

MA Spring 2016

Course: Information and Data Visualization

Tutor: Francesca

Members: DU Xiaojun, HAN Rui, JIANG Han and ZHANG Shiqi

DATA STORY



“Hi! This is the “Cut the Fruit” team, and our group members are Du, Rey, Han and Chichi. We are very grateful that we could have this class taught by Francesca, and we all think that we are in a team full of love and inspirations—although our journey was with many twists and difficulties—we appreciate the sweet and sour in our data story and believe that the whole experience of being

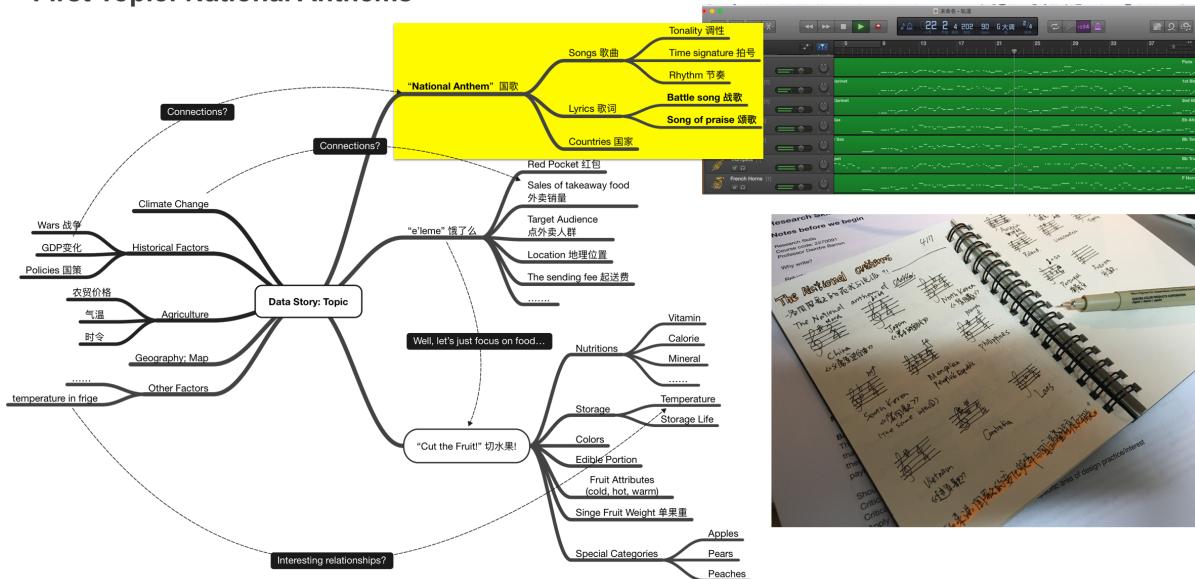
engaged in a completed process of data visualization is what we treasure most.”

The Difficulties of Choosing the Topic

It has been a quite long journey for us to decide our topic. We went from “National Anthem” to “e’leme”, and finally decided to do the topic of “fruit”.

1. National Anthem

First Topic: National Anthems



At the very beginning, we did brainstorming to think about different topics that we are interested in, and we reached an agreement that we all want to do something quite interesting. Du posted a link with the musical notations of the all the national anthems around the world. So, as music lovers, we decided to push it a little bit further to see if “national anthem” could be our topic.

The first step is about collecting the data. We began to set up our database by collecting all the musical notation as well as the midi files of the 201 national anthems all around the world. Then we extracted the information about the tonality, time signature, rhythm as well as the lyrics of each song, and organized the information into an excel form.

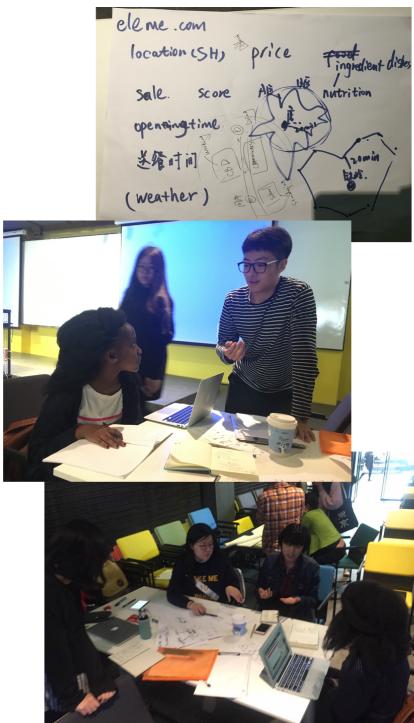
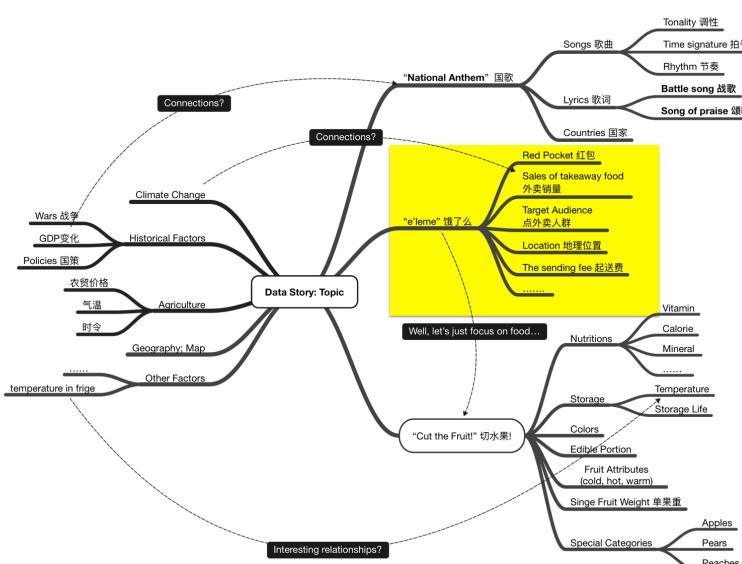
The second step is about finding connections. We wanted to see if there could be any connection between the tonality, time signature, rhythm and lyrics of the national anthems with some historical or economical factors of each country since we noticed that the 3/4 time songs are often used by European developed countries with fewer wars and higher GDP... on the contrary, the 1/4 or 2/4

time marching songs are often used by developing countries with a tortuous and difficult path in modern history...

But we stopped later since it was a little too subjective to analyze the data and find the connections. Also, the national anthems have been performed by different people with different speeds or tonalities or even genres, which means our data is not strong enough to support the design later, and that was why we decided to change the topic.

2. E'leme

Second Topic: E'leme

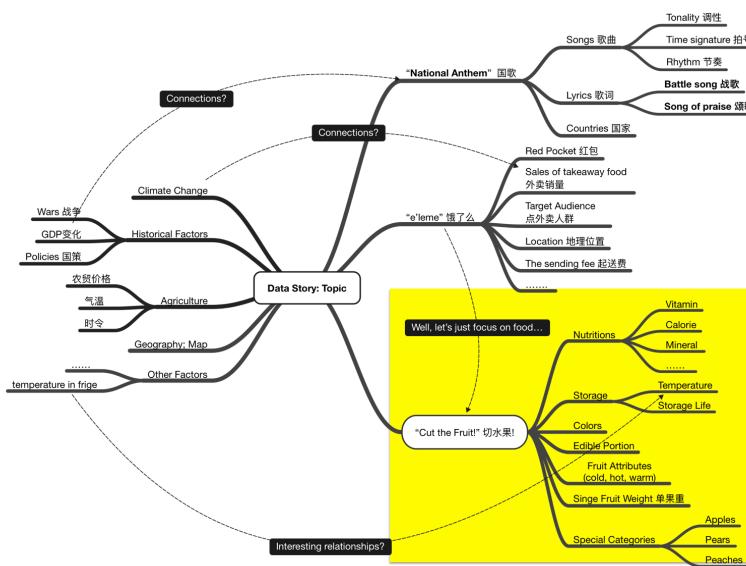


Being afraid of losing our way in the forest again, we decided to ask Francesca if she could provide us some big and reliable dataset. Francesca suggested us to meet with Neema, a very nice and cute exchange student from America who major in Statistics to help us to get some data from "e'leme", a very popular App for ordering takeaway food in China. The brainstorming with Neema helped us to produce many provoking ideas, and we were especially interested in the points of people's behaviors while using "e'leme" such as who how will the target audience use the "Red Pocket", also the sales and the sending fee food among different restaurants... We were also checking the data related to the environment background, such as the climate changes, the locations, and provinces, to see if there could be any connection between these data and "e'leme". However, as the beginners of data visualization, we did not know that how hard web data crawling was. The red pocket information could be the commercial secret of "e'leme" and the other data were also so hard to grab, which means, we might have to change the topic again! Well, although it would not be "e'leme", to save our time and effort, we agreed to just focus on foods this time.

