ASSIGNMENT 4 - REPORT

TEAM 7: Neeraj Kulkarni, Sahana Bhat, Vighnesh Misal

TABLE OF CONTENTS

Abstract

The objective of our project is to explore the design and evaluation issues involved for creating an interactive visualization. We aim at addressing a concrete problem derived from a unique data set, by building a novel creative visualization. The visualization is interactive and allows the user to explore the dataset to answer the questions and make decisions. Our Area of Interest was in the field of Candy industry. Confectionery is a big industry with high household penetration. Confectionery has an enormous economic impact, either directly in the form of economic output & wages, or indirectly as it supports various other industries to create jobs in its manufacturing, supply-chain, wholesale and retail. Apart from the evident economic impact, confectioneries create a strong emotional bond with the unique taste and alleviating properties. Understanding the candy consumption pattern can help executives make decisions on candy production operations. Through our interactive visualizations we look forward to answering questions about consumption of candies.

Introduction to topic of Interest

Candies are unique, fun, affordable and special as they create a strong emotional bond. Studies have proven that 88.9% of the time candies are consumed as treats. Most of the time candies are consumed to pacify our sweet cravings. The scientific reason for our craving for sweets is based on an evolutionary basis, back when our primitive ancestors were scavengers. Sugary foods are excellent sources of energy, so we have evolved to find sweet foods particularly pleasurable. Foods with unpleasant, bitter and sour tastes can be unripe, poisonous or rotting. So to maximize our survival as a species, we have an innate brain system that makes us like sweet foods since they're a great source of energy to fuel our bodies.When we eat sweet foods the brain's reward system, called the mesolimbic dopamine system, gets activated. Dopamine is a brain chemical released by neurons and can signal that an event was positive. When the reward system fires, it reinforces behaviours making it more likely for us to carry out these actions again.

However, eating too much sugar can have many negative health effects. An excess of sweetened foods and beverages can lead to weight gain, blood sugar problems and an increased risk of heart disease, among other dangerous conditions. Hence, the new generation of people have different opinions on consumption of candies. The patterns in candy consumption has changed due to dietary restrictions, allergies based on the content and even the brand name itself.

To understand the consumption of candies we have researched on various candies based on their contents and their popularity in each state in the U.S. We have also considered the popular preferences of people based on the number of votes received. Apart from comparing the candies themselves, we have also analyzed the consumption patterns over years and their trends to identify the stunning seasonality of consumption which peaks during the Halloween, Christmas and New Year celebrations.

Problem Statement

One of the major concerns we look forward to answer through our visualization is related to Candy Production. By understanding the patterns and trends in Candy consumption, we can support candy manufacturing executives make decisions for expanding their operations. This helps candy manufacturers observe states that their products are popular and ramp up production. By identifying the most popular candies in a particular state, the wholesale and retail seller can decide on their orders. Understanding the contents/ingredients involved in the most popular candies, the manufacturing executives can design new products which might have higher success rates.

Our visualization also aids candy consumers to choose the best candies when organizing an event. Knowing the preferred ingredients of the audience can help the organizers eliminate risks of wastage. Identifying ingredients that they are known to be allergic in certain candies can help find alternatives. Understanding the sugar content of each helps to provide replacements which are organic or sugar-free for the consumers who are having dietary concerns, within the budget of the organizers.

Questions to answer through the visualization:

- How is the candy production over the years?
- Is there any seasonality associated with candy production in a particular year? (Analyzing candy production trend over the months)
- How is the demand for candies distributed across the United States?
- What are the top three candies in every state?
- Which candies are more likely to be chosen by consumers?
- What are the most popular candies which have chocolate and caramel?
- Which are the candies with the highest Sugar % ?
- What are the chances of Reese getting sold out over M & Ms?
- What should be the content of a new candy for greater success? Should it contain Chocolate, Nougat, Peanuts or Fruity?
- What is the range of Win % for a candy with Crisped Rice Wafers?
- Is a candy with Caramel expected to have higher Sugar % than a candy with Peanuts?
- What are the candies that need to be avoided by consumers with Peanut allergy?

Dataset

Here we are using 3 different sources of data.

● The Ultimate Halloween Candy Power Ranking: This dataset was obtained from a Kaggle challenge. The data was collected by creating a website where participants were shown presenting two fun-sized candies and asked to click on the one they would prefer to receive. In total, more than 269 thousand votes were collected from 8,371 different IP addresses. Apart from the popularity of the candy, we also have information about the attributes associated with each candy such as does it contain chocolate, caramel. Nougat, peanuts, etc. We have the sugar %, price % and winning % for each candy within the dataset.

