

Project Team #1: Urban Beekeeping Laboratory, DS 4200 S20

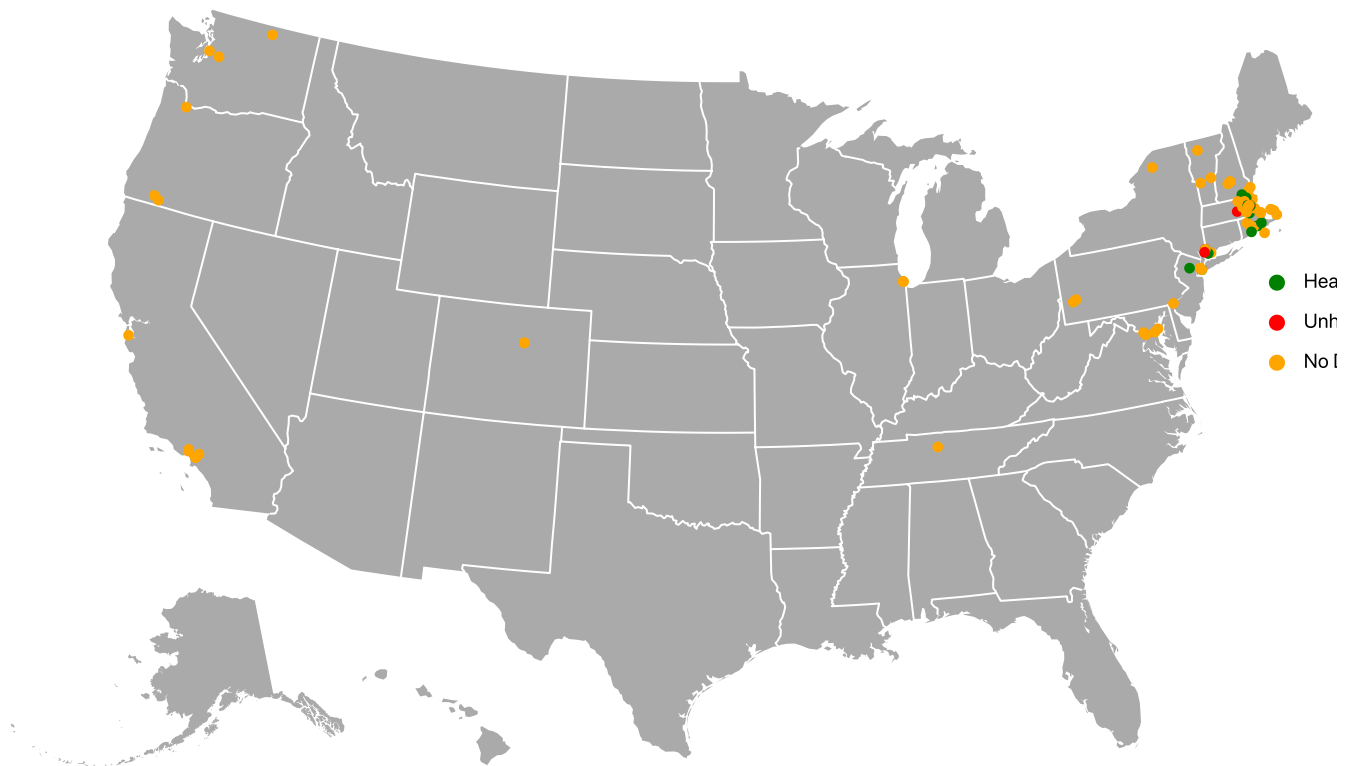
Tom Cairns, Sid Radhakrishnan, Ryan Liang, Lilly Olson

Service-Learning Course Project as part of [DS 4200: Information Presentation and Visualization](#), taught by [Prof. Dylan Cashman](#), [Data Visualization @ Khoury, Northeastern University](#).