Food is also a big subject with too much information. We were still confused about the things to do and the data sources. So, during one midnight, Han went to data crawling websites like GoSeeker and Bazhuayu, and spent several “2 yuan”, “5 yuan”, “10 yuan” to download the data grabbed by other people, and finally got a ten-yuan-dataset called “Chinese food ingredients (2014)” (《中国食物成分表(2014)》), which then became our first the data source for the final work.

3. “Cut the Fruit!”

Third Topic: Cut the Fruit!



In the dataset of “Chinese food ingredients (2014)”, we found out that the fruit part was very interesting. You could see the differences between fruits in vitamin, calorie and mineral nutrition. Also, the 24 kinds of pears, 17 kinds of apples, and 12 kinds of peaches normally seen in the market in China was a good point to explore too... So, we decided to visualize the data of fruit, and gave our group a name called “Cut the Fruit!” which means we would select the data of fruit and analyze them to reach a result.

Get the Raw Data and Information

Just as said before, we got our first data set on some data-crawling websites like Bazhuayu, GoSeeker and Dataduoduo. However, we thought that these were not enough. To make sure that the data are reliable, we searched for some similar dataset through official channels.

We tested the accuracy of our dataset by comparing what we have with a published book from the library “Chinese Food Ingredients (2002)” (the reason of why we didn’t use this directly is because

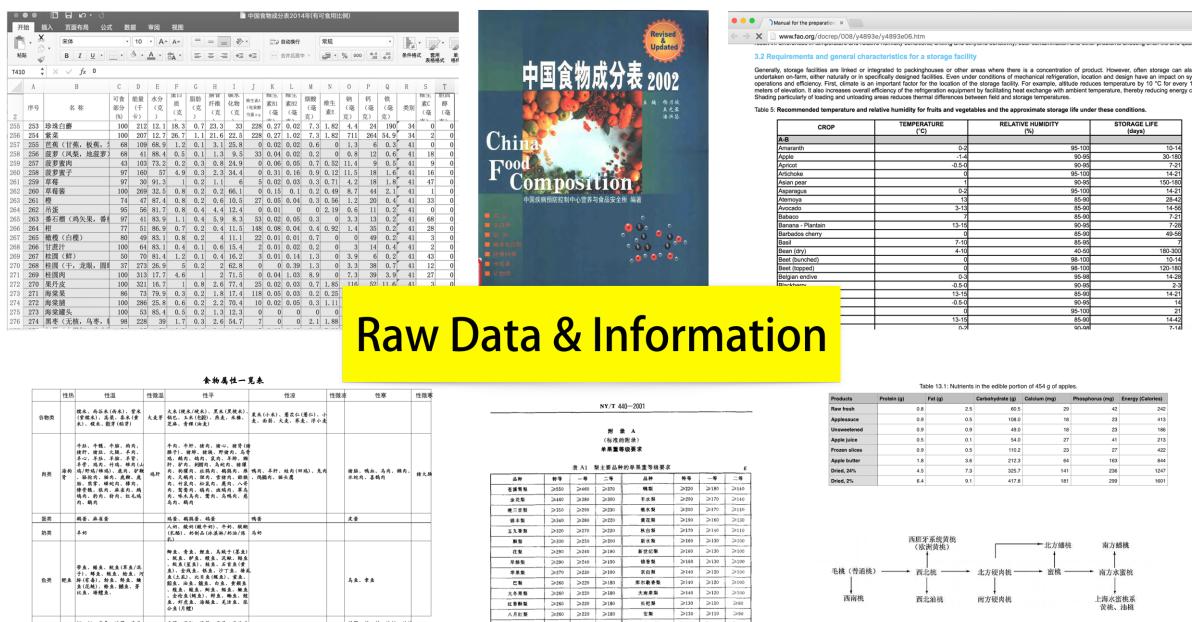
it is only paper version and might cause too much effort to retype the data by ourself), and we checked all the items of fruit and correct a few possible mistakes.

We also went to see the website of FAO (United Nations Food Agriculture Organization) and find the information about the most suitable storage temperature and storage life of each fruit.

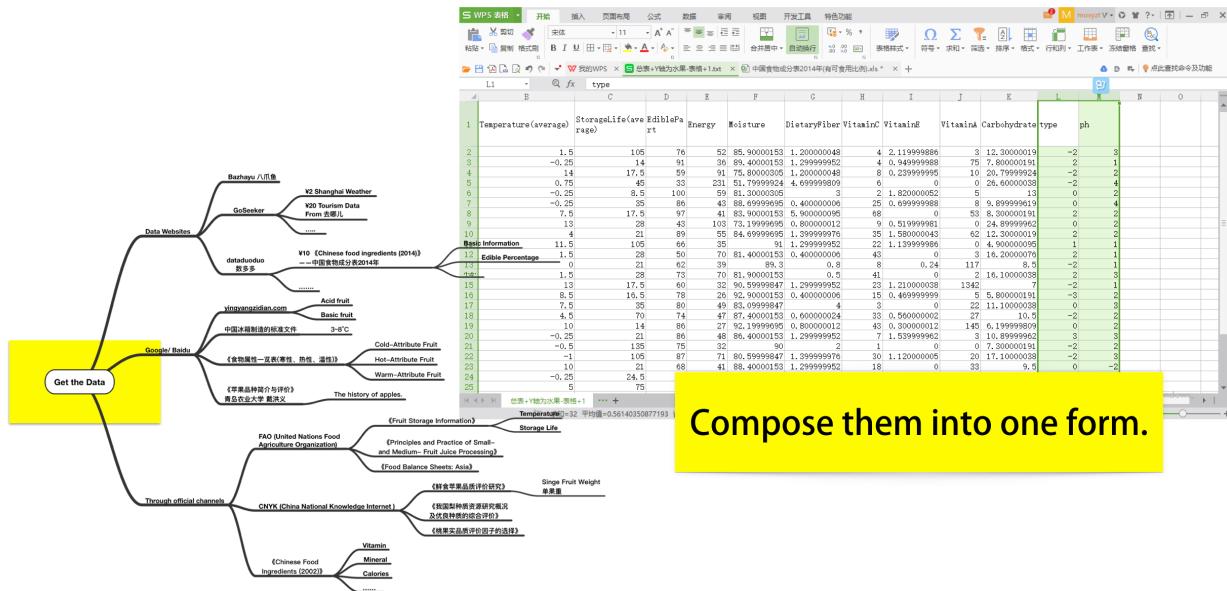
(Recommended temperature and relative humidity for fruits and vegetables and the approximate storage life under these conditions: <http://www.fao.org/docrep/008/y4893e/y4893e06.htm>).

The CNYK(China National Knowledge Internet) is also a reliable place to get the data. We researched about the 24 kinds of pears, 17 kinds of apples, and 12 kinds of peaches through CNYK, by reading the papers such as 《鲜食苹果品质评价研究 (The fresh apple quality assessment)》, 《我国梨种质资源研究概况及优良种质的综合评价 (The pear germplasm resources research in China)》, 《桃果实品质评价因子的选择 (The peach fruit quality evaluation factors)》 ...

Besides the official channels, searching engines like Google and Baidu helped us with our data/information collection too. We checked the 营养字典(yingyangzidian.com) and got the information of the ph value of each fruit, we checked the 《食物属性一览表(寒性、热性、温性)》 to see the attributes (cold, hot, warm) of fruits. We also searched for files like 《中国冰箱制造的标准文件 (China's refrigerator manufacturing standard file)》 in order to find about the relationships between the storage conditions and the fruits.

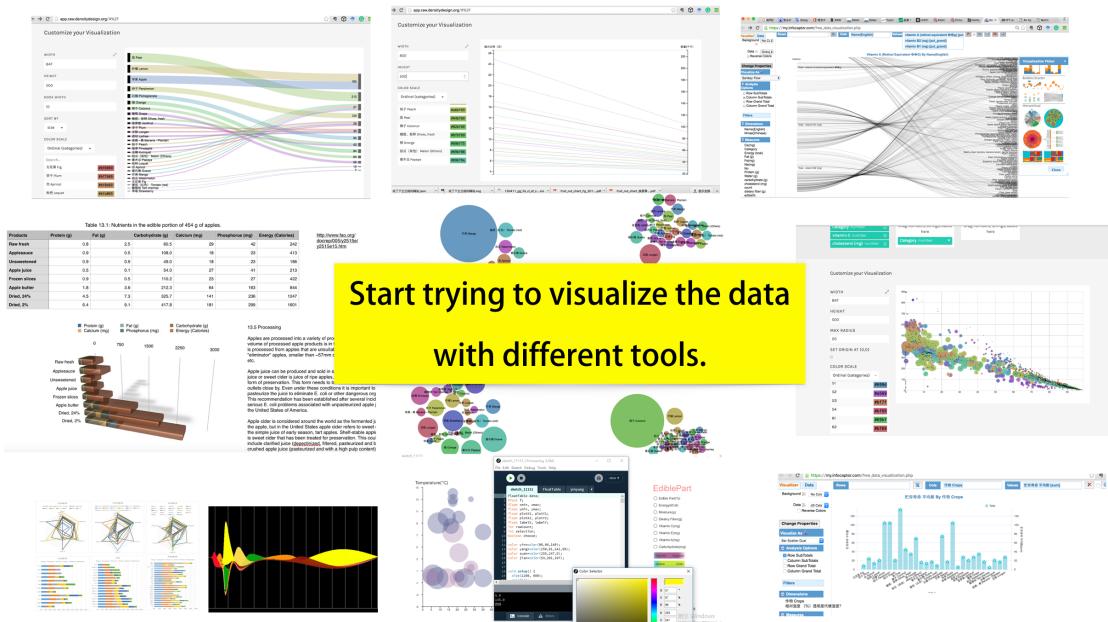


After all, we've got the nutrition information like vitamin, calorie and mineral contents and edible portions of fruits, we've got the suitable storage temperatures and storage life spans of fruits, we've got the colors and the cold/hot/warm attributes of fruits, and we've also got some more detailed information about the special categories of apples, pears, and peaches. We organized all these data and information from data-crawling websites, official channels and searching engines, and composed them into one excel form.