Source:
https://www.kaggle.com/fivethirtyeight/the-ultimate-halloween-candy-power-ranking

● State-by-State Favorite Halloween Candy: This dataset provides the top most popular candies in each state of the U.S. It also gives us the 2nd and 3rd most popular candies in the states. This provides the candy manufacturers and retailers an overview of candy consumption.

Source:
https://public.opendatasoft.com/explore/dataset/state-by-state-favorite-halloween-candy/export/

● US Candy Production by Month: This dataset consists of candy production along the timeline from 1972 to 2017. The dataset is retrieved from the Federal Reserve Bank of St. Louis. This represents the data approved by the Board of Governors of the Federal Reserve System (US), Industrial Production: Non Durable Goods, under the sub category of Sugar and Confectionary Product [IPG3113N].

Source: https://www.kaggle.com/rtatman/us-candy-production-by-month

Visualizations

The visualizations were built using Tableau. The final dashboard has been published in Tableau Public. Link to Tableau Dashboard:
https://public.tableau.com/views/Assignment4_Group7/Story1?:display_count=y&:origin=viz_share_link

A. **Top Candy Ranking by State:** The top candies for each state is represented with their iconos for each state. The candy icons act as visual cues so that the tooltip for candy names need not be read. Hovering on the icon displays the candy names in plain text format.

Filters are provided for the users to select State and Ranking (1st , 2nd , 3rd Rank) of interest. The icons over the states dynamically change based on the selections.

Insights:
- America's Top 10 candies are Skittles, Reese's cups, M&M, Snickers, Candy Corn, Starburst, Hot Tamales, Tootsie pop, Sour Patch Kids and Hershey's
- The Top Ranked candy in Washington state is the Salt Water Taffy

B. Candy Production Timeline: This visualization provides a graphical representation of candy production over the years from 1972 to 2017. The candy production value is measured using the Industrial Production (IP) Index which measures the real output of all relevant establishments located in the United States, regardless of the ownership, but not those located in U.S Territories. The line chart is represented based on candy production index over the months within the same year. The visualization has an interactive element which allows to select the year of interest and a Play button that shows the graphs like an animation.

Insights:
- The graph shows the stunning seasonality in the Industrial Production Index, during October, November, December and January, which explains the increased consumption during Halloween, Christmas and New Year celebrations.
- As the demand is higher during Autumn and Winter months, it's advisable to start production early in the Summer.
- Apart from production, if new products need to be launched, the months during the highest demand would have the best opportunity for success
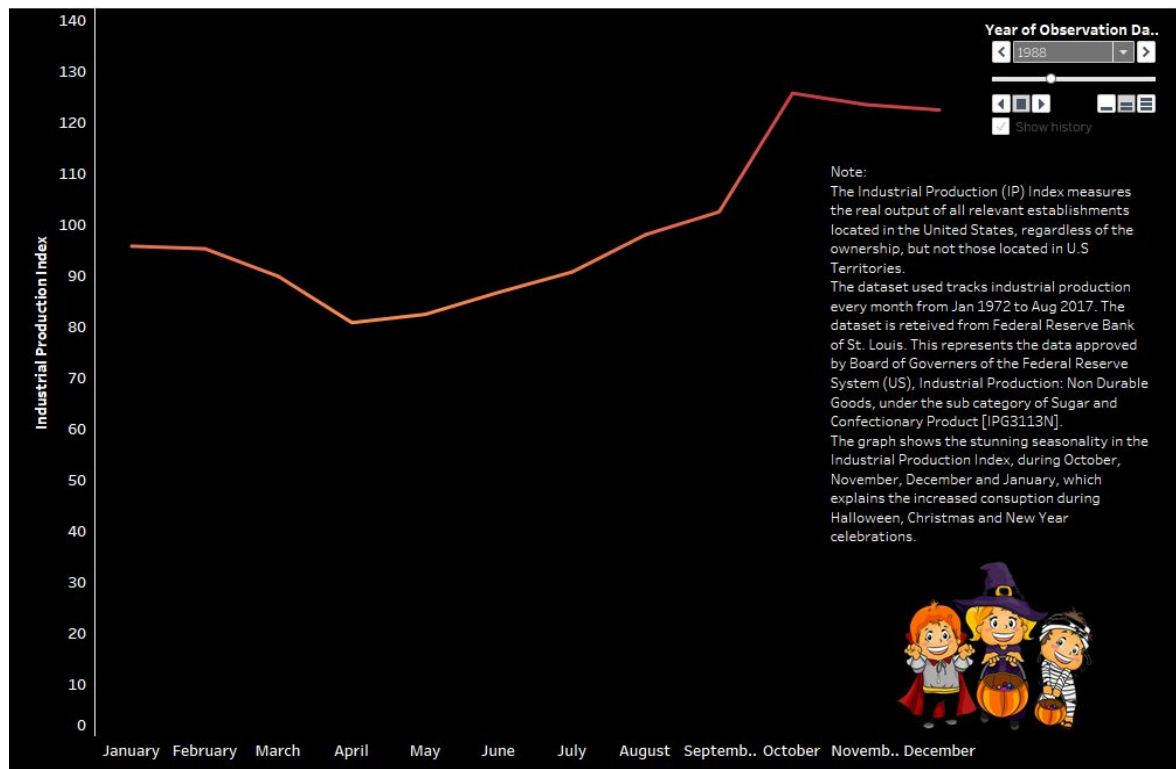
C. Candy Wars: This interactive visualization is a 'fun' take on displaying the demand for various candies. In a study, participants were asked to pick a candy from two options presented to them. The participants then chose a particular candy and their response was recorded. Similarly, every candy was pitted against the other, and the number of times each candy won was taken as a percentage of the total number of '1 v 1' candy matchups, and this variable was established as 'Win Percentage' of that candy.

