

Subject	Key Points and Takeaway	Presenter
<u>Introduction and Problem Statement</u>	Background, Audience, Impact Goal (so that...)	Andrew
<u>Project Workflow</u>	Source > Investigate > Cluster > Project > Package for a non-technical audience	Farah
<u>Data Sourcing</u>	Original Sources, Transform, Resample, Merge, Feature Selection	Farah
<u>EDA & Visualizations</u>		Andrew
<u>Clustering</u>	Purpose - Process, Iteration, Output (How many clusters created along how many var groups?)	Bryan
<u>Time Series & Loose Projections</u>	Perspective (2015 as a datapoint, extended). What do we have? Trends and Projections. Inferences.	Bob
Presentation via Streamlit	Process, Purpose (access for the masses)	Jennafer
Streamlit Demo	Click in! Get a county from the audience	Jennafer
Recommendations and Questions	Additional slices, end user testing, publication	Andrew

Visualizing Water Usage in an Evolving Climate

Bringing water usage discourse closer to home with accessible, contextualized, locally relevant information.

Water Usage in the United States

How much water does the average American use every day?

Water Usage in the United States

How much water does the average American use every day?



82 Gallons*

16 Buckets or 10 Kitchen Trash
Cans...
per person
per day



*Source: U.S. Environmental Protection Agency

**Image Source: hgtv.com (7.9 gallon trash can pictured)

*** Image Source: [Uline.com](#)

Problem Statement

Background

Climate change has emerged as a pressing global challenge, with significant implications for water resources. As temperatures rise and climates become increasingly erratic and unpredictable, the task of **monitoring and managing water usage** grows more complex

Objective

The objective of this project is to use machine learning to build a **clustering model** to better understand, and be able to compare and contrast, state-county level water supply and consumption.

- By providing locally-relevant information, consumers of this information may **identify areas for improvement and efficiency**, and industries may adjust consumption patterns through awareness and advocacy.

Project Workflow

01 Data Sources

- Water
- Temperature
- Drought
- Income
- County Weather

U.S. Geological Survey
Nature.com
Nat'l Drought Mitig. Center
Census.gov
Nat'l Weather Service

02 Data Exploration & Viz

- County-level trends for temperature, drought, water usage, population, etc.
 - Barplots
 - Scatterplots
 - Choropleths

03 Data Modeling

- KMeans Clustering of U.S. counties
- Time-Series Analysis of 2000-2020 temperature and drought data

04 Interactive Application

- Streamlit display for clustering insights
- Tableau choropleth with various levers and automatic filtering

Our Well of Data



U.S. Geological Survey

Water withdrawal and consumption data broken down by:
Public Supply | Domestic | Industrial | Irrigation | Thermoelectric

For forming insights on water patterns and conservation efforts



Nature Scientific Data Study

Hourly temperature metrics calculated using reanalysis data across CONUS from January 2000 to December 2020

For time-series analysis on long-term climate trends, extreme weather events, and climate-change related impacts



