Midterm Report

I am starting with cleaning the data and then separating it into 4 data sets: California census, California survey, Florida census, and Florida survey.

California

All the data cleaning for California is done in this section.

Census Data

In this section, I'm cleaning the census data for California. This is where I am getting the information for the operations vs producers section.

Survey Data

I am cleaning the survey data for California in this section. Here is where I am discovering the information I use for California's chemical analysis.

Florida

All the data cleaning for Florida is done in this section.

Census Data

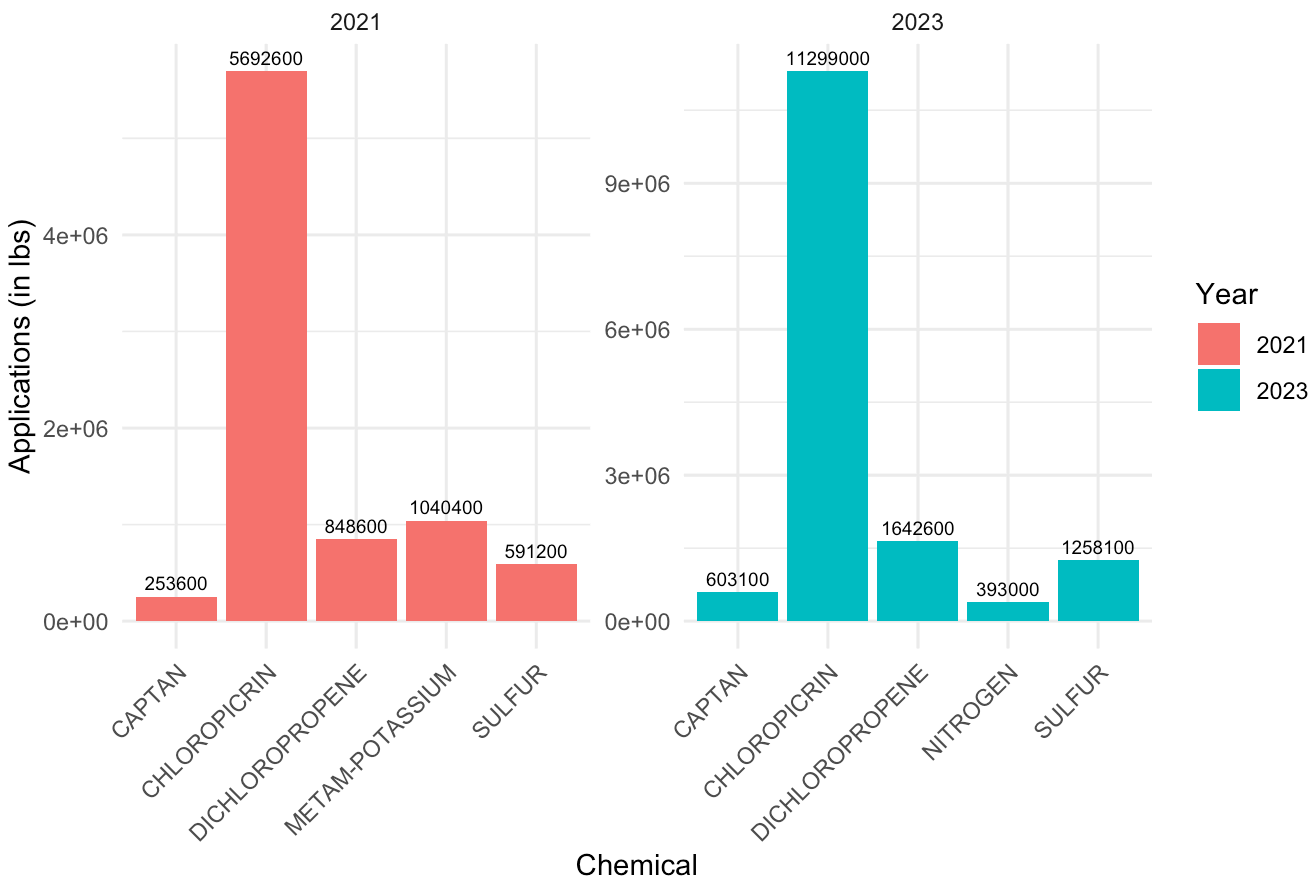
In this section, I'm cleaning the census data for Florida. Similarly to California, this is where I am getting the information for the operations vs producers section.

