The University of British Columbia

CPSC 547 - Information Visualization

Dr. Tamara Munzner

Country vs. Country:

Food & Allergy Edition

Update

Jose R. Carvajal

joc08@mail.ubc.ca

Lucca Siaudzionis

lsiau@cs.ubc.ca

Alireza Iranpour

alireza.iranpour@cs.ubc.ca

Contents

Contents 1

Abstract 2

Introduction 2

Related Work 3

Data Abstraction: Update 4

Task Abstraction 5

- R I - Encode allergen-safety of a country. 6
- R II - Support search for “allergen-safe” countries 6
- R III - Handle incomplete data (missing values) 6
- R IV - Easily add new food attributes 6

Use Cases 6

- Use case 1: “I’m dangerously allergic to X, Y, and Z, and I really want to travel” 6
- Use case 2: Explore - Enjoy 7
- Use case 3: Discover/Derive 7

Implementation 8

Exploration 8

Milestones 9

References 10

Abstract

Mindful of the significant segment of the world population that have food allergies [1, 2, 3], the rise in air travel for leisure [4], and the convergence of food consumption habits in the face of globalization [5], the aim of this project is to design a tool that is simultaneously user interactive, visually informative, and helpful to guide users in their decision making. Informing our research is the question of: where is it safe to travel in light of my allergies or significant dietary restrictions?

Introduction

The purpose of this project is to create an interactive visualization tool, in the shape of a website, that allows users to answer the question of: where is it safe for me, or my children, to travel to in light of our food allergies? Informing the motivation for this project is the lack of an existing tool that serves a similar purpose, the large portion of the population that is affected by food allergies, especially children, and the rise in globalized food consumption habits and leisure travel.

Presently, according to the World Allergy Organization, 220 to 520 million people around the world live with food allergies [3]. According to a study by Loh and L.K. Tang, the percentage of the world population challenged by food allergies can be as high as 10%, with a great prevalence noted amongst young children [2]. To provide context, in the United States, this accounts, yearly, for 2,000 hospitalizations and 200 deaths per year [3]. Although, at present, 170 foods have been known to cause allergic reactions due to consumption, eight main allergens: milk, eggs, fish, crustacean shellfish, tree nuts, peanuts, wheat, and soybean are responsible for the majority of adverse food reactions [1].

At a global scale, while the mosaic of world cuisines is complex, at a national level it is reasonable to say that certain ingredients compromise the basis of that nation's cuisine. For instance, for historical, colonial, and environmental reasons, well beyond the scope of this project, traditional Costa Rican food consists of the following ingredients: rice, beans, coffee, potatoes, plantains, bananas, pork, maize, coconut, beef and fish [10]. As such, despite being a famous travel destination due to its natural resources, it also boasts of a national traditional cuisine that is relatively allergen friendly and bland: in terms of spice.

Our hope is that in mind of our globalized world which has seen a dramatic expansion in leisure air travel [4] and in the consumption of non local goods due to greater interconnectedness of goods, cultures and capital [5], that information on food allergens is not unattainable or overwhelming when making travel decisions. As such, the present authors ask themselves the following questions: in mind of my severe food allergies, and with the available knowledge of the most common ingredients used in each country's cuisine, what countries is it relatively safe for me to travel to? And, where can I access this information for free, in an easy to consume, easily personalized, and informative tool?

The aim of this tool is to inform end users in their decision to travel and how to prepare if they decide to travel in light of their dietary restrictions. The tool is not meant to homogenize cultures or to make destinations unfriendly. Rather, as an end user, it should help me decide what measures I should keep in mind if I decide to travel to a country with a high presence of allergens that affect me in their traditional cuisine.

Related Work

Presently, the authors have found no similar work or comprehensive tool that addresses this issue. Despite the existence of certain blogs or articles that show best practices on how to travel despite your food allergies [6, 7, 8, 9] there exists no data base or interactive tool that links together the eight most common food allergens and their prevalence or presence in traditional national cuisines or food consumption habits.

However, drawing from these articles and blogs, it is clear that despite the rather unique problem this presents for people, "can I travel in mind of my allergies", allergens *are* a consideration when making decisions to travel for people with severe food or environmental allergies [9, 8]. For instance, in Scott [9], the author devotes a whole section of his website to the issue of travel, leisure and allergies. When travelling, due to his daughter's severe allergies to shellfish, peanuts, and eggs, he makes clear that distance to hospital, local food labelling practices, and ability to fluently speak the local language inform his decisions to travel. For users like Scott, Laura, Missy and Megan [8, 9] when making travel plans for leisure, an important consideration is the predominance of allergens in the national consumption habits of their destination. As a blog said, you make the decision to travel to Italy *despite* being allergic to gluten [9]. As such, we find this project relevant in its ability to provide a useful service that is currently being unmet. We point to the first imperative for travelling with allergies, as outlined in all of these blogs and web articles: research where you're going [6, 7, 8, 9,] as the ideal end user for our tool.

