x

Final Deliverable

Our Team


Jessica Liu


Jenny Chung


Nicholas Huang


Aadil Jamari


Alivia Ding


Khushboo Teotia


Cristina Prieto


Avery Hipolito


Michael Murphy


Sona Wyse


Kelly Hu


Lizbeth Velazquez


Jeffrey Gao


Catherine Wang


Dhruv Syngol

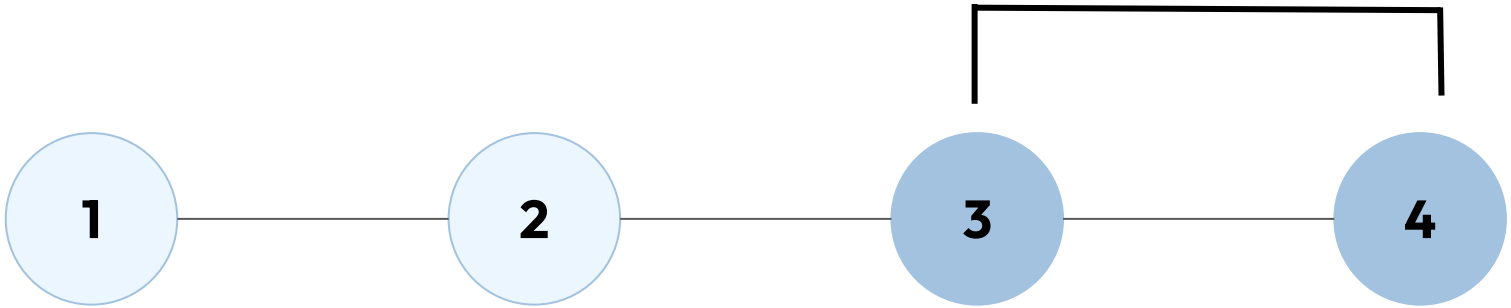

Sanay Bordia

Agenda

1. Introduction
2. Time Series
3. Geospatial Mapping
    a. Overlaying Census Data
    b. Key Trends & Takeaways
4. Questions

Project Timeline

Project Objective: Investigate the temporal and spatial behavior of PFAS chemicals to identify potential relationships between PFAS and other environmental, geological, and geographical factors



Recap

From mid-semester deliverable:

- Time Series Analysis and Forecasting ✅
- Initial Geospatial Mapping and Analysis ✅

After mid-semester deliverable:

- Move forward with Prophet for Time Series
- Implementation of census data
- Drawing cohesive conclusions from time series and spatial analyses

Time Series

Since the Mid-Semester Check In

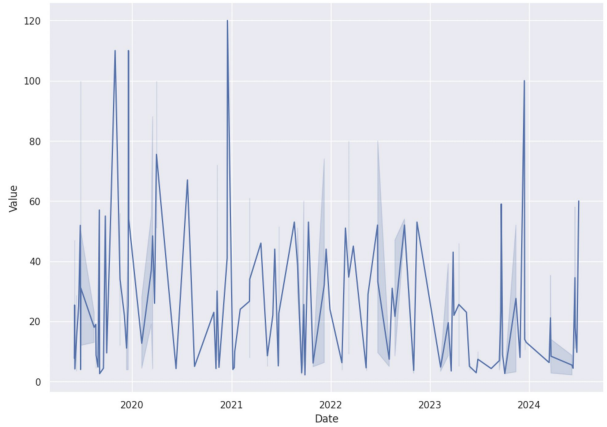
- Continuing to utilize Meta's Prophet Time Series Model to predict future trends of chemical concentrations
- Focused on specific problem sites
  - Cities and counties
  - Water systems
- Troubleshooting accuracy and approach to predicting trends

Areas of Focus

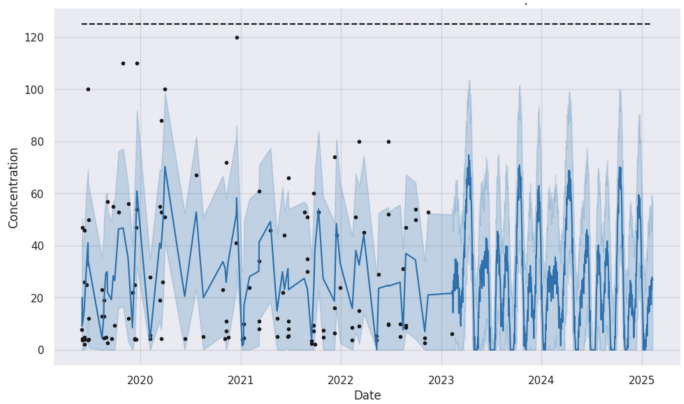
1. San Luis Obispo (SLO county)
    - PFOS
    - PFHxS
2. Red Bluff (Tehama County)
    - PFOA
3. Livermore (Alameda County)
    - PFOS

Results and Troubleshooting

PFOS concentration in SLO over time



Existing trends over time
(with a few outliers removed)



Forecasted concentration
over time

Problem = VERY poor model/prediction accuracy

Issues with Forecasting

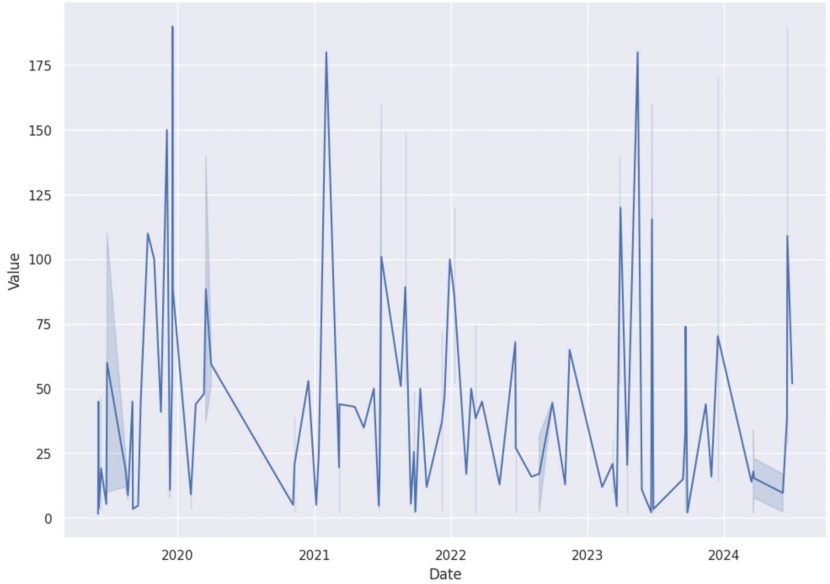
1. Irregular sampling, no continuous sampling
    - Continuity over time is key in building forecasts

2. Too many unpredictable events
    - COVID, wildfires, etc

3. Too little data in these areas
    - Combined with factors (1) and (2), makes building accurate models very difficult

Results cont.

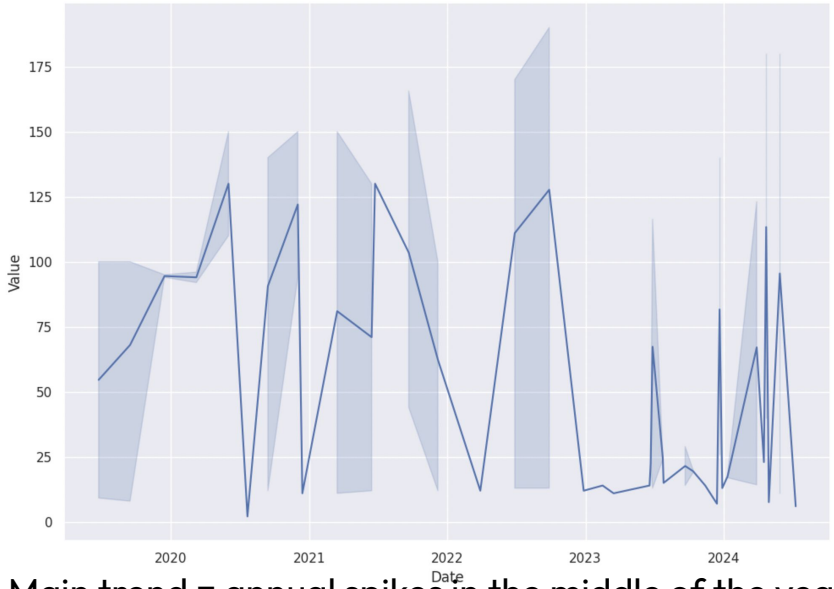
PFHxS concentration in SLO over time, from existing data



Main trend = annual spikes at the end of the year

Results cont.

PFOA concentration in Red Bluff (Tehama County) over time, from existing data

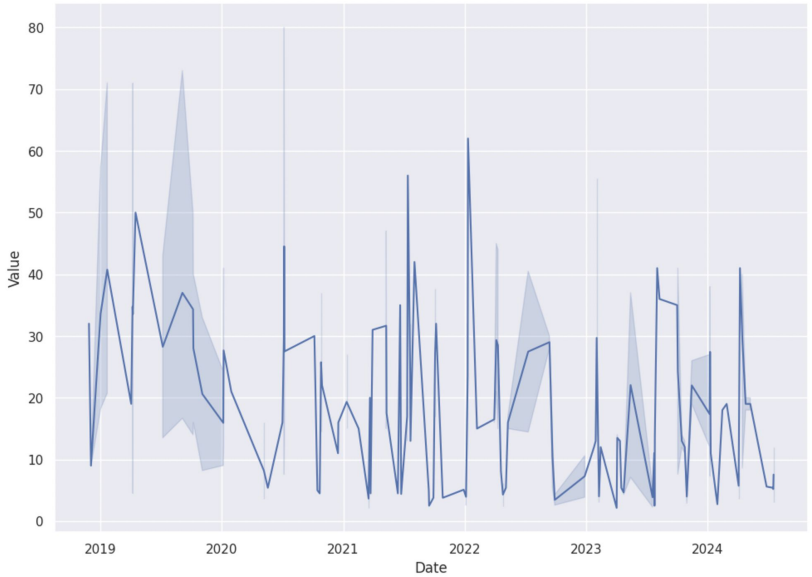


Main trend = annual spikes in the middle of the year

- Trend upwards at the beginning of the year, curving downward at the end of the year

Results cont.