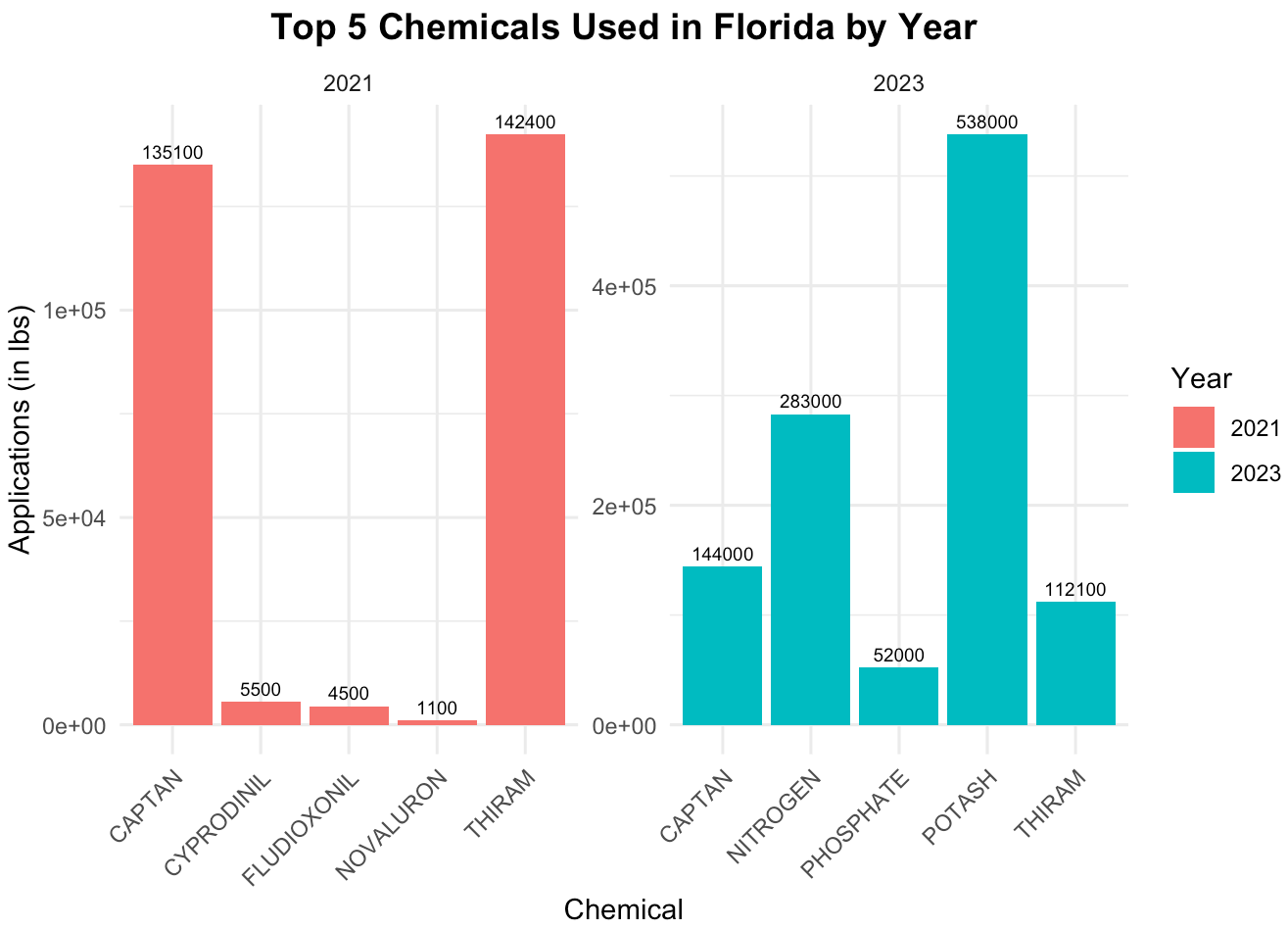
Survey Data

I am cleaning the survey data for Florida in this section. Like with California, here is where I am discovering the information I use for Florida chemical analysis.

