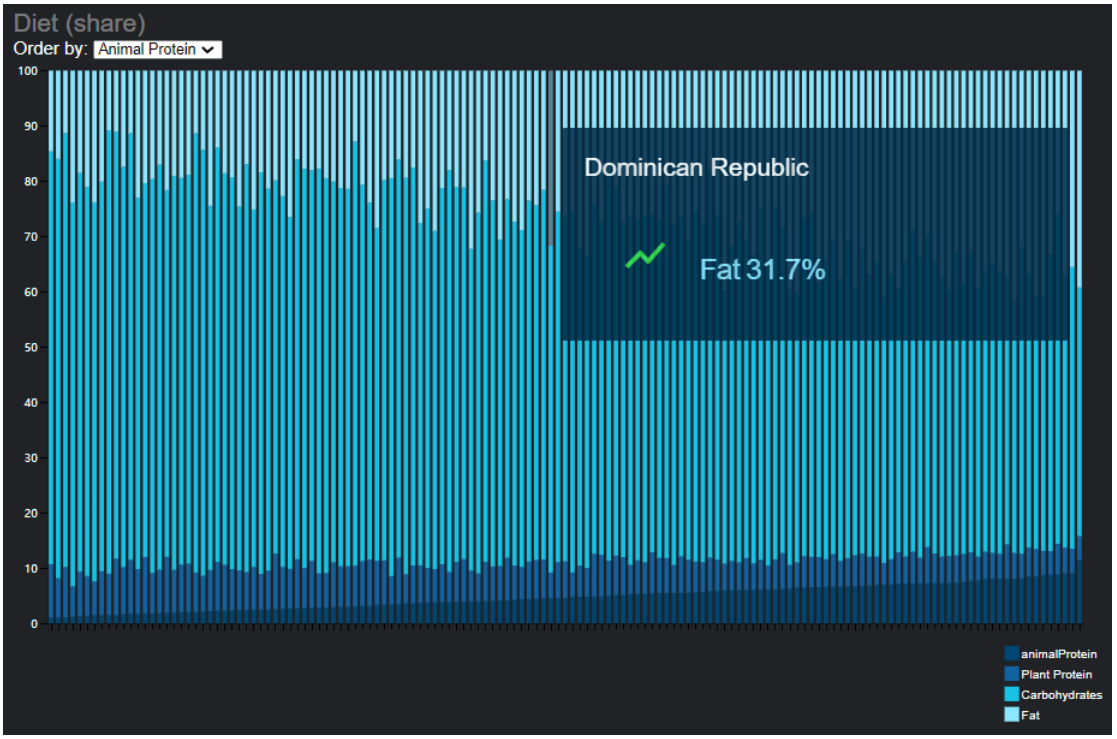

Figure-16

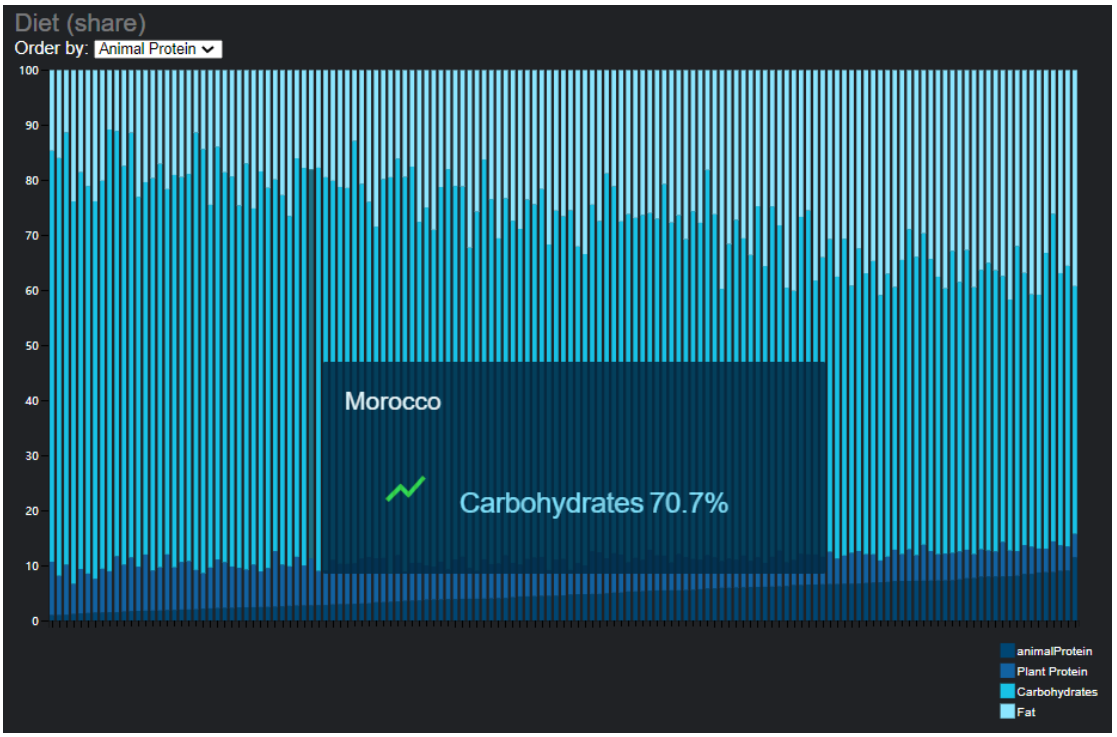

Figure-17

The legend is interactive too. If we mouse over the rectangular of one nutrient, the stacked bar chart will highlight the selected nutrient, as we can see in figure-18. Figure-18 shows the effect when we mouse over 'fat'.