PFOS concentration in Livermore (Alameda County) over time, from existing data



No seasonality or big trends, just many irregular spikes in concentration levels

In general, the top 5 most prevalent chemicals seem to have annual spikes in many identified regions

Geospatial Mapping

Overlaying Census Data

Overlaying Census Data

- Datasets:
  - CalEnviroScreen 4.0
  - American Community Survey

- Aggregated drinking wells dataset and census datasets on longitude and latitude to match county and combine with census tract data

OEHHA
California Office of Environmental Health Hazard Assessment

Overlaying Census Data

- Notable variables:
  - CalEnviroScreen 4.0
    - Poverty (2x below federal poverty level) percentile
    - Cardiovascular disease percentile
    - Racial demographic distribution
  - ACS
    - Income and Benefits (2022-inflation adjusted)
    - Percentage of families and people whose income in the past 12 months is below the poverty level
    - Unemployment rate

Key Trends and Takeaways

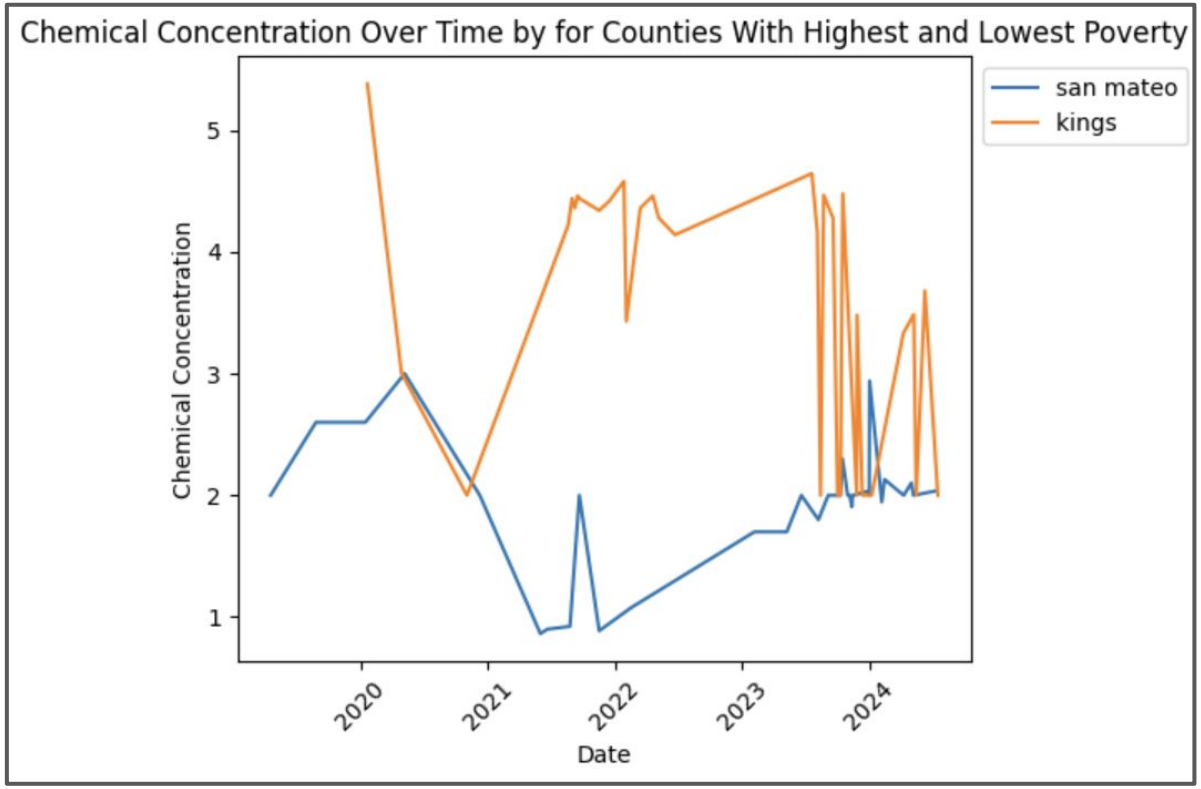
General Analysis

	Poverty	Education	Unemployment
Highest	Kings	Madera	Imperial
Lowest	San Mateo	Plumas	Mono

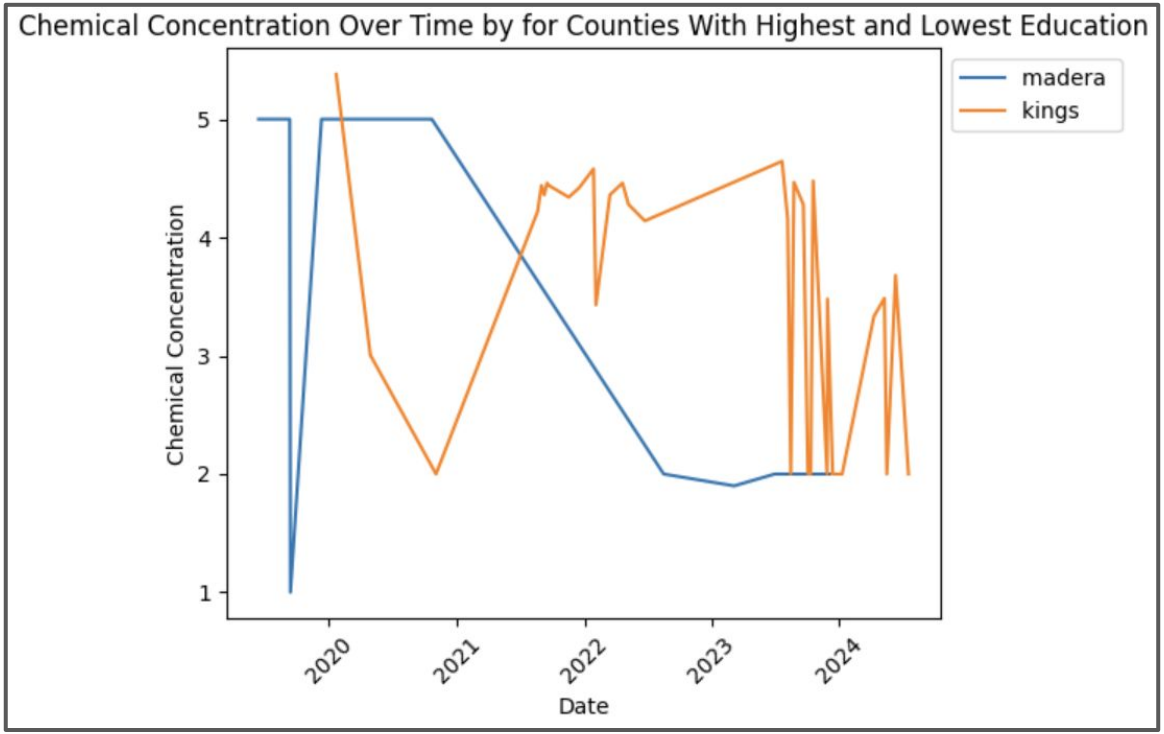
Goal:

- Look at PFAS concentration distribution for these areas.

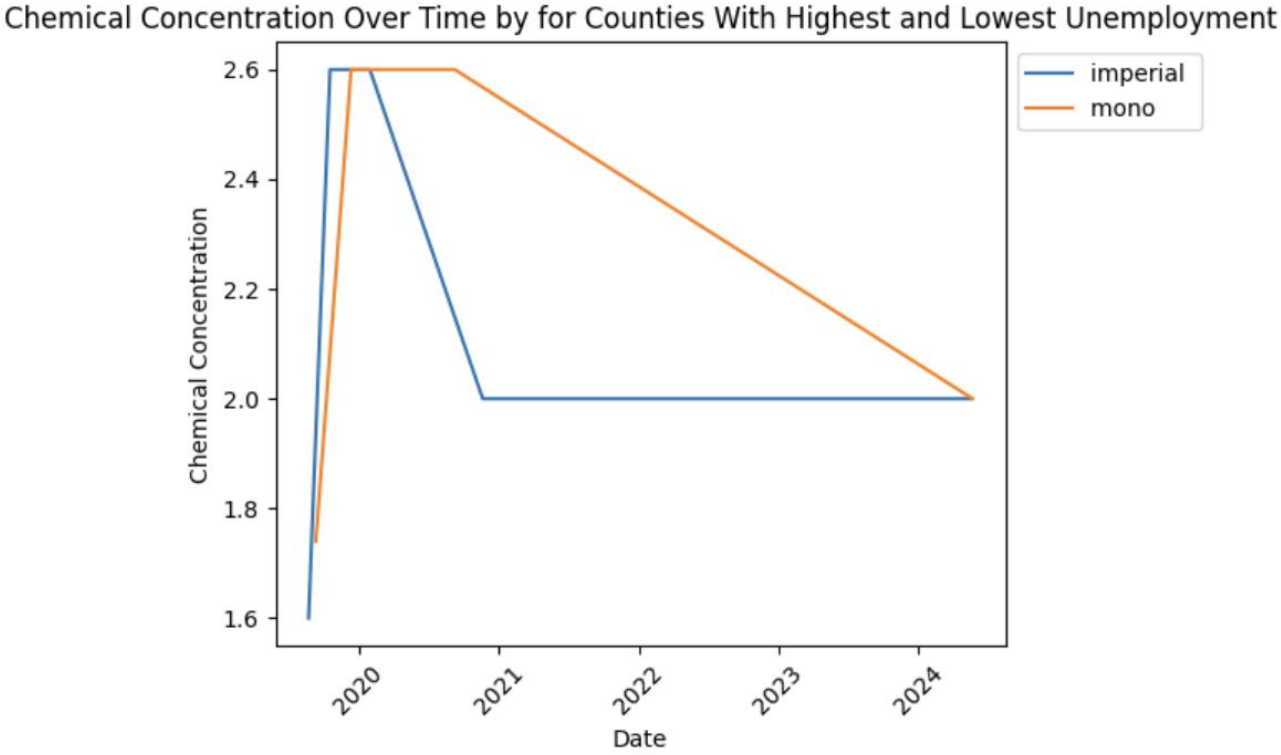
Poverty


Chemical Concentration Over Time by for Counties With Highest and Lowest Poverty
san mateo
kings
Chemical Concentration
1
2
3
4
5
2020
2021
2022
2023
2024
Date

Education


Chemical Concentration Over Time by for Counties With Highest and Lowest Education
madera
kings
Chemical Concentration
1
2
3
4
5
2020
2021
2022
2023
2024
Date

Unemployment



Multiple Linear Regression - Fresno

	Poverty	Education	Unemployment
Coefficient	-0.017498	0.017417	0.024327

Mean Squared Error	3.6316
R-Squared	-0.0246

Goal:

- Determine what socioeconomic factors (poverty, unemployment, and education), had the most impact on PFAS values in Fresno. Notice MSE and R-Squared, maybe these factors aren't as impactful as we think.

Multiple Linear Regression - Overall

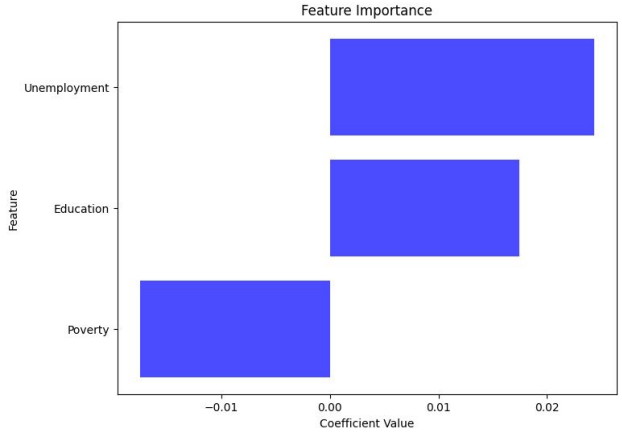
	Poverty	Education	Unemployment
Coefficient	-0.003745	-0.024436	-0.027

Mean Squared Error	146.606
R-Squared	0.00133

Goal:

- Determine what socioeconomic factors (poverty, unemployment, and education), had the most impact on PFAS values in all counties.