To inform our research for the development of this tool, in mind of the niche nature of the task, we intend to significantly review the literature on geospatial data visualization to inform our choice of idioms. On the issue of geospatial visualization of food data, we have have identified four relevant studies to inform our research Johnson, Cozart and Isokpehi [11], Lunterova, Spetko, and Palamas [12], Maličká [13], and Zhihua [14]. More broadly, to inform our consideration on the right choice of idioms for geospatial data we draw from Munzner [15], Karimi and Karimi [16] and Speckmann and Vaarbeek [17] on their respective research into geospatial data visualization. At this time, the literature review is ongoing and relevant related work will be integrated, when appropriate, as relevant scholarly work is identified. Findings, insights and considerations from this lit. review will discussed below.

Data Abstraction: Update

We have completed our dataset for development purposes. There are three major changes from the abstraction:

1. **Focus on consumption only.** Consumption data is more important than purely knowing which allergens are present in the popular "traditional" dishes in a country. If a user is worried about cross-contamination, a very high consumption rate of, say, shellfish, is much more worrying than simply knowing what the popular dish is. All consumption is calculated per capita.
2. **Time information.** We have gathered each country's consumption of each allergen over several years.
3. **One table per allergen.** The reason behind this change is purely technical convenience. The addition/deletion of allergens in our early designs for the tool consist now of commenting in/out a line of code. Each of these tables consist of *at least* four columns: two to represent the country (name and entity code), one for the year, one for the consumption rate. Of course, not all data we found was in this format. A lot of work went into cleaning the data to conform to this.

Something worth noting is that all our design respects R IV from the task abstraction (following section), meaning that the effort required to add another allergen *once data is gathered* is very low. So far, we have data for the following:

1. Fish and shellfish
2. Egg
3. Milk
4. Rice
5. Wheat
6. Peanut
7. Some types of tree nuts
8. Beef and buffalo. Though this is not a popular allergen, it is a popular food restriction, including for one of the authors, and thus it was added

Each of those is stored as a .csv file, representing a **two-dimensional table**. This two-dimensional table could be transformed into a multi-dimensional one, since either the "country" and "year" columns could be grouped. With the current two-dimensional version, each table consists of approximately 4 columns and approximately 10,000 rows (~200 countries times ~50 years).

To illustrate this, the following is a screenshot of part of our egg-consumption data:

	A	B	C	D
1	Entity	Code	Year	Eggs - Food supply q
2	Afghanistan	AFG	1961	0.94
3	Afghanistan	AFG	1962	1.01
4	Afghanistan	AFG	1963	1.04
5	Afghanistan	AFG	1964	1.06
6	Afghanistan	AFG	1965	1.1
7	Afghanistan	AFG	1966	1.11
8	Afghanistan	AFG	1967	1.24
9	Afghanistan	AFG	1968	1.12
10	Afghanistan	AFG	1969	1.17
11	Afghanistan	AFG	1970	1.04
12	Afghanistan	AFG	1971	0.94

In some cases, such as peanuts, we have consumption data from two different sources, leading to 5 columns instead of 4. Also, in the case of peanut-consumption, we only have data for 5 years, rather than 50, leading to 10 times less rows.

	A	B	C	D	E
1	Entity	Code	Year	PeanutConsumption	EstPeanutConsumpti
2	China	CHN	2012	12.36	13.73
3	China	CHN	2013	12.28	13.64
4	China	CHN	2014	12.14	13.49
5	China	CHN	2015	12.05	13.39
6	China	CHN	2016	12.11	13.45
7	India	IND	2012	3.83	3.71
8	India	IND	2013	4.31	4.79
9	India	IND	2014	3.35	3.72
10	India	IND	2015	3.12	3.47
11	India	IND	2016	4.25	4.72
12	Nigeria	NGA	2012	11.59	12.88

Task Abstraction

The following is a tentative set of requirements that we aim to implement.

R I - Encode allergen-safety of a country.

For each given allergen, we will encode how "safe" a country is in relation to it. We will use the information of an ingredient being popular in said country as a proxy for this information. When available, we will also use the per-capita consumption to create a score. If an allergen is common/heavily consumed in a country, then that place will be deemed unsafe for people with that allergy.