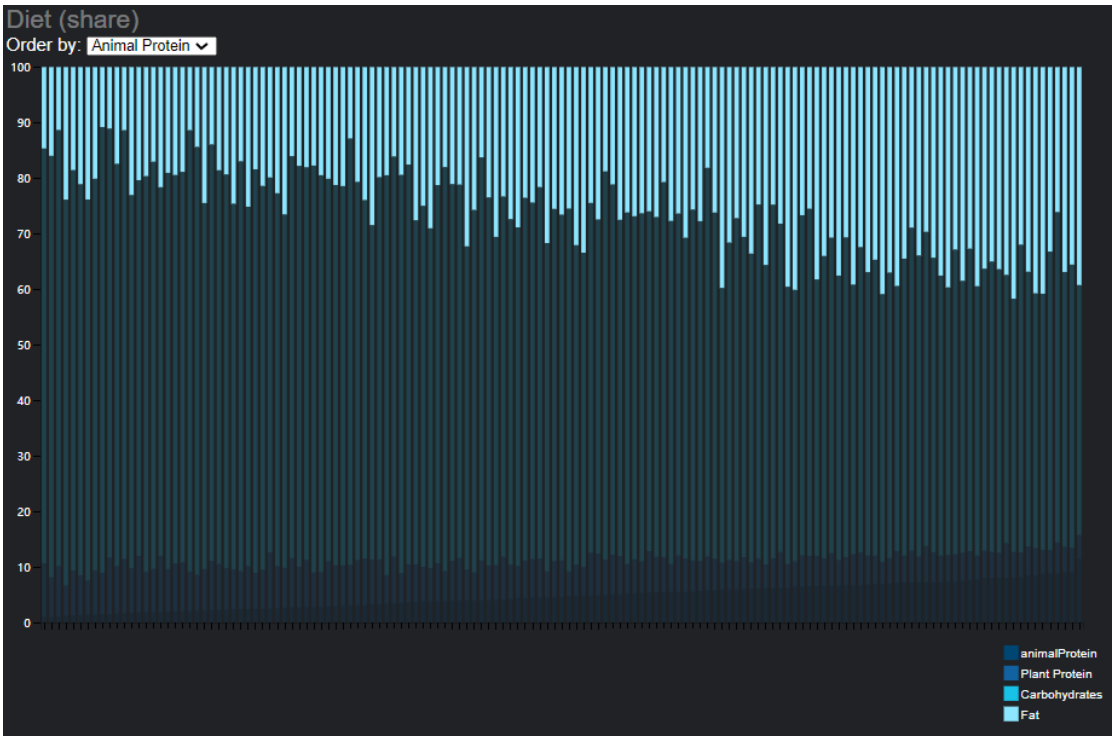

Figure-18

Evaluation

In the world map, I selected the countries that have the high overweight rate from 2003 to 2013. The stacked bar chart shows the share of macronutrients of these countries, as we can see in figure-19.

