

Midterm project

Group 10: Kayla Choi, Kosuke Sasaki, Clare Zike Tan, Franky Yifan Zhang

2021/11/03

1. Introduction

Our research questions about the strawberry production and chemical use in the US are:

- i) Does the type of chemical have a correlation to the amount of strawberries produced (Measured in lb/acre/ application)?
- ii) Do different states use different chemicals?
- iii) How do the amount of strawberries produced differ by state?

To answer these questions, we looked up two data sets, cleaned and merged them into one data set, and created graphs, tables and maps for Exploratory Data Analysis (EDA).

2. Methods

i) Data set

We used two data sets, USDA strawberry data and pesticide data. For the strawberry data set, the provided one in the class does not seem to be enough for our research in terms of information on each state production, so we looked up the data on the USDA web site again and got the data set which include in total 10215 observations of strawberry data in 10 states.

For the pesticide data, we adopted the provided data.

Finally, we combined these two data sets, deleted rows with no information, and selected the variables needed for EDA.

ii) Exploratory Data Analysis

EDA by chemical type

First we analyzed the data set making bar chart, violin and box plot, and density graph by chemical type to see if there is any difference in the amount of strawberry production between the two chemical types.

EDA by year and state

We also analyzed the data in terms of the amount of strawberry production by year and state where we make use of shiny application which allows us to explore strawberry production trends in each state and year more easily.

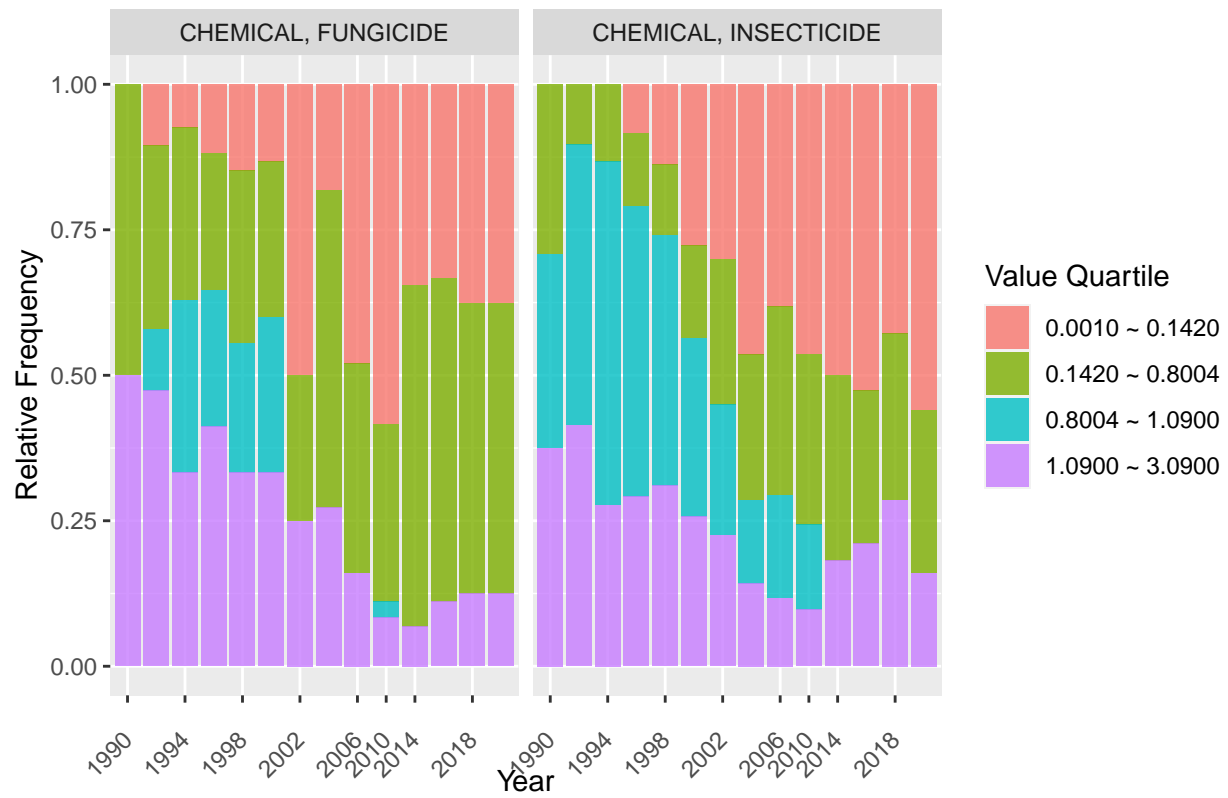
EDA by toxicity level

Finally we analyzed the data in terms of the amount of strawberry production by toxicity level.