National Drought Mitigation Center

Weekly drought monitoring using severity and coverage index in CONUS by county

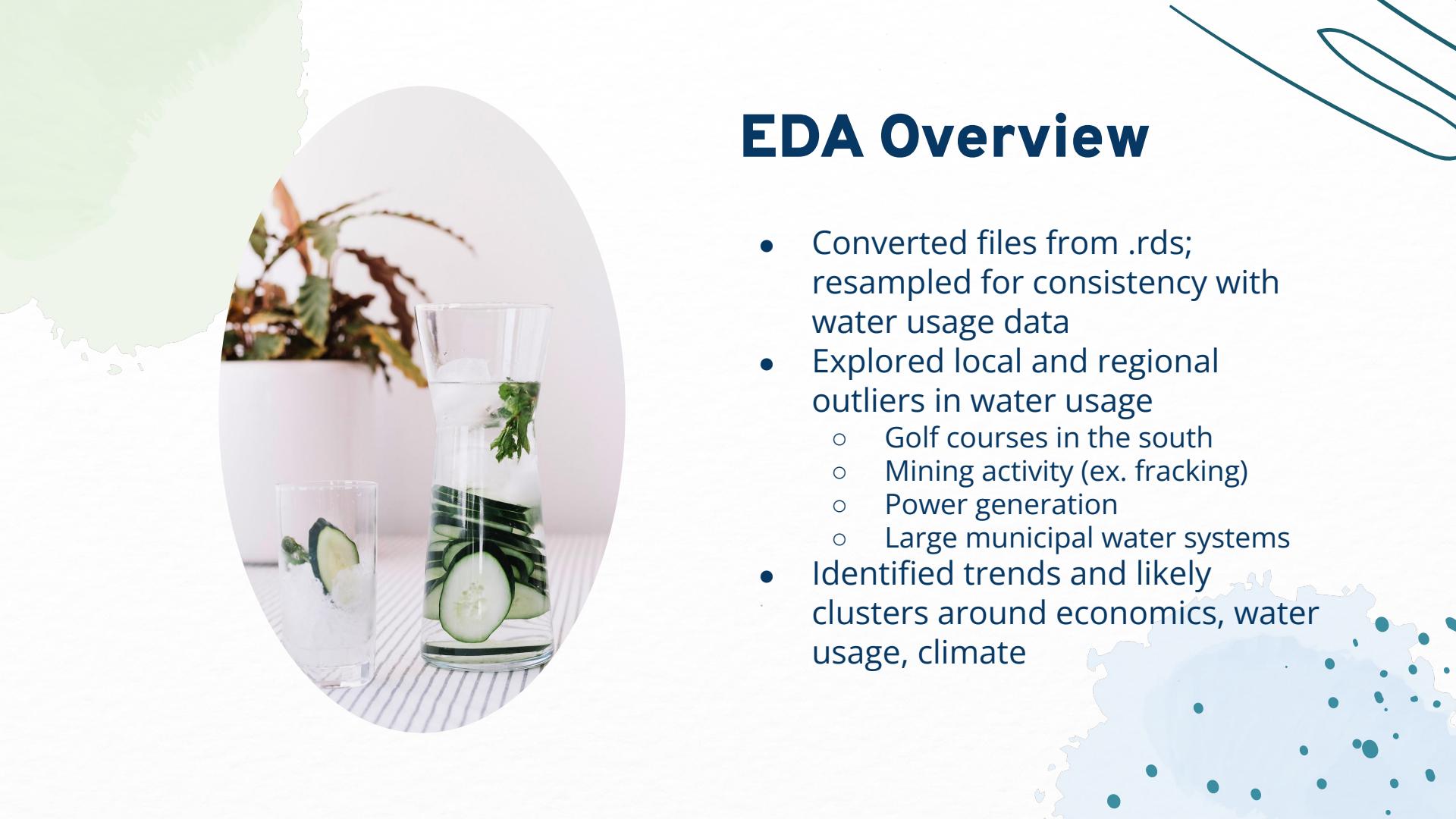
For data exploration and insights into water resources and management, agricultural and irrigation practices



