Nutrition Bytes: Visualizing Food Content

Shuai He* Daniel Kerrigan† Ronald Metoyer‡

University of Notre Dame

Abstract

Food choice is a more confusing task than ever before given the amount of information that is accessible to the everyday consumer. One component of information is the food label which includes nutrient and ingredient information. While a limited subset of such information is provided on the physical labels of food products, there is much more information available to consumers. In this project, we explore the design of an interactive visual representation of nutrient and ingredient information to support food decision making. Our design integrates (1) representation of nutrients in the form of Daily Intake Percentages, (2) representation of ingredients and sentiment surrounding those ingredients, and (3) small multiples to support comparison.

Index Terms: H.5.2 [Information Interface and Presentation (e.g., HCI)]: User Interfaces—Interaction styles (e.g., commands, menus, forms, direct manipulation)

1 Introduction

Obesity and other diet-related issues have raised concerns in society and stimulated burgeoning research to support healthy eating behaviors. According to a recent survey by the International Food Information Council Foundation [2], consumers are confused given the overwhelming amount of information around food choices. In particular, consumers tend to struggle with interpreting ingredient information and comparing multiple foods. It is known that consumers tend to consider the ingredients and the length of the ingredient list as an indication of healthfulness [2]. To further complicate things, the ingredients list often contains elements that the consumer is simply not familiar with. Additionally, consumers are often interested in comparing multiple foods, but this is often difficult to do given numerical values only.

Despite the open exploration opportunities, little research has been dedicated to the intersection of nutrition and visualization. However, the nutrition research community has placed a growing emphasis on a technology supported approach to address the nutrition communication challenge [1,3]. While some initial research [5] has recognized the need, current visualization designs are limited to showing nutrients and other important information such as ingredients are not considered.

In this paper, we explore the design space for representing foods with a focus on ingredient and nutrient information. We further augment this information with sentiments surrounding ingredients in order to indicate general beliefs regarding that ingredient.

2 Data

The data is collected from the Food Composition Database provided by the United States Department of Agriculture: Agricultural Research Service [1]. The USDA database is the original data source for labels on consumer food items and it contains over 183,995 foods. Each food has a set of nutrients that are divided into the following groups: proximates, minerals, vitamins, lipids, and other. Examples of Proximates are carbohydrates, fiber, and sugar while examples of minerals are calcium and iron. The dataset also contains a list of ingredients for each food item. In total, the data consists of ingredient names (nominal), nutrient names (nominal), nutrient group (nominal), the relative amount of an ingredient (ordinal), and nutrient percentage daily intake values (quantitative).

3 Visualization Design

We employed Munzner's nested model to guide our design decision making [6] throughout our user-centered process. In the following sections we discuss the domain characterization, data and task abstraction, and encoding decisions.

3.1 Domain Characterization and Abstraction

Nearly all consumers (96%) in a recent survey seek out health benefits from the foods they eat and drink [2] and it suggests that users display increasing interests in learning more about the food they are eating. This survey as well as other resources have indicated that consumer food choices are largely affected by other people's opinions, especially those of their families and friends [8]. This research also indicates that consumers are eager to track their shopping behavior to monitor their nutritional intake. Additionally, when buying foods at a grocery store, many consumers compare food items by nutrients. [7]

Based on the existing literature about consumer food behaviors as well as interviews with consumers, we determined the following questions and associated abstract tasks to be important for food choice decisions:

- What are the ingredients and nutrients (and associated daily intake percentages) in specific food items? This requires that consumers be able to *search* nutrient and ingredient information as well as find values for intake percentages.

- How do the nutrients and ingredients of one food compare to another? Consumers should be able to compare two food items across ingredients and nutrient and look for outliers (e.g. a high percentage intake value or a large number of ingredients.

- What is the general consensus about this food? Consumers want additional information about a food - what are others saying about it and/or it's ingredients? Consumers should be able to, for example, look up particular ingredients and *derive* the general opinions of that ingredient.

  To support this last task, we derive sentiment data for each food ingredient. Sentiment analysis has been used widely in the text mining field [4] to gauge options about particular topics. We utilize the Microsoft Bing API to search for and collect the five most popular results for a given ingredient and we run a sentiment analysis algorithm on the text of these top results to calculate the positive sentiment percentage for the ingredient (quantitative).