Insights:

- The 'Win percentage' helps us understand which candies are most likely to be picked by consumers, and thus an assessment can be made for that candy's demand.
- For example: In the screenshot below, Haribo Gold Bears has higher win % than Swedish Fish.

D. Candy Content Filter: The Candy Content filter mainly has two use-cases:

Sugar % vs Win %:
The Sugar % refers to the percentile of sugar it falls within the data set and Win % is the overall win percentage according to 269,000 matchups. The graph shows a scatter plot representing the points by candy icons. Thus, with a quick glance at the graph one can identify the candies with the highest Win %. Also, the Sugar % can be used to filter out candies that are high in sugar content for people with diabetes or sugar-conscious people in general.

Insights:
- Reese Stuffed with pieces is the candy with the highest amount of Sugar % and Reese miniatures has the highest Win %.
- When filtered over candies that are Bar shaped and containing Caramel, Twix has the highest Win % and Milkyway Simply Caramel has the highest Sugar %

Allergen Filter:
The visualization also provides an interactive filter to select the attributes associated with the candy such as its contents: Chocolate, Nougat, Crisped Rice Wafer, Peanut-y/Almondy, Fruity, Caramel. This information can be used by people allergic to a particular content by selecting the relevant image and filtering out those candies which have that content. The filter also provides some options for filtering between various candy forms: Bar, Hard or Pluribus.
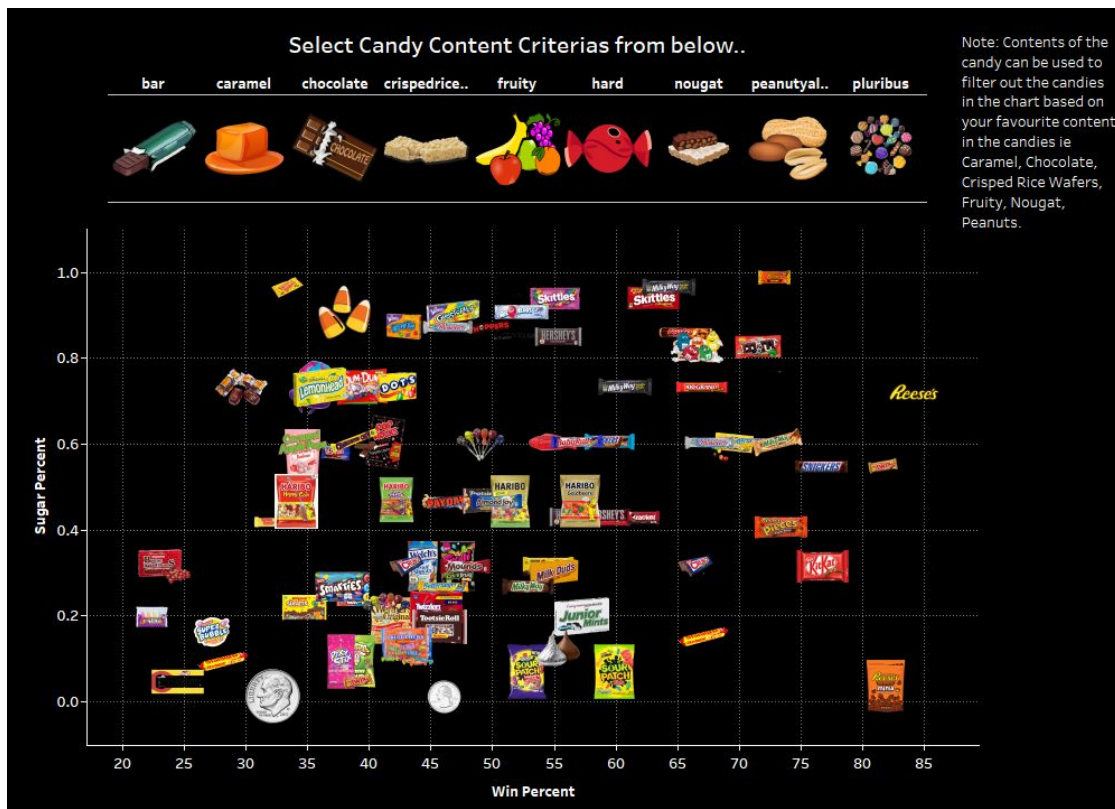
Challenges