Multiple Linear Regression

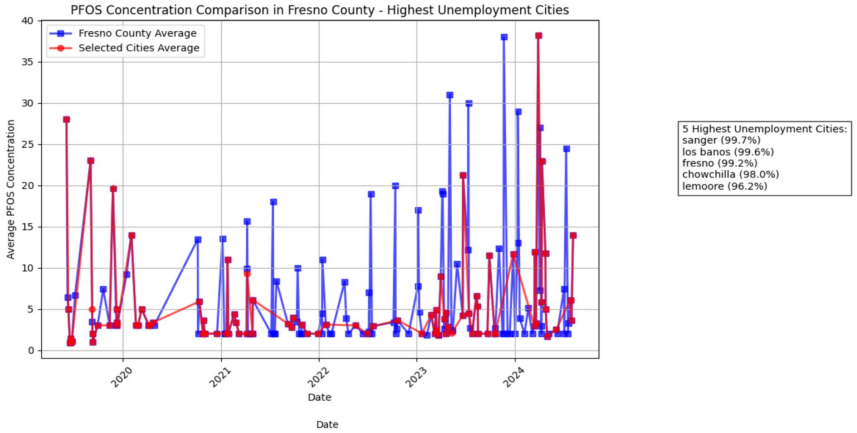


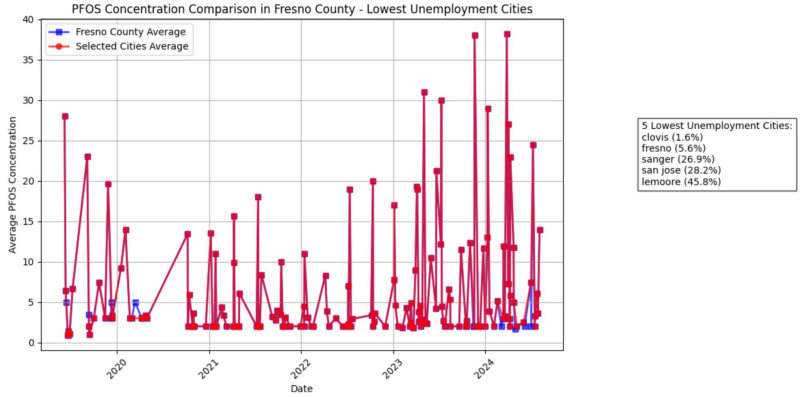
Fresno

- In Fresno and all counties, poverty has a negative effect on PFAS, while unemployment and education seems to have a positive correlation.
- Still skeptical about true effect of these factors due to the small correlation coefficients.

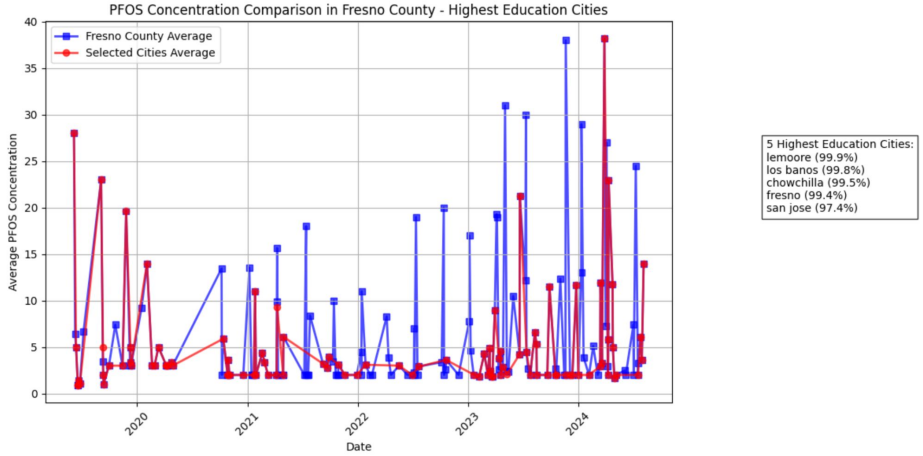
Looking into PFOS in Fresno County

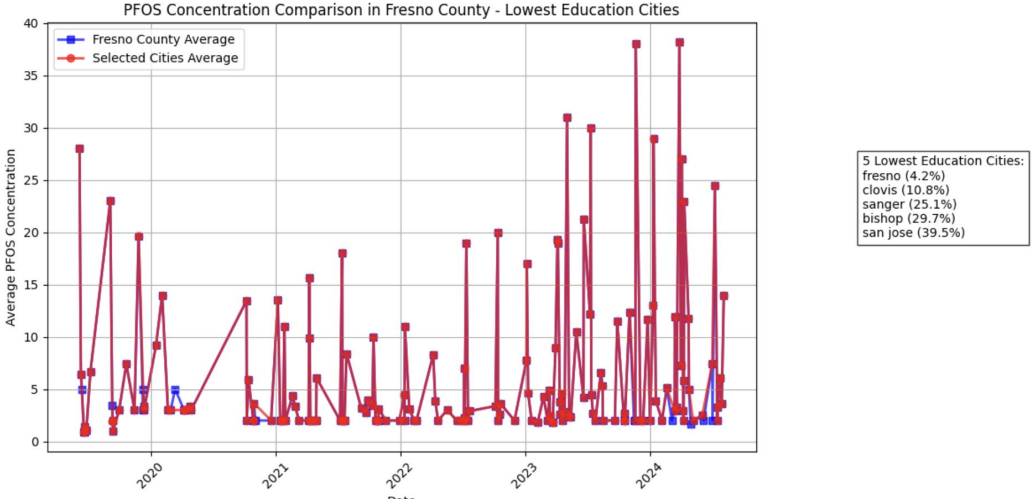
	Unemployment
Highest	Sanger, Los Banos, Fresno, Chowchilla, Lemoore
Lowest	Clovis, Fresno, Sanger, San Jose, Lemoore



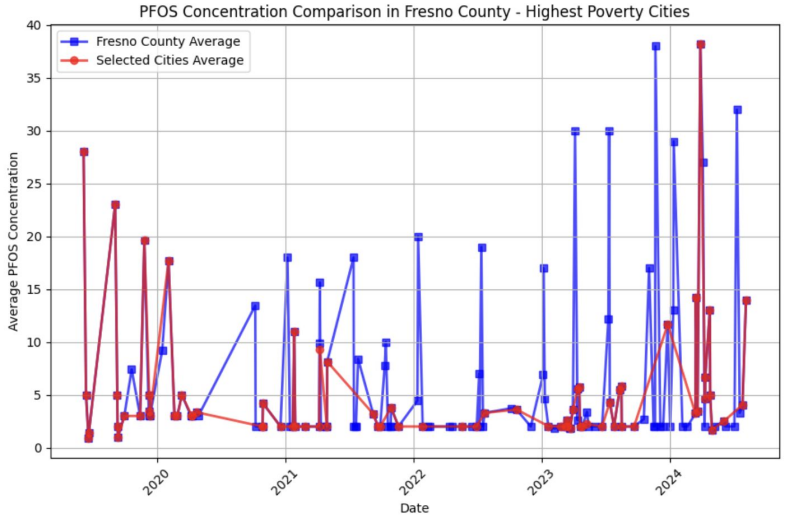


	Education
Highest	Lemoore, Los Banos, Chowchilla, Fresno, San Jose
Lowest	Fresno, Clovis, Sanger, Bishop, San Jose

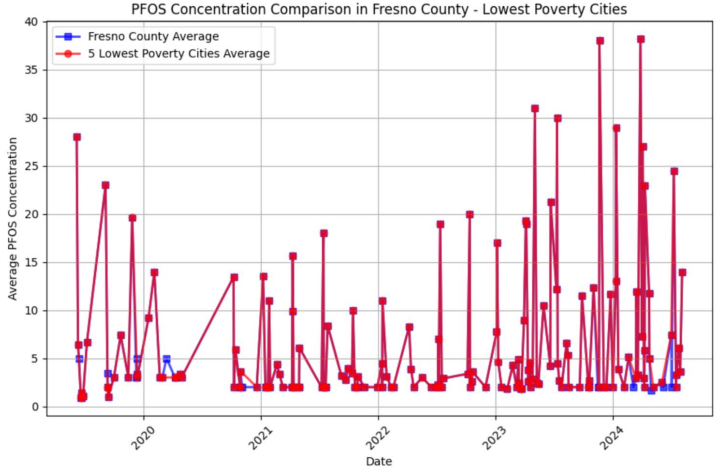




	Poverty
Highest	Sanger, Los Banos, Fresno, Chowchilla, Lemoore
Lowest	Clovis, Fresno, Sanger, Bishop, San Jose



5 Highest Poverty Cities:
sanger (99.7%)
los banos (99.6%)
fresno (99.2%)
chowchilla (98.0%)
lemoore (96.2%)



5 Lowest Poverty Cities:
clovis (0.6%)
fresno (6.1%)
sanger (17.5%)
bishop (22.9%)
san jose (50.9%)

Socioeconomic Trends vs. Concentration

Alameda County

(6.25% people below poverty level; $162288.25 average income)

Cities	PFOS	PFOA	PFNA	PFHXS
Fremont +25.75% income -48.42% poverty	-37.51%	-13.4%	-23.47%	-47.79%
Livermore +20.53% income -73.14% poverty	+34.44%	+8.84%	N/A	+41.31%
Pleasanton +39.97% income -63.4% poverty	+51.02%	+7.67%	+2.93%	+65.08%

Summary/Trends : Alameda County

- Analyzed Fremont, Livermore, and Pleasanton in Alameda County.
- Fremont shows a strong correlation between **rising income, reduced poverty**, and **lower levels of environmental pollutants.**
- Livermore and Pleasanton, cities with **higher average income**, actually had **worse (more than) average concentrations of chemicals** compared to county average.

