Group Project Final Report

• the difficulties that you have encountered;

First, we have many interesting ideas about the food topic, for example, the relation between food habit and weather, the relation between fruit’s color and popularity. However, we find it’s very hard for us to find a detailed dataset in a specific aspect.

Second, when we collect some relevant datasets, the format of datasets varies in lots of styles, it’s unlikely to directly use these raw datasets in some visualization tools (e.g. Tableau or XDAT). We have to normalize the dataset by programming to modify the file.

Third, different visualization tools need different formats of data, so if we want to use more than one visualization tools on one dataset, we have to modify the format to fit the different tools, respectively.

Fourth, some tools have friendly UI and are easy to visualize a given dataset, others don’t. For example, Google Chart and Rgraph have no UI to use, we have to learn how to use these kind of tools and to code to realize the visualization.

• the different visualization approaches that you have tried;

Tools: Tableau, XDAT, Google Charts, Rgraph, Gephi, D3.js, Python Plot

Algorithm: k-means++,

• the visualization techniques that are realized in your results;

• the explanation for your choice of techniques;

• how your visualization can help answer the queries or questions you have on the data set;

Question 1: For a specific kind of food, which county love to eat it most?

Question 2: Is there any apparent relation between eating habit and health?

Question 3: Is some stereotype in our mind correct? (e.g. Alcohol is bad for health)

Question 4: What’s every countries eating habits? Is there any similarity among countries, around the world?

Question 5: Why do some clusters of countries hold similar eating habits? Because of weather, culture or geography?

• any new insights on the data set that you found using your visualization;

1. Obesity is less relevant to Fats intake, but more relevant to Annual insolation.

2. It is a positive correlation between HALE and Energy, Protein, Fats, Animal Products, Sugar.

3. There is no significant correlation between HALE and Alcohol, Fruit, Vegetable.

4. There is no significant correlation between cholesterol and egg, seafood, offal

5. Clustered Food Habit summary.

• any visualization or functionalities that you think should have included, but have not

done so.

D3.js to realize some novel but intuitive visualization.