USDA Food Composition Visual Browser

Ankit Nagpal *(an1641)*
an1641@nyu.edu

Preeti Patel (*pp1569*)
pp1569@nyu.edu

Varun Elango (*ve365*)
ve365@nyu.edu

Vignesh Ramesh (*vr839*)
vr839@nyu.edu

Project page (on Github): https://github.com/orgs/NYU-CS6313-Fall16/teams/group-9

Video: https://vimeo.com/196914006

Working demo: https://nyu-cs6313-fall16.github.io/USDA-9/index.html

What is the problem you want to solve and who has this problem?

Information about the nutritional value of food is not easily accessible. Nutrition labels attempt to convey this information, however they are difficult to comprehend in relation to the larger context of a consumer's diet and do not offer the information in a very simple and accessible view. Also it can be confusing in the sense that it doesn't tell much about whether something is good or bad for a consumer.

The goal of this project is to make food data more accessible to the world and to help users answer the question **"What's in my food?"** and to assist them in making healthy choices about what to eat. The application will give the user a high-level understanding of the nutrient composition of the food as a whole, with the option to drill down to more detailed information if needed. The project will make use of USDA Food Composition Database as the source of food composition data for the United States.

What are the driving analytical questions you want to be able to answer with your visualization?

The visualization answers the following questions:

- What is the nutrient composition of a food X?
- What foods contain a specific composition of nutrients?
- What are the alternative food options to a specific selection of food(x)?
- What are the top foods containing a combination of nutrients?
- In a specific food group, Is nutrient X correlated to nutrient Y?

What does your data look like? Where does it come from? What real-world phenomena does it capture?

The food data comes from USDA Food Composition Database. It includes information about nutrient content of a large collection of food items.

Attribute Name	Attribute Type	Description	Value Range	Derived
NDB No	Categorical	Unique Identification number for each product/food	Integer Eg: 45130770	No
Description	Categorical	Product/ Food name	String Eg: Abiyuch, raw	No
Food Group ID	Categorical	Unique Identification number for each food group	Integer Eg: 1500	No
Food Group	Categorical	Food category	String Eg: Baby Foods	No
Nutrient ID	Categorical	Unique Identification number for each Nutrient	Integer Eg: 324	No
Nutrient Name	Categorical	Name of Nutrient	String Eg: Vitamin D	No
Nutrient Unit	Categorical	Name of Nutrient Unit	String Eg: g, μg,OZ	No
Nutrient Value	Quantitative	Quantity of Nutrient Unit	Integer Eg: 18.867 g	No
Nutrient Group	Categorical	Category of nutrient	String Eg: Vitamins	No

What have others done to solve this or related problems?

1. **Sage Project** (https://sageproject.com/)

The project turns nutrition labels into data visualizations to make better food/eating decisions. Their aim is to make food data accessible to the world by making food data smart, simple, and personalized. Through the site, you can selects brands or specific foods by brands and receive a full breakdown of not only its nutritional value and macronutrient breakdown, but you also get a breakdown of just how much exercise it would take to burn it off by the kind of exercise of your choice.

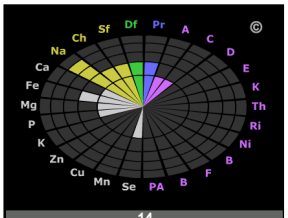
2. **Self Nutrition Data** (https://nutritiondata.self.com/)

Nutrition Data is another application which provides a database of nutrition information about common foods, including glycemic load, inflammation factor, and breakdown of carbs, fat, and protein. The site provides tools for entering recipes and tracking and analyzing your own nutrition using this information. There are also standalone tools for calculating your recommended daily amounts, recommended BMI, and searching the food database based on nutrient amounts. Additionally, the site contains numerous guides on nutrition related topics including how to understand the numeric data, recipes, dieting, and how to make food choices.

Caloric Ratio Pyramid What is this?

61%	33%	6%
Carbs	Fats	Protein

Nutrient Balance Score What is this?

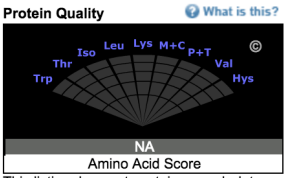


14
Completeness Score

Nutritional Target Map What is this?

1.7	1.4
Fullness Factor	ND Rating

Protein Quality What is this?



NA
Amino Acid Score

This listing does not contain enough data on individual amino acids to determine protein quality.