Socioeconomic Trends vs. Concentration

San Diego County

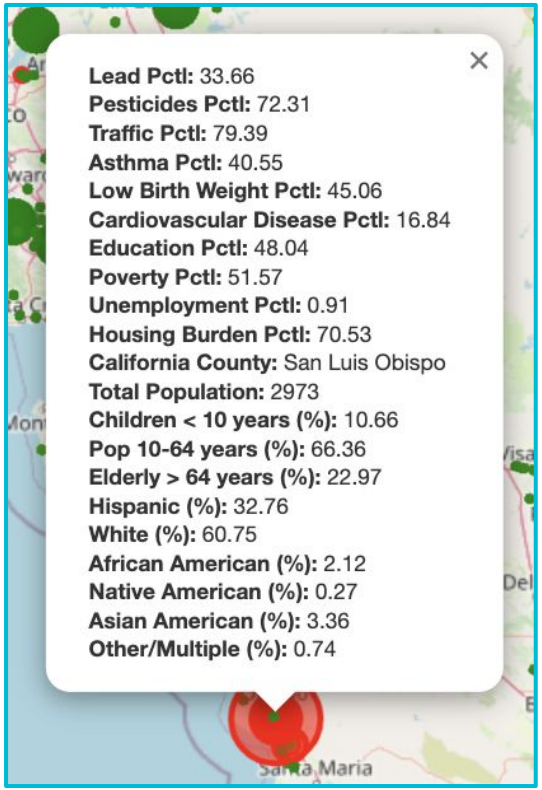
(7.06% people below poverty level, $130262.69 average income)

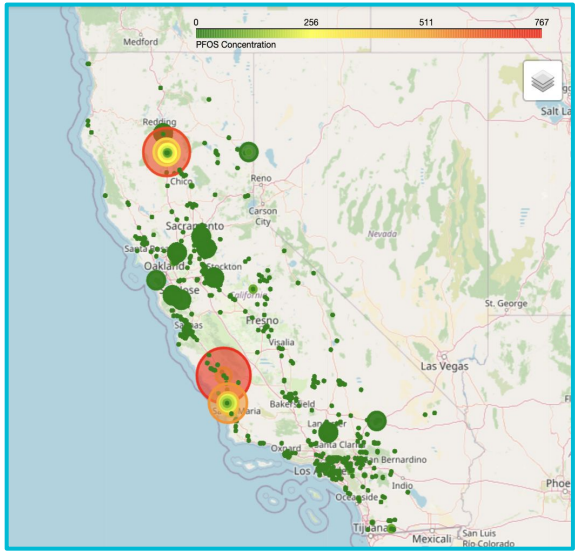
Cities	PFOS	PFOA	PFNA	PFHXS
San Diego +21.99% income, +1.35% poverty	+72.1%	+131.8%	+88.9%	+131.1%
Chula Vista -21.65% income, +0.32% poverty	-26.1%	-59.2%	-41.9%	-60.8%
Oceanside -41% income, +8.93% poverty	-46.0%	-72.6%	-47.0%	-70.3%

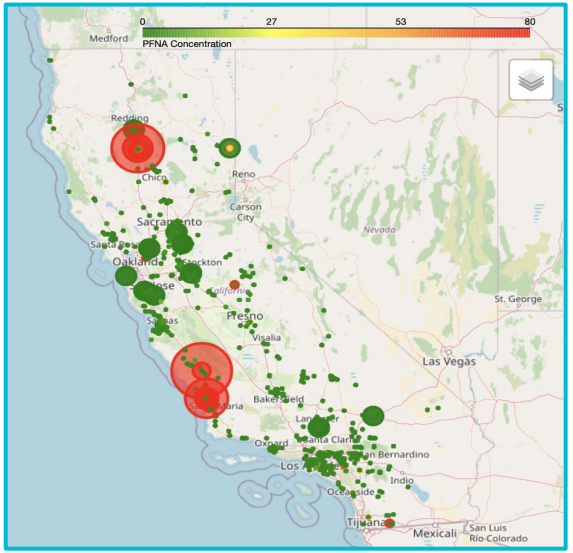
Summary/Trends : San Diego County

- Analyzed 3 specific cities in San Diego County
- We noticed that San Diego, which dominates most of the census tracts in the county and has **higher levels of average income**, actually had **worse (more) average concentrations of chemicals** compared to the county average
- Chula Vista and Oceanside, cities with **lower average income**, actually had **better (less) average concentrations of chemicals** compared to county average
- Findings suggest that **PFAS chemicals are linked to higher income** in San Diego County specifically
- “Income in the general population is positively correlated with PFAS concentrations” - Nature 2024

Overview of Concentration Maps



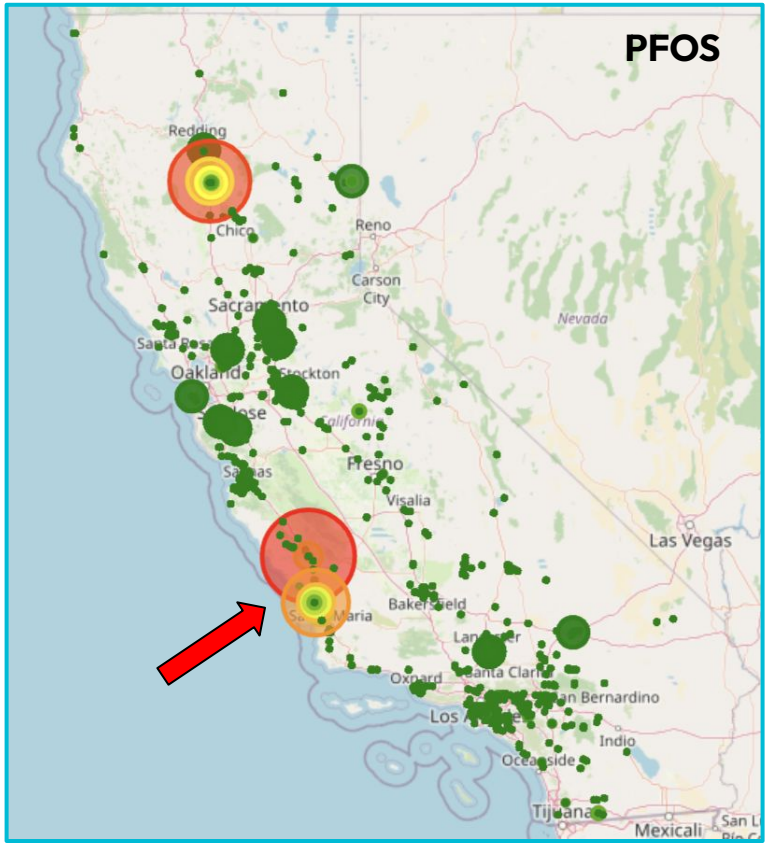




Clicking on a point reveals detailed information, including:

- California County, Environmental Indicators, Health Metrics, Socioeconomic Data, Racial Demographics

Trends + Insights : Demographics



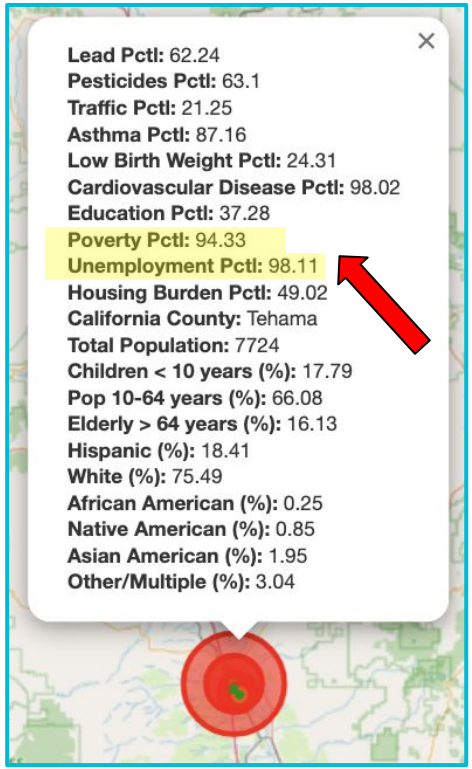
What Stuck Out?

- **San Luis Obispo**: Predominantly **White** population.
- **However…Hispanic communities** in SLO face **higher exposure** to PFAS contamination

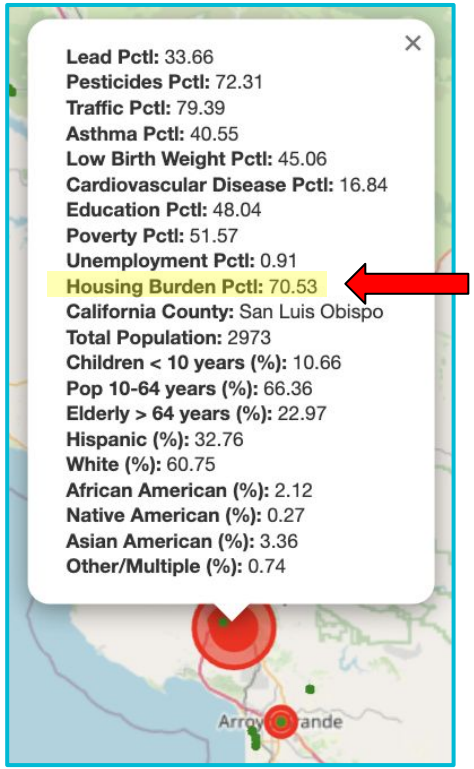
Hispanic (%): 40.8
White (%): 54.52
African American (%): 0
Native American (%): 0.54
Asian American (%): 1.27
Other/Multiple (%): 2.86

Hispanic (%): 20.96
White (%): 70.82
African American (%): 0.15
Native American (%): 1.05
Asian American (%): 2.1
Other/Multiple (%): 4.93

