

# INFO 3300 Project 1 Final Report

Jordyn Vu (jnv8), Joanna Zhang (jaz76) & Julie Jeong (sj598)

## Visualizations

Figure 1: Maps of honey-producing bee colonies in the United States in 1998 and 2012

We want to see if the number of honey-producing bee colonies has changed over time, and if this number varies by state. This can be used to more efficiently direct conservation efforts.

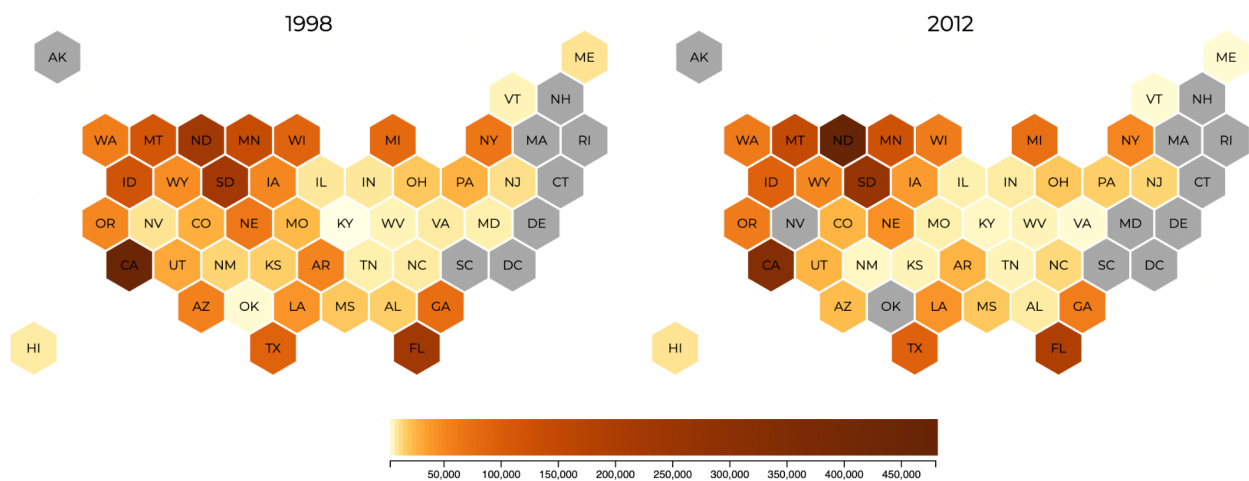
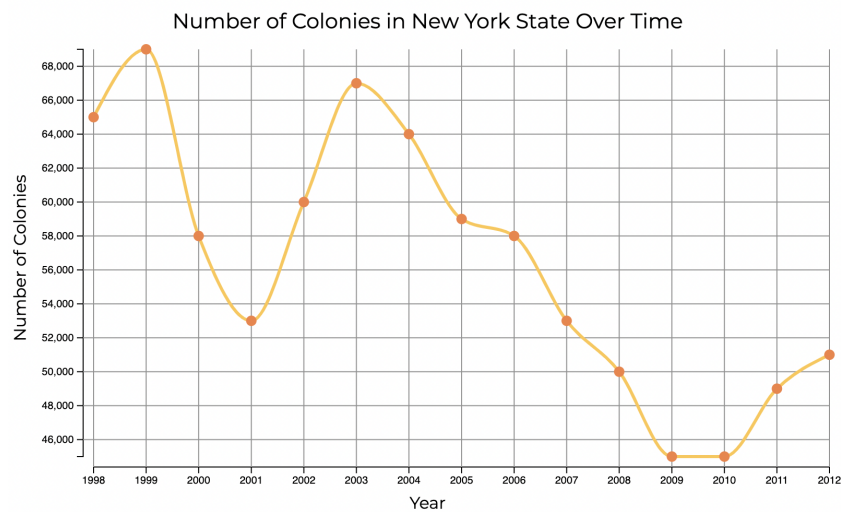


Figure 2: Line graph of honey-producing bee colonies in New York from 1998 to 2012

We want to see how the number of honey-producing bee colonies changed over time for the state of New York.



# Data Description

We got our data “Honey Production in the USA (1998-2012)” from <https://www.kaggle.com/datasets/jessicali9530/honey-production?resource=download> on Kaggle. It has the following columns: `state`, `numcol`, `yieldpercol`, `totalprod`, `stocks`, `priceperlb`, `prodvalue`, `year`. For both visualizations, we wanted to show the trends in the number of bee colonies across the United States over time. Therefore, we decided to use the variables `state`, `numcol` (number of honey producing colonies), and `year`.