Estimated Glycemic Load

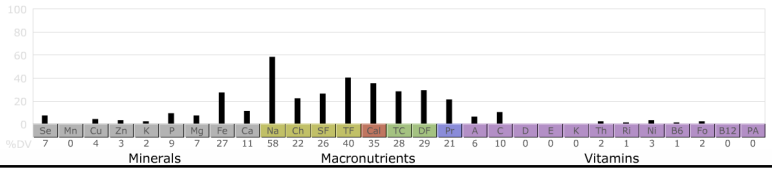
44

0 250

Typical target total is 100/day or less

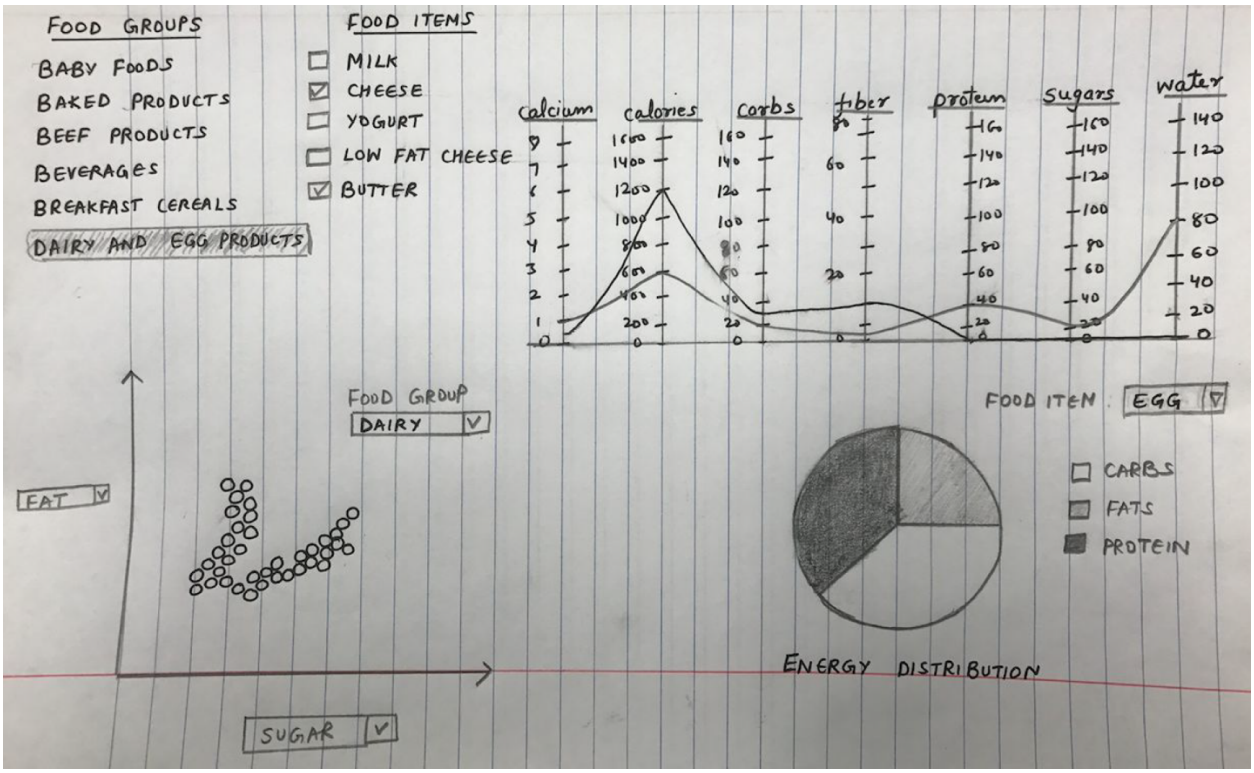
What is this?

Daily Values Chart Set your daily values What is this?