3. Results

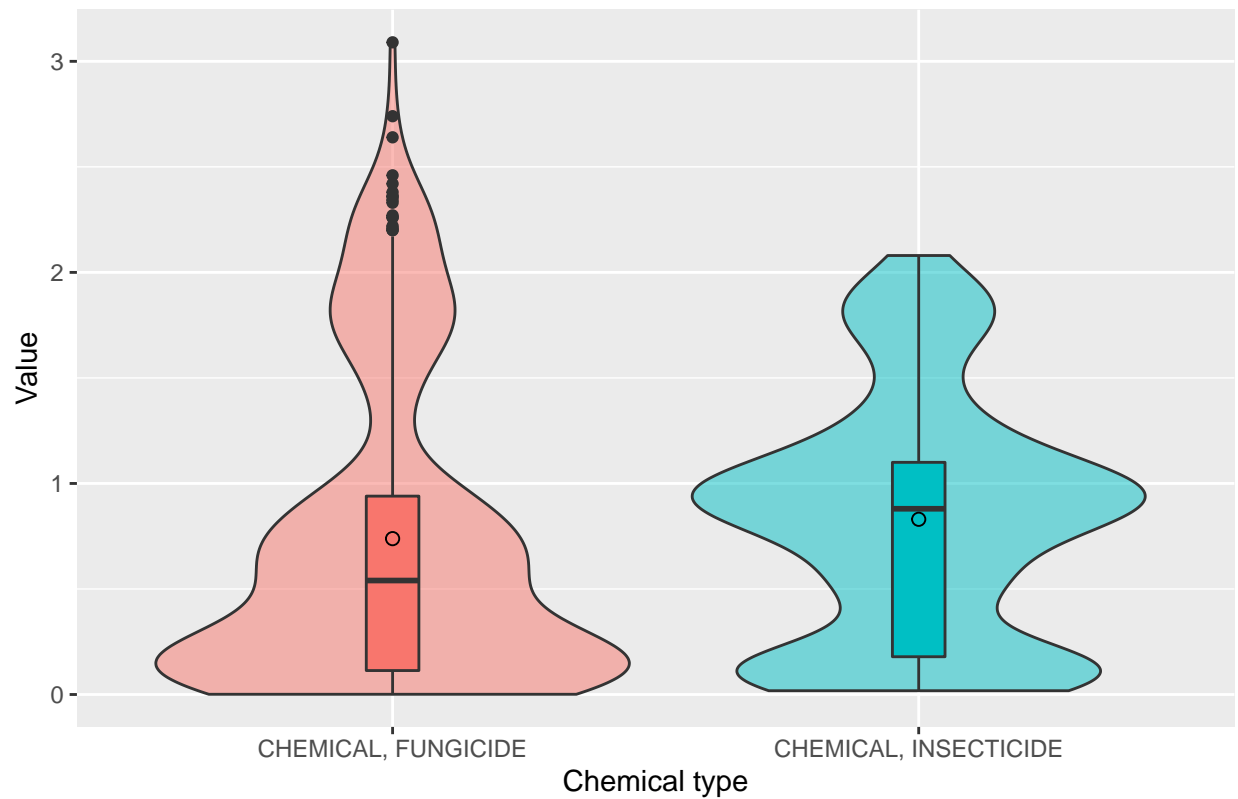
EDA by chemical type

fig.1 Relative Frequency of Chemical use per year



For both fungicide and insecticide, the value generally goes down every year, which we can see from the increasing 1st quartile (pink) and decreasing 4th quartile (purple).