R II - Support search for "allergen-safe" countries

This is an interactive feature. A user should be able to input their personal set of allergens and easily visualize which countries are considered "safe to travel." The encoding mentioned in the previous requirement will be vital for this task. Comparison a country's safety with that of its geographical neighbours should be simple.

R III - Handle incomplete data (missing values)

As we gather data for our dataset, it is inevitable that many values will be missing. In fact, that has already happened several times since we started the process. The likely solution we will take is to omit those values with, for example, grey colour-coding. However, other solutions will be explored as well.

R IV - Easily add new food attributes

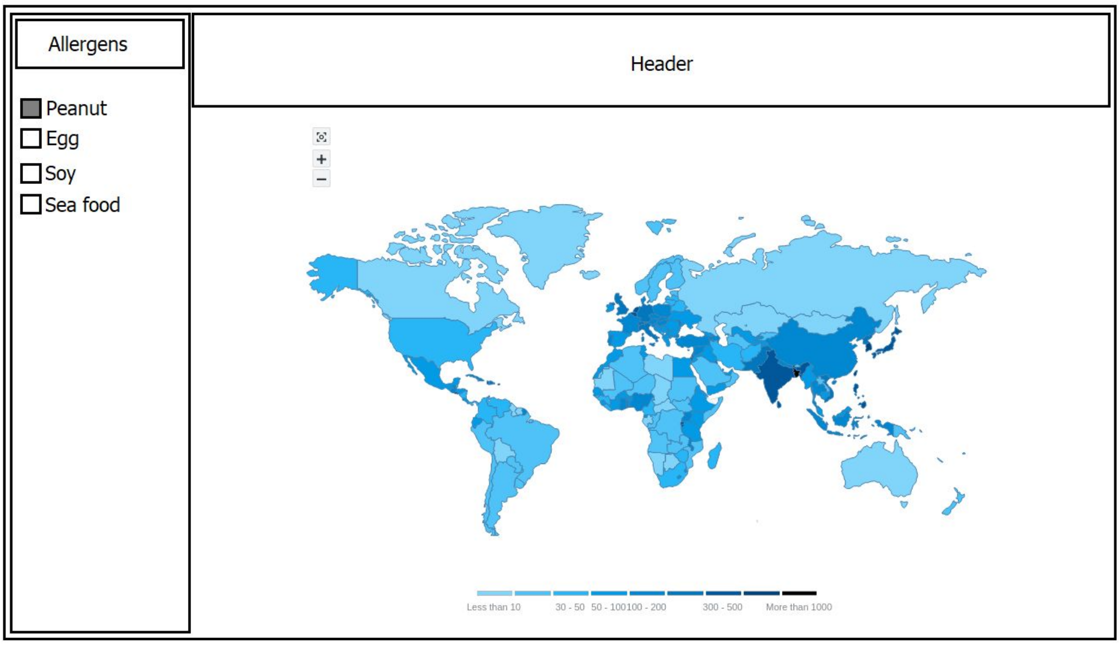
The implementation for our solution should be scalable. It should be easy for the maintainer to add support for new allergens when the data for such is provided.

Use Cases

There are several features for the tool that we want to implement. We frame these in terms of use cases, or how a user would use our tool.

Use case 1: "I'm dangerously allergic to X, Y, and Z, and I really want to travel"

The first use case is designed for people who want a **quick**, **visual** idea of the safety of countries in terms of their allergens. Each country will receive a safety score given a user's *choice* (the tool is interactive) of allergens. We already developed a way to encode this safety score, but we omit it from this report. Though this has not been finalized, we will likely visualize this as a choropleth map, such as the one in the mockup below. We will also explore other geographical visualizations.



UI Mockup

Use case 2: Explore - Enjoy

Following Munzner [15], we similarly anticipate an end user that is driven by curiosity and finds this web tool by accident or simply as someone who has no allergies, food restrictions, or 'search' related queries. To this end, anticipating this end user driven by an 'explore' or 'enjoy' imperative, we intend the tool to be both accessible and enjoyable simply beyond its intended task. To accommodate for this user, while staying true to our original task and in mind of data we have available, we have increased our scope of consumption data beyond the eight main allergens and have included time series data, when available, for the consumption of users who may be curious about historical and comparative trends in food consumption habits. This way, any user who stumbles on the web tool may interact with the data and gain insight into historical and geographic trends in overall food consumption habits.