Design Iterations

Initial Mockup



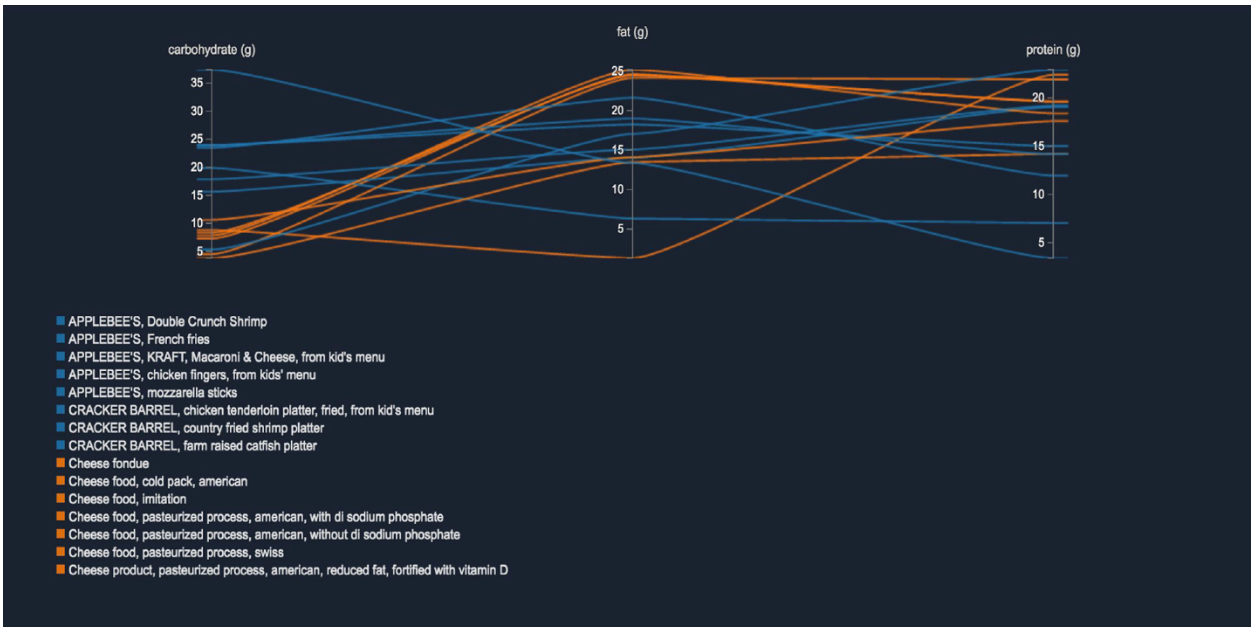
(Figure 1) - *Displays the initial mockup of the project.*

The original idea was to use

1. A parallel coordinates chart to visualize the nutrient content of the selected food groups and items.
2. A scatter plot to visualize correlation between two nutrients for a particular food group
3. A pie chart to show the energy distribution of the selected food item.

According to this mockup, a user would first select the food groups and items of his interest. The selected food items will be displayed on the parallel coordinates. To find correlation between two nutrients for a particular food group, user will select the nutrients as X and Y axis of the scatterplot as well as the food group. An example of this scatter plot would be a correlation between fat and sugar content of all the dairy products available. The pie chart will show the energy distribution (quantities of carbs, fats and protein) of an individual food item.

Figure 2 displays the partial implementation of this mock up. We used a sample of 15 food items and 3 nutrient values (carbs, fats and protein) to visualize for the first implementation. With this visualization, we could clearly see the amount of nutrient content present in the sample food items. We used color to distinguish food from separate groups.

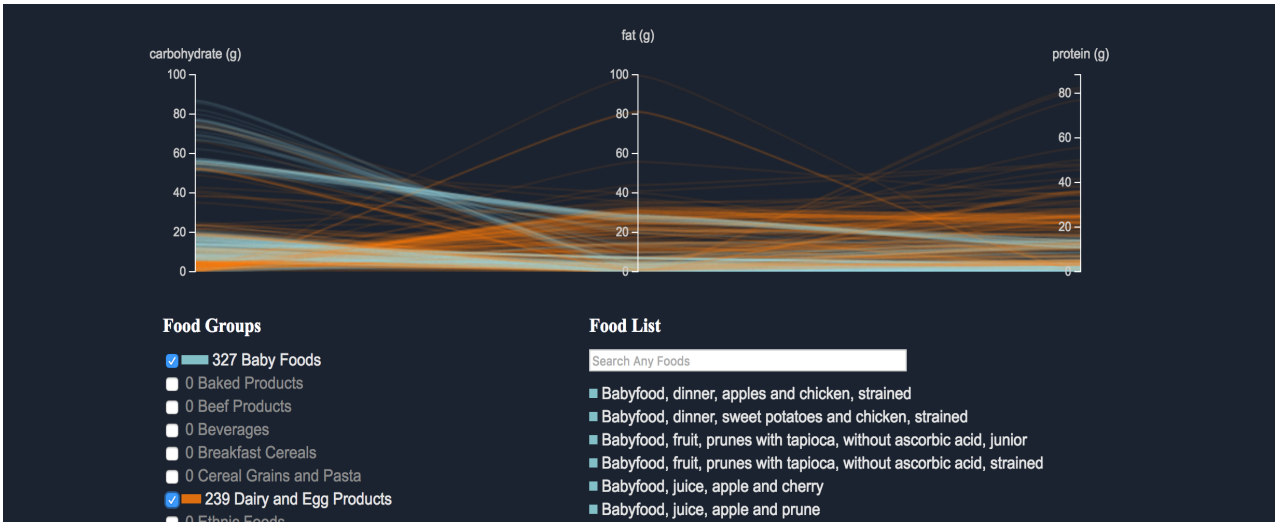


(Figure 2) - *Partial Implementation of the Initial Mock up*

The issues with this implementation and mock up were:

1. The axes values showing nutrient quantity were hardcoded in the initial implementation but the nutrient quantity was varying for different food items
2. Auxiliary views (scatter plot and the pie-chart) were disconnected with the main chart
3. The use of scatterplot for correlating nutrients was not justified. We wanted a way to find correlation between nutrients by using the parallel coordinates
4. Sample dataset was used for implementation (not the complete dataset). A lot of overlap of food items in the visualization was making it cluttered and difficult to read.