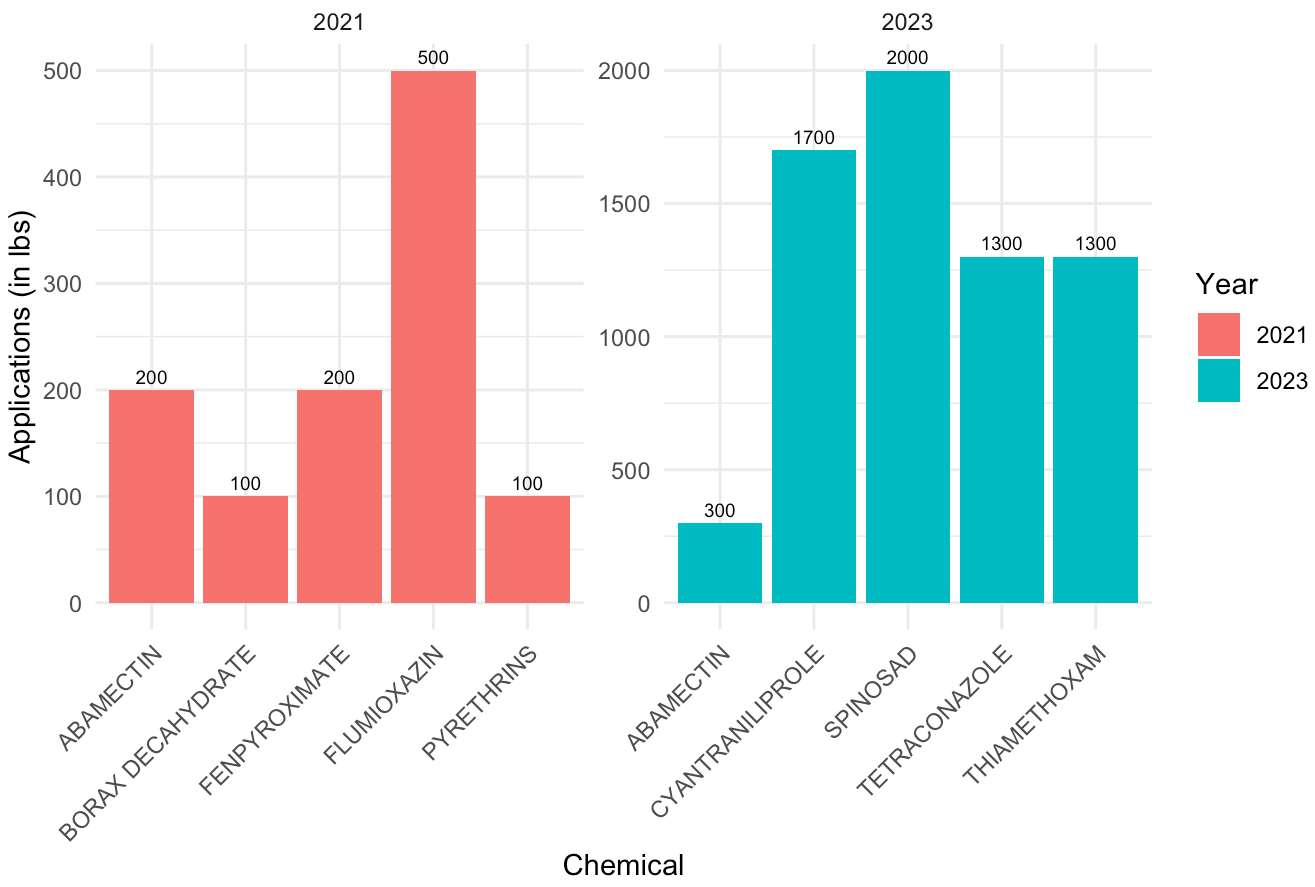
Chemical Analysis

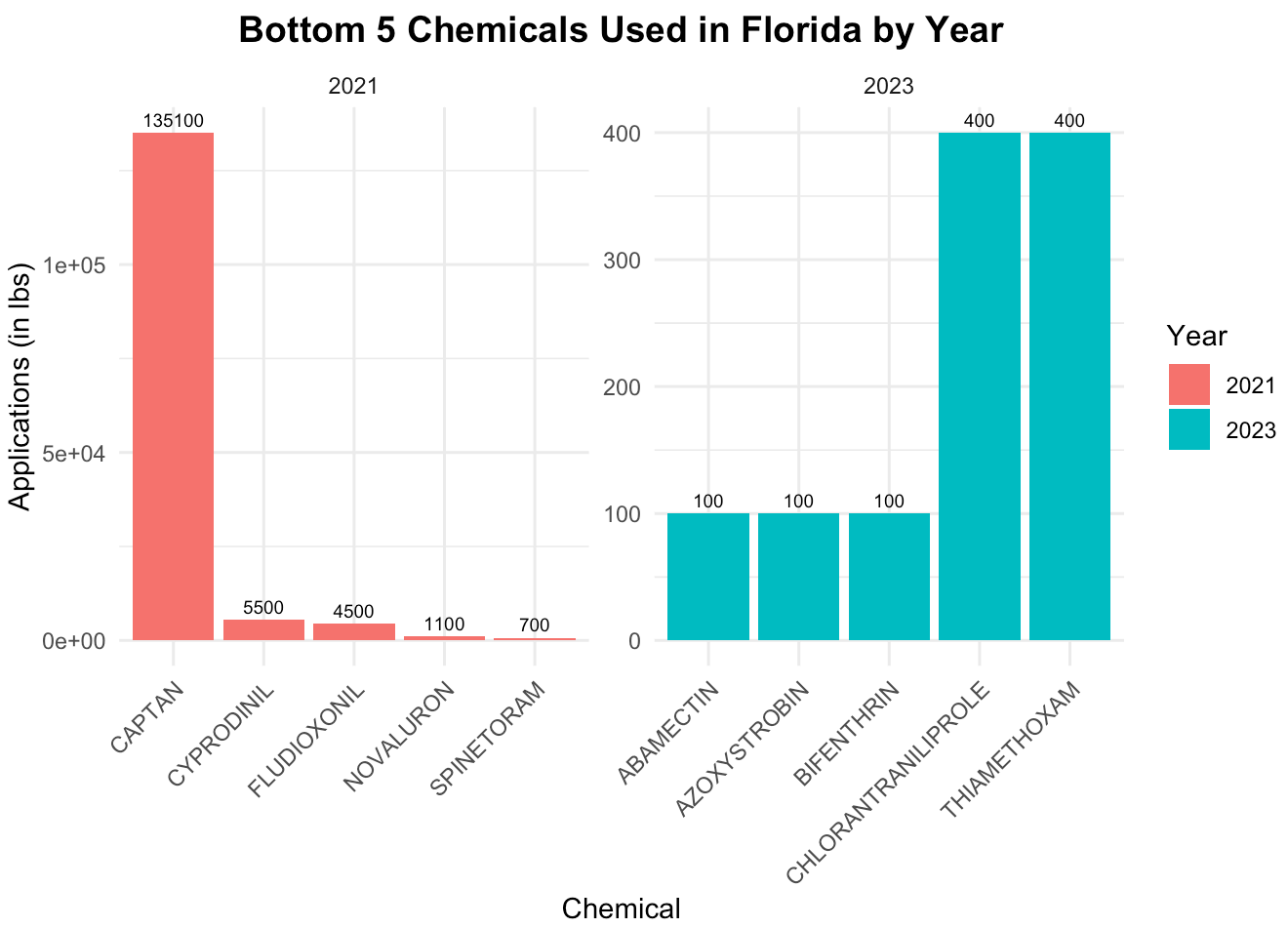
First, I'm comparing the top 5 chemicals used in each state.