Use case 3: Discover/Derive

Lastly, we also anticipate our tool to be useful for further research. To this end, we keep in mind the 'discover' imperative of an end user who may be interested in the data and the data visualization for purposes beyond the search for travel destinations or personal exploration for leisure purposes. In mind that such tool does not currently exist, and aware that food consumption data is available in isolation and in its raw form but not integrated into one accessible tool, we also intend for this design to be useful for a user to derive from it macro trends in food consumption habits globally, historically, and with available data for geographic exploration into these trends.

For instance, while the tool is designed to make informed travel decisions in mind of a person's allergies, it is also a useful tool for someone interested in the dietary composition of food consumption habits by country. Questions that we anticipate the tool may help answer are: "Are there important historical trends in food consumption habits that may inform a sustainable global food policy agenda?" and "what transnational economic arrangements are rendered visible through the geographic exploration of national food consumption habits?"

Implementation

Our goal is to implement the tool as a highly interactive Jupyter Notebook, in which the user does not need to alter any code. Implementation is already under progress. We will use Python to perform our data wrangling and data processing on our local machines.

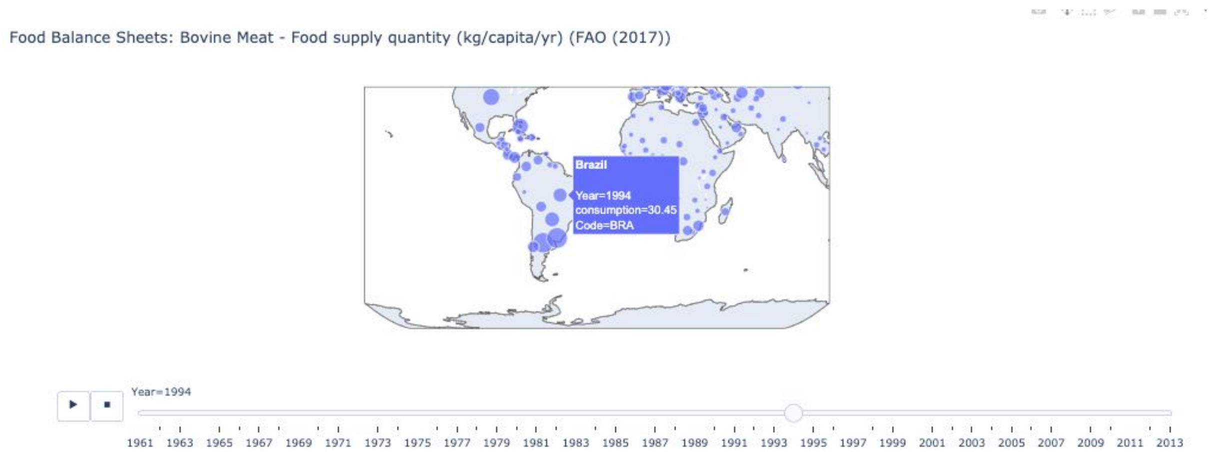
We expect any external user to be able to access our platform through the web.

Exploration

The following are some screenshots of our work in progress.

Choropleth map for beef consumption.

Time-series comparison of beef consumption in Canada (blue) with the United States (red). Note that a user can change the allergen and the countries chosen.



Map visualization of beef consumption in South America. The user can change the year chosen. Hovering their mouse on top of a country shows more information. The user can navigate through the map and explore other regions as well.

Milestones

All the times are **combined** times. People assigned are not expected to work equally.

- Gather and clean data
    - Assigned: Lucca, Jose
    - Time: 30 hours
    - Date: Nov 17th
- Related work

    - Assigned: Jose
    - Time: 30 hours
    - Dates:
        - Annotated bibliography by Nov 17th
        - Literature review by Nov 26th
- Explore use cases
    - Assigned: Lucca, Jose, Alireza
    - Time: 15 hours
    - Date: Nov 17th
- Study which idioms we *can* encode
    - Assigned: Lucca, Alireza
    - Time: 20 hours
    - Date: Nov 21st
- Tool design
    - Assigned: Lucca, Jose, Alireza
    - Time: 30 hours
    - Date: Dec 1st
- Implement of the visualization tool (MVP)
    - Assigned: Lucca, Alireza
    - Time: 30 hours
    - Date: December 5th
- Deploy of the interactive tool
    - Assigned: Lucca, Alireza
    - Time: 20 hours
    - Date: December 7th
- Prepare demo and presentation
    - Assigned: Jose, Alireza
    - Time: 10 hours
    - Date: Dec 10th
- Write report (including explanation of all idioms we did not include)
    - Assigned: Lucca, Jose, Alireza
    - Time: 35 hours
    - Date: Dec 14th

Working overhead: 20 hours.