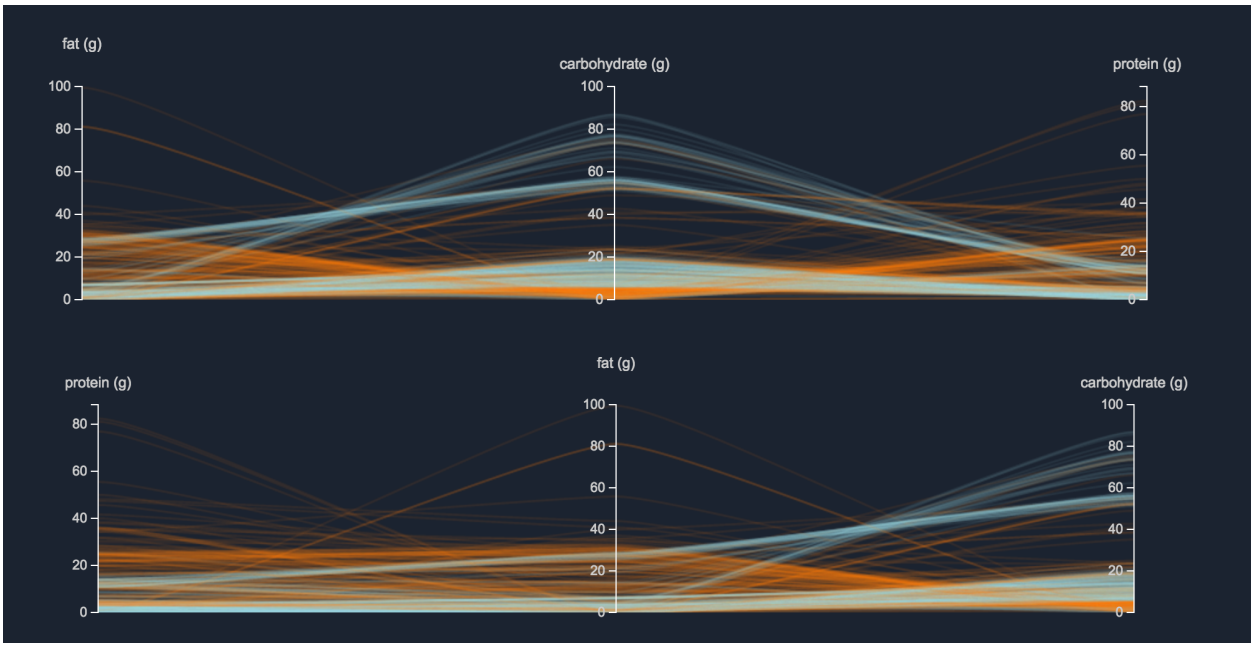
Figure 3 shows the next iteration of the visualization. We used a larger number of food items for this implementation but the number of nutrient values were still three. To use this visualization, a user could select food groups and food items of interest to visualize them. Additionally, a user could search for a particular food item out of the selected food groups to visualize only a subset of food items.