Top 5 Chemicals Used in California by Year
2021
2023
Applications (in lbs)
0e+00
2e+06
4e+06
6e+06
9e+06
3e+06
253600
5692600
848600
1040400
591200
603100
11299000
1642600
393000
1258100
CAPTAN
CHLOROPICRIN
DICHLOROPROPENE
METAM-POTASSIUM
SULFUR
NITROGEN
Chemical
Year
2021
2023

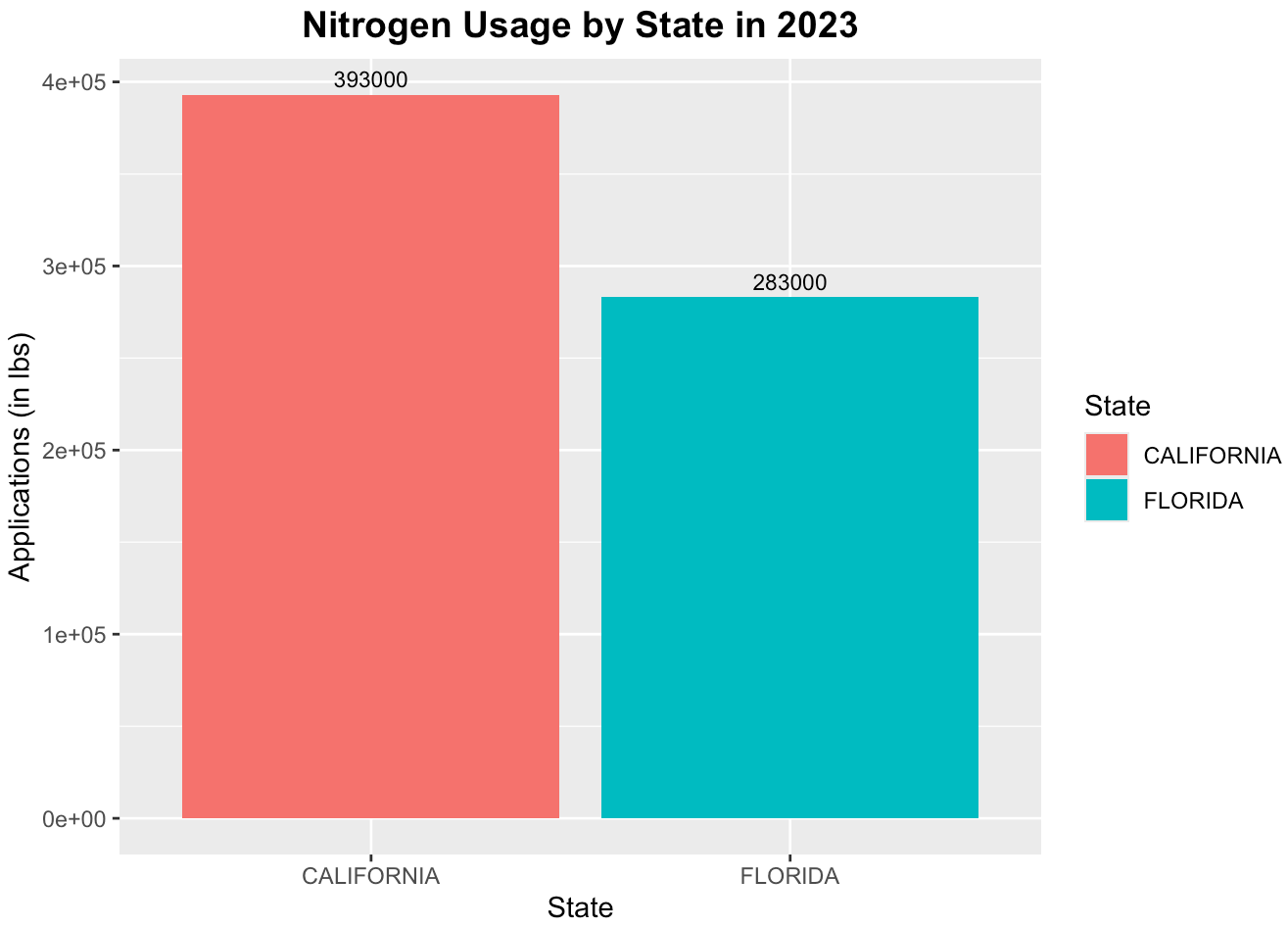


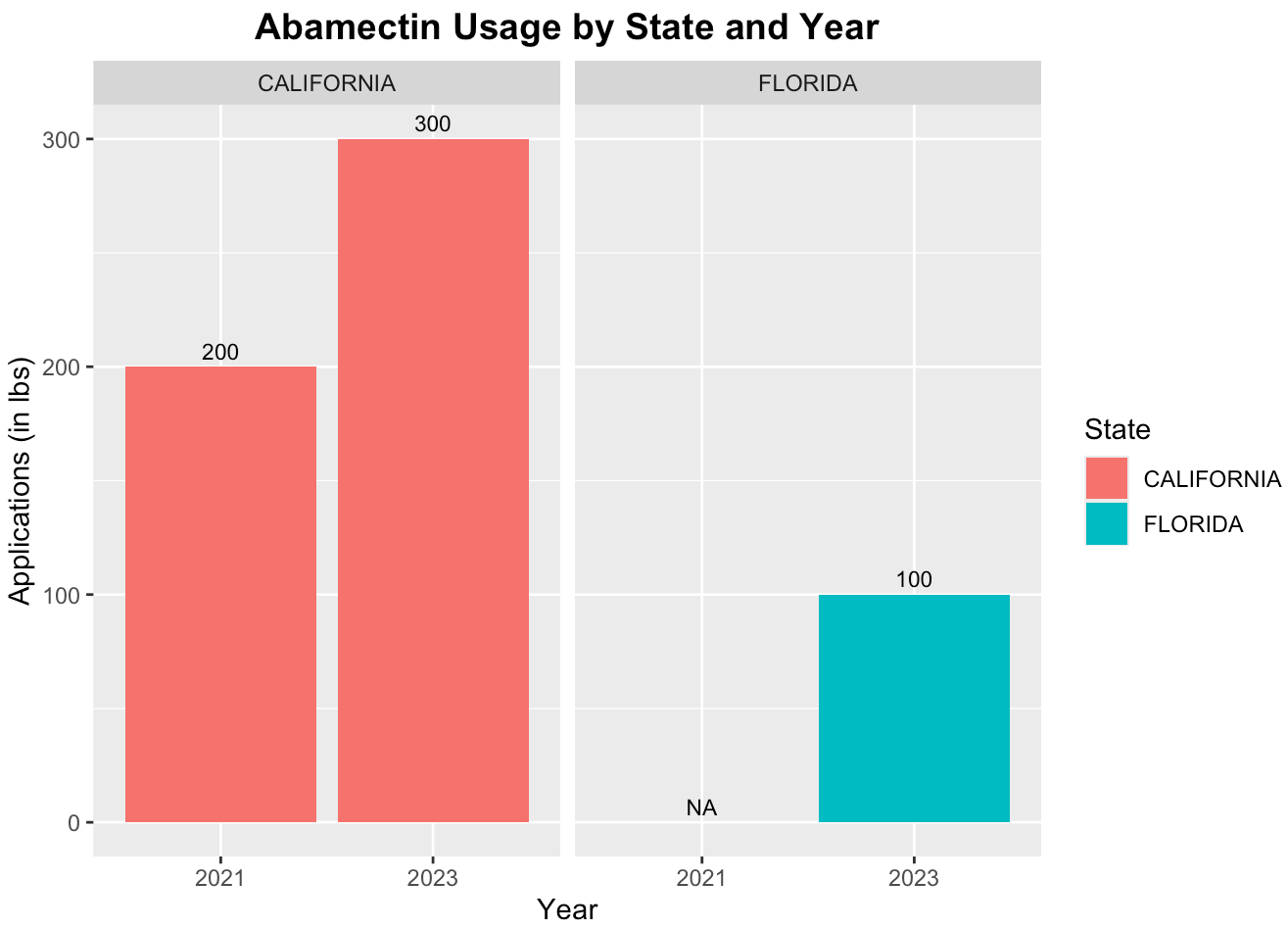
Now I'm comparing the the bottom 5 chemicals used in each state.