For **visualization 1**, we created a simplified version of the U.S. map with each state being represented as hexagons. To create this map, we used the data “us\_states\_hexgrid” from [https://team.carto.com/u/andrew/tables/andrew.us\\_states\\_hexgrid/public/map](https://team.carto.com/u/andrew/tables/andrew.us_states_hexgrid/public/map) and got each state’s coordinates to create hexagon shapes. From our honey production dataset, we chose to show the maps for 2 distinct years 1998 and 2012, thus we filtered the data by year and created two subsets. However, to create the scale that’s used on both graphs, we also created a list of `numcol` from both years to pass it to the color scale.

For **visualization 2**, we created a line graph of the number of colonies in NY over time from 1998 to 2012. In order to create a subset that only contains data points for NY, we filtered the data by using `const numColNY = data.filter(d => d.state=="NY")` for. Also, to represent the years without the comma separating (e.g. 1,998), we used a time parser to format it nicely.

# Design Overview

## For the map visualizations:

*To show change over time, we made two different maps following the same procedures.*

Marks were the hexagons representing states, which we made all the same size because we wanted them to only vary by our saturation channel, not by size, as some of the smaller states would be harder to see in that case. A trade-off of the hexagon mark was that it displaced the states into odd positions, despite us using coordinates from a geoJSON file, making it potentially confusing for a viewer used to the traditional representation of the United States on a map. The way we addressed this was by labeling states by their acronyms.

The channel we chose was saturation, with states with less colonies (taken from the `numcol` column in our `honeyproduction.csv` file) being less saturated in color (pale yellow) and states with more colonies being more saturated (brown). States with no data were colored a neutral gray. We decided to use a sequential color scale to represent the number of bee colonies in each state. To see not only the difference across states but also the change in trends over time, from 1998 to 2012, we applied the same color scale for both graphs. And, since a few states had significantly higher numbers in bee colonies, we used log values to even out the color distribution. By doing so, we were able to compare the relative number of bee colonies in states in 1998 and 2012.

### **For the line graph visualization:**

Marks were the circles representing years while channels included horizontal aligned position and vertical aligned position. We chose to make a line graph so the viewer could follow along the line to see how the number of bee colonies declined from 1998 to 2012, though there was an increase from 2001 to 2003. We curved the line to make the viewing experience more fluid.

## Story

From our two map visualizations in 1998 and 2012, we can see that both maps show that the west coast consistently has more honey-producing bee colonies than the east coast. In general, there has been a decrease in the number of honey-producing bee colonies across the US states with a couple of exceptions. The states that saw a dramatic decrease in bee colonies are: ME, VT, WI, NY, PA, NE, MO, VA, UT, NM, KS, AR, AZ, AL. The states that saw an increase in bee colonies are: ND, OH, NJ, NC, HI. It's interesting to see that California was the predominant bee colony producer in 1998, and California, North Dakota, and South Dakota had the highest honey-producing bee colonies in 2012.

For our line graph for New York from 1998 to 2012, we can see an overall decrease in the number of honey-producing colonies. We found the dip in 2001 and the rise after 2010 to be interesting. We found out beekeeping became legalized in NY in 2010, which could have contributed to the increase in bee colonies after 2010.

These visualizations clearly show that environmental changes such as global warming are impacting the number of honey-producing bee colonies. With the increased burning of fossil fuels, our carbon footprint has dramatically increased throughout the years. According to the Milken Institute School of Public Health, as temperatures rise, pollen and nectar decrease in availability, which happens to be

honeybees' main source of nutrition.<sup>1</sup> However, global warming is not the only factor causing the decline of bee colonies - habitat destruction, increased use of pesticides, and loss of natural abundance of flora and fauna are also contributing to the decline in bee colonies. A third of the world's food production relies on pollination from bees, and bees also provide a source of income for many rural communities.<sup>2</sup> Through our project, we hope to inform people and raise awareness about the bee colonies.

## Contribution

### Collaboration

Coded the map visualization (~2.5 hours):

- Jordyn Vu (created initial hexagonal cartograms)
- Julie Jeong (created scales and mapped data to map)
- Joanna Zhang (added styling to maps)

*All three group members participated in the planning and coding stages of the map visualizations, though Joanna did pre-planning so the process would flow smoother*

Coded the line graph visualization (~1.5 hour)

- Jordyn Vu
- Julie Jeong

### Individual

*The project write-up was split amongst the three group members according to the outline below.*

Jordyn Vu

- Wrote the design overview/rationale (~20 minutes)

Julie Jeong

- Wrote the data description (~30 minutes)

Joanna Zhang

- Planned and sketched the visualizations to be coded (~20 minutes)
- Wrote the story; researching and summarizing the project topic (~45 minutes)

---

<sup>1</sup>

<https://onlinepublichealth.gwu.edu/resources/colony-collapse-climate-change-public-health/#:~:text=As%20temperatures%20rise%20and%20the.crops%20can't%20produce%20fruit.>

<sup>2</sup> <https://www.unep.org/news-and-stories/story/why-bees-are-essential-people-and-planet>