**Not the Bees!**

**Data:** <https://www.kaggle.com/kevinzmith/honey-with-neonic-pesticide/home>

I liked the bees because it goes back to how I feel about the environment and protecting it. I’ve heard news about the decline of bees and wondered if pesticides are a factor of the declining population of bees.

**Variables:**

Looking at the dataset, it seems the main variables would be:

*year state pesticides (kg) totalprod (lbs)*

The sub-variables that I may use are:

*numcol, yieldpercol, and StateName*

**Cleaning Up:** I looked at the dataset and much of it was already neatly written without me having to refine it too much. I ended up removing rows that had any empty columns in it. I also removed columns I felt were unnecessary. These included the four sub neonic *pesticides*, *FIPS*, *Regions*, *prodvalue*, *priceperlb*, and *stocks*.

I removed the four sub pesticides because I wanted to just focus on the overall effect of *pesticides*. I removed *prodvalue*, *priceperlb*, and *stocks* because I wasn’t going to be working with the prices of honey. I didn’t know what *FIPS* was, but I removed it since it didn’t seem relevant. *Regions* was also irrelevant as I wanted to focus on the state rather than the regions of it.

**Data Analysis**: After working on the code, I decided to change what was going to be on the y and x axis. I realized that it didn’t make since to have *totalprod* on the x-axis and *pesticides* on the y-axis. The vertical bars I would have would make no sense of the relationship between the two. I knew I wanted to show the effects of pesticides on bees.

I looked back at the dataset and tried to create a relationship between the columns. I couldn’t just use *numcol* because that didn’t seem as though it was the total amount of the colonies in the state. If *numcol* wasn’t the total amount of colonies in the state, then the effects of pesticides on bees would seem invalid. This led me to use *totalprod* instead which is the total output of honey from the *numcol* x *yieldpercol*. Then I thought about using the *years* variable. If I were to change the x-axis to years and the y-axis to *totalprod*, then the user could see the *totalprod* output per year for that that state. Perhaps this way, the user can see the *totalprod* trends in history for the state and make predictions of what is to come for the next unknown year. I didn’t want to leave the *pesticides* variable out from my graph as I felt that was an important variable that may play a role in for the state’s *totalprod* in a certain year.

I decided to add different fill colors to the bars to represent the amount of *pesticides* used during a certain year as well. The user will be able to get the exact amount of pesticides used for the state during a certain when they mouse over a bar. The more pesticides used, the browner the bar will be. For lower amounts of pesticides used, the bars will fill to a light yellow. The contrast between the light yellow and brown will hopefully help the user visually see and compare how much pesticides were used between different years.