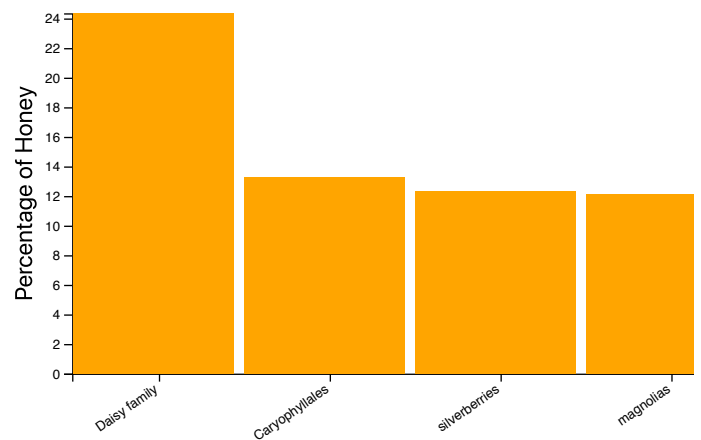
Motivation

Our group worked with the Best Bees Company in Roxbury. Their mission as a company is to improve bee health. To do this they set up and care for hives around the country. They invented a technique to analyze the plants that make up the honey produced by the hive called HoneyDNA. They are interested in seeing if there are certain species of plants are correlated with hive health. To do this they wanted a visualization that made it easy to compare the health of the hives with the plants found in the honey.

Visualization



Hive ID	City	Health
S011821	Acton	N/A
S011893	Altadena	N/A
S005728	Arlington	Good
S005757	Arlington	Good
S005764	Arlington	N/A
S011813	Bedford	Troubled
S005715	Bedford	Troubled



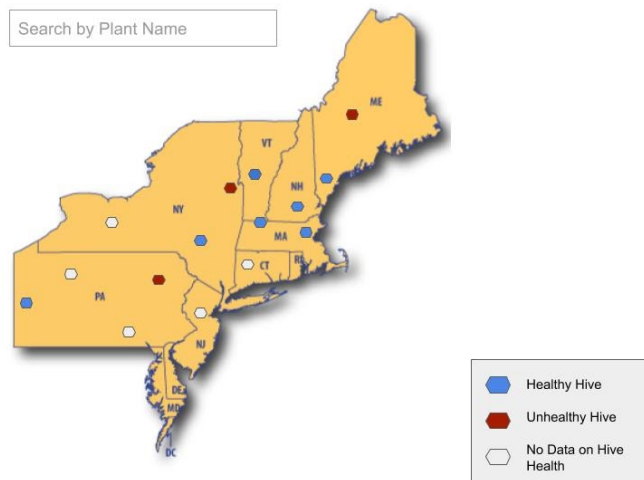
Demo Video

Embedded MP4 demo video using the HTML5 <video> tag. For example, this screen recording Prof. Cody Dunne made of Mike Bostock's flexible transitions in D3 slide:

0:00 / 0:15

Visualization explanation

One of the main goals of our project was to identify "blue zones" or areas where hives were doing well. This was to try to narrow down areas that might have healthy bee hives and compare them to areas that had unhealthy hives. Therefore we wanted to use a map as one of our visualizations. Our other task was to look at the presence of different plants in the honey of different hives. We decided the best way to visualize this was with a bar chart for each hive.



As can be seen in the image above it is easy to see areas of these "Blue Zones". To use this visualization the first interaction involves clicking on one of the circles on the map, representing a hive or a group of hives. This will update the bar chart with the HoneyDNA of the hive selected.

Data Analysis

We used a couple different data sets to compile this visualization. We were given an excel file containing the HoneyDNA breakdown and location for every hive Best Bees had data on. Each sheet represented a different hive. We split this data into individual csv sheets to be more easily accessed and referenced. The location of each hive was given by city and state so we used a

Python script to find the latitude and longitude coordinates of each location. We were also given a couple different excel files containing the condition of the hives for the last four years. Using these files we created a single file containing the condition of the hive for the year the HoneyDNA sample was collected. We also used a JSON file provided during an online lecture containing the data necessary to create a map of the United States.