Play with The Data by Using Different Tools

Now we've got our dataset! The next step is to play with the data by trying different tools.



We first tried all the online visualization tools provided by Francesca, such as RAW.

Through exploring by online tools, we've got some interesting facts among fruits, such as coconut contains the maximal calories and the minimal edible portion, mango contains the highest vitamin A, and the longan is full of vitamin B... Also, interestingly, there are differences in nutrition facts among the 24 kinds of pears, 17 kinds of apples, and 12 kinds of peaches.

Our Data Visualization of Fruit

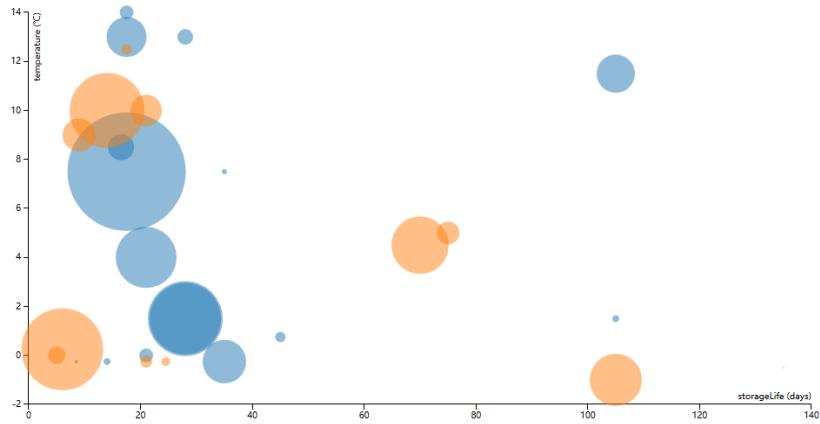
In our process of data visualization, Han is mainly responsible for collecting and integrating the data and creating the basic charts for the later design with tools, Chichi and Rey are responsible for the coding part and they designed the first and second visualization by Processing (the Storage+X Bobble Chart, and the Storage-Color Rectangle Chart), and Du is mainly responsible for the overall style of our work and also the design of the third part with Adobe Illustrator (the Specific Fruit Radar Chart).

Part 1. Storage+X Bobble Chart

Since we've tried our data with websites like RAW (<http://raw.densitydesign.org>) and Google Charts (<https://developers.google.com/chart/?hl=en>), we've got a basic understanding of our dataset, which contains a large member of factors of the fruit such as the energy, carbohydrate, Vitamin A/B/C/E, dietary fiber, edible portion, PH value, Yin&Yang and so on.

作物 Crops	Crops	Temperature	StorageLife	EdiblePart	Energy(kCal)	Moisture(g)	DietaryFiber(g)	VitaminC(mg)	VitaminE(mg)	VitaminA(ug)	Carbohydrate(g)	type	ph
2 苹果 Apple	Apple	1.5	105	76	52	85.90000153	1.20000048	4 2.119999886	3	12.30000019	-2	1	
3 杏 Apricot	Apricot	-0.25	14	91	36	89.40000153	1.299999952	4 0.949999988	75	7.800000191	-2	4	
4 香蕉 - 焦 Banana - Plantain	Banana(Plantain)	14	17.5	59	91	75.80000305	1.20000048	8 0.239999995	10	20.79999924	-2	3	
5 椰子 Coconut	Coconut	0.75	45	33	231	51.79999924	4.699999809	6 0	0	26.60000038	-2	3	
6 无花果 Fig	Fig	-0.25	8.5	100	59	81.30000305	3	2 1.82000052	5	0	13	0 3	
7 葡萄 Grape	Grape	-0.25	35	86	43	88.69999695	0.40000006	25 0.699999988	8	9.899999619	0	2	
8 番石榴 Guava	Guava	7.5	17.5	97	41	83.90000153	5.90000095	68 0	53	8.300000191	2	3	
9 芒果 Jackfruit	Jackfruit	13	28	43	103	73.19999695	0.80000012	9 0.519999981	0	24.89999962	0	3	
10 金橘 Kumquat	Kumquat	4	21	89	55	84.69999695	1.39999976	35 1.58000043	62	12.30000019	2	3	
11 柠檬 Lemon	Lemon	11.5	105	66	35	91	1.29999952	22 1.13999996	0	4.90000095	1	3	
12 龙眼 Longan	Longan	1.5	28	50	70	81.40000153	0.40000006	43 0	3	16.20000076	2	2	
13 枇杷 Loquat	Loquat	0	21	62	39	89.3	0.8	8 0.24	117	8.5	-2	2	
14 荔枝 Lychee	Lychee	1.5	28	73	70	81.90000153	0.5	41 0	2	16.10000038	2	2	
15 芒果 Mango	Mango	13	17.5	60	32	90.59999847	1.29999952	23 1.210000038	1342	7	-2	2	
16 甜瓜(其他) Melon (Others)	Melon(Others)	8.5	16.5	78	26	92.00000153	0.40000006	15 0.469999999	5	5.800000191	-3	2	
17 嫩枝(新鲜) Olives(fresh)	Olives(fresh)	7.5	35	80	49	83.09999847	4	3 0	22	11.10000038	2		
18 橙 Orange	Orange	4.5	70	74	47	87.40000153	0.80000024	33 0.560000002	27	10.5	-2	2	
19番木瓜 Papaya	Papaya	10	14	86	27	92.19999695	0.80000012	43 0.30000012	145	6.199999809	1		
20 桃子 Peach	Peach	-0.25	21	86	48	86.40000153	1.29999952	7 1.539999962	3	10.89999962	3	2	
21 梨 Pear	Pear	-0.5	135	75	32	90	2	1 0	0	7.300000191	-2	1	
22 柿子 Persimmon	Persimmon	-1	105	87	71	80.59999847	1.399999976	30 1.120000005	20	17.10000038	-2	2	
23 菠萝 Pineapple	Pineapple	10	21	68	41	88.40000153	1.299999952	18 0	33	9.5	0	2	
24 李子 Plum	Plum	-0.25	24.5	91	36	90	0.899999976	5 0.74000001	25	7.800000191	0	2	
25 石榴 Pomegranate	Pomegranate	5	75	57	64	78.69999695	4.90000095	13 3.72000029	0	14.5	1	3	
26 草莓 Strawberry	Strawberry	0.25	6	97	30	91.30000305	1.10000024	47 0.709999979	5	6	-2	2	
27 酸樱桃 Tart cherries	Tart cherries	0	5	80	46	88	0.30000012	10 2.22000029	35	9.899999619	3	2	
28 番茄(红色) Tomato (red)	Tomato(red)	9	9	97	19	94.40000153	0.5	19 0.569999993	92	3.5	-2	1	
29 西瓜 Watermelon	Watermelon	12.5	17.5	56	25	93.30000305	0.30000012	6 0.10000001	75	5.5	-3	2	

Rey and Chichi knew some coding before, so we thought that we could let them push their coding skills a little bit through this projects and decided that we would do some interactive visualization.