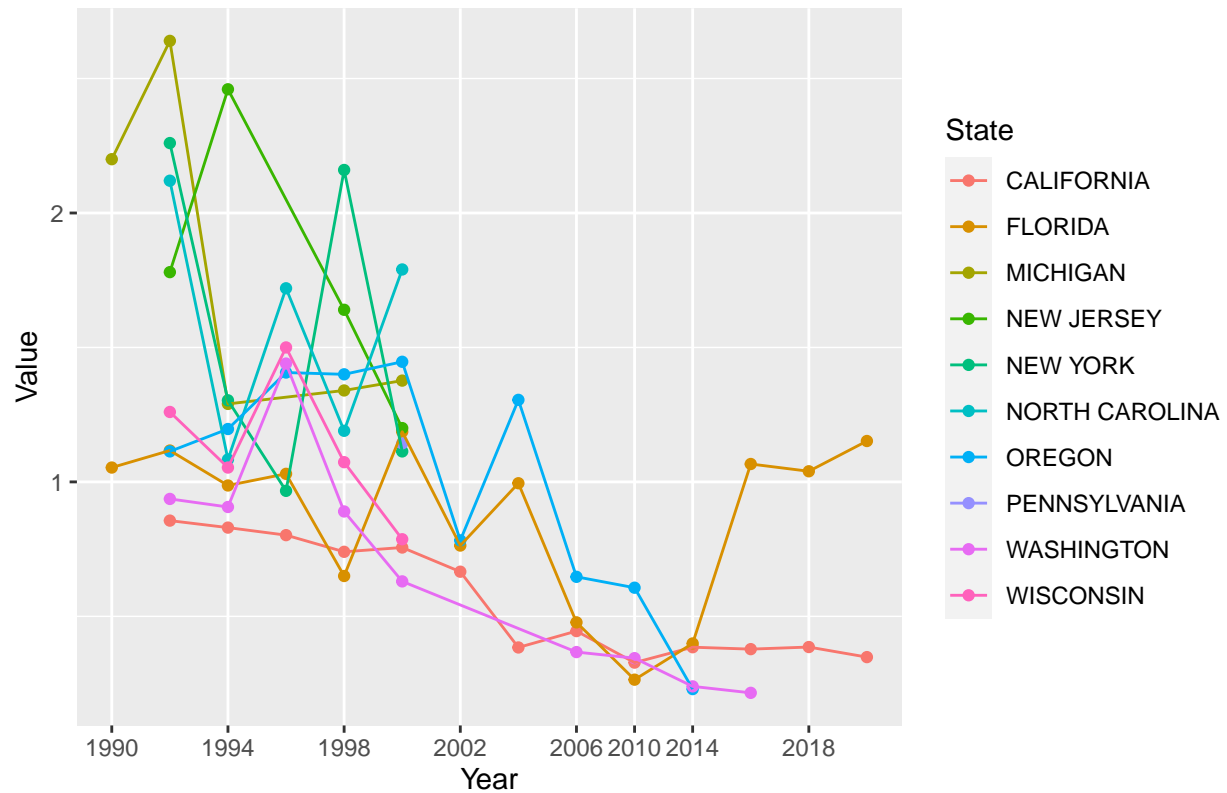
fig.2 Distribution of Chemical use and corresponding strawberry production



Insecticide has higher mean and median value, but fungicide has a larger spread. Fungicide corresponds with lower strawberry production but has larger spread. Insecticide has higher values.