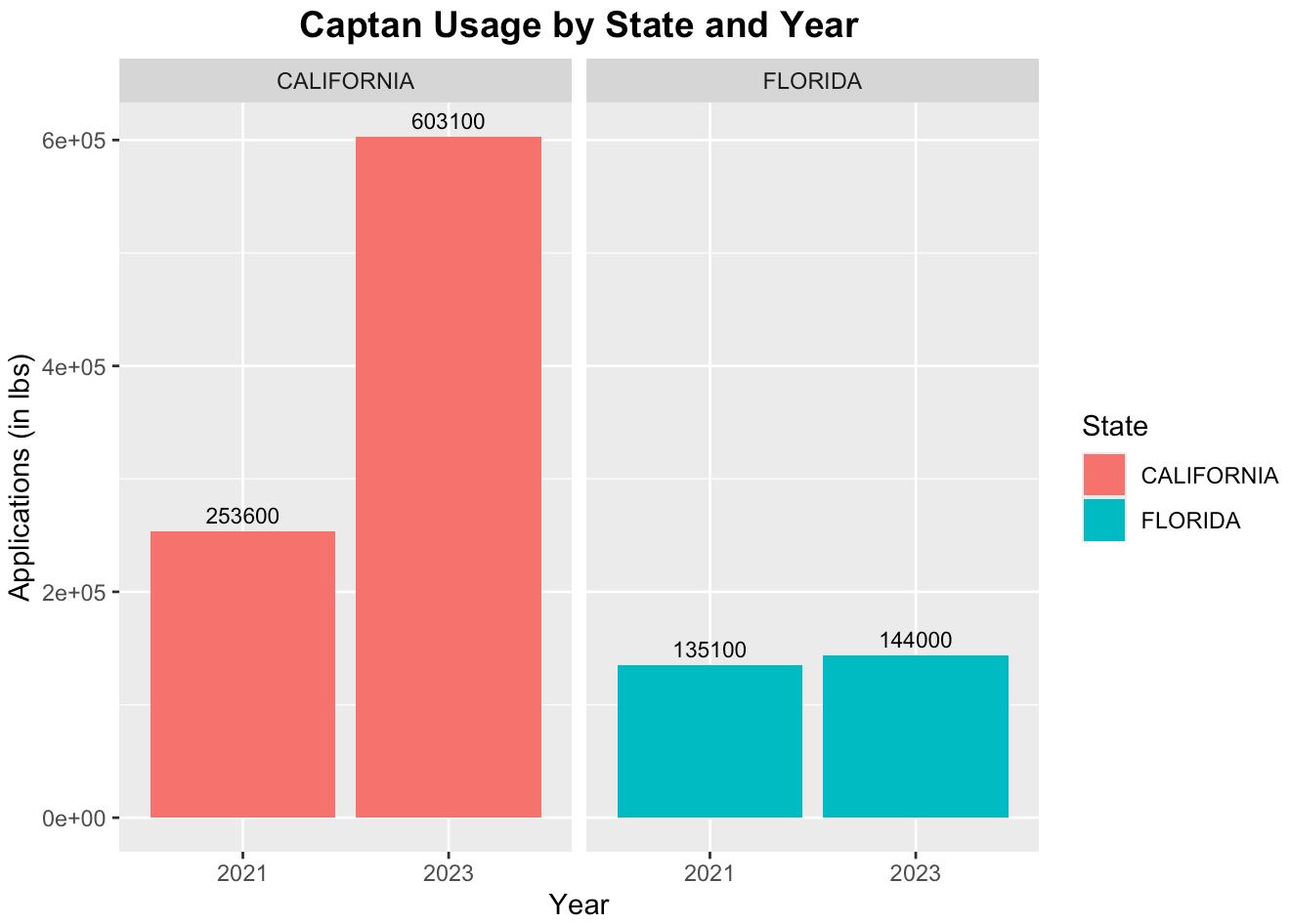

Bottom 5 Chemicals Used in California by Year
2021
2023
Applications (in lbs)
0
100
200
300
400
500
500
1000
1500
2000
200
100
200
500
100
300
1700
2000
1300
1300
ABAMECTIN
BORAX DECAHYDRATE
FENPYROXIMATE
FLUMIOXAZIN
PYRETHRINS
ABAMECTIN
CYANTRANILIPROLE
SPINOSAD
TETRACONAZOLE
THIAMETHOXAM
Chemical
Year
2021
2023



Based on the above graphs, shared chemicals between the top 5 chemicals in each state include captan and nitrogen, and the bottom 5 chemicals in each state include abamectin and thiamethoxam. Using this information, I will compare the usage of captan, nitrogen, and abamectin between California and Florida.


