## Introduction to the dataset

Global concern was raised in 2006 over the rapid decline in honeybee population when large numbers of hives were lost due to Colony Collapse Disorder, which is when disappearing worker bees result in the remaining hive colony collapsing.
This disorder may be due to hive diseases and pesticides harming the pollinators, though this has not been proved. Some industries have shown recovery over the last couple of years but the American honey industry is still struggling.

I found a tidy dataset on Kaggle "honeyproduction.csv" that provides insight into honey production supply and demand in America by state from 1998 to 2012. Metadata on the column names are provided below:

- *numcol*: Number of honey producing colonies. Honey producing colonies are the maximum number of colonies from which honey was taken during the year. It is possible to take honey from colonies which did not survive the entire year
- *yieldpercol*: Honey yield per colony. Unit is pounds
- *totalprod*: Total production (numcol x yieldpercol). Unit is pounds
- *stocks*: Refers to stocks held by producers. Unit is pounds
- *priceperlb*: Refers to average price per pound based on expanded sales. Unit is dollars.
- *prodvalue*: Value of production (totalprod x priceperlb). Unit is dollars.

## Summary

I used the dataset described below to see how honey production has changed over the years. I also investigated how honey production was affected by Colony Collapse Disorder pre and post 2006.

**Main finding:**

Total honey production decreases over the years and for most states average honey production as well as average yield per colony is higher before 2006 than after.

**Additional findings:**

North Dakota is the main producer of honey, ranging between first and second place each year. South Dakota, California, and Florida are also big contributors to total honey produced each year in the USA.

Combined (for all states) yield per colony has a decreasing trend over the years.

As total honey produced decreases, price per pound increases.

Stock decreases when prices increase and vice versa.

For  North Dakota, the number of honey producing colonies have been increased over the years, this could explain why honey production is higher after 2006 than before 2006 for North Dakota, despite Colony Collapse Disorder and decrease in yield per colony.

## Design decisions

I combined my Tableau sheets and dashboards into a story, for which I have provided a link below:

Here is the first visualization: link1

In the first slide I show a line graph. The line graph shows change in total honey produced over the years. I added a trend line to clearly show the decrease in overall honey production over the years.

In the second slide I used a map. The map uses a color visual encoding to indicate the amount of honey produced per state. I added a table with detail on the five highest producing states. I included a sliding filter for time.
Sliding the filter and moving the date along changes the colors on the map to display the total production of honey at that time, and updates the table. A color legend was included to make it easier to identify which states have higher honey production relative to each other, based on color.

The third slide show line graphs.
For these line graphs, color was used to clearly distinguish between lines in each graph. Legends were also provided for clarity. Dual axis were used to make it easier to see the relationship between the features that were compared. Clear labels were used for the axis labels, and units were also included. I used colors that can easily be distinguished between if the viewer is color blind.

For the fourth slide I used the Group method to create two groups based on if data was before or after 2006 - when the Colony Collapse Disorder first started to become a global concern. I illustrated average honey produced before and after 2006 using a bar graph and two different colors.

The final slide is a scatter plot. It aims to show the change in average yield per colony before and after 2006. I used two different colors to distinguish between these time periods.
I used a scatter plot rather than a line graph for yield per colony data since I was not showing yield per colony over time, but rather per state.

**Design changes based on feedback**

Based on feedback I added another two slides to the story.

For the first slide I added I used line graphs to show the relationship between stock and price over time.

For the second slide I added I show change in number of colonies over time, filtered specifically for North Dakota. A trend line was added, clearly illustrating the increasing trend.

Here is the new visualization: link2

## Feedback

## I received the following feedback:

Provide information on stock change with price change over time to see the relationship between these two.

Show the change in number of colonies for states over time.

## Resources


https://www.kaggle.com/jessicali9530/honey-production

http://www.abc.net.au/news/2017-05-08/colony-collapse-ten-years-after-crisis-what-is-happening-to-bees/8507408

https://www.bloomberg.com/news/articles/2017-08-01/good-news-for-bees-as-numbers-recover-while-mystery-malady-wanes

https://www.washingtonpost.com/news/wonk/wp/2015/05/14/dying-bees-could-mean-the-end-of-american-honey/?noredirect=on&utm_term=.c366ef917c16