U.S. Census Bureau

Median household income data by state-county

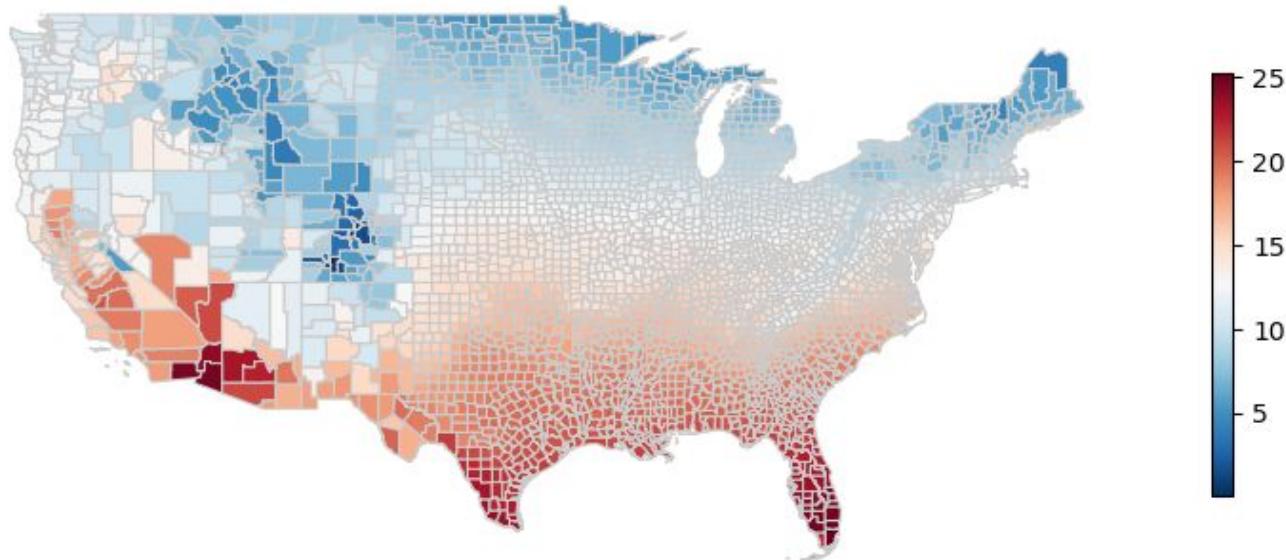
For baseline understanding of socioeconomic considerations, water demand, and resource management