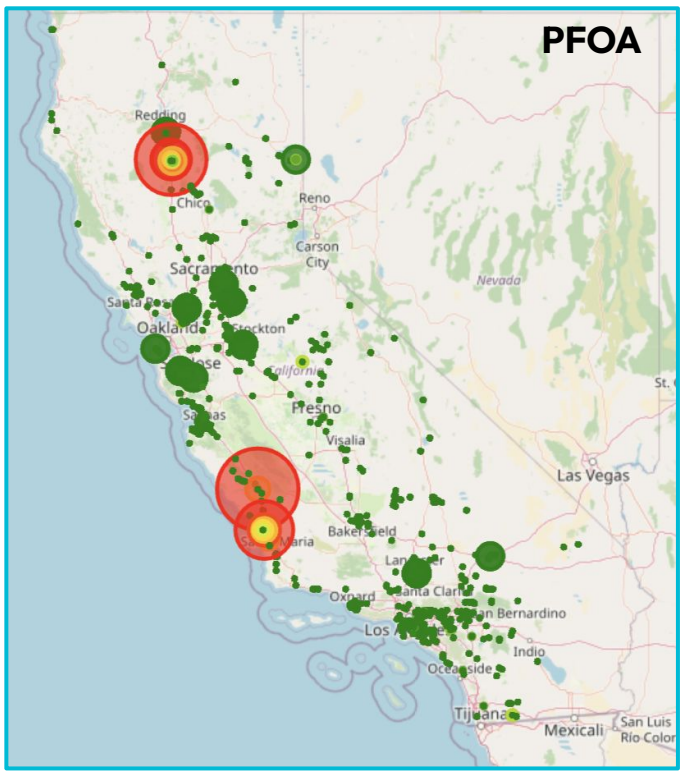
Trends + Insights : Socioeconomic



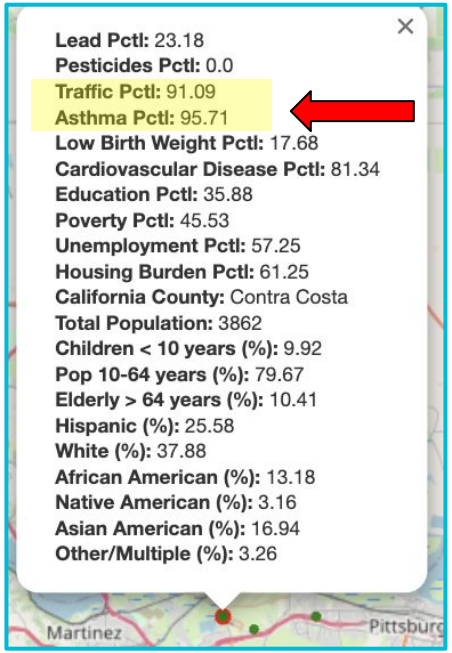
Tehama



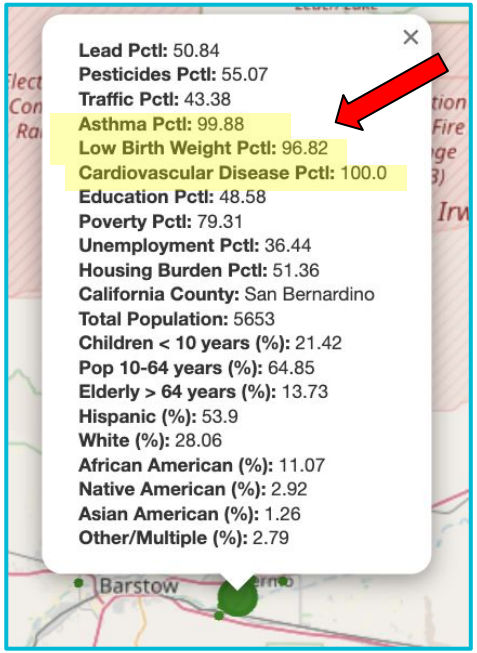
San Luis Obispo



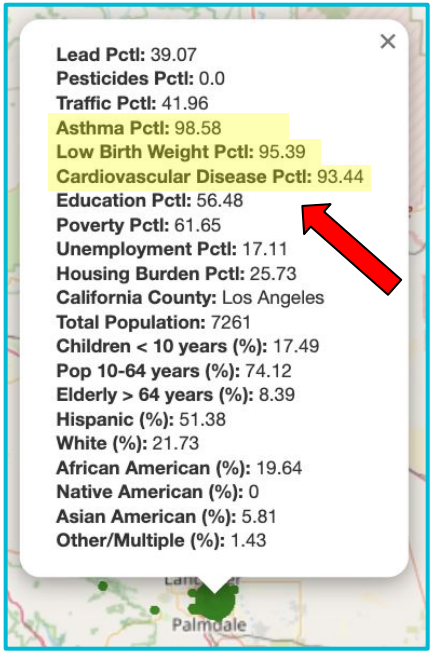
Trends + Insights : Health



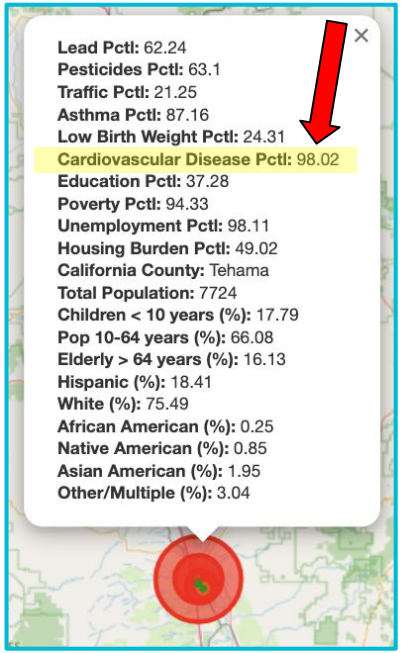
Contra County



San Bernardino

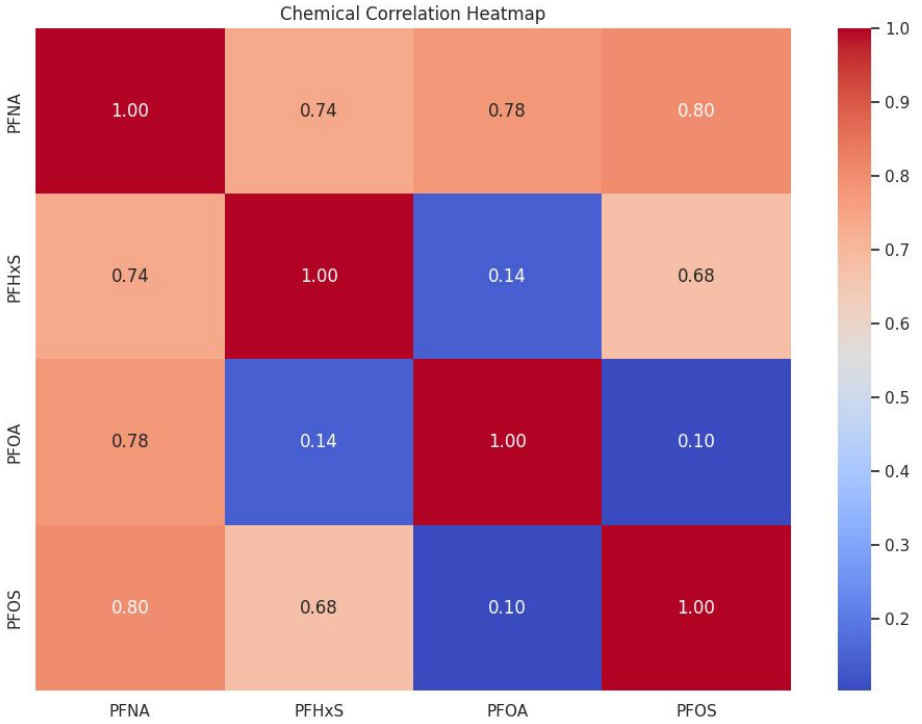


Los Angeles



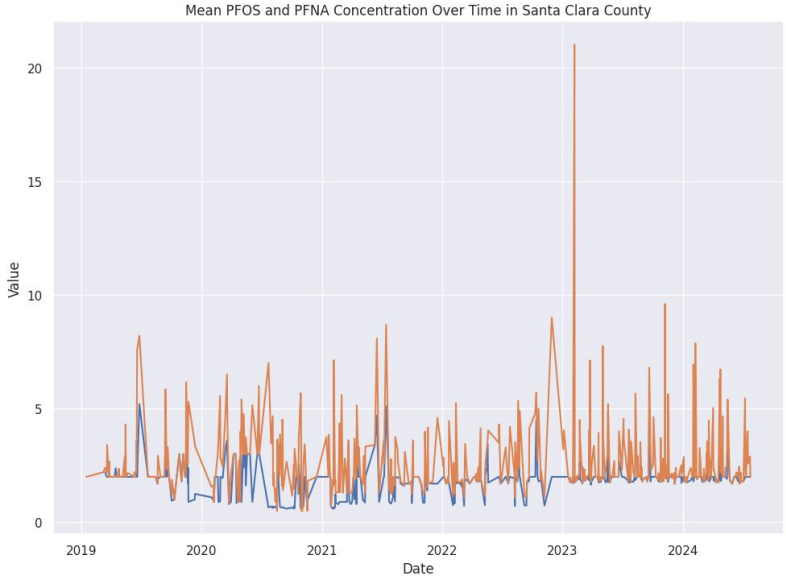
Tehama

Trends + Insights : Chemical Correlation

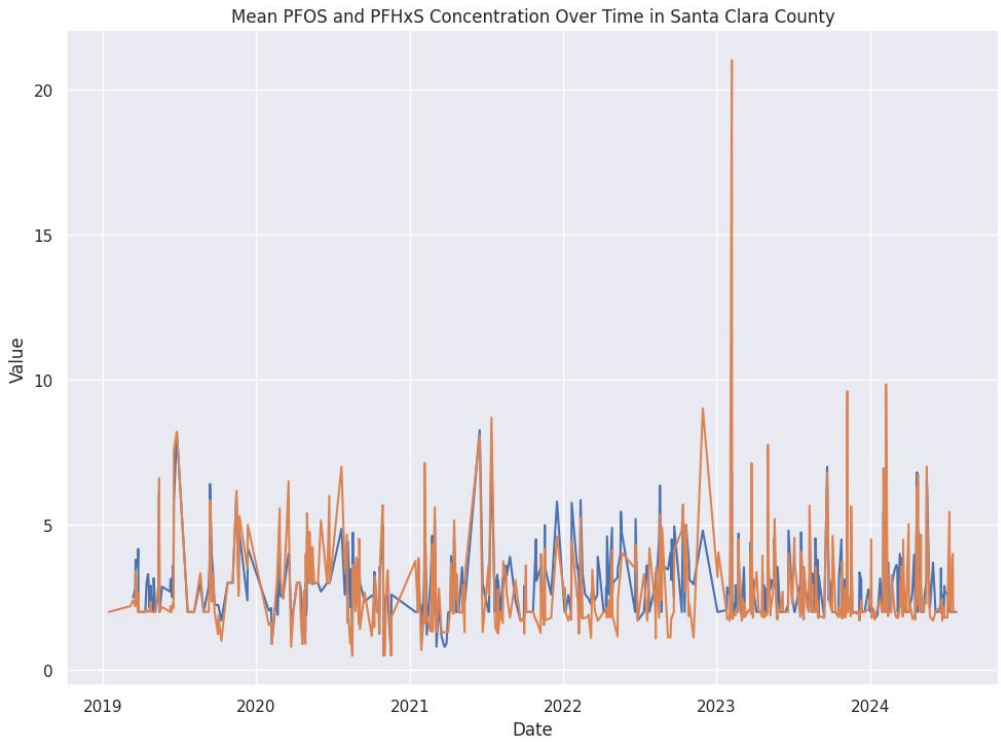


Insights

- **PFNA** has **high correlation** between all other chemicals

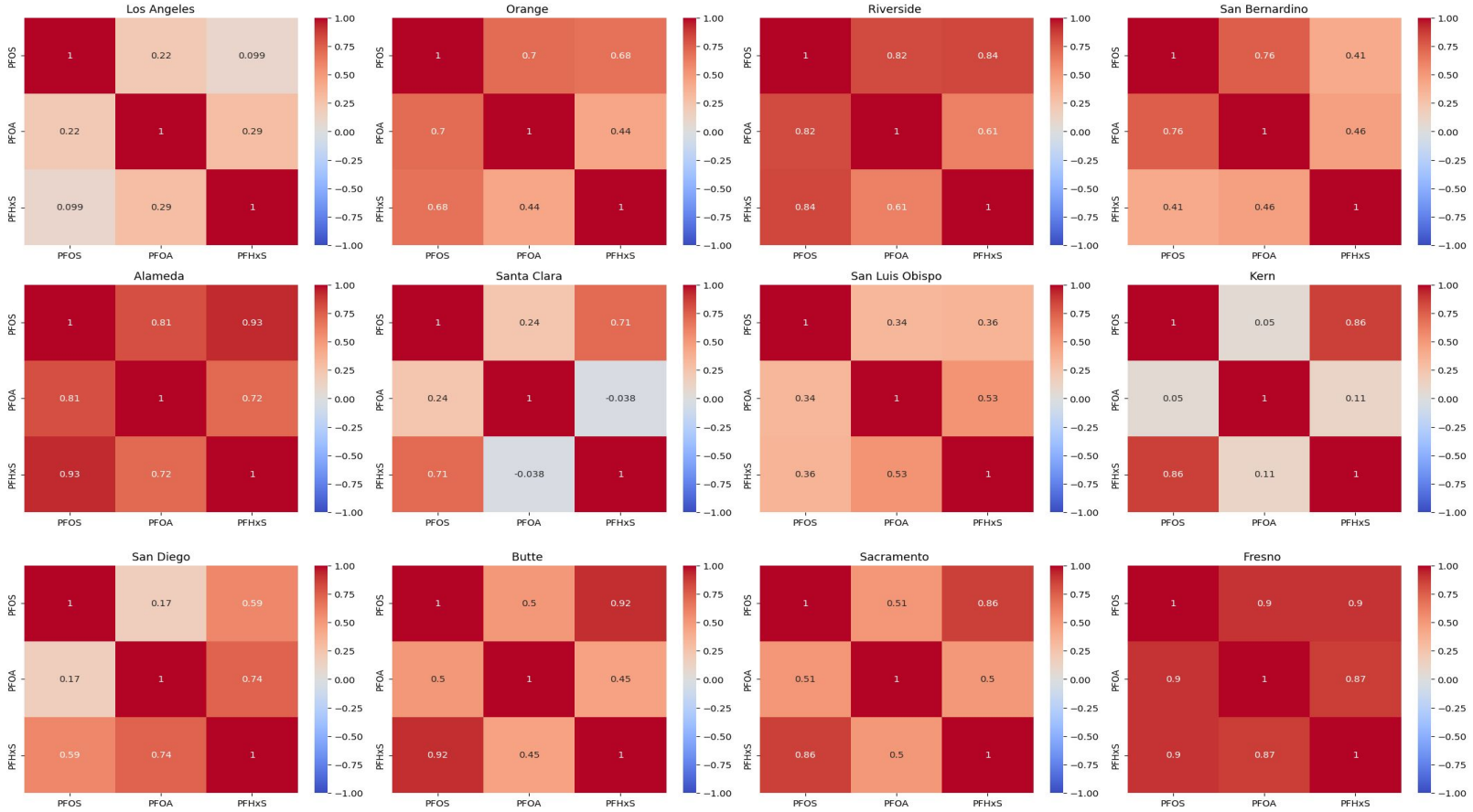


Trends + Insights : Chemical Correlation



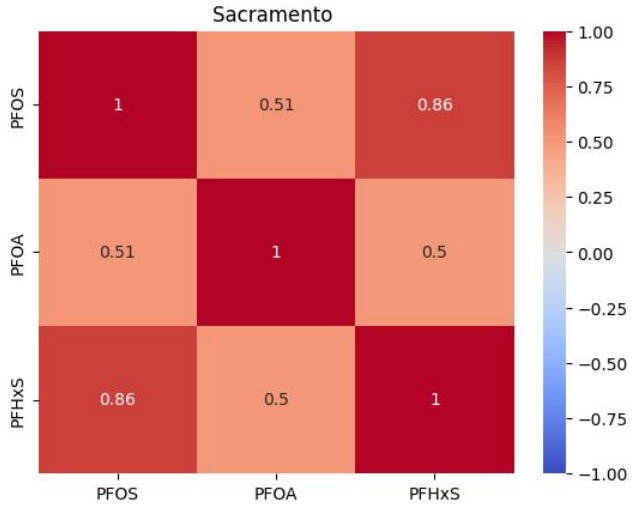
- San Jose is shown to have the highest concentration of PFOS in Santa Clara County