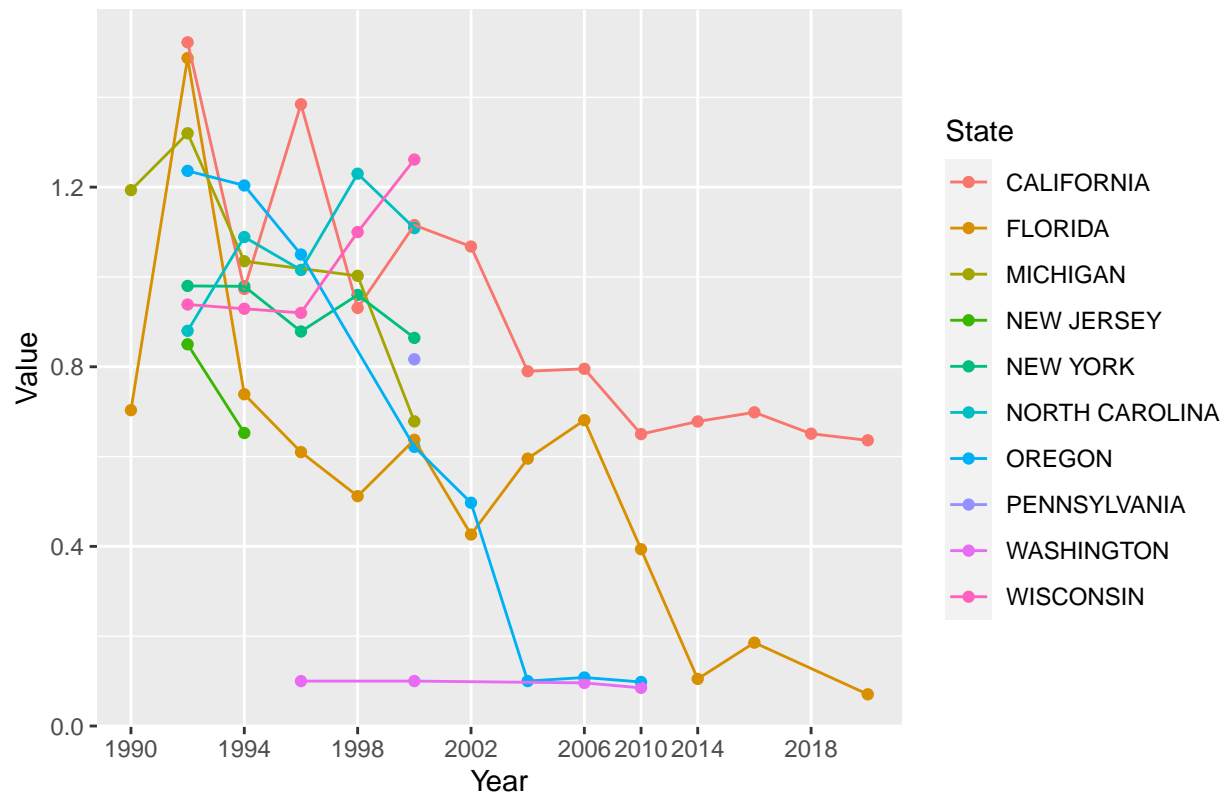
EDA by year and state

fig.3 Strawberry production per acre per fungicide application over time by st:



The trend of strawberry production per acre per application decreases over time, which corresponds to the fig.1 result.

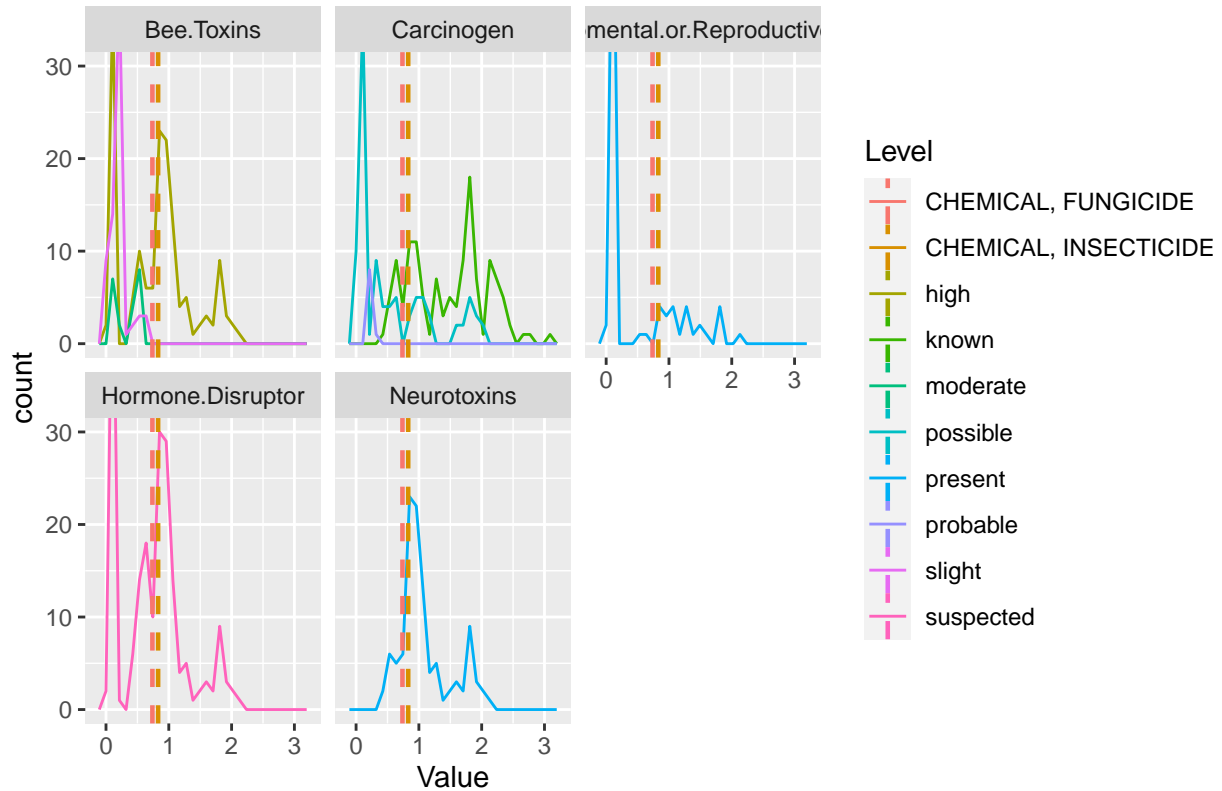
fig.4 Strawberry production per acre per fungicide application over time by s