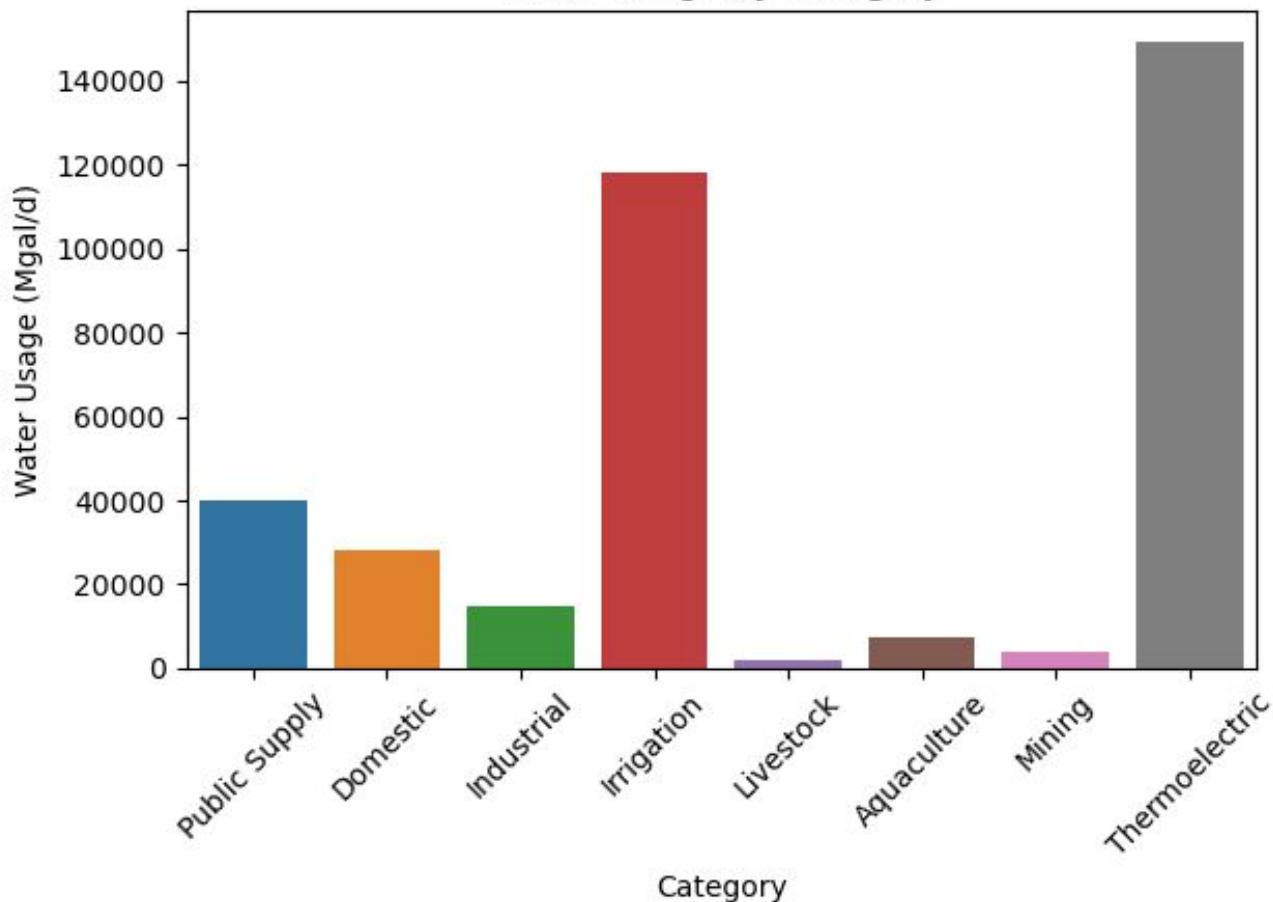
EDA Overview

- Converted files from .rds; resampled for consistency with water usage data
- Explored local and regional outliers in water usage
 - Golf courses in the south
 - Mining activity (ex. fracking)
 - Power generation
 - Large municipal water systems
- Identified trends and likely clusters around economics, water usage, climate

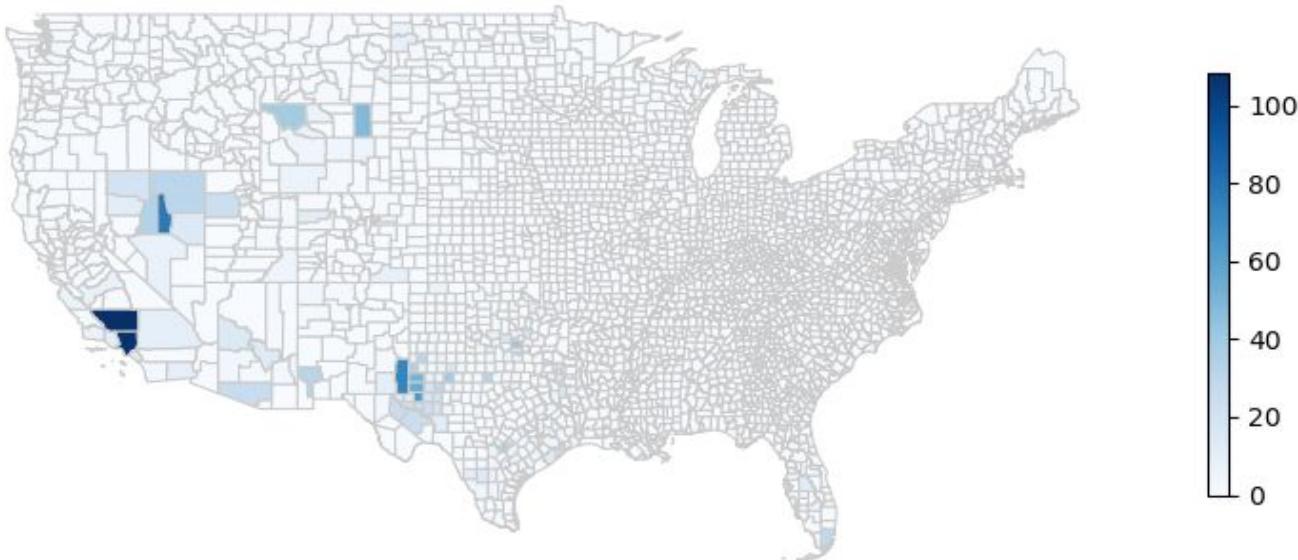
Mean Temperature ($^{\circ}\text{C}$)