- PFOS and PFHxS shown to have highest correlation in San Jose with correlation coefficient of 0.79

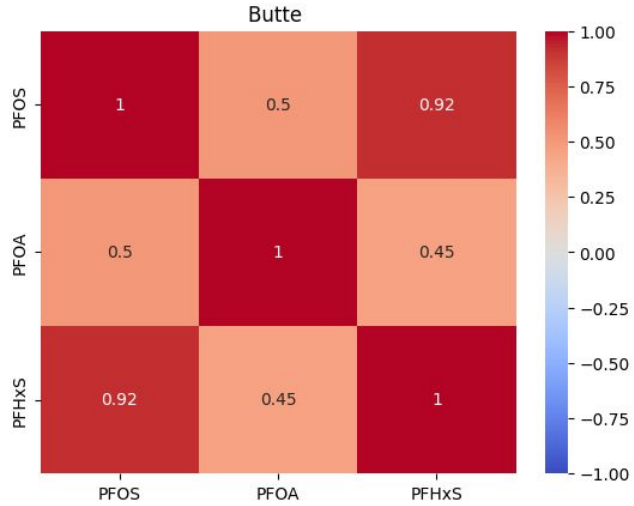
Chemical Correlation by County



Chemical Correlation by County

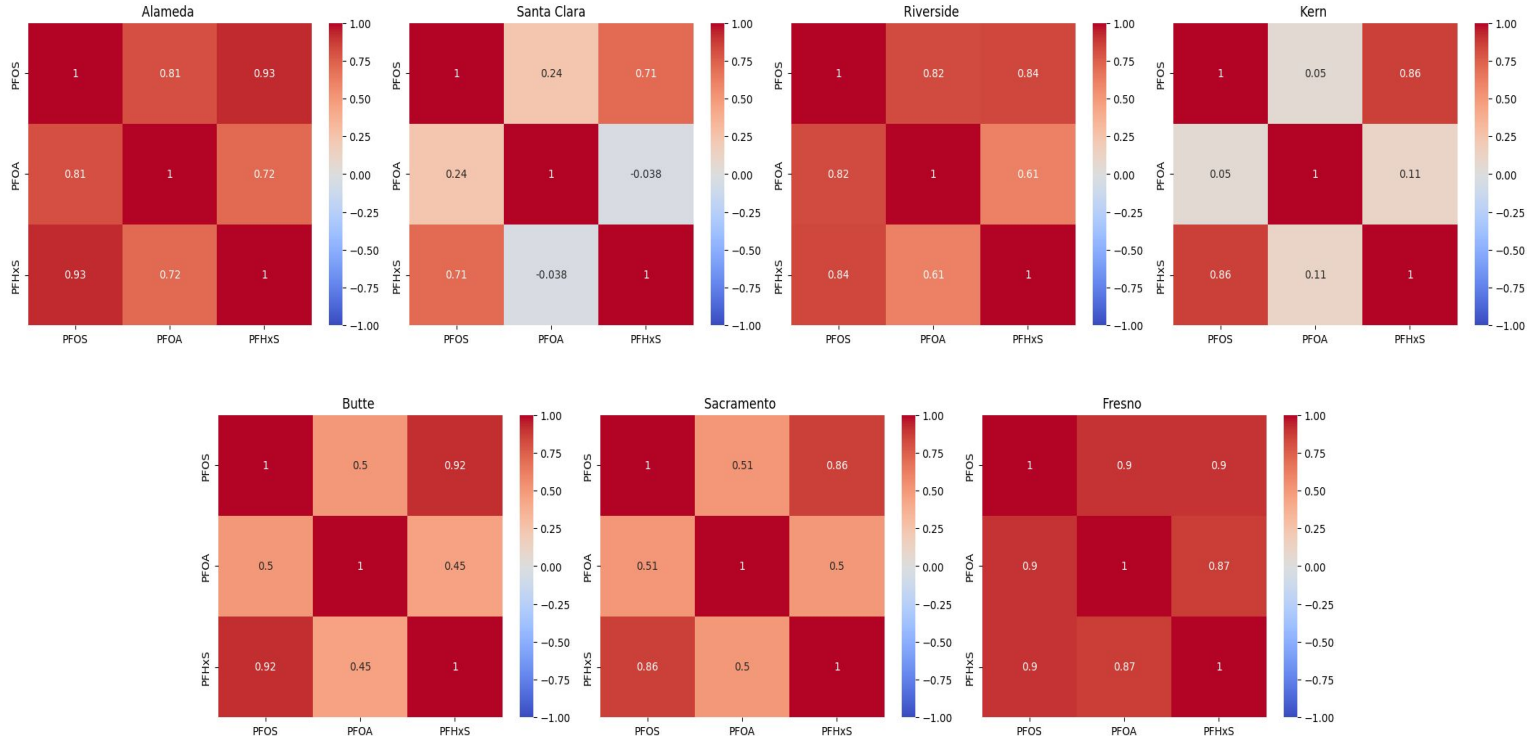
- Sacramento and Butte County show almost identical chemical correlations
- Most likely due to their close geographical proximity





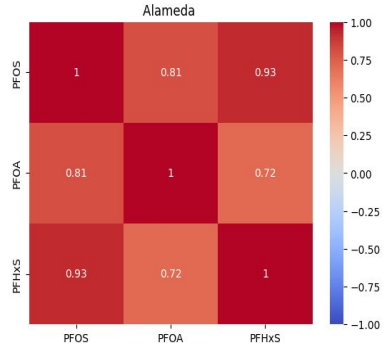
Chemical Correlation by County

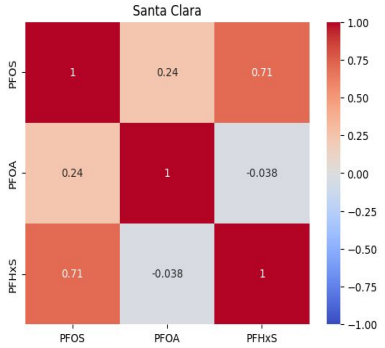
7 out of 12 most sampled counties show a strong correlation (r > 0.7) between PFOS and PFHxS

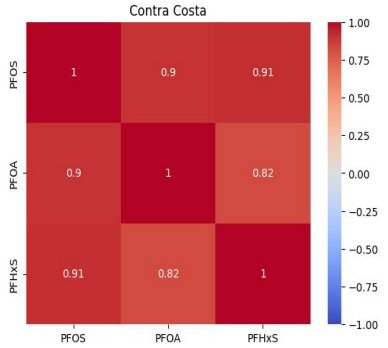


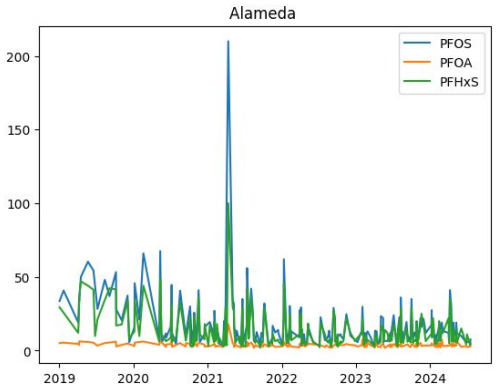
Chemical Correlation in the Bay Area

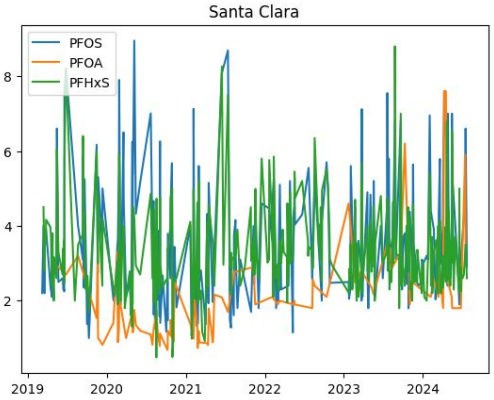
The 3 most populous Bay Area counties all show a strong correlation between PFOS and PFHxS