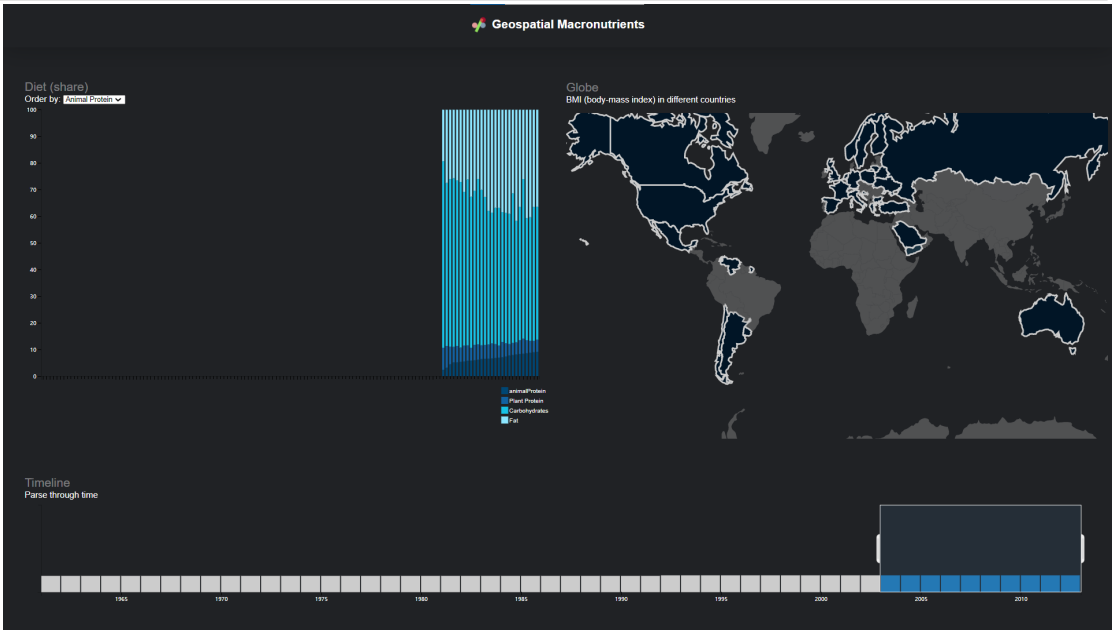

Figure-19

Then I selected the countries that have the low overweight rate from 2003 to 2013. The stacked bar chart shows the share of macronutrients of these countries, as we can see in figure-20.

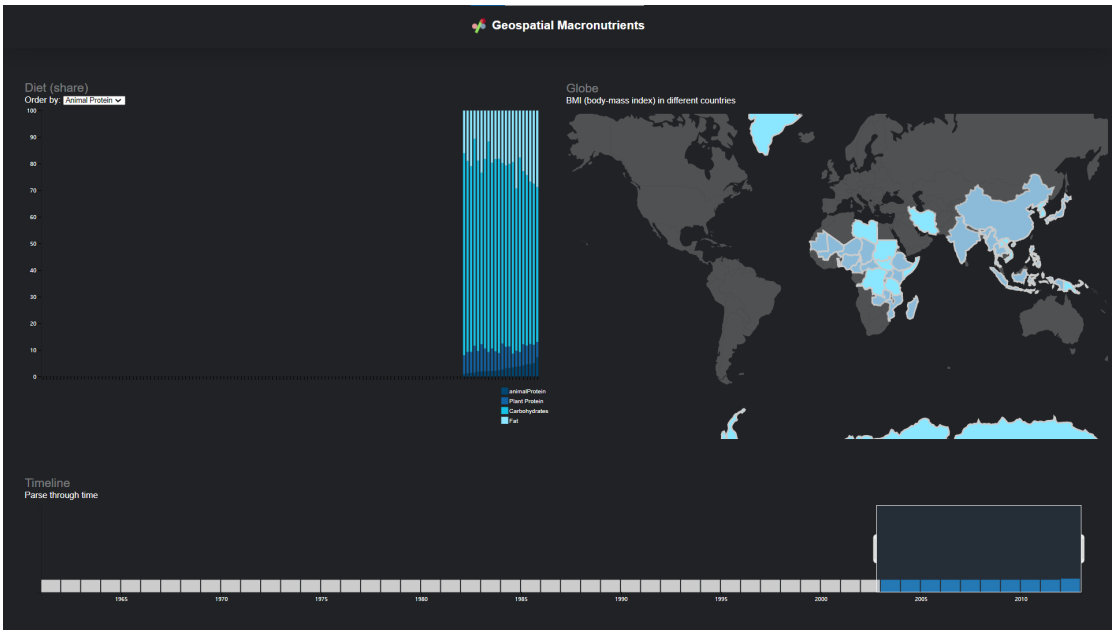

Figure-20

If we put the two stacked bar charts together and make a comparison, we will see the following figure.