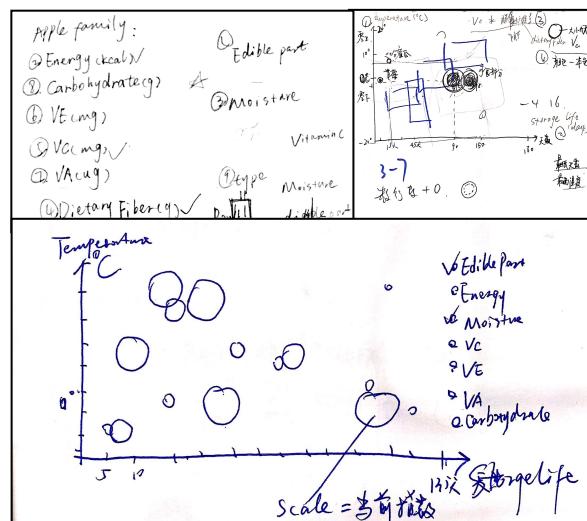
Rey tried our data with D3 first, but she met problems like failing to add the names of that fruit to each circle represents the fruit for some reason... However, Rey didn't give up exploring and succeeded to finish the visualization by using Processing.

```

C:\Users\admin\Desktop\visualization02\fruit\data.csv - Sublime Text 2 (UNREGISTERED)
File Edit Selection Find View Goto Tools Project Preferences Help
02.html
1 name,temperature,storageLife,vc,water,fiber,species
2 苹果,Apple,1.5,105,4,85.90000153,1.200000448,ying
3 杏,Apricot,-0.25,14,4,89.40000153,1.29999952,ying
4 香蕉,Banana,Plantain,14,17.5,8,75.8000035,1.20000048,ying
5 莴苣,Cabbage,-0.25,17.5,45,53.10000024,0.699999809,ying
6 芒果,Fig,-0.25,8.5,25,81.30000353,3,ying
7 葡萄,Grape,-0.25,35,25,88.69999695,0.40000006,ying
8 奇石榴,Guava,7.5,17.5,68,83.90000153,5.90000095,ying
9 俊木瓜,Jackfruit,13,28,93,73.1999997,0.80000012,ying
10 梨,Lemon,4,21,8,54.10000024,1.29999976,ying
11 柑橘,Lemon,11.5,105,22,91,1.29999952,ying
12 龙眼,Longan,1.5,28,43,81.40000153,0.40000006,ying
13 桃吧,Loquat,0,21,8,89.3,89.3,ying
14 桃,Lychee,4.5,28,18,80.30000153,1.2999995,ying
15 芒果,Mango,13,17.5,23,90.59999847,1.29999952,ying
16 甜瓜,{'其他': Melon},Melon (Others),8.5,16.5,15,92.90000153,0.40000006,ying
17 橄榄,{'其他': Olives},Olives, fresh,7.5,35,3,83.09999847,4,ying
18 橙,Orange,4.5,105,30,87.10000024,0.69999976,ying
19 木瓜,Pear,18,14,21,92.10000024,0.80000012,ying
20 桃子,Peach,-0.25,21,7,86.40000153,1.29999952,ying
21 梨,Pear,-0.5,135,1,90,2,yang
22 桃子,Persimmon,105,30,80.59999847,1.29999976,yang
23 桃子,Persimmon,10,18,80.40000153,1.29999952,yang
24 李子,Plum,-0.25,24.5,5,98,0.89999976,yang
25 石榴,Pomegranate,0.25,5,75,13,78.69999695,4.90000095,yang
26 草莓,Strawberry,0.25,6,47,91.30000035,1.100000024,yang
27 酸枣,Tart Cherry,0,10,88,80.40000153,1.2999995,yang
28 西红柿,Tomato (red),9,19,49,40000153,1.5,yang
29 西瓜,Watermelon,12.5,17.5,6,93.30000035,0.30000012,yang
30

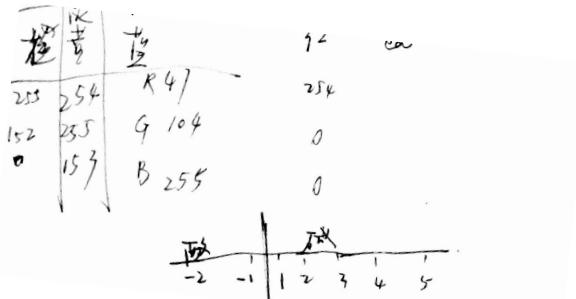
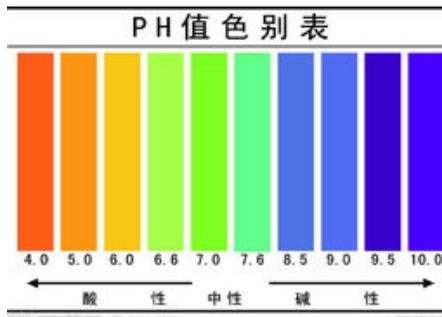
```

We have got the raw data of 28 kinds of fruits, while we looked at our raw data and tried to figure something out for all those numbers. The data are complex and diversity, so if we visualized them in one interface, it might be confusing. So we decided to guess the relations of these attributes, and Neema helped us to analyzed the data with "R"... Well, we realized that the possible relations of the attributes are weak, which means, instead of trying to find the connections so hard, we'd better present the fruit attributes in a more direct way.

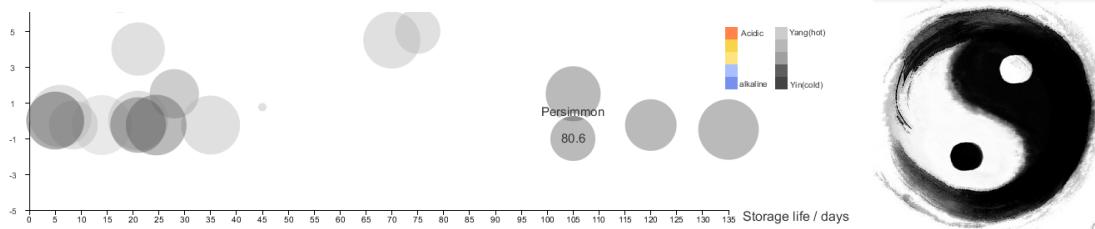


We chose to use the colors of the circles to represent PH value, which is popular in Western countries, and the Yin&Yang on the contrary, which is popular in Eastern countries, especially in the field of Chinese medicine.

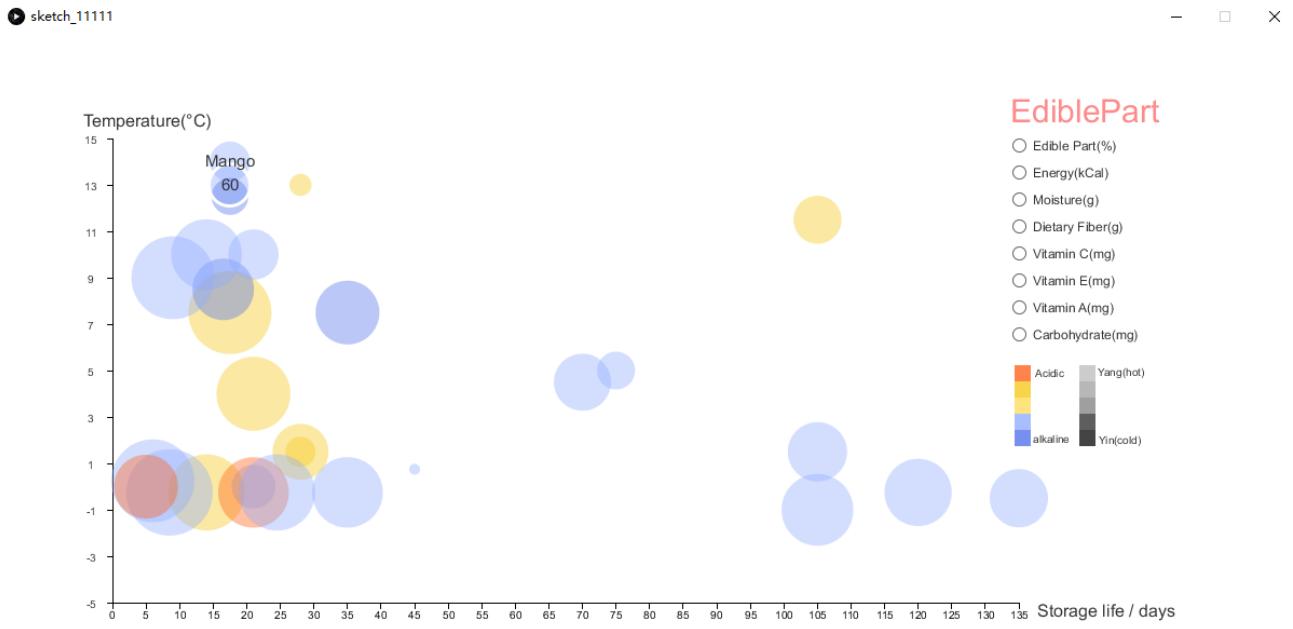
Also, we realized the most fruits are alkaline(basicity), and we've got their PH values which could be put into our visualization. However, if we use the standard colors of the pH values (like the picture below), the colors would all fall in the range of blue. In order to make it more apparently, we define the colors by ourselves, with three colors from orange to Yellow, and then to blue.



According to Eastern traditions, the forces of Yin&Yang are energetic qualities that shape everything in the universe, including our health. The Chinese symbol for Yin is coolness, dampness, and darkness, relative to the Yang qualities of warmth, dryness, and light. Yin foods tend to be cooling or moistening, on the other hand, Yang foods tend to be warming and drying. This has less to do with the actual temperature or moisture of the food and more to do with its "energetics". Boiled spinach, for example, is cooling and moistening, as is baked tofu. Chilled wine is warming, as is roast beef. Toast, while dry to the touch, actually moistens the body. The effects of such food qualities on health have been observed for thousands of years. Your acupuncturist is trained to balance your body's constitution. By observing your body and understanding the energetics of food, you can make food and activity choices to speed your body's healing progress. Since the theory is from "Tai Chi(太极)", Rey defined to use black and white to represent Yin&Yang.