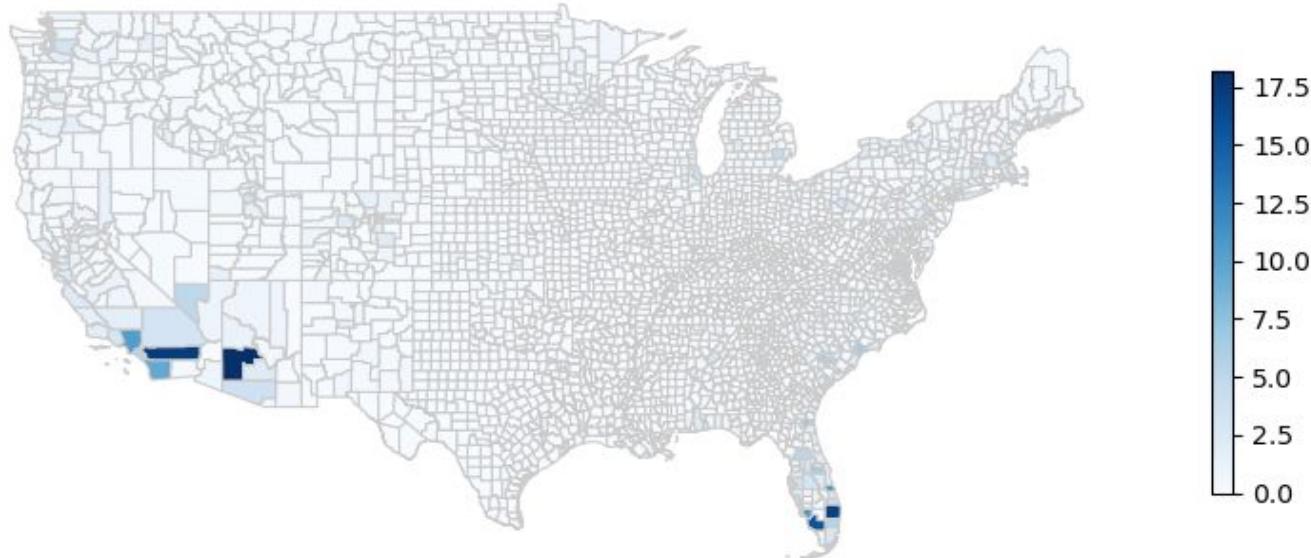
Water Usage by Category



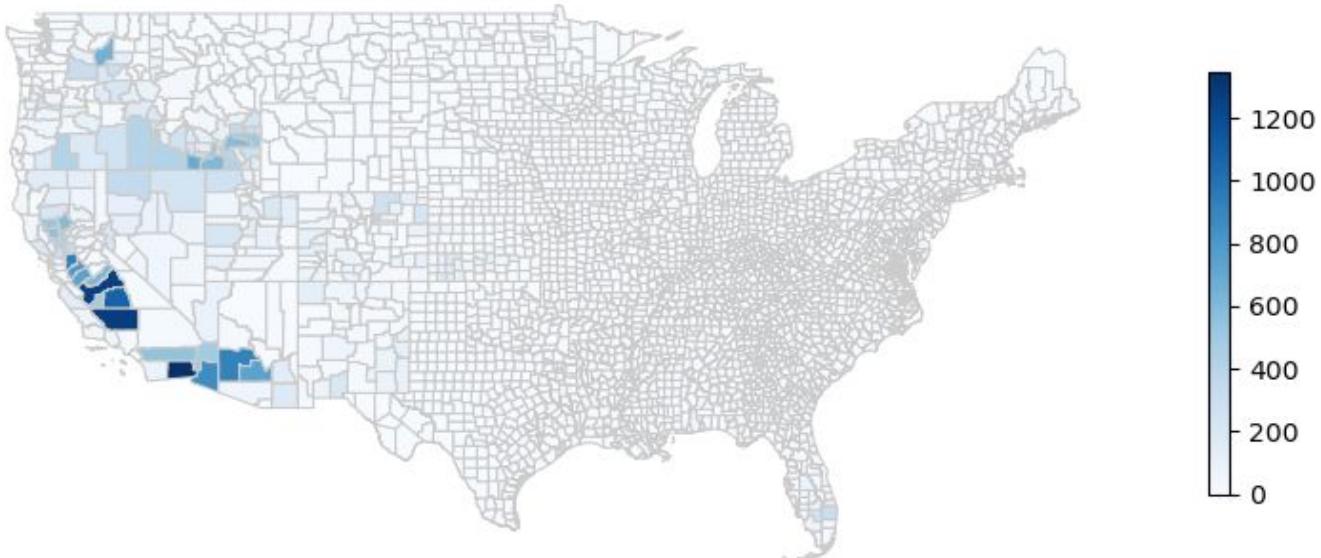
Mining, groundwater