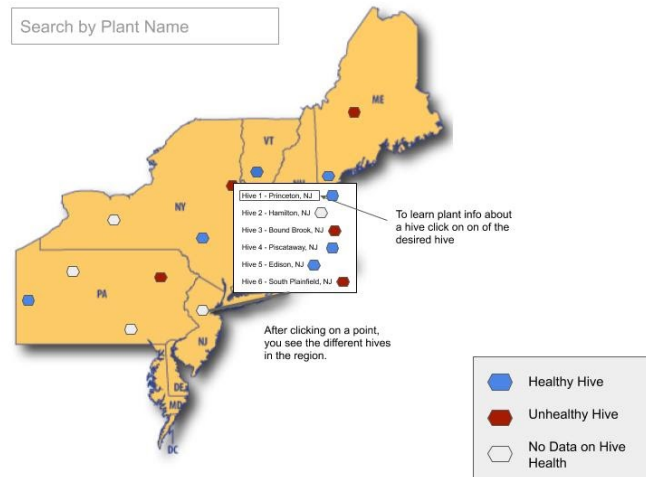
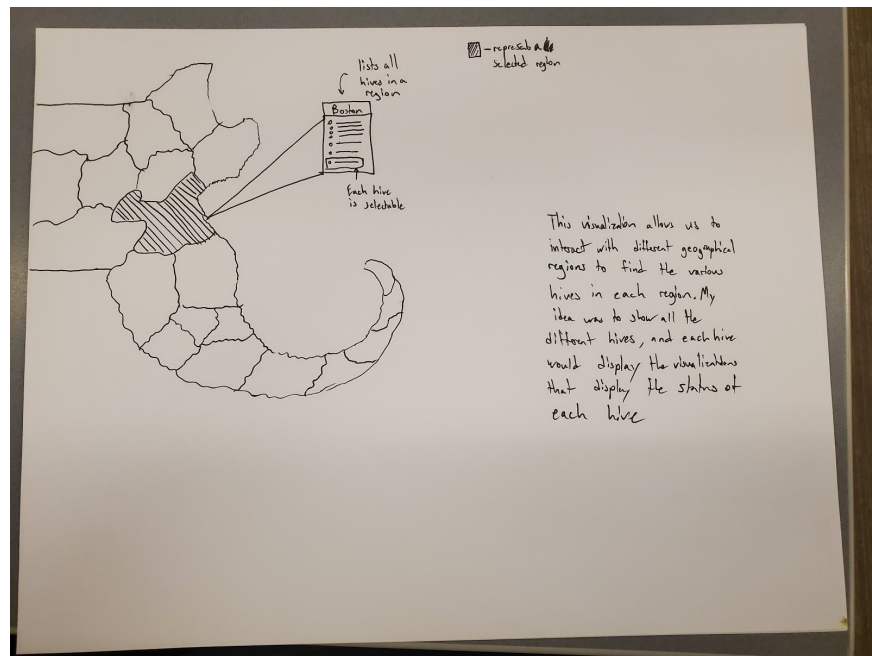
Task Analysis

The first goal was to show areas of good hive health. We also wanted to be able to identify the plants that makeup the honey of each hive. Using this we want to see if there are any correlations between plants and hive health.