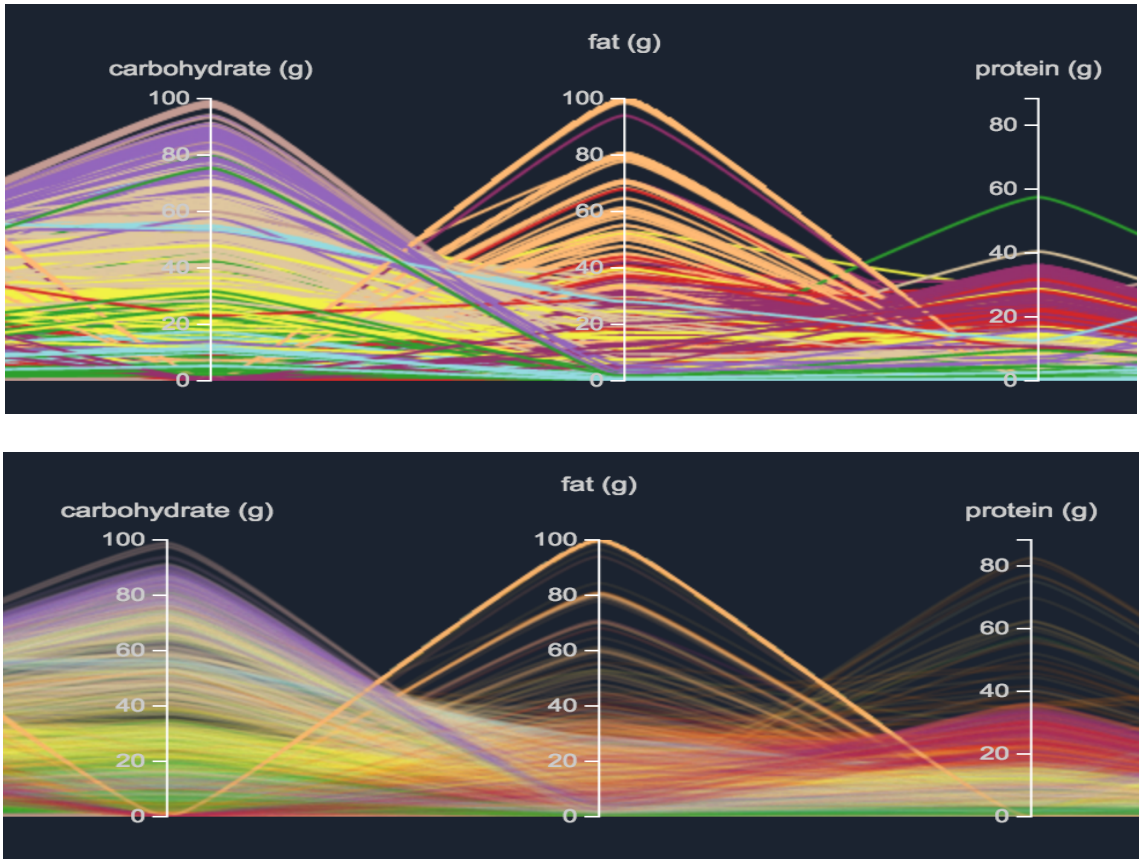
(Figure 3) *- Next Iteration*

We solved the following problems with this implementation:

- The scale(extent) of each nutrient was identified based on the entire food selection
- Removed the scatterplot by allowing the users to rearrange axes to visualize correlation between two or more nutrients. Figure 4 demonstrates this.



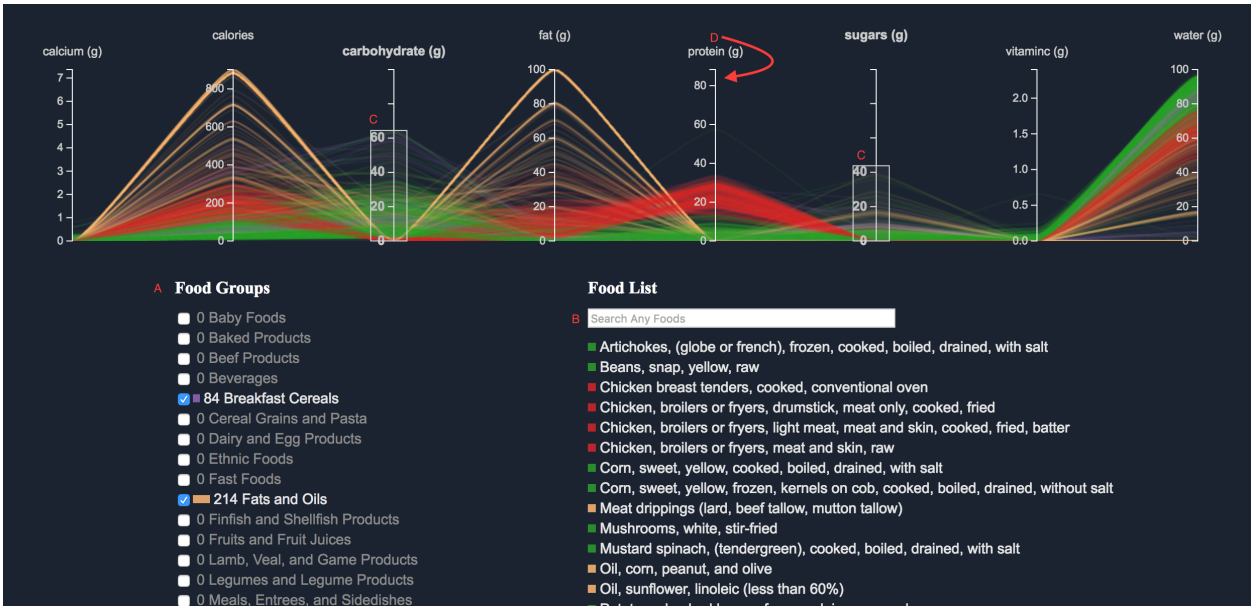
(Figure 4) *- Rearrangement of axes*



(Figure 5) - *Use of Opacity*

- Used opaque lines to visualize a large number of food items without making it difficult to read because of the overlap. Figure 5 demonstrates the difference between opaque and non-opaque lines.