*e-mail: she1@nd.edu
†e-mail: dkerriga@nd.edu
‡e-mail: rmetoyer@nd.edu

[1] https://ndb.nal.usda.gov/ndb/

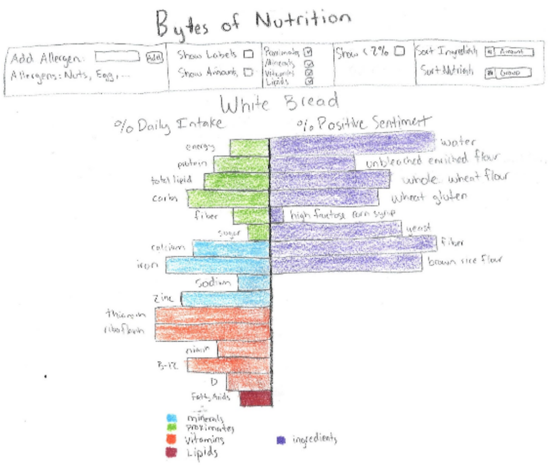

Figure 1: Design Sketch

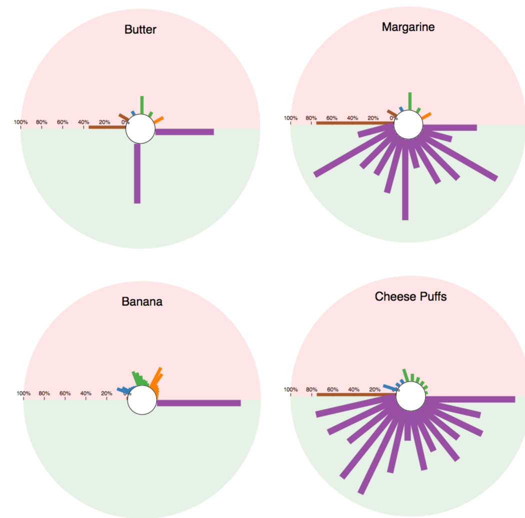

Figure 2: Iteration 2

3.2 Visual Encoding

3.2.1 Design Iteration 1

Our initial design encoded the two quantitative values (nutrient dailty intake % and positive sentiment %) with a bar length encoding. Color encodings distinguish nutrient groups as well as ingredients from nutrient. We also visually separated nutrients from ingredients using two sides of a vertical axis (See Figure 1 The bar length encoding on a common scale on the left side represents the % daily intake for a particular nutrient. On the right side, it represents the % positive sentiment for an ingredient. The intent was to provide an *overview* of the total number of nutrients and ingredients as well as the % daily intake and % positive sentiment for each respectively. Text labels are used to communicate nutrient and ingredient name.

Users are provided several interaction mechanisms. They can turn labels on/off, filter out nutrient groups, order nutrients by amount, and order ingredients by either amount (the default) or by sentiment. Additionally, users can flag specific allergens to be highlighted if they exist in the food.

Finally, the representation employs small multiples in a grid layout to allow consumers to compare multiple foods. Selecting a food item in the small multiple graphs will switch the main chart to the selected food item, providing details on demand. [9].

Evaluation: An informal paper prototype study indicated that users were able to answer the questions above using the design. One problem, however, is that the shared vertical axis between nutrients and ingredients led to confusion. Users were inclined to think that there was an apparent comparison between the left (nutrients) and right (ingredients) sides. This design, unfortunately, violated the expressiveness principle.

3.2.2 Design Iteration 2

To resolve the expressiveness violation, we removed the vertical axis and replaced it with a radial layout, keeping all other major design decisions from Iteration 1. This design avoids the shared axis that encouraged the unwanted comparison across ingredients and nutrients. To visually separate ingredients from nutrients, all nutrients are rendered in the top half of the chart while ingredients are shown in the bottom half. The updated design is shown in Figure 2 as implemented with the D3.js javascript library [2]. This design also includes small multiples to allow for comparison of multiple foods.

[2] https://d3js.org

4 Discussion and Future Work

Pilot users of the visual representation found several interesting insights.

One assumption of consumers is that the number of ingredients and healthfulness of foods are negatively correlated (e.g. highly processed foods) whereas the number of vitamins and minerals are positively correlated to healthfulness of foods. According to this theory, the banana is the more healthy food (large amount of blue (minerals) and orange (vitamins) and a single purple bar(ingredient)) when compared to cheese puffs.

Users were also able to compare two foods that they expected to be quite similar. For example margarine is considered to be a butter replacement. In Figure 2 we see that they both contain almost exactly the same amount of sodium, fat, energy, and Vitamin A. Margarine successfully imitates butter in this regard. The main difference in terms of nutrients is trans fatty acids vs. saturated fatty acids. Margarine also contains more ingredients with some of them, such as vegetable monogclycerides, having a very low positive sentiment value. Overall, users were interested in ingredient sentiments, especially for ingredients that were unfamiliar to them.

While this initial design is a promising first step to representing nutrient and ingredient information, it is far from complete. Future work includes a more thorough treatment of sentiment analysis to disambiguate opinion of taste from healthfulness, integration of sentiment for specific nutrients, and a thorough user study of the representation.

References

[1] J. Brug, A. Oenema, and M. Campbell. Past, present, and future of computer-tailored nutrition education. *The American journal of clinical nutrition*, 77(4):1028S–1034S, 2003.

[2] FoodInsight. 2017 food and health survey: ä healthy perspective: Understanding american food values, may 2017.

[3] K. D. Lewis and B. M. Burton-Freeman. The role of innovation and technology in meeting individual nutritional needs. *The Journal of nutrition*, 140(2):426S–436S, 2010.

[4] B. Liu and L. Zhang. A survey of opinion mining and sentiment analysis. In *Mining text data*, pp. 415–463. Springer, 2012.

[5] C. Mah, R. Pusch, S. Huron, and S. Carpendale. A visualization fingerprint: Comparing nutrient data visually.

[6] T. Munzner. A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics*, 15(6), 2009.

[7] R. M. Nayga, D. Lipinski, and N. Savur. Consumers' use of nutritional labels while food shopping and at home. *Journal of Consumer Affairs*, 32(1):106–120, 1998.

[8] H. Schäfer. Personalized support for healthy nutrition decisions. In *Proceedings of the 10th ACM Conference on Recommender Systems*, RecSys '16, pp. 455–458. ACM, New York, NY, USA, 2016. doi: 10.1145/2959100.2959105

[9] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pp. 336–343. IEEE, 1996.