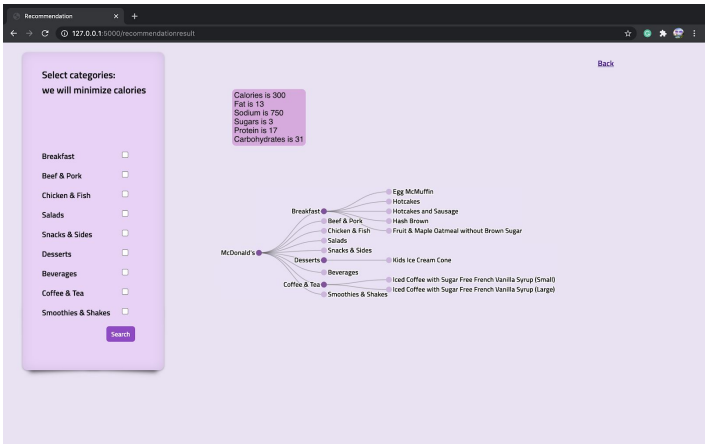
Information Visualization 2020 Fall - Final Report

Ke Xu (kx417) & Liancheng Gong (lg2984)

1. Overview

a. Screenshot of your visualization



b. Summary of your project briefly

Take McDonald's menu and nutrition data as a source, we propose a visualization to present various nutrition indexes such as calories, fat, Cholesterol, etc. for each product in McDonald's, divide those indexes into healthy and unhealthy categories according to the user's input of his nutrition standard initially, and recommend several personalized optimal sets of meals for the user.

2. Data

a. description of your data (e.g., dataset type, scale/cardinality)

Our dataset contains 260 products of McDonald's in the type of table from the Kaggle dataset. The dataset consists of 9 categories including Breakfast, Beef & Pork, Chicken & Fish, Salads, Snacks & Sides, Desserts, Beverages, Coffee & Tea, Smoothies & Shakes. Each product within that category has its corresponding serving size and 21 nutrition indexes, including calories, total fat, cholesterol, sodium, etc. Every attribute is quantitative.

b. Include a URL linking to the source of your data

https://www.kaggle.com/mcdonalds/nutrition-facts

c. Briefly describe your current data preprocessing pipeline, if there is one.

We use pandas to preprocess the data, first, we drop all irrelevant data from the dataset and divide those items with specified sizes into Items and subitems for a clear layout of the treemap. Also, we create a new table for the parent-children relationships to fit the d3.stratify(). In addition, we calculate the difference between user-inputted data and real data for the bar chart in javascript. The preprocessing process is recorded in the IV Final Project Data Preprocessing.ipynb. For the recommendation part, we use Flask to combine HTML and python, and use the pulp package to run simple linear programming to minimize the calories of the set of meals.

3. Goals and tasks

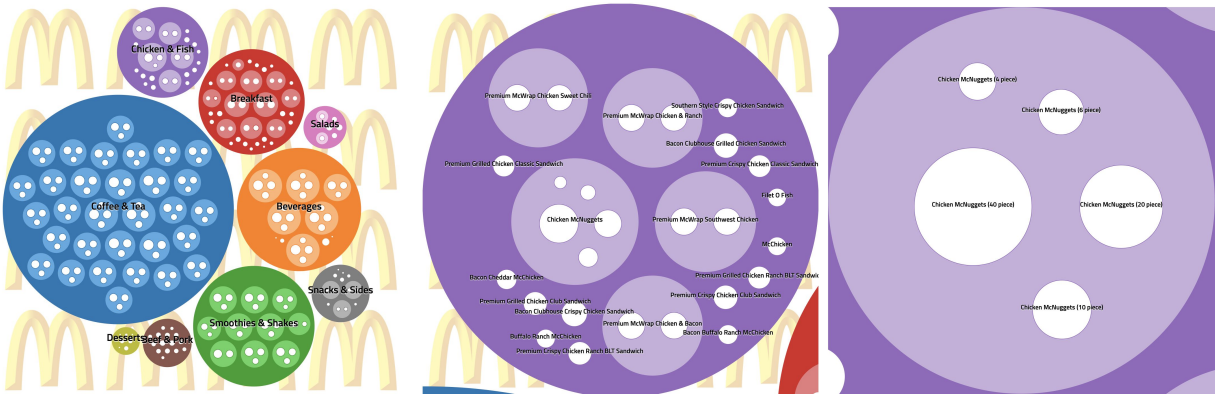
a. Description of your intended task(s).

For usage scenarios, we hope the user can compare his preferred nutrition standard with each item Mcdonald's provides when he browse the menu to have a better understanding of what he decides to eat. Then based on his standard, we use our algorithm to provide a recommended set of meals for reference.

For visualization language, on the homepage, we construct a zoomable treemap to represent the menu, as it fits the containment link. When the user selects each category, the related bar graph with index differences will show, as the aligned position is best for qualitative data comparison. We specified different colors for categorical data. On the recommendation page, we draw a collapsible tree for the recommendation set of meals, as it can show the result for the user at the first glance while maintaining the containment.