Final Visualization

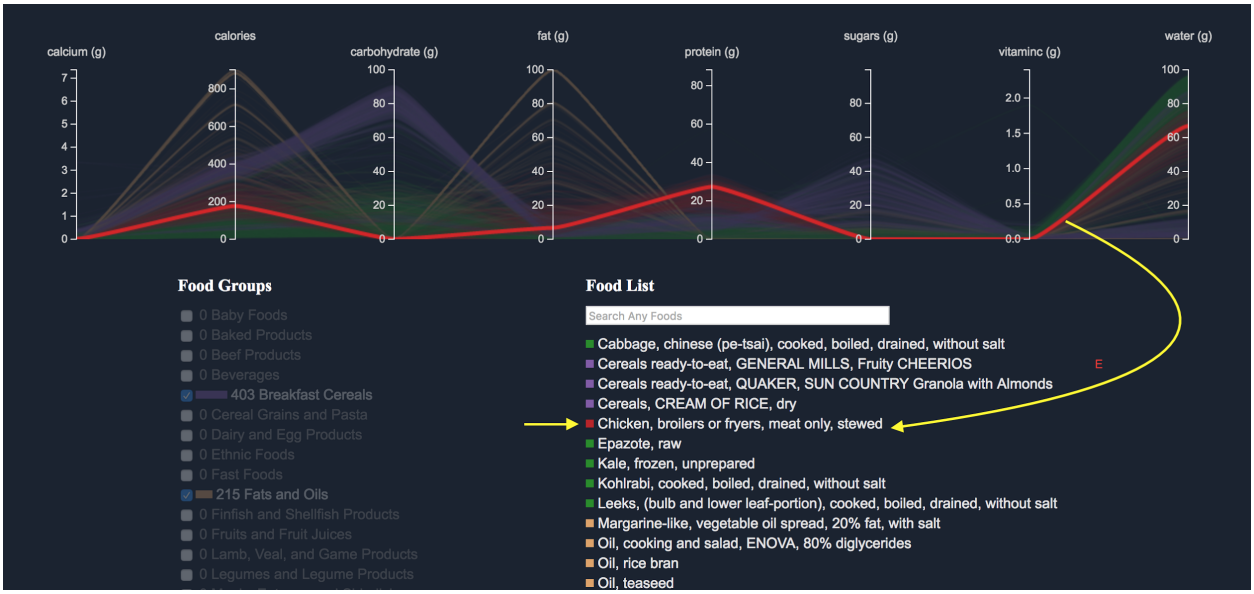


(Figure 5)- *Final Visualization*

For the final visualization, we selected 8 most relevant nutrient values to visualize out of the available 190.

Following are steps to use the visualization:

A. Select the food groups to visualize. The Food List (B) will be updated based on this selection. Each food group is represented by a different color.
B. Search food items present within the selected food groups. A random list of 20 food items is shown by default.
C. Drag over any number of nutrient axes to filter food items based on the range selected. The food groups (A) and the food list (B) will be updated based on the selection. Click anywhere on the axis to remove the selection.
D. To compare two nutrients, drag a nutrient axis to move it adjacent to any other nutrient axis.