Our final output of the Part 1 looks like the picture below.

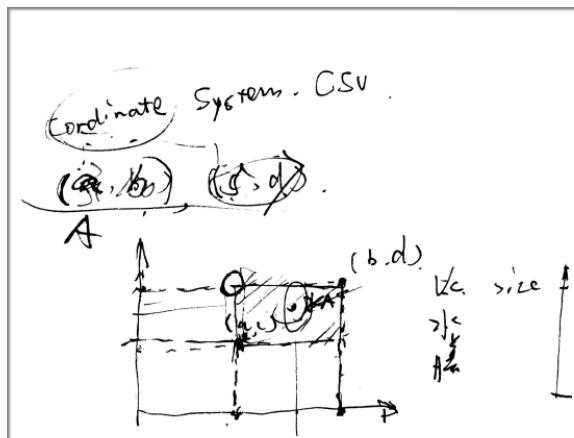
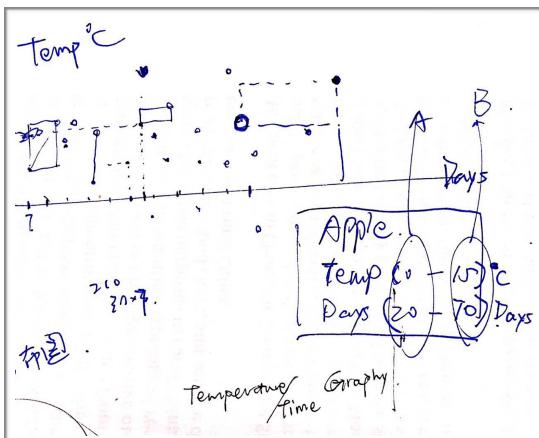
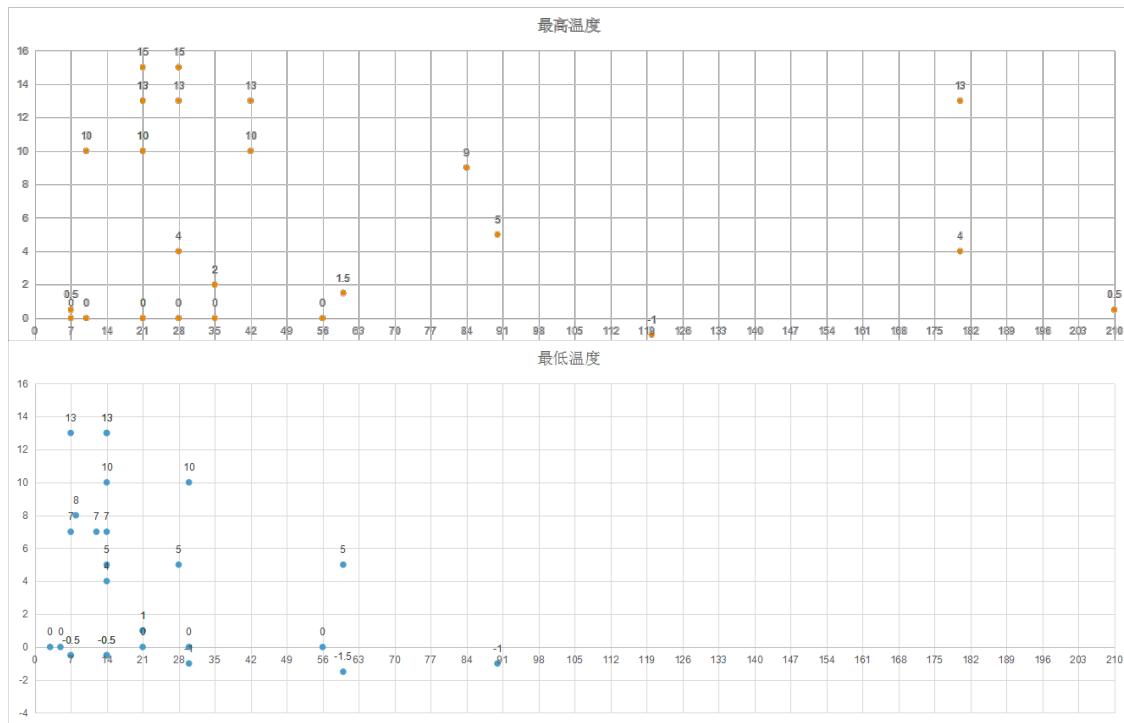


Part 2. Storage-Color Rectangle Chart

Chichi is mainly responsible for the creation of part 2. In this part, we created the work by using a part of our dataset, which contains 8 parameters about the storage information of around 70 kinds of fruits come from FAO, from which we noticed that the humidity condition data were quite similar and the more differences are in temperature condition and storage life. The “temperature” and “storage life” might be considered as the relation of “cause and effect”, so we decided to visualize these two parts.

FAO CORPORATE DOCUMENT REPOSITORY		Produced by:	Agriculture and Consumer Protection
Title: Manual for the preparation and sale of fruits and vegetables...		PDF version	More details
Español	Français		
Table 5: Recommended temperature and relative humidity for fruits and vegetables and the approximate storage life under these conditions.			
CROP	TEMPERATURE (°C)	RELATIVE HUMIDITY (%)	STORAGE LIFE (days)
A-B			
Amaranth	0-2	95-100	10-14
Apple	-1-4	90-95	30-180
Apricot	-0.5-0	90-95	7-21
Artichoke	0	95-100	14-21
Asian pear	1	90-95	150-180
Asparagus	0-2	95-100	14-21
Atemoya	13	85-90	28-42

Then we asked ourselves what can be seen from the two categories? Our answer is that it could probably tell us the capability of different fruit can keep fresh so we can choose the best way to preserve them. So Chichi firstly thought about scatter plot may be helpful if we set the horizontal axis as “storage life” and the vertical axis as “temperature” like below. We worked on two scatter plot charts “max temperature-max fresh life”/ “min temperature-min fresh life”, then overlapped both of them and used the full circles and empty circles to make the difference. But the output seemed really confusing.



Fruit color(outside)



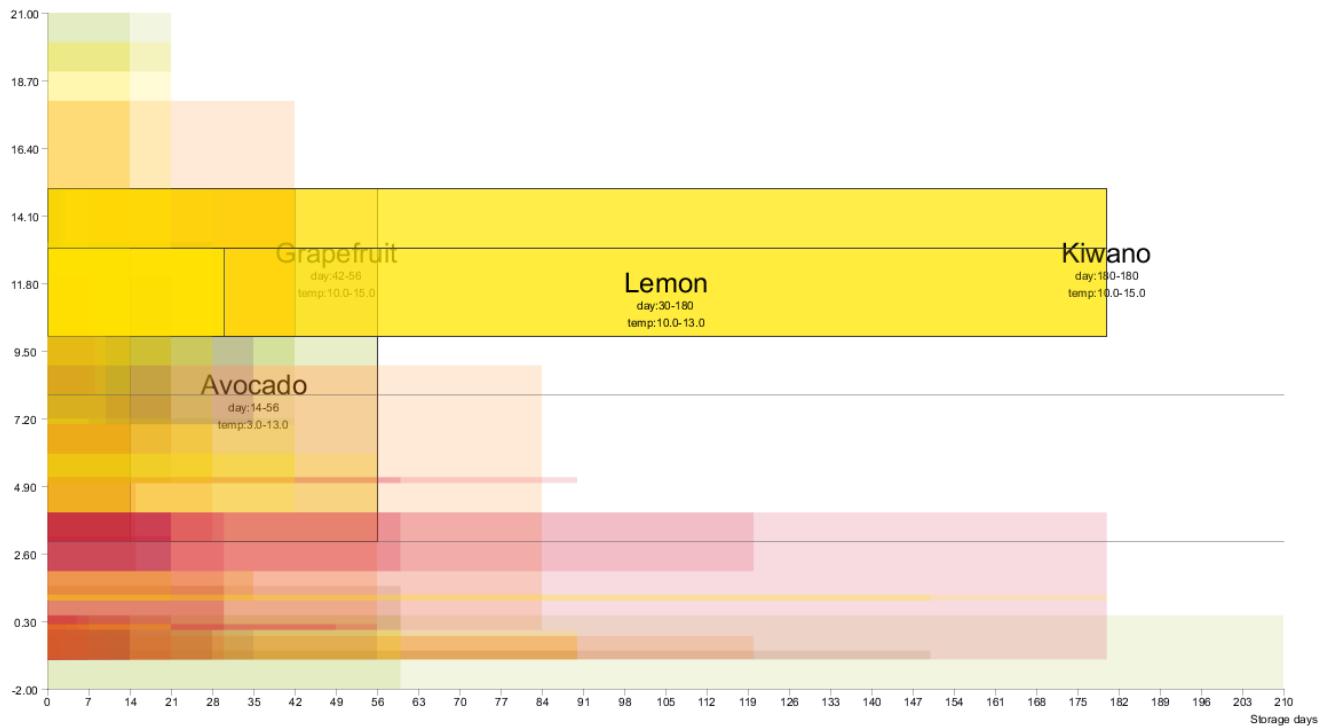
yellow green pink orange red brown purple