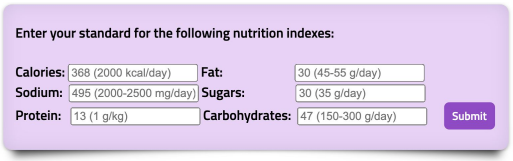
4. Visualization

In the first HTML page, we designed two views including a circular treemap and a horizontal bidirectional bar chart.

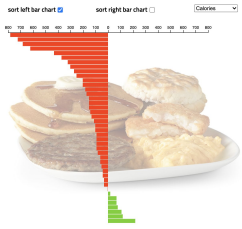


The circular treemap allows users to view the nine categories of McDonald's without zooming (above-left), to view specific products within those categories when zooming into the second layer (above-middle), to view the serving size of that product when zooming into the third layer (above-right). The size of the nodes depends on the serving size of the products.

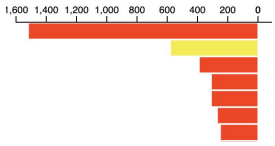
The right view includes a selection box and a horizontal bidirectional bar chart.

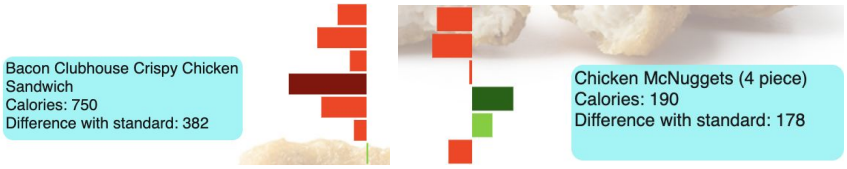


The selection box allows the users to input their ideal intake value for each product and we also set the default value as shown in the placeholder without the user's input. However, don't click submit after inputting your value, it will be used in further applications! The bidirectional bar chart allows the user to view the exact value of each nutrition index as shown in the following interaction part, to select the nutrition index 

, to sort  in both ascending and descending order to find out the healthiest or the least healthy product, to view the difference between the actual value and the value users' set or the default value also as shown in the interaction part. The left bar chart with the color red represents values exceeding set values; the right bar chart with the color green represents the values under set values.

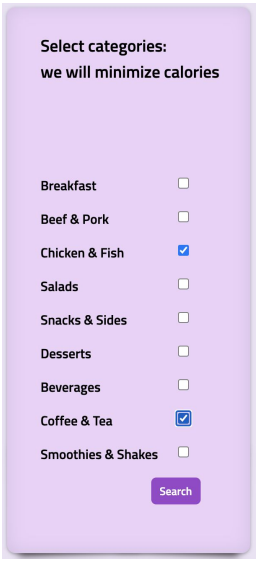
For the interaction between two views, when the mouse hovers over the tree node of the circular treemap, a tooltip will appear on the top left corner 

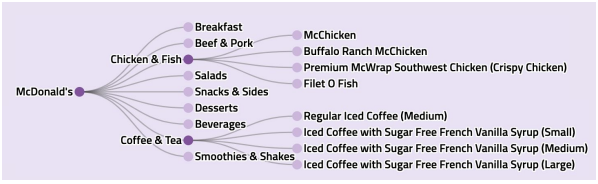
, and the bar representing the selected product will be highlighted by yellow 

. When the mouse hovers over the bars, the selected bar will be highlighted by dark red or dark green 

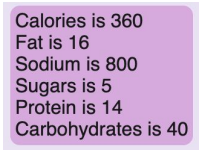
, a tooltip beside the bar will be displayed, and the corresponding node in the treemap will also be highlighted by yellow 

.





In the second HTML page, there will be another selection box shown on the left, which allows the users to choose the categories they want to enjoy. By clicking the search button, a collapsible tree with the recommended products will be displayed as shown above. By hovering the mouse onto the products, a tooltip will be displayed on the top left corner.



5. Reflection

a. Describe how your project has developed from your initial proposal to your final product.

Basically, we are satisfied with the progress we made, as we almost completed all functions we proposed. We substitute the collapsible tree for the circles and stack bar graph on the recommendation page, because the nutrition indexes have different units and a stacked bar graph might not be the best choice. Also, the initial circles are not the best for the containment link and the collapsible tree is much clearer. In addition, we added the images for the bar graph background to better visualize the category of the items.

b. How have your visualization goals changed?

We focused more on the layout of the whole web page and we abandoned the scrollytelling transition in the proposal, instead, we designed a separate webpage for recommendation to make the flask work stably. For further improvements, we might add more transitions when we switch the bar graph.

c. How have your technical goals changed?

We might further improve the recommendation algorithm with the help of machine learning, but here our main goal is information visualization, so we limit the time for the algorithm part and choose linear programming.