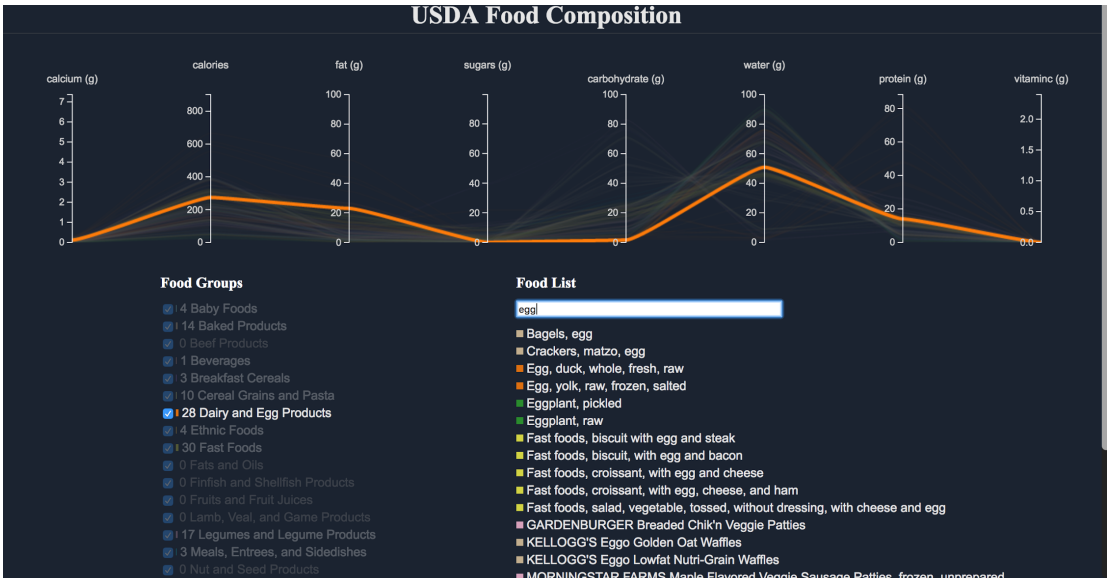
(Figure 6) - *Final Visualization*

E. Hover over a food item in the food list to highlight it in the visualization

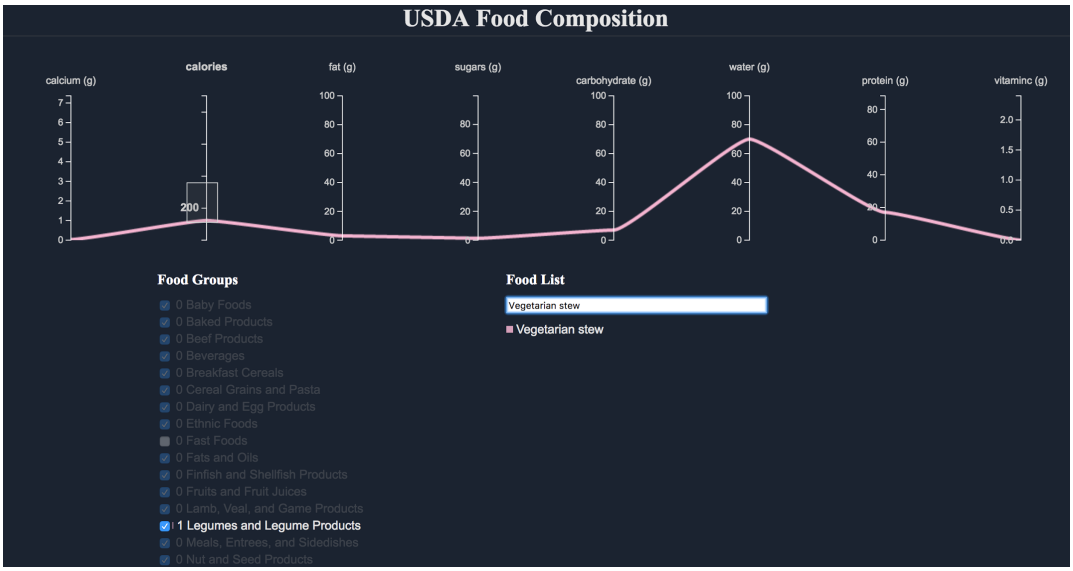
Findings

Below is the analysis of a Non-vegetarian person looking to become a vegetarian.