2	Apple	.1	4	5	1.5	30	180	150	105
3	Apricot	-0.5	0	0.5	-0.25	7	21	14	14
4	Banana - Plantain	13	15	2	14	7	28	21	17.5
5	Coconut	0	1.5	1.5	0.75	30	60	30	45
6	Fig	-0.5	0	0.5	-0.25	7	10	3	8.5
7	Grape	-0.5	0	0.5	-0.25	14	56	42	35
8	Guava	5	10	5	7.5	14	21	7	17.5
9	Jackfruit	13	13	0	13	14	42	28	28
10	Kumquat	4	4	0	4	14	28	14	21
11	Lemon	10	13	3	11.5	30	180	150	105
12	Longan	1	2	1	1.5	21	35	14	28
13	Loquat	0	0	0	0	21	21	0	21
14	Lychee	1	2	1	1.5	21	35	14	28
15	Manoo	13	13	0	13	14	21	7	17.5

We realized that “temperature” and “storage life” were present in pairs, which means a pair of the full circle and empty circle belong to a specific fruit. We should connect them instead of leaving them separately. At this time, the form of rectangles seems better than the form of circles! The areas of the rectangles represent the capacity of storage between different fruits.

Chichi also thought the colors of the rectangles should have their meaning so he searched for all these fruits in Wikipedia and collected the outside colors of all the fruits then divided them into groups.

However, Chichi still thought that there could be something missing... and luckily, we've got a standard file of Chinese refrigerator manufacturing and noticed that the refrigeration temperature is between 3 to 8 degrees, which is suitable for only some kinds of fruit. Thus, it could be really interesting if we add two lines to represent the refrigeration temperature so that people could know what kinds of fruit is should be put into the fridges while others are not. The final output of Part 2 then came out.



Part 3. Specific Fruit Radar Chart

The issue on specific fruits has been put forward since we want to provide a scientific guideline for the public to know the exact difference among those fruits. You may recognize that Fuji apples are crisp and refreshing while Red Star apples are soft due to your own experience. But you will never try hard to see how they vary in nutritional ingredients and have a clear recognition. Therefore, the third part of our fruits project aims at making a visible comparison among those fruits.

Concerning the limited time we had, we selected three categories of fruits (apple, pear and peach) as the design targets from the database. A doubt later raised. Should we visualize all the data we had about those fruits? (i.e. energy, moisture, fiber, carbohydrate, vitamin and chemical elements) Chichi believed that all the data should be shown in order to form a comprehensive understanding. But Du and Han argued that only representative data could be put into the visualization, which would help maximize the differences and provide a direct identification about where they exist. The argument finally won by the two girls. After consulting her friend who was studying nutriology abroad, Du determined the issues that were usually used to evaluate the fruits. (i.e. dietary fiber, carbohydrate and vitamin A/C/E) Also, the data about energy and edible part were taken into consideration.

- Apple *17
- Pear * 24
- Peach *12



Name (Chinese Name/English)	Name	Energy %	Carbohydrate %	Vitamin E	Vitamin C	Vitamin A	Dietary fiber	edible%	Energy (Kcal)	Dietary fiber (g)	carbohydrate (g)	vitamin A (microg)	vitamin C (mg)	vitamin E (mg)
1 苹果 (香柰) Apple (Apf)	苹果 (香柰苹果) Ap	100.00	57.53	100.00	58.82	34.69	69.00	59.00	1.70	14.00	10.00	0.00	0.84	
1 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	83.05	82.14	54.11	66.67	17.65	44.90	88.00	49.00	2.20	11.50	3.00	4.00	0.84
1 苹果 (金元) Apple (Apf)	苹果 (金元帅苹果) Ap	84.75	87.14	41.78	66.67	88.24	22.45	78.00	50.00	1.10	12.20	15.00	4.00	0.61
1 苹果 (黄妃) Apple (Apf)	苹果 (黄元帅苹果) Ap	93.22	92.14	14.38	66.67	88.24	36.73	80.00	52.00	1.80	12.90	15.00	4.00	0.21
1 苹果 (国光) Apple (Apf)	苹果 (国光苹果) Ap	91.53	83.83	7.58	58.82	19.41	42.40	78.00	54.00	0.80	12.00	10.00	0.00	0.11
1 苹果 (红富士) Apple (Apf)	苹果 (红富士苹果) Ap	83.05	81.29	23.34	50.00	100.00	18.37	87.00	69.00	0.90	11.10	3.00	3.00	0.37
1 苹果 (红肉) Apple (Apf)	苹果 (红肉苹果) Ap	100.00	97.86	1.37	50.00	41.18	12.24	84.00	59.00	0.60	13.70	7.00	3.00	0.02
1 苹果 (红肉) Apple (Apf)	苹果 (红肉苹果) Ap	76.27	80.87	40.63	53.33	48.22	42.40	88.00	62.00	2.10	9.00	10.00	2.00	0.46
1 苹果 (伏花) Apple (Apf)	苹果 (伏花苹果) Ap	76.27	75.71	10.27	33.33	0.00	24.49	86.00	65.00	1.20	10.60	0.00	2.00	0.15
2 苹果 (秋红) Apple (Apf)	苹果 (红星苹果) Ap	77.97	75.87	4.79	33.33	11.76	30.61	86.00	46.00	1.80	11.10	2.00	2.00	0.07
2 苹果 (红脆) Apple (Apf)	苹果 (红脆苹果) Ap	96.61	95.43	14.38	16.67	11.76	16.33	85.00	87.00	0.80	13.60	2.00	1.00	0.21
2 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	84.75	80.71	0.00	16.67	47.00	88.00	80.00	80.00	1.70	11.30	4.00	0.00	0.00
2 苹果 (国贵) Apple (Apf)	苹果 (国贵苹果) Ap	74.58	73.71	0.00	0.00	17.65	100.00	90.00	44.00	4.90	9.00	3.00	0.00	0.00
6 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	72.88	71.43	0.00	0.00	11.76	95.92	84.00	43.00	4.70	10.00	2.00	0.00	0.00
7 苹果 (秋红) Apple (Apf)	苹果 (秋红苹果) Ap	59.32	58.57	0.00	0.00	0.00	75.51	85.00	35.00	3.70	8.20	0.00	0.00	0.00
8 苹果 (不) Apple (Apf)	苹果 (不) Apple (Apf)	50.85	47.86	0.00	0.00	0.00	34.69	96.00	30.00	1.70	6.70	0.00	0.00	0.00

2. Establish the function:

$$D2 = K2 / \text{MAX}(K2:K18) * 100$$

1. Selected items from raw data.

Name (Chinese Name/English)	Name	Energy %	Carbohydrate %	Vitamin E	Vitamin C	Vitamin A	Dietary fiber	edible%	Energy (Kcal)	Dietary fiber (g)	carbohydrate (g)	vitamin A (microg)	vitamin C (mg)	vitamin E (mg)
1 苹果 (香柰) Apple (Apf)	苹果 (香柰苹果) Ap	100.00	57.53	100.00	58.82	34.69	69.00	59.00	1.70	14.00	10.00	0.00	0.84	
1 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	83.05	82.14	54.11	66.67	17.65	44.90	88.00	49.00	2.20	11.50	3.00	4.00	0.84
1 苹果 (金元) Apple (Apf)	苹果 (金元帅苹果) Ap	84.75	87.14	41.78	66.67	88.24	22.45	78.00	50.00	1.10	12.20	15.00	4.00	0.61
1 苹果 (黄妃) Apple (Apf)	苹果 (黄元帅苹果) Ap	93.22	92.14	14.38	66.67	88.24	36.73	80.00	52.00	1.80	12.90	15.00	4.00	0.21
1 苹果 (国光) Apple (Apf)	苹果 (国光苹果) Ap	91.53	89.29	7.53	66.67	58.82	16.33	78.00	54.00	0.80	12.60	10.00	4.00	0.11
1 苹果 (红肉) Apple (Apf)	苹果 (红肉苹果) Ap	83.05	81.43	24.66	50.00	100.00	18.37	87.00	69.00	0.90	11.10	3.00	3.00	0.37
1 苹果 (红肉) Apple (Apf)	苹果 (红肉苹果) Ap	100.00	97.86	1.37	50.00	41.18	12.24	84.00	59.00	0.60	13.70	7.00	3.00	0.02
1 苹果 (伏花) Apple (Apf)	苹果 (伏花苹果) Ap	76.27	80.87	40.63	53.33	48.22	42.40	88.00	62.00	2.10	9.00	10.00	2.00	0.46
1 苹果 (伏花) Apple (Apf)	苹果 (伏花苹果) Ap	76.27	75.71	10.27	33.33	0.00	24.49	86.00	65.00	1.80	12.90	15.00	4.00	0.31
2 苹果 (秋红) Apple (Apf)	苹果 (红星苹果) Ap	77.97	75.87	4.79	33.33	11.76	30.61	86.00	46.00	1.80	11.10	2.00	2.00	0.07
2 苹果 (红脆) Apple (Apf)	苹果 (红脆苹果) Ap	96.61	95.43	14.38	16.67	11.76	16.33	85.00	87.00	0.80	13.60	2.00	1.00	0.21
2 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	84.75	80.71	0.00	16.67	47.00	88.00	80.00	80.00	1.70	11.30	4.00	0.00	0.00
2 苹果 (国贵) Apple (Apf)	苹果 (国贵苹果) Ap	74.58	73.71	0.00	0.00	17.65	100.00	90.00	44.00	4.90	9.00	3.00	0.00	0.00
6 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	72.88	71.43	0.00	0.00	11.76	95.92	84.00	43.00	4.70	10.00	2.00	0.00	0.00
7 苹果 (秋红) Apple (Apf)	苹果 (秋红苹果) Ap	59.32	58.57	0.00	0.00	0.00	75.51	85.00	35.00	3.70	8.20	0.00	0.00	0.00
8 苹果 (不) Apple (Apf)	苹果 (不) Apple (Apf)	50.85	47.86	0.00	0.00	0.00	34.69	96.00	30.00	1.70	6.70	0.00	0.00	0.00