withdrawals, total, in Mgal/d.



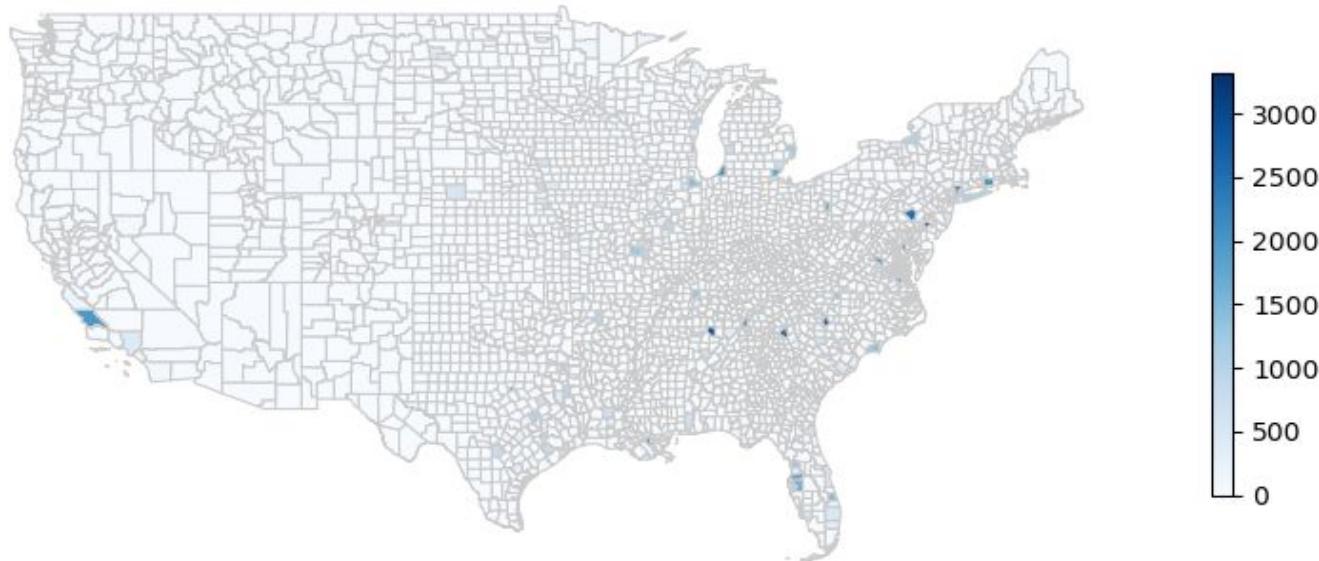
Irrigation-Golf, acres irrigated, total, in thousand acres.