This plot also shows the down trend, but the value for California levels off after 2010.

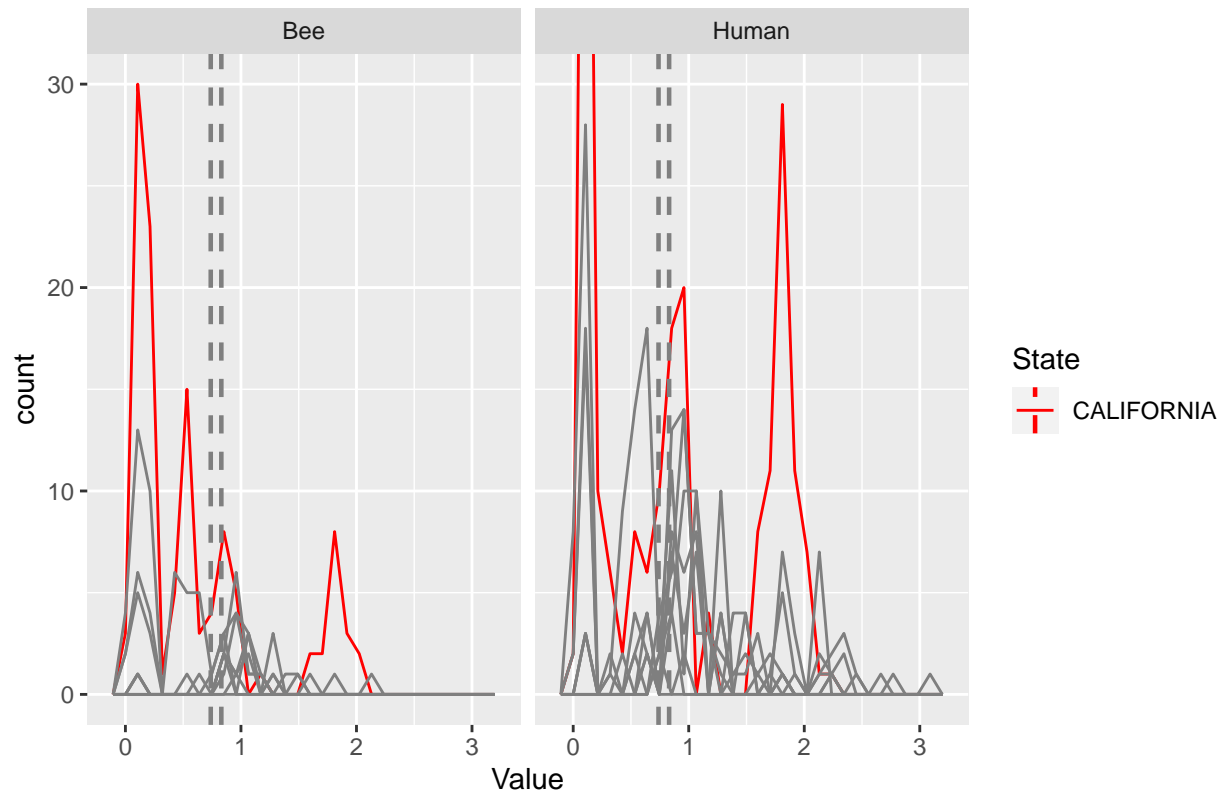
EDA by toxicity level

fig.5 Chemical use by Toxicity level



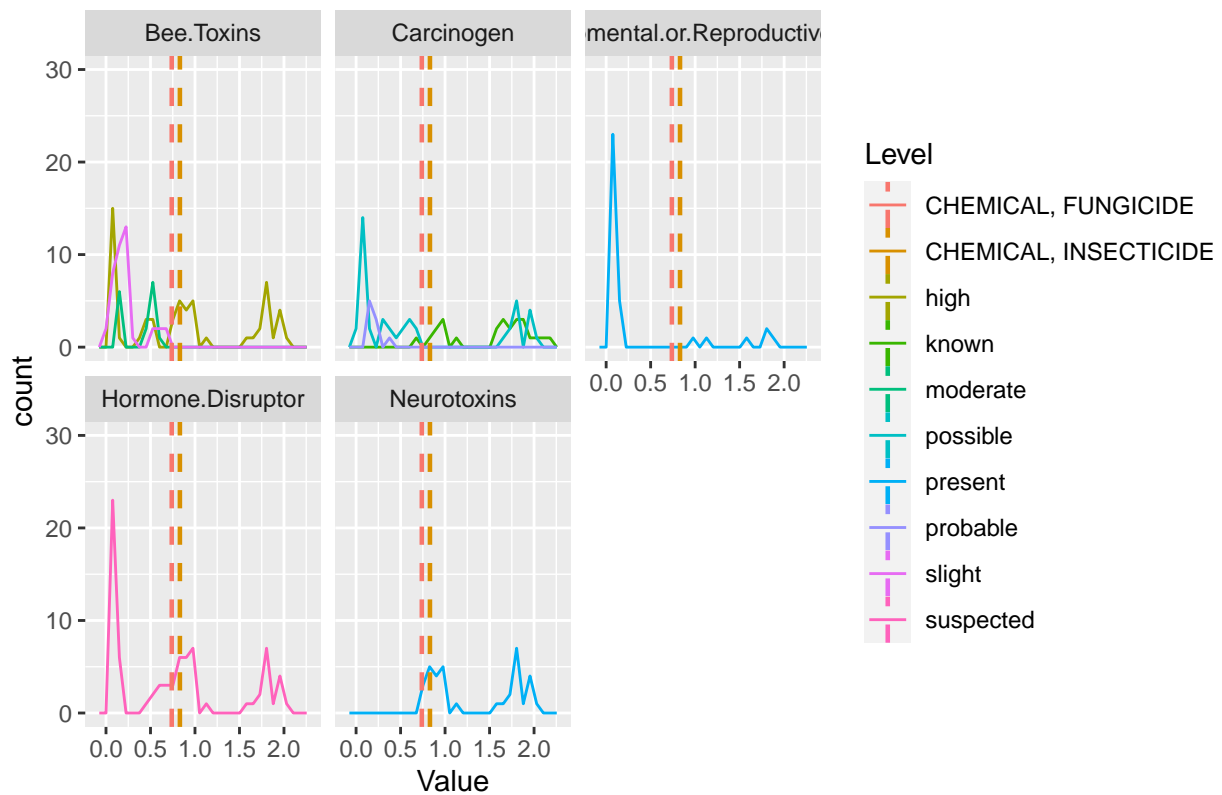
These graphs show the distribution of the amount of strawberry production per acre per application by toxicity type. For the chemical toxic to bee, there seems to be a correlation between high toxic chemical and the amount of strawberry production. For the other chemicals toxic to human, know carcinogen, hormone disruptor and neurotoxins seem to have high strawberry production per acre per application.

fig.6 Chemical use in California compared to other states



These graphs show the distribution of the amount of strawberry production per acre per application by state, focusing on California. For the both chemical types, California seems to have high amount of strawberry production per acre per application, especially around the value of 2 (lb/ acre/ application).

fig.7 Chemical use by Toxicity level in California



These graphs show the distribution of the amount of strawberry production per acre per application by toxicity type in California. Based on the previous two sets of graphs, there seems to be some correlation between high bee toxic chemicals, known carcinogen, and some human toxic chemicals and the amount of strawberry production as well as California and the amount of strawberry production. So we look at the distribution of the amount of strawberry production by toxicity type in California.