The challenges in implementing the visualization were as follows:
- Restructuring the data in the required format to create visualizations
- Associating the Candies to their respective icons where real images were used was time consuming
- Synchronizing connected visualizations to change the information based on user input
- Using color and visual cues to output as much information without making the user read text

Usability Testing

Usability testing is a method of testing the functionality of a website, app, or other digital product by observing real users as they attempt to complete tasks on it. Here, our goal was to evaluate the usability with respect to our Tableau dashboard with visualizations. Sometimes the visualization might be stuck with a design process that the team members understand, but the target audience do not. Hence, the usability testing helps reveal areas of confusion and uncover opportunities to improve the overall user experience.

The usability testing helps to:
- Confirm if the visualization is able to convey the accurate information with ease
- Identify issues in user action workflows
- Recognize minor errors that might be neglected

Usability Test Design

Our Test plan was to make sure the interactive features employed in the visualizations were well utilized to derive the information required. The users were guided with tasks to perform. We made sure to stay very close to the user and watch their actions as we did not use any devices to record them. The questions asked were open ended, i.e the users were guided by the tasks to meet the goals, but were free to make their own moves to complete the tasks. We also made sure to know the background of the user, their familiarity with UX and the candy data sets itself.

All the users were asked 4-6 questions and their feedback regarding improvements were noted down. After which a google form was provided to rate the visualization based on usability, interactivity and design intuition. The usability test tasks that were given to all the users are mentioned below:

1. In the visualization representing Top ranked candies by states in the U.S Map, identify the Top ranked candies (i.e 1st, 2nd and 3rd Rank) in the state of Washington.
    Choose from the drop down control provided at the right most corner to update the visualization to 2nd and 3rd Ranked Candies. To zoom in on the required state, you may scroll or swipe over the map.
2. Which are the candies that occur frequently in the Top Rankings across multiple states?
3. In the 2nd visualization on Candy Production Timeline, do you identify any trend in the production during a certain period or season of the year? Please use the controls provided, the Play button to start an animation and the Stop button to pause.
4. For the year of 2015, what is the minimum and maximum number of industrial production index measured?
5. In the Candy Wars Dashboard, considering that the candy with higher Win% is the champion, which candy would win when the players are Haribo Gold Bears and Swedish Fish? What are their Win %?
6. In the Candy content filter chart, filter the candies in the chart based on Caramel and identify the Candy with highest Sugar % and Win %. Finally, search for your favourite candy to know its contents and Win% over other candies.

Test Script

"Hi, _____ . Our names are Neeraj, Sahana, and Vighnesh. If you could spare a minute of your time to talk to us about candies, it would be awesome!

We're working on a UX Project based on candy-data. We are going to be asking you some questions, don't worry, it's not an exam, they're fun questions based on candies! There are no wrong answers, we just want your opinion to understand some statistics.