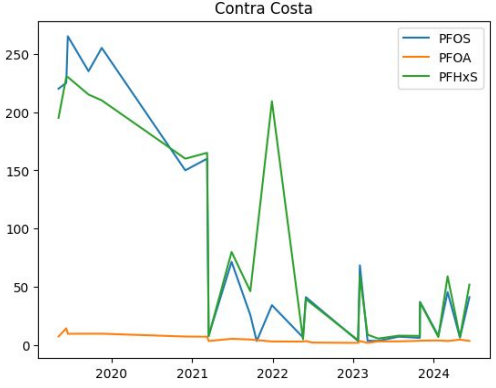










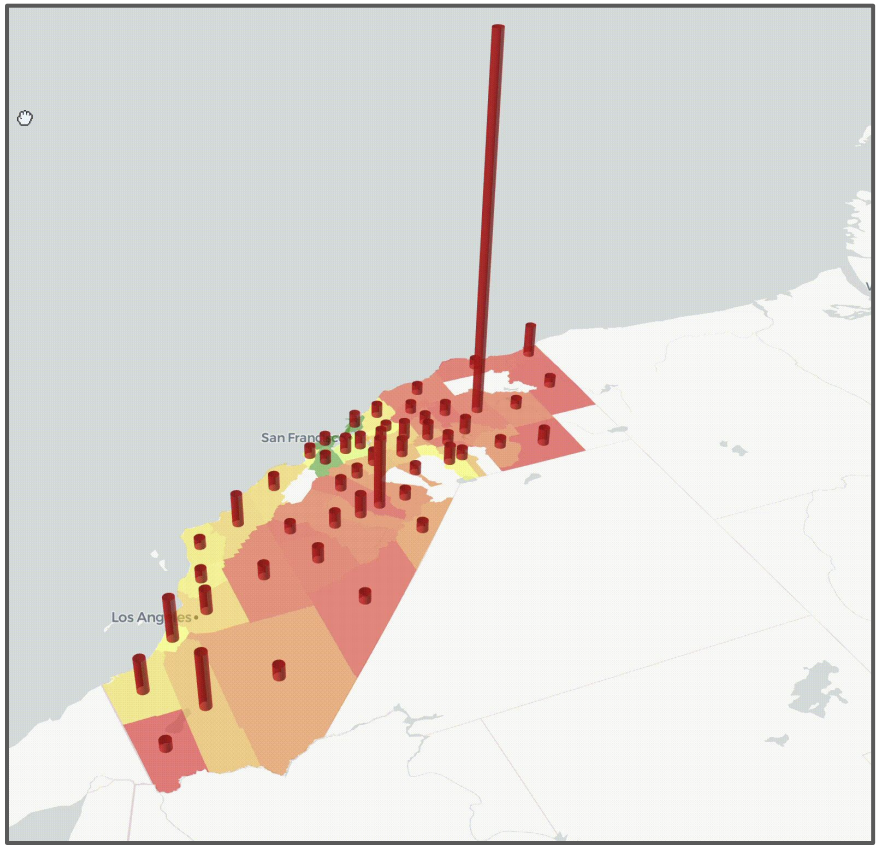


PFOS and PFHxS Similarity

- Sulfonic Acids (contain the functional group -SO3H)
- Bioaccumulative
- Common causes:
  - Firefighting foams
  - Industrial waste
  - Stain repellents

The similar chemical properties and sources for these PFAS likely contribute to their high correlation in California counties.

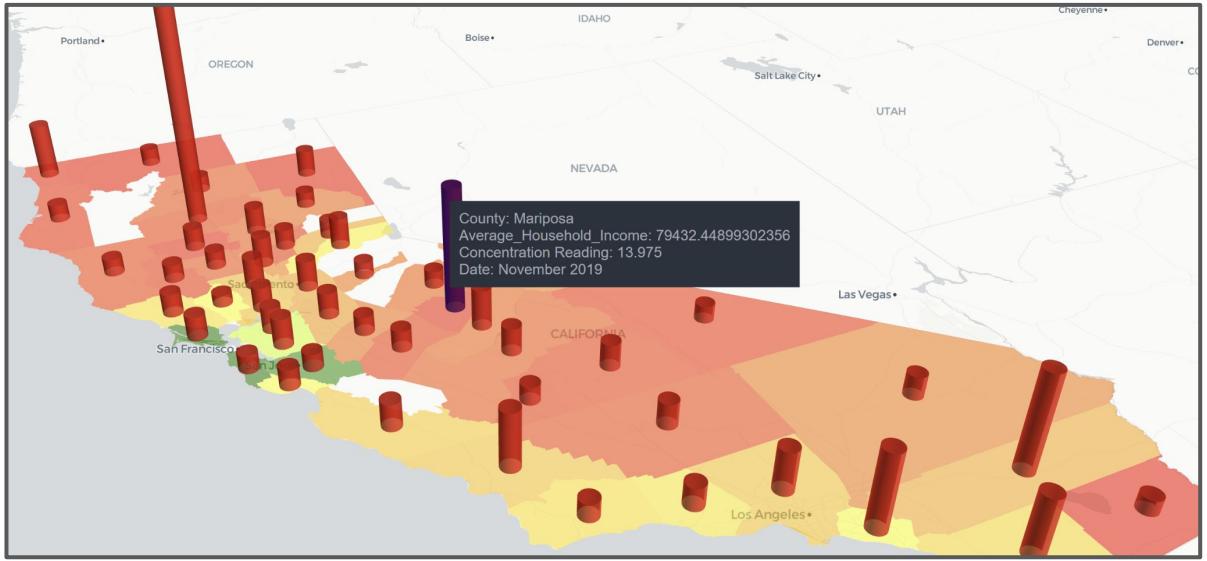
3D Geospatial Visualization of PFAS vs Income



- Boundaries in the map represent **counties**, and are shaded based on **mean household income.**
- Mean household income is represented by a color spectrum from **green** to **red.**
- **Green** represents **higher mean household income** (relative to other counties)**; red** represents **lower.**
- The **bars** represent the concentration of a specified PFAS Chemical, for a specified month.
- Pipeline we created can be **generalized.**

3D Geospatial Visualization of PFAS vs Income

- Hovering over a county displays its census and PFAS concentration data.
- Counties of concern are those with **darker shades of red** and **taller bars**. This indicates that the mean household income is low, and that the water supply is more contaminated.
- **Note**: No PFAS water well data was available for Trinity, Sierra, El Dorado, Alpine, Calaveras and San Benito counties (colored white).



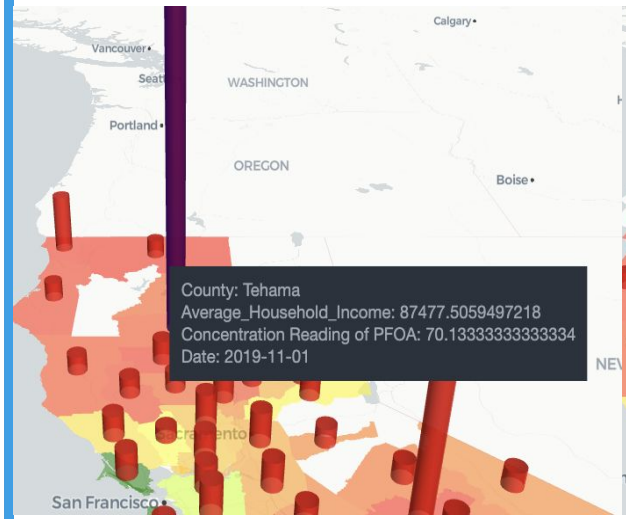
Why these dates?

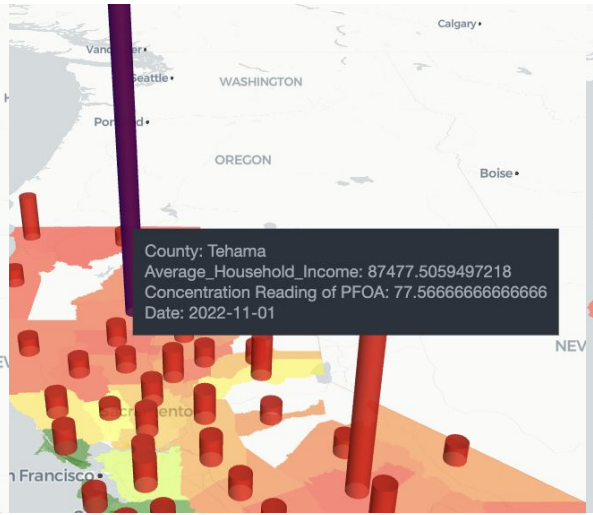
We analyzed the changes over time in PFOA, PFOS, and PFHxS for the following 6 dates:

1. **11 / 01 / 2019 :** In October 2019, an order was issued that required the determination of 25 PFAS chemicals at Chrome Plating Facilities in CA
2. **03 / 01 / 2020:** CA is about to enter lockdown for the COVID-19 pandemic
3. **01 / 01 / 2021:** Enactment of National Defense Authorization Act
4. **12 / 01 / 2022:** EOY 2022
5. **12 / 01 / 2023:** EOY 2023
6. **10 / 01 / 2024:** In April 2024, PFAS chemicals were designated under CERCLA

Visual Representation of PFAS over time

PFOA in Tehama County:

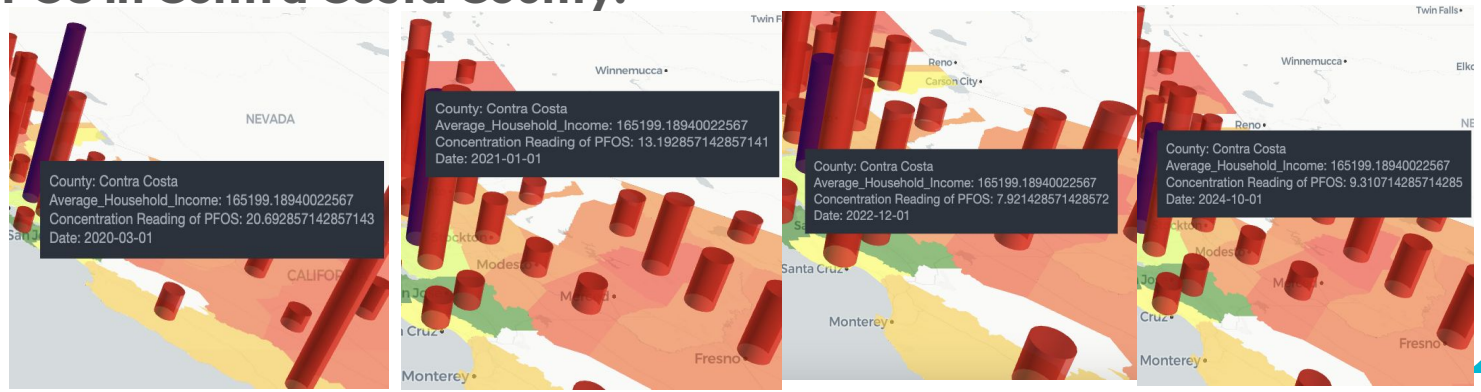




Visual Representation of PFAS over time

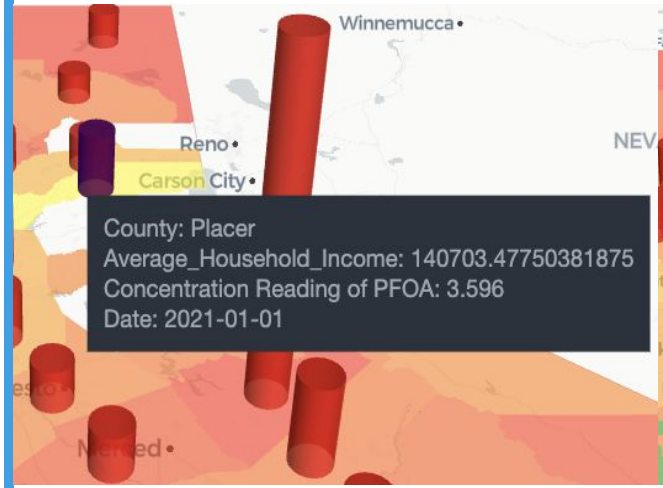
PFHxS in Contra Costa County:

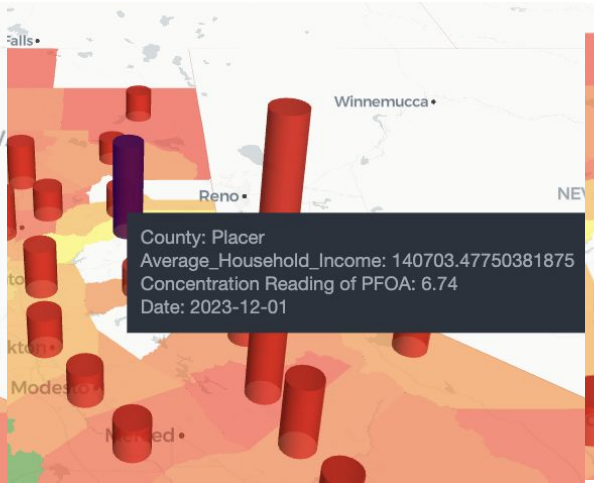
PFOS in Contra Costa County:

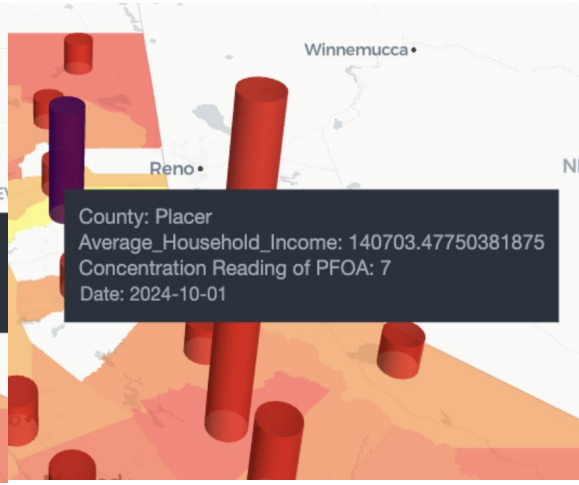


Visual Representation of PFAS over time

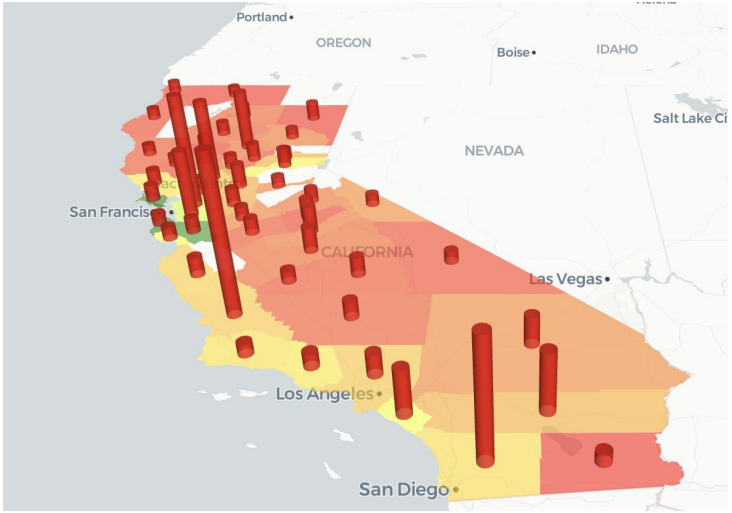
PFOA in Placer County:

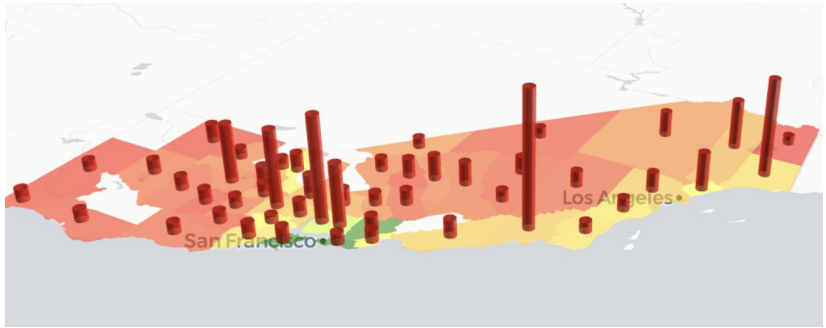






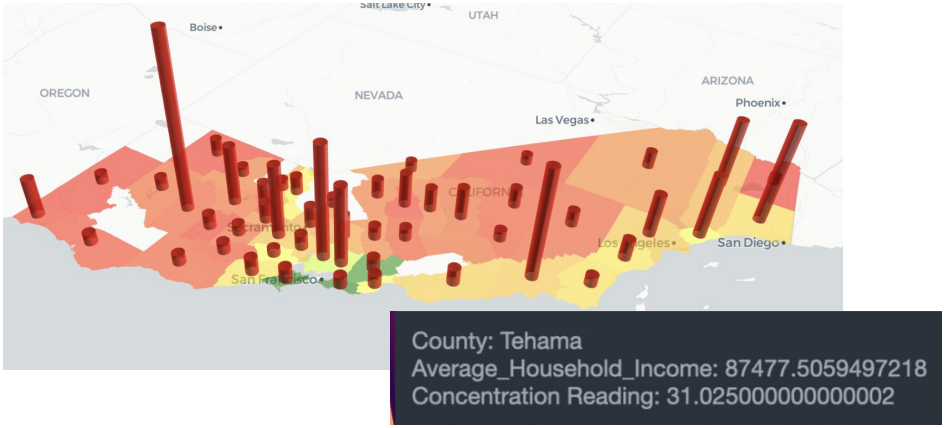
PFHxS x Income Findings

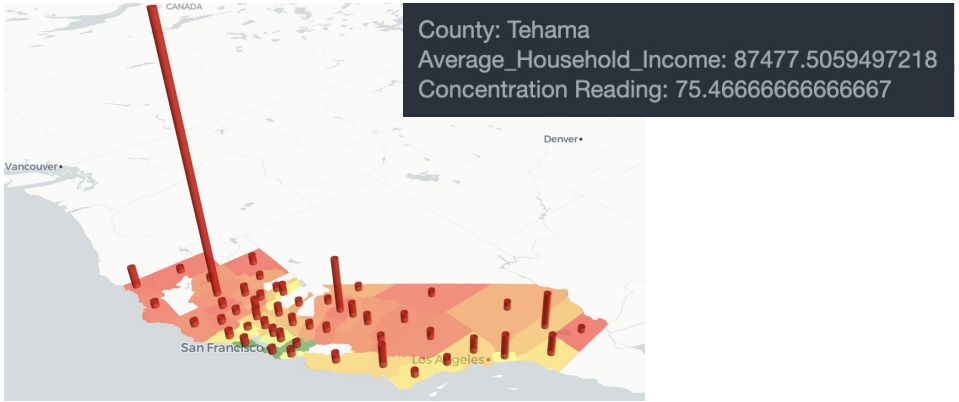




- Counties in the **middle class** (yellow) show **elevated** PFHxS concentrations compared to low-income and high-income counties
- Includes:
  - Alameda (10.74)
  - Yolo (13.13)
  - San Luis Obispo (23.87)
  - San Diego (17.86)
- From a geographical point, higher concentrations tend to be near the coast.

PFOS, PFOA x Income Findings





- **Highest** concentrations found in Tehama, a **middle-income** county
- **Middle-income** counties generally had higher levels of PFOA concentrations
- Both PFOS and PFOA display **low** concentrations in **high-income** (green) counties and **lower low-income** (darker reds) counties
- Very wealthy areas may benefit from better infrastructure, water treatment systems, and environmental regulations while lower-income areas in far Northern and Eastern california may be more rural.

Summary and Future Steps

Summary

- Forecasting not accurate
  - **Non-continuous samples and heavy impact from unpredictable events** → focus on analysis on existing data

- *County-level:* correlations between socioeconomic factors with PFAS concentrations
  - **Lower poverty levels** appear to be correlated with **lower levels of PFAS concentrations**

- *City-level (with respect to county):* Correlation between socioeconomic variables and chemical concentrations
  - ***Alameda + San Diego***: Positive correlation with industrial activity and negative correlation with greater urban development
  - ***Tehama + SLO***: Correlations between varying racial groups, high housing burden %, health issues, and poverty levels and high PFAS concentration levels

Summary

- Correlations between chemicals in the Big 5 (PFOS, PFOA, PFNA, PFHxS, HFPO-DA)
  - PFNA w/ all other chemicals
  - PFOS and PFHxS
  - **Geographical proximity** may affect concentration distributions

- Geospatial map overlaying census data
  - **Low PFOS and PFOA concentration levels** in both high and low income counties
  - **Greater concentration of PFAS chemicals** observed in middle-class counties such as ***Tehama, Alameda, SLO***

Moving Forward

- **Technical Handoff:**
  - Compile all code for 3D geospatial mapping (pipeline that takes in chemical, socioeconomic factor, and time to output a snapshot of the map at the inputted time)

- **Future Potential Directions:**
  - Conduct further statistical analysis and testing to isolate relationships between various variables potentially affecting PFAS concentrations
    - Examples:
      - Geospatial/geographical variables with PFAS sources
      - PFAS sources with socioeconomic factors (like poverty)
  - Explore impacts of governmental environmental policies.
  - Focus analysis on certain drinking water wells to implement upstream solutions targeting issues contributing to sources of PFAS chemicals

Thank You!

Questions?