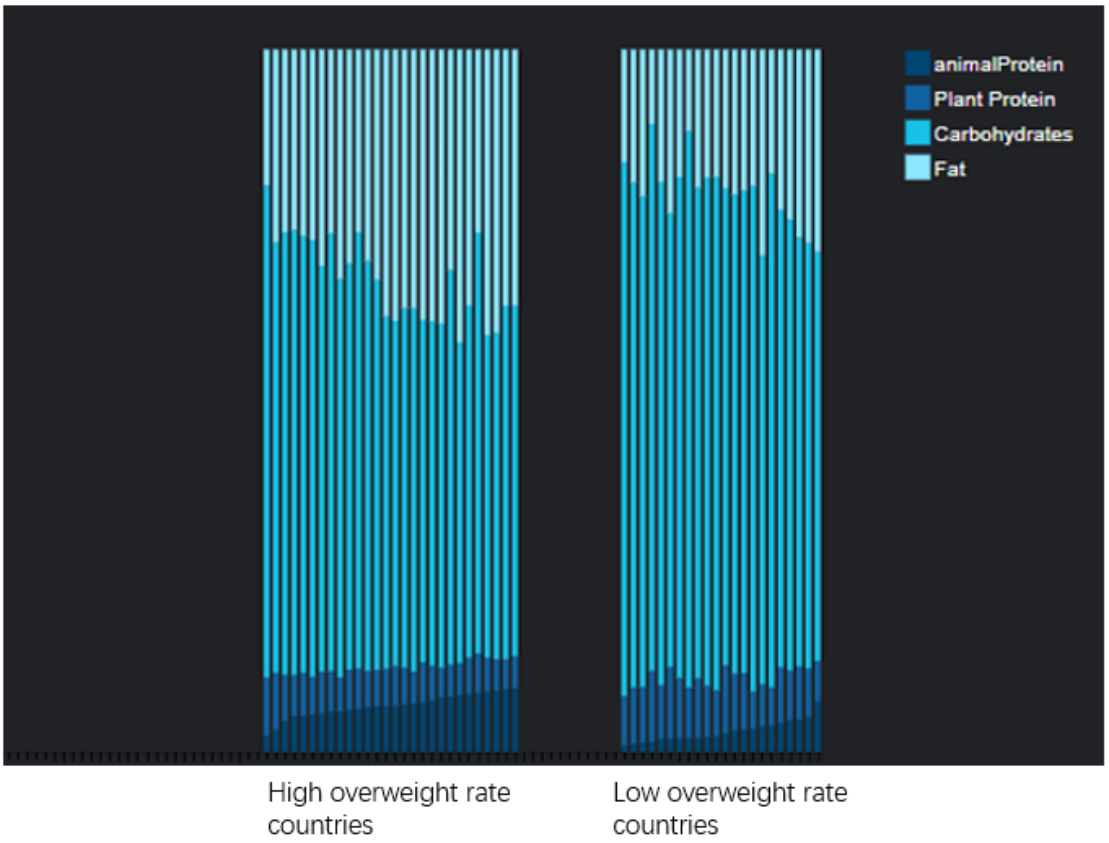

Figure-21

It refers that people in the countries that have high overweight eat more fat and fewer carbohydrates than those who live in the countries that have low overweight rate. The amounts of protein are close in the two datasets, but people in the countries that have high overweight have more animal protein and less plant protein.

Thus, it's reasonable to consider that more fat fewer carbohydrates, or more animal protein and less plant protein could be the key to getting overweight. This result does answer our first question about '*How do macronutrients influence people's overweight rate?*'.

Since I don't have any level of expertise in nutrition or global health studies, but I still enjoyed our visualizations and have some conclusions on how does macronutrients influence people's overweight rate globally. This is the answer to our second question about '*How to make new users, who do not have any level of expertise in nutrition or global health studies, are able to consistently explore the data using our visualization and draw meaningful conclusions?*'. Our visualization design is a good approach to let new users explore and manipulate data, it works great to help users observe information and draw conclusions.

For future work, I think we can add some control tests to figure out which nutrient influences the overweight rate most. Control group 1 can be more fat fewer carbohydrates and same animal protein same plant protein. Control group 2 can be more animal protein less plant protein and same fat same

carbohydrates. After we collect enough data points, we can repeat the analysis we did above.

References

[1] Popkin, B. M. (2006). Global nutrition dynamics: the world is shifting rapidly toward a diet linked with noncommunicable diseases–. The American journal of clinical nutrition, 84(2), 289-298.
[2] https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweigh
[3] Riera-Crichton D, Tefft N. Macronutrients and obesity: revisiting the calories in, calories out framework[J]. Economics & Human Biology, 2014, 14: 33-49.
[4] UN Food and Agricultural Organization FAOstat food balance sheets. Available at: http://www.fao.org/faostat/en/#home [accessed 31st July 2017].
[5] Chapter 3: Calculation of The Energy Content of Foods - Energy Conversion Factors. Food and Agriculture Organization of the United Nations. Available at: http://www.fao.org/docrep/006/Y5022E/y5022e04.htm [accessed 31st July 2017].
[6] Munzner T. Visualization analysis and design[M]. CRC Press, 2014.
[7] Alissandrakis A, Nake I. A new approach for visualizing quantified self data using avatars[C]//Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. 2016: 522-527.