We'd like you to look at the visualizations we have made on some candy-data. We would appreciate it if you take a look at them, and play around with them since they're interactive.
While you're doing that, we would like to record your responses/behavior that express your feelings about the visualization, with your permission of course!

Once you're done interacting, we would like to ask you some questions, regarding how easy it is to navigate around and interact with visualizations, are there any problems in particular you're facing during the interaction, and also how useful the visualization is from an insights perspective.

Treat this as a fun survey, and we assure you that your responses will be completely anonymous!"

Execution

The study intended to collect information about two use-cases: Professional & Personal
However, a usability testing sample should be quite diverse for it to give an effective representation of generality. So, we got a good, diverse mix of survey respondents in the following manner:
We went to Fosters School of Business to survey people with business backgrounds and an entrepreneurial mindset to gather data for the professional front. Next stop was candy stores and grocery stores, and interacted with people who were looking to purchase candy (were in the candy aisles) to get the 'targeted customers' of the industry. We also talked to other people in the stores who weren't necessarily looking to buy candy, some people on the streets, and other students in the University.
We were able to survey people of different ages, from young kids to the older generations.
While surveying the kids, we made sure they were accompanied by a parent(s), and explained to their parent(s) about the survey, before finally asking the kids relevant questions with their parents' permission of course.

Observations and Findings of usability test

- We had 4 users from UX background and 19 from Non-UX background.
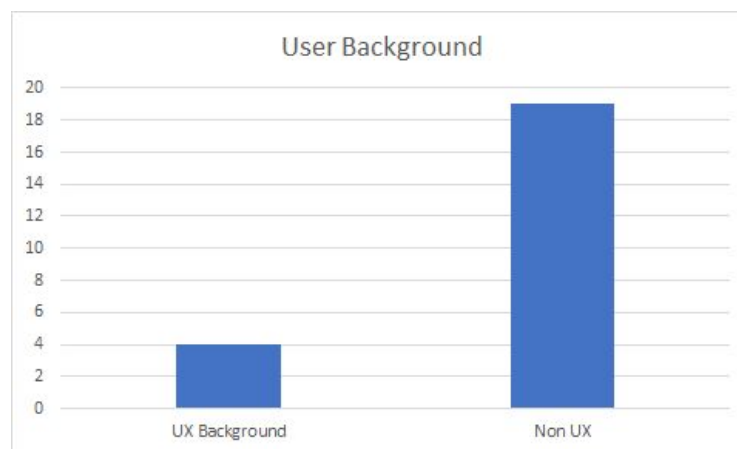

- **Initial user thoughts on usability:** A large percentage of the users that we interviewed were not proficient in UX Research/Design. As such when we asked them the question regarding what they thought of a visualization with candy data, their responses were quite tepid. They didn't have any clue regarding the use case for such a visualization and shrugged it off as redundant. When we explained our use cases to the people, they were a lot more responsive to taking part in the usability testing. We then tried explaining the two major cases that we were aiming for with the visualization:
1. Helping executives in a candy manufacturing organization make executive decisions related to expanding their businesses

2. Helping people who are allergic to certain ingredients in the candy, avoid such candy without paying a premium price or an unnecessary visit to a hospital.
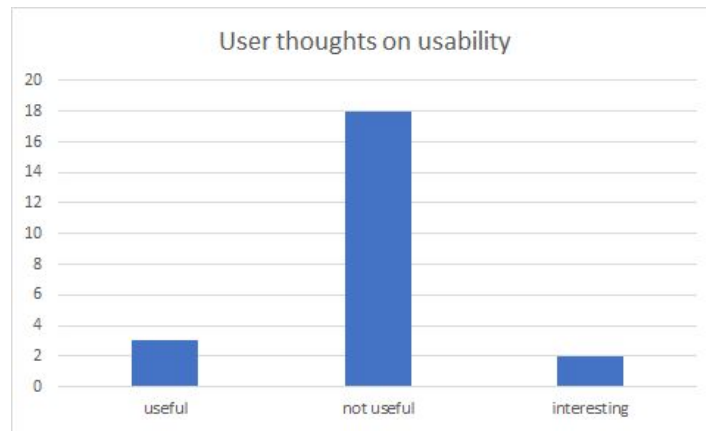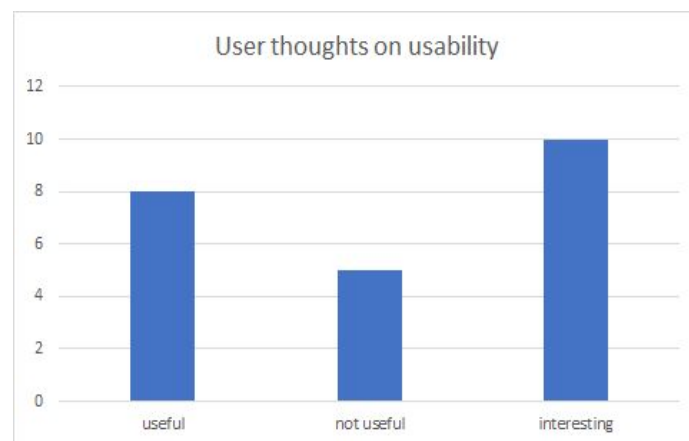