1. What is Nutrient Composition of eggs?

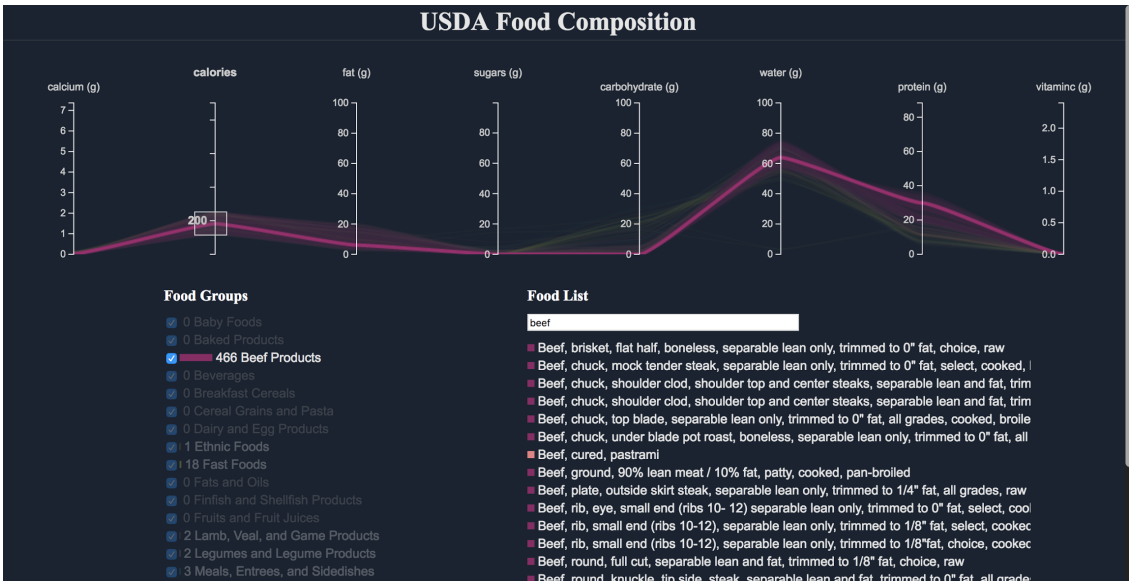


2. What foods are similar to eggs?
   Tofu or Vegetarian Stew could be an alternative to eggs as it has a similar nutrient build up



   Below is the nutrient composition of Tofu, which is similar to egg.

3. What are my alternatives to beef?

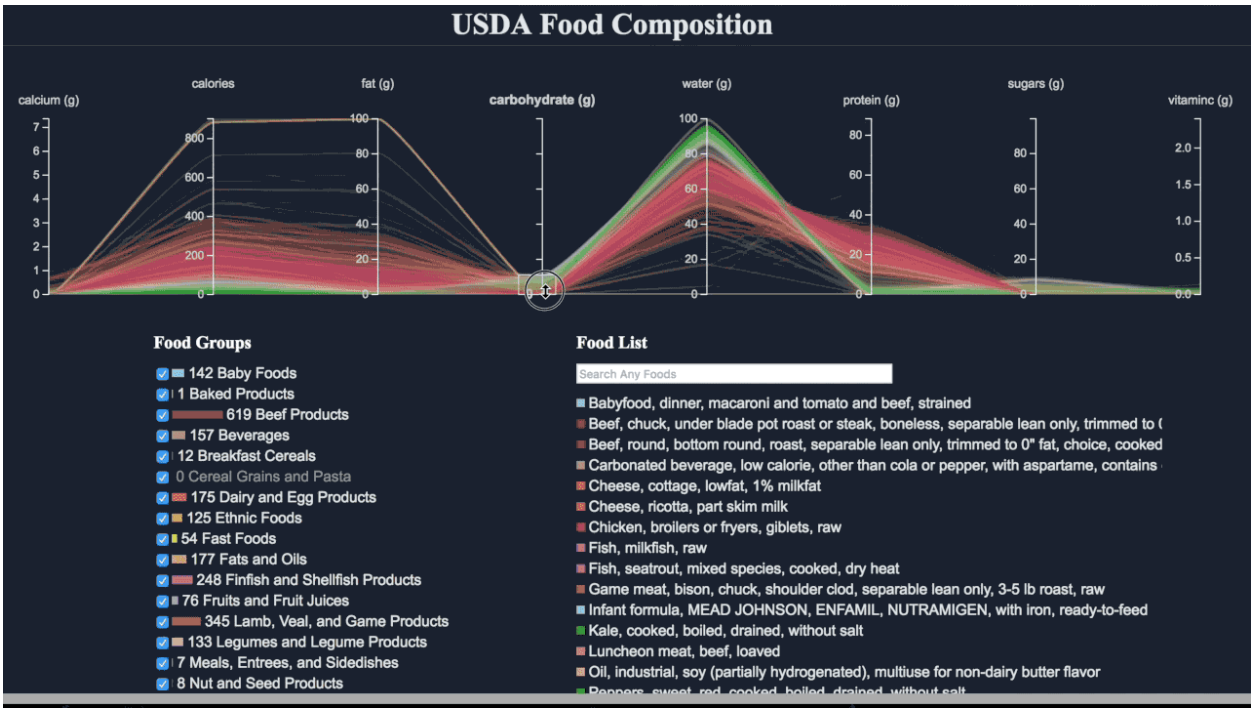


Soy has a similar to composition to beef

Other interesting findings

4. There is a **direct** correlation between fats and calories for most food groups. For most foods the calorie count increases as the fat content is increased

5. There is an inverse correlation between carbohydrates and water, whereas there is a direct correlation between carbohydrates and sugar



Limitations and Future Work

- Using the current application, a user cannot select two or more food items to compare their nutrient composition
- The application can display a similar distribution of vitamins and minerals for food items in addition to current nutrient list
- Possibility of adding/removing nutrient axes to focus only on the desired nutrients rather than all of the available ones
- The current application can be used to calculate or show the nutrient content of popular food recipes