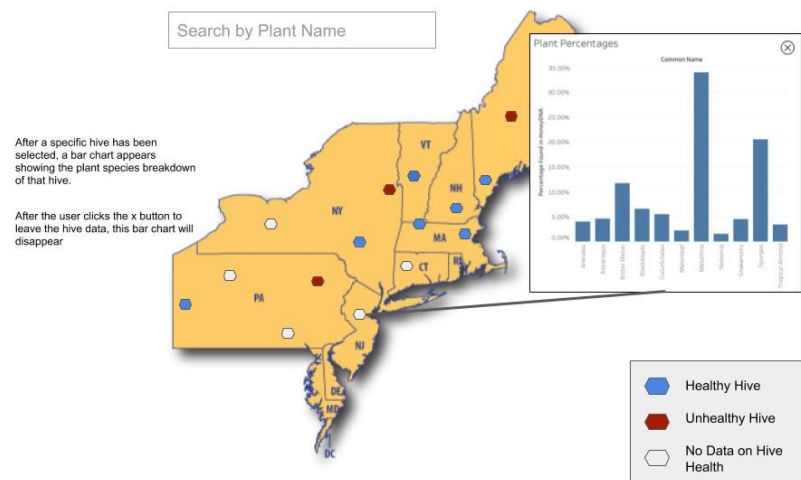
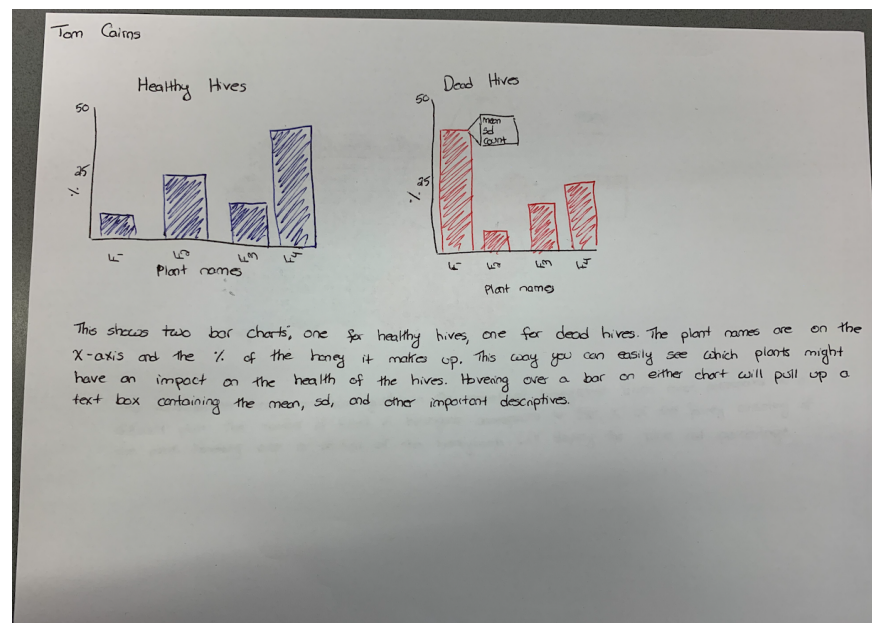
The areas of good bee health were visualized on the map with the color blue to help represent "Blue Zones". By looking at the map it is easy to see if an area has good bee health or not just from the color encodings. Hovering over a point on the graph will update the bar chart in the visualization so the plants found in the HoneyDNA can be seen visually. This way it is easy to see if the hive gets most of their pollen from one type of plant or from a variety of different plants. This can then be compared to other healthy hives to see if there is any correlation between the plants used by the bees and the health of the hives.

Design Process

Our designs started off more ambitious either planning on dividing the data by county or just using dots to represent the location of different hives. This later changed to having dots representing locations of hives. An area with multiple hives would have one dot representing the location of all the hives in that area which could then be clicked on to pull up multiple hives. The color encoding was also changed to easily see the difference between a healthy hive or a dead hive. And another color was added to show hives that did not have health data.



We also had a design for different bar charts representing the average percentage of plants found in healthy hives and unhealthy hives. This was later changed to be visualized from each hive instead of finding the average. Our final design included a box that popped up on top of the map, but this would hide part of the map visualization so we decided to move it to underneath.



One of the major improvements suggested from the usability testing was to make sure to include axes labels on the bar charts.

Conclusion

We were able to plot points on a map of the US to show the location of hives with HoneyDNA data. We adjusted the color of these points to be able to tell where there were healthy hives, dead hives, or hives with no health data. The bar chart underneath the map updates with the hover over a hive to show the HoneyDNA of that specific hive.

There is more work that can be done for the future. We would like to add in an ability to highlight the hives that contain a certain species of plant. This way it can be seen if a lot of healthy hives contain that plant and thus may be used to identify correlation of plants and hive health. This can lead to scientific research into the plant species that may offer the biggest improvement to hive health. This can be used to encourage people to plant

bee-friendly species to improve the overall condition of bees. The Urban Beekeeping Laboratory would like to add this visualization to their website so that it can be accessed by people who are interested in where bees have the best health and which plants might provide benefit to the overall health.

Acknowledgments

Some code for our visualization including rendering the map and adding points was taken from the in class programming assignment d3 tutorial.

- [D3: Data-Driven Documents](#) by Mike Bostock.
- [Pure CSS responsive "Fork me on GitHub" ribbon](#) by Chris Heilmann.
- [Adding labels to axes](#) by Pheobe Bright.
- [Adding a vertical menu](#) by W3schools.com