1. **User Ratings for interactivity in visualization:**

Great	3
Good	7
Ok	9
Bad	4
Horrendous	0
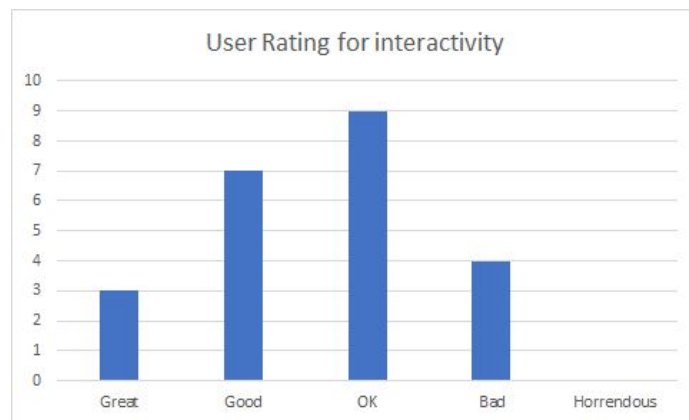
2. **Users thoughts on usability:**

Useful	8
Not Useful	5
Interesting	10

3. User thoughts on Design Intuition

Easy	11
Medium	10
Difficult	2
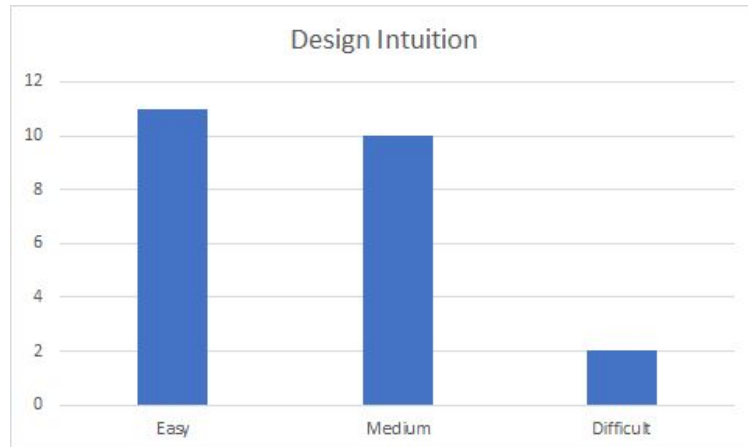
Summary of Results

We then allowed the users to interact with our visualizations, fully aware of the context and they said that they liked the seamless interactivity between the visualizations. We also wanted to know what they thought of the general understanding concerning the interactive elements without us giving them cues, and the users thought that it was a well designed interface. However, they recommended that we use more color (our was too black) and adjustable fonts for people with impaired vision. They also felt that it would be better if we could highlight areas of interest with a brighter, more prominent color. It would greatly improve the visualization.

Recommendations for Improvement

Apart from performing the usability test itself, we also made sure to receive recommendations for improvement. Here are the following suggestions provided by users:
- Improve category filters for candy content by adding more options
- Make a version for people who are visually impaired
- Use of color and pop out elements to highlight selected areas
- Improve on tooltip to provide more information

References

- https://www.usda.gov/oce/forum/2019/speeches/Larry_Wilson.pdf
- Top Candies per state:
https://public.opendatasoft.com/explore/dataset/state-by-state-favorite-halloween-candy/export/
- Halloween candy ranking:
https://www.kaggle.com/fivethirtyeight/the-ultimate-halloween-candy-power-ranking
- Candy Production vs Year:
https://www.kaggle.com/rtatman/us-candy-production-by-month