fig.8 Map of the strawberry production in 2000

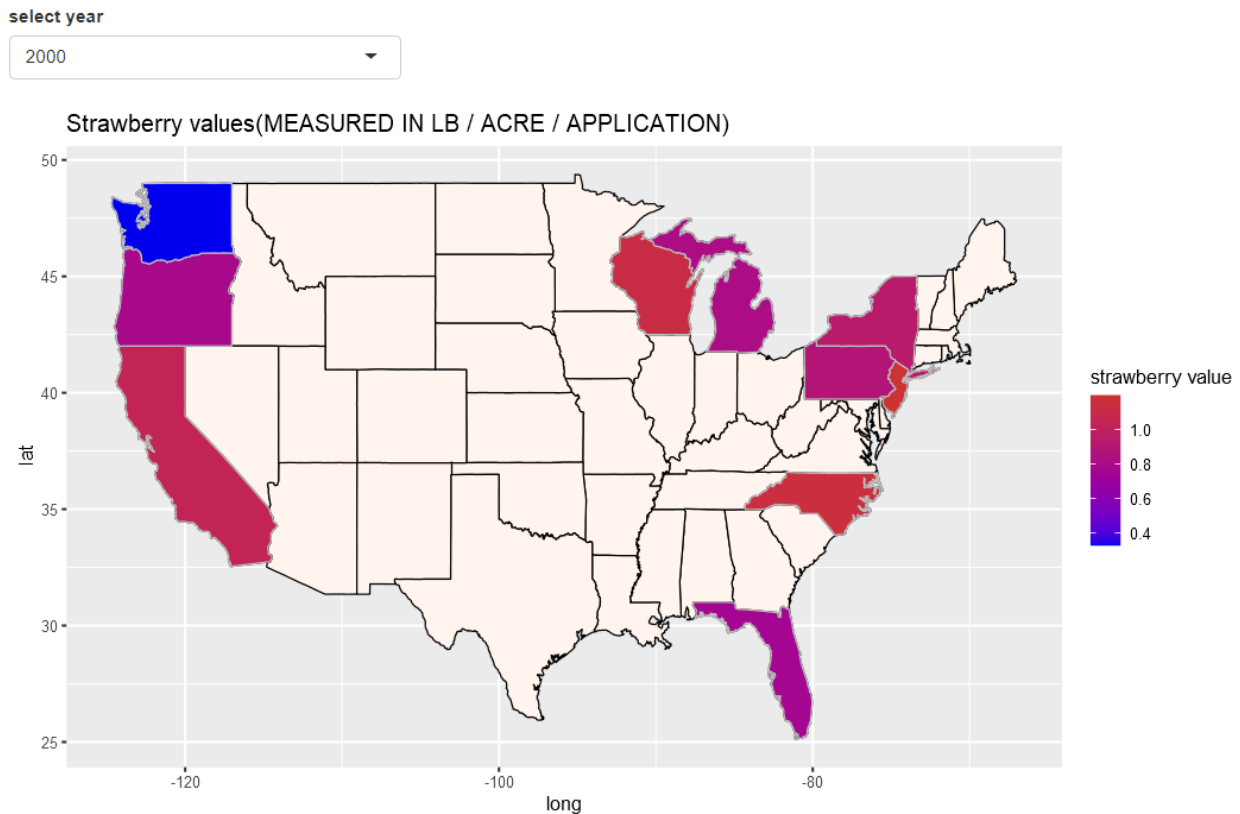


Figure 1: map of the strawberry production in 2000

California shows the most strawberry amount per acre per application among the states.

4. Conclusion

We saw a mostly downward trend in the strawberries/acre/application over time, but California has not decreased their insecticide application. This makes sense because in the 2010s, the lygus bug, which is a major insect pest of strawberries, started becoming a bigger problem in Southern California due to warmer and drier climates.¹

We also saw some correlation between high bee toxic chemicals, known carcinogen, and some human toxic chemicals and the amount of strawberry production even within California, which suggests there seems to be some correlation between the chemical types and the amount of strawberry production when adjusted for states.

Reference

1. Summary of Pesticide Use Report Data 2015