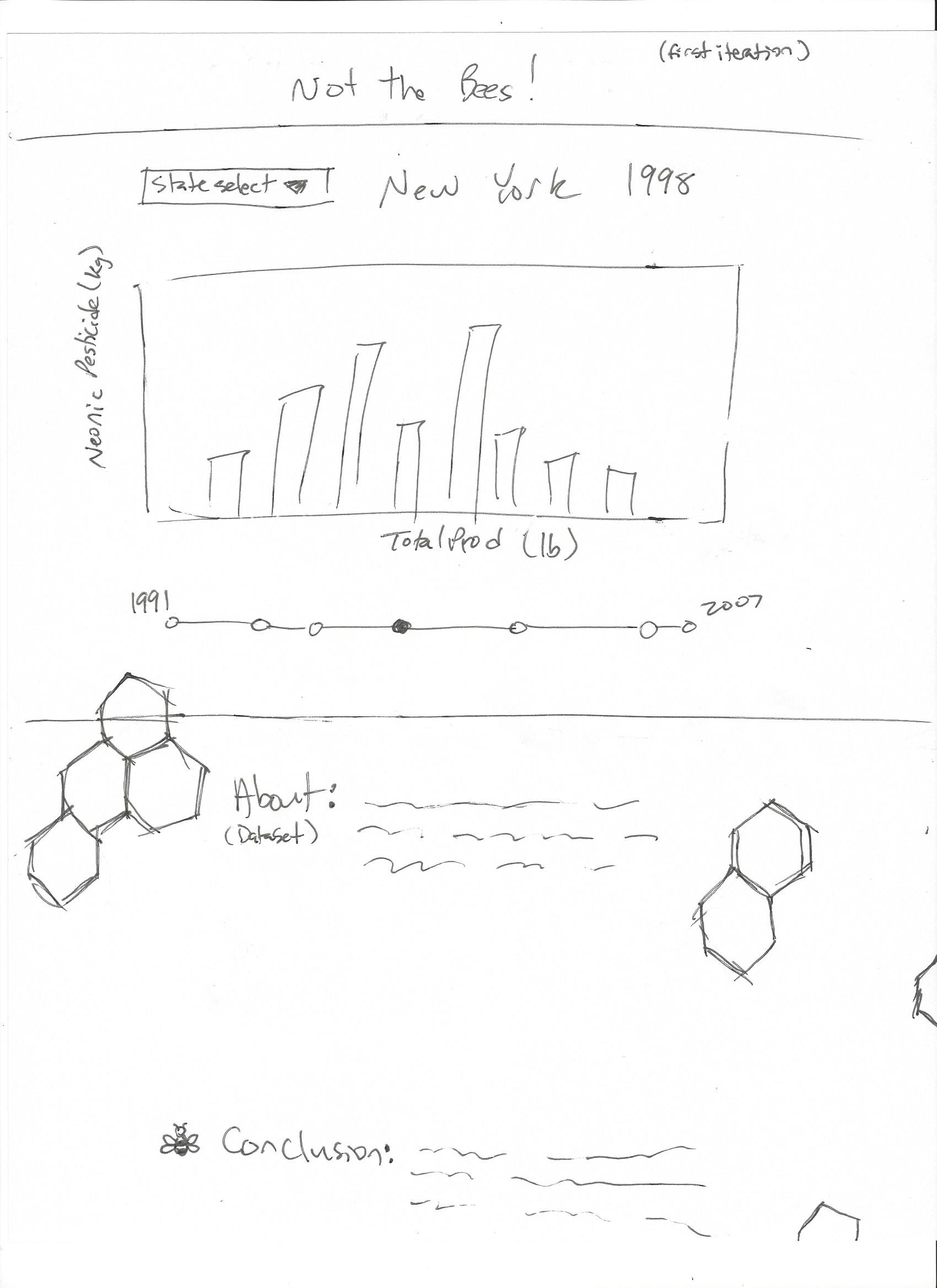
**Challenges:** I didn’t encounter too many problems. Most of my problems were because I got attr or styles mixed up and didn’t parse the values from the dataset as int or float. A lot of what I wanted to do which included animations, horizontal grids, and mouse events, all I had to do was look back at past coding example and or use Google to look at how others implemented what I wanted.

**Story:** I think it’s easy to see a story within each dataset for each state. Each state has its *totalprod* output for each year and the amount of pesticides used that year.

For example, in the state of Idaho, between 1994 and 2006, the *totalprod* output is bit higher than the *totalprod* between 2007 and 2016. As the viewer, you can see that within 1994 and 2006, lower amounts of *persticides* were used where as between 2007 and 2016 a higher amount of *pesticides* were used. The story here tells the viewer that pesticides has affected the *totalprod* out in Idaho. This type of pattern can also be seen in the state of Florida, Alabama and other states.

In the state of Minnesota, the effects of *pesticides* are very similar to that of Idaho and Florida, however, the years after 2014, Minnesota used much less pesticides during those year than the years before hand. The *totalprod* after the years after 2014 have been slowly increasing. The viewer can further conclude that the reduce use of *pesticides* have a positive effect on *totalprod*.

There is another story that pops up to the viewer. If the viewer views several states, they can see that most of the states listed all started to increase their use of pesticides around the same years which is between 2011 and 2014. After 2014, the states listed have reduced the use of pesticides.

**Visualization:** I created a quick sketch of how I thought the webpage should look like:

It will be a one-page website with 3 sections: the title, the data and graph, and the text/paragraph section.

In the Data and graph section, there will a drop-down selection of different states for the user to look at. The graph will have the *pesticides* on the y-axis and the *totalprod* on its x-axis. There will be a timeline where the user can click on each dot to change the *year* they are looking at. The user will also be able to hover over each bar on the graph to get the exact values of the *totalprod* and pesticides used. The bars will animate and change in height when the user changes what state to look at.

I also looked up some color palettes to use for the website:



<http://colorpalettes.net/color-palette-3655/>



<http://colorpalettes.net/color-palette-3809/>

I wanted the website to have the same colors of a beehive. The off-white will be used as the neutral background-color for the data and graph section. The yellow will be used for the title and text/paragraph section.

I ended up not creating the *year* timeline in which the user could change the dataset for the state by *year* because the *year* variable is now used for the x-axis.