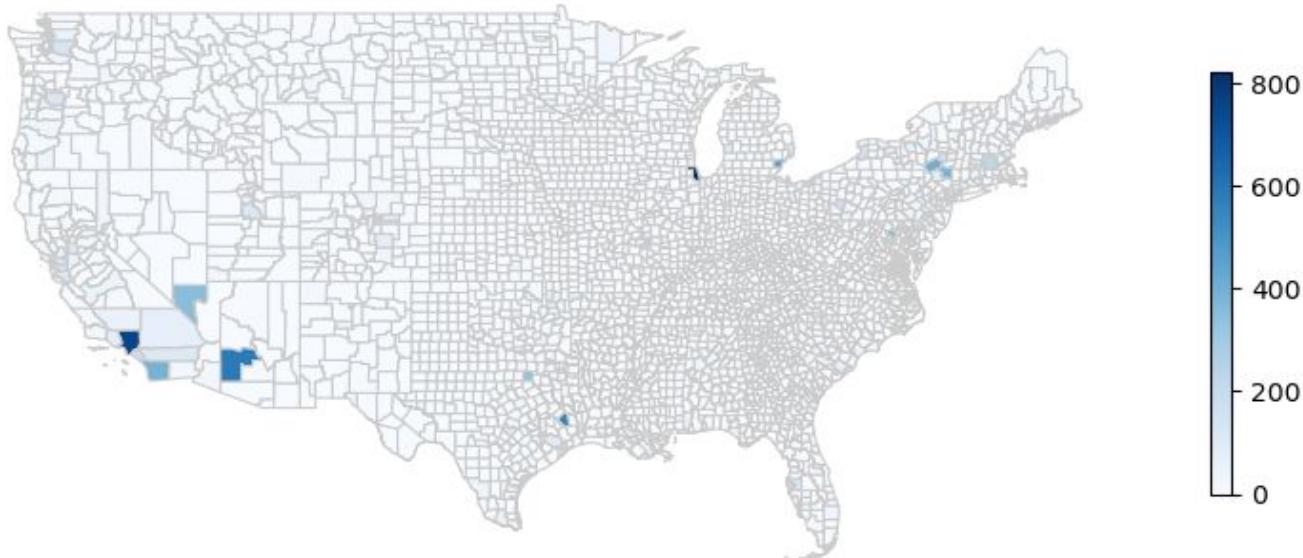
Irrigation-Crop, consumptive use, fresh, in Mgal/d.