Nitrogen Usage by State in 2023
4e+05
3e+05
2e+05
1e+05
0e+00
Applications (in lbs)
393000
283000
CALIFORNIA
FLORIDA
State
State
CALIFORNIA
FLORIDA


Abamectin Usage by State and Year
CALIFORNIA
FLORIDA
300
200
100
0
Applications (in lbs)
200
300
NA
100
2021
2023
2021
2023
Year
State
CALIFORNIA
FLORIDA



Nitrogen was only used in 2023 in both California and Florida. There is a 110,000 lb difference in application between states. I wonder if this is due to California producing more strawberries than Florida. Also, why is there no nitrogen usage in 2021 for both states?

In regard to abamectin, both California and Florida used more in 2023 than 2021. Although Florida in 2021 doesn't have any data, I will assume that there was little to no usage. If I assume there is no usage, then both California and Florida use 100 lbs more in 2023 than in 2021.

However, the difference in captan usage between states is drastic. There is a 349,500 lb difference between 2021 and 2023 in California and a 8,900 lb difference between 2021 and 2023 in Florida. What happened between 2021 and 2023 in California for the dramatic increase in captan usage? And why is there such a small difference in captan usage from 2021 to 2023 in Florida while there's a huge difference in California?

Operations vs Producers

Comparison of California's Producers and Operations Net Income in 2022 by Area Operated

Area Range (acres)		Producers	Operations	Difference
Low	High	Value ($).x	Value ($).y	Difference in y-x ($)
1	9.9	-46056000	-33586000	12470000

Area Range (acres)		Producers	Operations	Difference
Low	High	Value ($).x	Value ($).y	Difference in y-x ($)
10	49.9	357332000	376405000	19073000
50	69.9	134158000	150061000	15903000
70	99.9	236584000	253660000	17076000
100	139.0	191307000	223317000	32010000
140	179.0	253627000	285520000	31893000
180	219.0	282964000	282080000	-884000
220	259.0	126638000	129012000	2374000
260	499.0	1321748000	1355260000	33512000
500	999.0	1847524000	2007260000	159736000
1000	1999.0	1892878000	1927889000	35011000
2000	NA	5075486000	5147340000	71854000

Comparison of Florida's Producers and Operations Net Income in 2022 by Area Operated

Area Range (acres)		Producers	Operations	Difference
Low	High	Value ($).x	Value ($).y	Difference in y-x ($)
1	9.9	4045000	3887000	-158000
10	49.9	192224000	232432000	40208000
50	69.9	108651000	115499000	6848000
70	99.9	63683000	75146000	11463000
100	139.0	66343000	85037000	18694000
140	179.0	44154000	47097000	2943000
180	219.0	82161000	84339000	2178000
220	259.0	79781000	87509000	7728000
260	499.0	123693000	131097000	7404000
500	999.0	189513000	198473000	8960000
1000	1999.0	306493000	309366000	2873000
2000	NA	971452000	984558000	13106000

Based on the tables above, we see that the net income of operations is significantly higher than the net income of producers for both California and Florida. This suggests that larger agricultural businesses (operations) are more profitable due to factors such as better market access. However, producers likely face higher costs, which leads to a lower net income.

We also see that California tends to see a higher profit than Florida in most of the area ranges. This suggests that California grows more strawberries than Florida, which can be further validated by this data set.

Further Analyses

There were many topics explored above in regard to chemical analyses and operations/producers analyses. However, there is always more exploration to be done in a data set as large as this one. I suggest further analyses into the chemicals, specifically why some chemicals are used in one state and not in the other. Additionally, I'm curious to know what happened between 2021 and 2023 for there to be such an increase in chemical usage. On top of that, it would be good to know why there is no chemical information in 2020 and 2022.

Regarding the operations/producers analyses, it would be interesting to learn the specifics behind why there is such a large difference in profit between the states. Is it due to only one reason, such as California producing more strawberries than Florida, or are there other reasons that have yet to be explored or seen in this data set?

There are plenty more questions to be answered from this strawberry data set. I suggest the further exploration of the analyses above, in addition to other questions that an analyst may have.