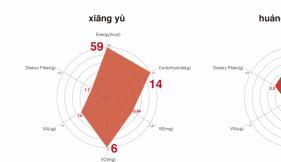
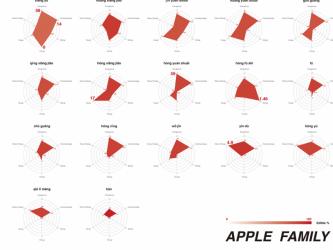
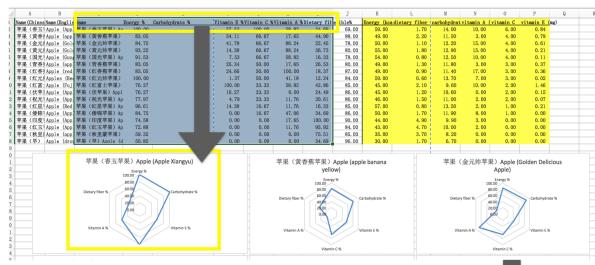
Name (Chinese Name/English)	Name	Energy %	Carbohydrate %	Vitamin E	Vitamin C	Vitamin A	Dietary fiber	edible%	Energy (Kcal)	Dietary fiber (g)	carbohydrate (g)	vitamin A (microg)	vitamin C (mg)	vitamin E (mg)
1 苹果 (香柰) Apple (Apf)	苹果 (香柰苹果) Ap	100.00	57.53	100.00	58.82	34.69	69.00	59.00	1.70	14.00	10.00	0.00	0.84	
1 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	83.05	82.14	54.11	66.67	17.65	44.90	88.00	49.00	2.20	11.50	3.00	4.00	0.84
1 苹果 (金元) Apple (Apf)	苹果 (金元帅苹果) Ap	84.75	87.14	41.78	66.67	88.24	22.45	78.00	50.00	1.10	12.20	15.00	4.00	0.61
1 苹果 (黄妃) Apple (Apf)	苹果 (黄元帅苹果) Ap	93.22	92.14	14.38	66.67	88.24	36.73	80.00	52.00	1.80	12.90	15.00	4.00	0.21
1 苹果 (国光) Apple (Apf)	苹果 (国光苹果) Ap	91.53	89.29	7.53	66.67	58.82	16.33	78.00	54.00	0.80	12.60	10.00	4.00	0.11
1 苹果 (红肉) Apple (Apf)	苹果 (红肉苹果) Ap	83.05	81.43	24.66	50.00	100.00	18.37	87.00	69.00	0.90	11.10	3.00	3.00	0.37
1 苹果 (红肉) Apple (Apf)	苹果 (红肉苹果) Ap	100.00	97.86	1.37	50.00	41.18	12.24	88.00	62.00	2.10	9.00	10.00	2.00	0.46
1 苹果 (伏花) Apple (Apf)	苹果 (伏花苹果) Ap	76.27	80.87	40.63	53.33	48.22	42.40	88.00	65.00	1.80	12.90	15.00	4.00	0.31
1 苹果 (伏花) Apple (Apf)	苹果 (伏花苹果) Ap	76.27	75.71	10.27	33.33	0.00	24.49	86.00	62.00	1.50	11.30	2.00	2.00	0.46
2 苹果 (秋红) Apple (Apf)	苹果 (红星苹果) Ap	77.97	75.87	4.79	33.33	11.76	30.61	86.00	38.00	0.80	13.60	2.00	1.00	0.21
2 苹果 (红脆) Apple (Apf)	苹果 (红脆苹果) Ap	84.75	80.00	0.00	16.67	47.06	34.69	86.00	50.00	1.70	11.10	8.00	1.00	0.00
2 苹果 (红玉) Apple (Apf)	苹果 (红玉苹果) Ap	76.58	70.71	0.00	0.00	17.65	100.00	95.00	43.00	4.90	9.00	3.00	0.00	0.00
2 苹果 (国贵) Apple (Apf)	苹果 (国贵苹果) Ap	72.88	71.43	0.00	0.00	11.76	95.92	84.00	43.00	4.70	10.00	2.00	2.00	0.46
2 苹果 (不) Apple (Apf)	苹果 (不) Apple (Apf)	58.32	57.82	0.00	0.00	12.24	95.92	83.00	35.00	3.70	8.20	0.00	0.00	0.00
8 苹果 (不) Apple (Apf)	苹果 (不) Apple (Apf)	50.85	47.86	0.00	0.00	34.69	96.00	30.00	1.70	6.70	0.00	0.00	0.00	0.00

Data were ready and then we started to think about the design. It still remained a huge question about how to visualize the data and convey our ideas effectively to the public at that moment. We had three categories of fruits and each had various kinds while the evaluation contained different

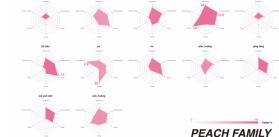
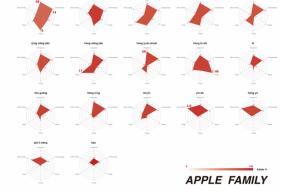
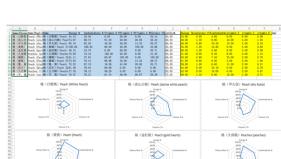
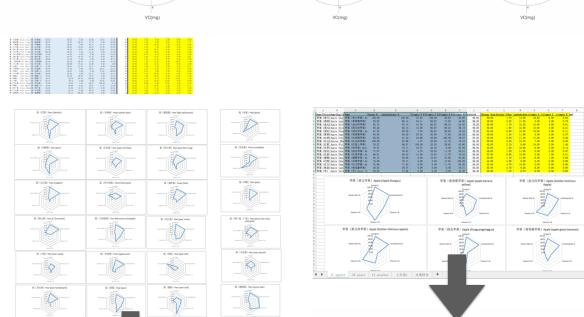
items. Han first tried the cumulative bar chart, gathering all the items together. But it turned out to be a failure since we could hardly read any detailed information from those graphs, especially the maximum value in each item. Fortunately, we thought of radar chart later, a useful tool to compare multiple data. Du did the attempt subsequently. Each axis represented an item. (dietary fiber, energy, carbohydrate, vitamin E, vitamin C, vitamin A). The edible parts data, considered somehow subjective, were visualized in the form of opacity. It seemed that all things were going well.

However, we soon realized that we lacked a criterion for the data. To be more scientific, it required some rules setting that the maximum of each item should be put in the outermost circle on each axis. Efforts were restarted. Han worked on Excel to produce some basic radar charts related to the issue and Du went on designing the new visualization. The differences among those fruits could be seen easily from the final output.