References

[1] Food and Allergy Research & Education. Facts and Statistics. [Internet] [Cited 2020 Oct 23]. Available from: https://www.foodallergy.org/resources/facts-and-statistics

[2] Loh, W. and Tang, M. L. K. The Epidemiology of Food Allergy in the Global Context. International Journal of Environmental Research and Public Health 15 (9). 2018. [cited Cited 2020 Oct 23]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6163515/

[3] World Allergy Organization. White Book on Allergy. [Internet] 2011. [cited Cited 2020 Oct 23]. Available from: https://www.worldallergy.org/UserFiles/file/WAO-White-Book-on-Allergy_web.pdf

[4] Banister, D. The Tip of the Iceberg: Leisure and Air Travel. Built Environment 26 (3); 226-235. 2000. [cited Cited 2020 Oct 23]. Available from: http://www.jstor.org/stable/23288402

[5] Hawkes, C. Globalization and the nutrition transition: A case study. Cornell University: Ithaca, NY; 2007. [cited Cited 2020 Oct 23]. Available from: https://ebrary.ifpri.org/digital/collection/p15738coll5/id/1229

[6] Traveling With Asthma and Allergies | AAFA.org [Internet]. Aafa.org. 2020 [cited 23 October 2020]. Available from: https://www.aafa.org/traveling-with-asthma-allergies/?gclid=CjwKCAjw_sn8BRBrEiwAnUGJDjoc3_kMtFPt5TWF4K16DETMEnOK1IUL5fZmtFw_ePta-gBYxNtjMBoCL8EQAvD_BwE

[7] Living with allergies [Internet]. Food Allergy Canada. 2020 [cited 23 October 2020]. Available from: https://foodallergycanada.ca/living-with-allergies/ongoing-allergy-management/managing-in-different-environments/travel/

[8] Traveling Abroad With Food Allergies is Totally Doable | Tortuga [Internet]. Tortuga. 2020 [cited 23 October 2020]. Available from: https://blog.tortugabackpacks.com/traveling-with-food-allergies/

[9] McKenzie S. Scott's Rules for Traveling with Food Allergies - Allergic Living [Internet]. Allergic Living. 2020 [cited 23 October 2020]. Available from: https://www.allergicliving.com/2010/06/30/allergies-travel-rules-for-food

[10] Foley E, Cooke B, Ryckman T. Costa Rica. Tarrytown, NY: Marshall Cavendish Benchmark.;.

[11] Johnson, M. O., Cozart, T., & Isokpehi, R. D. (2019;2020;). Harnessing knowledge for improving access to fruits and vegetables at farmers markets: Interactive data visualization to inform food security programs and policy. Health Promotion Practice, 21(3), 152483991987717-400. doi:10.1177/1524839919877172

[12] Lunterova, A., Spetko, O., & Palamas, G. (2019). Explorative visualization of food data to raise awareness of nutritional value. (pp. 180-191). Cham: Springer International Publishing. doi:10.1007/978-3-030-30033-3_14

[13] Maličká, L. (2020). Consumption of food in the EU by the degree of urbanization: Data visualization and cluster analysis of the EU sample. Potravinarstvo, 14, 343-350. doi:10.5219/1282

[14] Zhihua, L., Gong, Y., Holmes, M., Pan, X., Xu, Y., Zou, X., & Fernandes, A. R. (2019). Geospatial visualisation of food contaminant distributions: Polychlorinated naphthalenes (PCNs), potentially toxic elements (PTEs) and aflatoxins. Chemosphere (Oxford), 230, 559-566. doi:10.1016/j.chemosphere.2019.05.080

[15] Munzner, T., Taylor & Francis eBooks A-Z, & O'Reilly for Higher Education. (2015;2014;). Visualization analysis and design (1st ed.). Boca Raton: CRC Press, Taylor & Francis Group, CRC Press is an imprint of the Taylor & Francis Group, an informa business. doi:10.1201/b17511

[16] Karimi, H. A., Karimi, B., & Taylor & Francis eBooks A-Z. (2018). *Geospatial data: Science, techniques and applications*. Boca Raton: CRC Press,Taylor & Francis Group, CRC is an imprint of Taylor & Francis Group, an informa business.

[17] Speckmann, B., & Verbeek, K. (2010). Necklace maps. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 881-889. doi:10.1109/TVCG.2010.180