3. Draw the radar charts with excel.



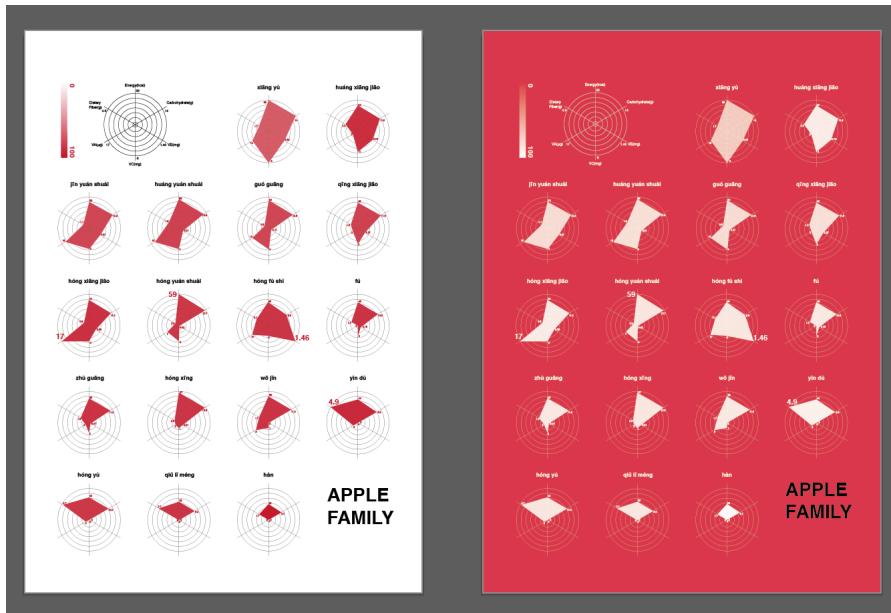
4. Create our design with Ai.



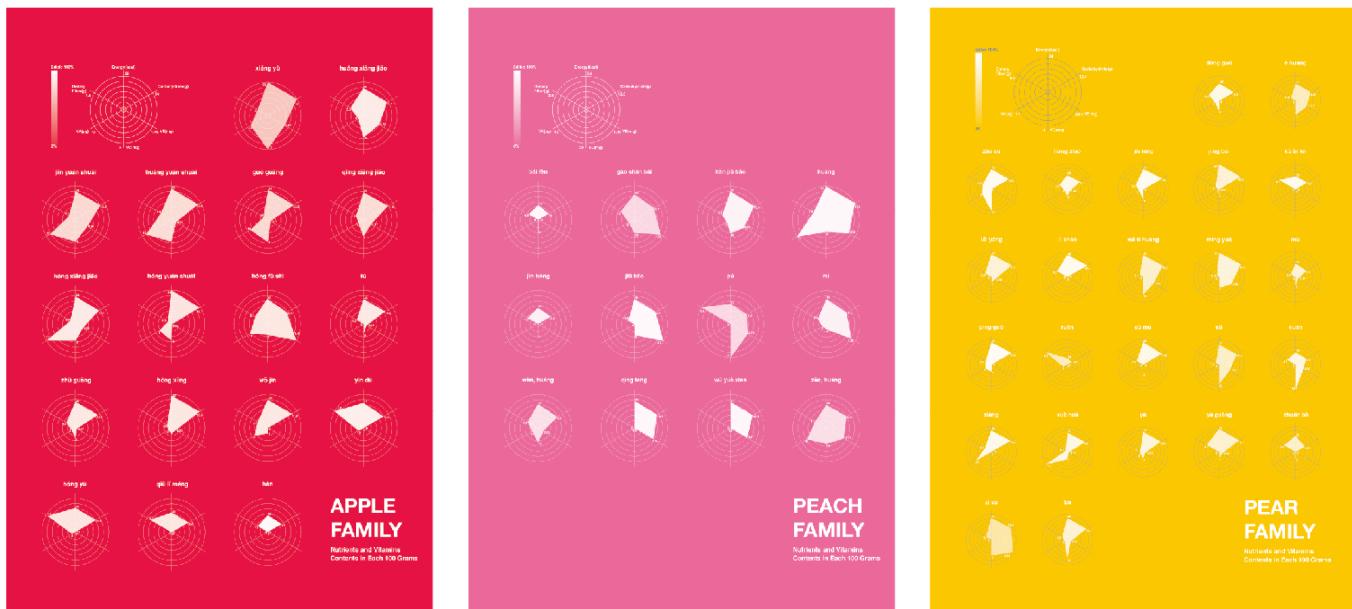
5. Create the other two.

PEAR FAMILY

Also, we've been thinking about the right colors to use. We made another version and finally chose to use the fruit's colors as the background colors (like the picture shown on the right), since it



makes the center part of our radar chart to be like "a bite on the apple/ pear/ peach", and that's fun! Our final output of the Part 3 looks like the pictures below.



Reflection

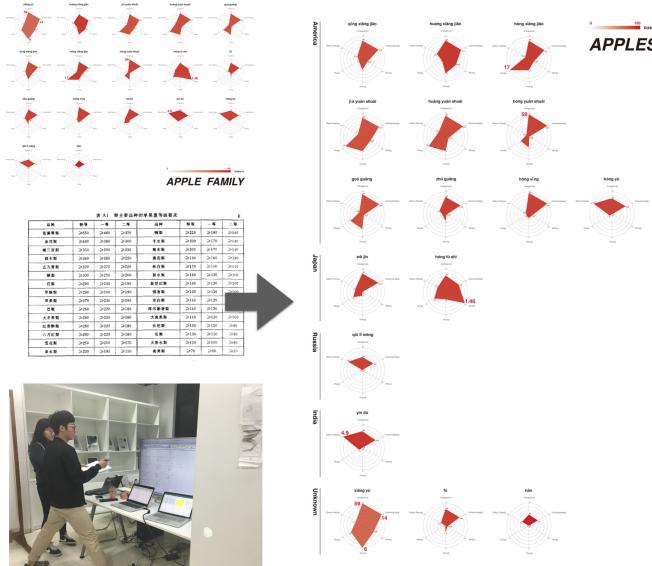
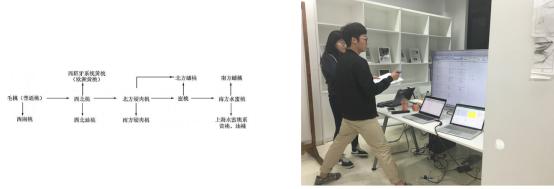
We know that our works are not perfect and we've been looking for ways to do better.

For Part 1, there might be a bug since the audience could not compare the Yin&Yang and the PH values at the same time. Maybe we would improve the reading experience by canceling the two option buttons and show the two factors directly on the circles. Maybe we would still use color as the expression for PH values, but for Yin&Yang, we could think about another solution such as adding symbols like "+" and "-" or simple shapes in the centers of the circles.

For Part 2, we set the horizontal axis with days every 7 days, however, we should just use “1 week” directly as the unit, instead of using “7 days”. Another problem is that there are large differences in storage days among fruits, some fruit with short storage life can not be shown detailed with a 210-days horizontal axis, so we need to divide the data graph into 3 scales, probably “1 weeks”, “1 month”, and “3 months”.

For Part 3, we then began to doubt if there was any connection to the places of origin. We found the relevant data about apples and divided the radar charts according to their location. The results seemed not optimistic, for there was no evidence from the visualization, supporting our assumption. Despite the disappointment, we will go on and try to make more fun out of the data.

表4 各个品种桃果实品质的单因子合理满意度 Table 4 Single factor rationalization-satisfaction index of quality characteristic of peaches						
品系	品种	单果质量	硬度	水分含量	固酸比	风味
北京39号	0.540	0.770	1.000	0.350	1.000	
北京24号	0.656	0.015	0.538	0.600	0.313	
桃王一号	0.387	0.490	0.591	0.750	0.438	
大久保	0.998	0.005	0.679	0.572	0.563	
水蜜桃	0.767	1.000	0.887	0.554	0.875	
蟠桃王仙桃	0.998	0.005	0.679	0.572	0.563	
桃	0.003	0.990	0.538	0.543	0.000	
早久保	0.034	0.000	0.275	0.454	0.250	
北京3号	1.000	0.790	0.075	0.000	0.750	
晚蜜	0.617	1.000	0.885	1.000	0.500	
瑞阳	0.622	0.692	0.000	0.157	1.000	
南国桃	0.721	0.000	0.909	0.585	0.585	
蟠桃	0.000	0.462	0.709	0.455	0.899	
蟠桃49	1.000	0.692	0.233	0.586	0.182	
碧桃	0.811	1.000	0.442	1.000	0.000	
早白凤桃	0.406	0.462	0.000	0.000	0.182	
蟠桃	0.337	1.000	0.000	0.000	1.000	
南国桃18号	1.000	0.000	0.000	0.221	0.300	
桃	0.600	0.917	0.877	1.000	0.000	



We also want to say something about our data journey with twists and difficulties. We are often in a dilemma, where if we've already got a good idea, we may still fail for lack of a good dataset, vice versa. So we were thinking about the best way to make our design process smoother: maybe we

could firstly find a reliable, official dataset, and secondly find a small but meaningful point from the dataset, thirdly we could expand and detail the topic even beyond the original data, and finally we finish our data visualization by revising our design constantly with the help of sketching, coding, layout, etc.

At last but not least, we are so glad that Francesca encouraged us to develop a good habit of keeping our data diaries, which is very helpful for not only this course. We appreciate our audience's comments and discussions. If you have some thoughts about our design and the process, please do not hesitate to contact us!