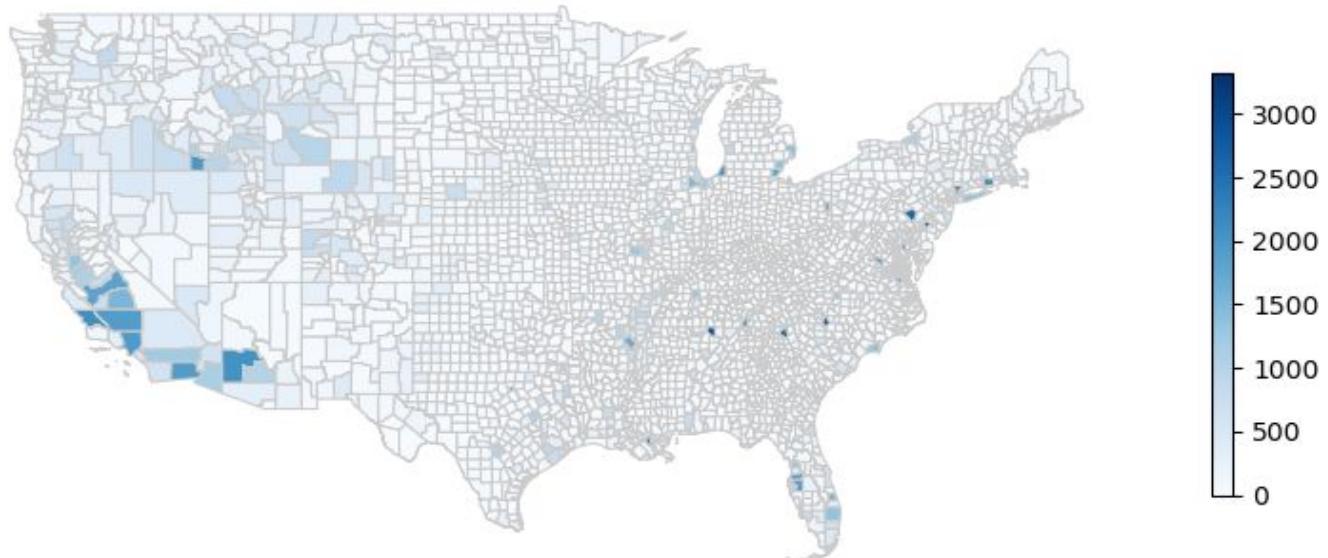
Thermoelectric, total withdrawals, total (fresh+saline), in Mgal/d.



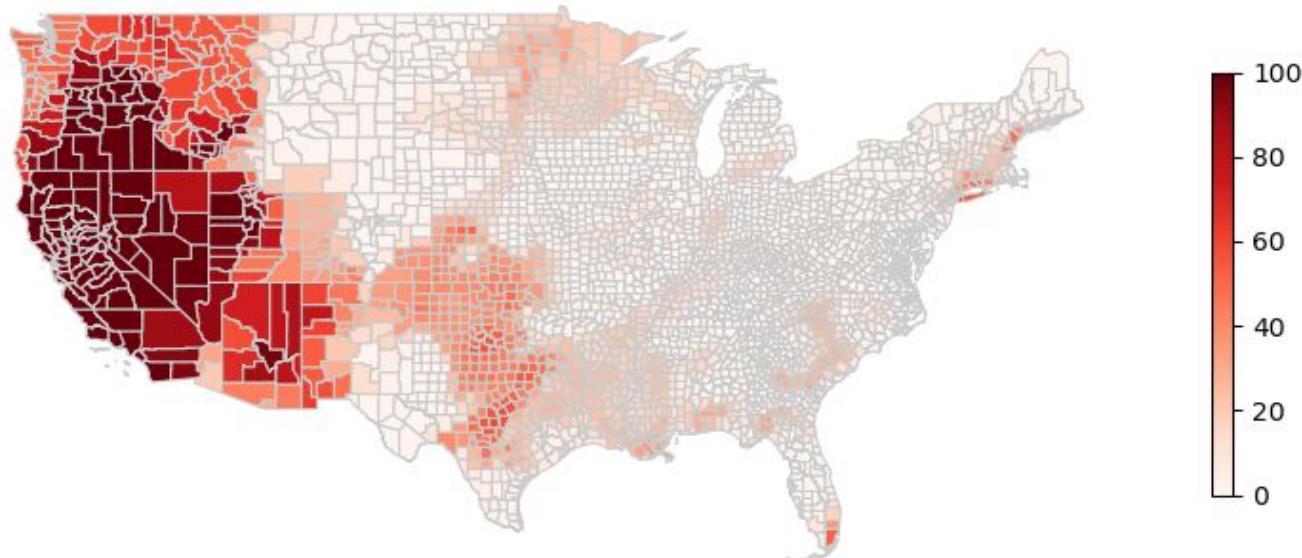
Public Supply, surface-water withdrawals, total, in Mgal/d.



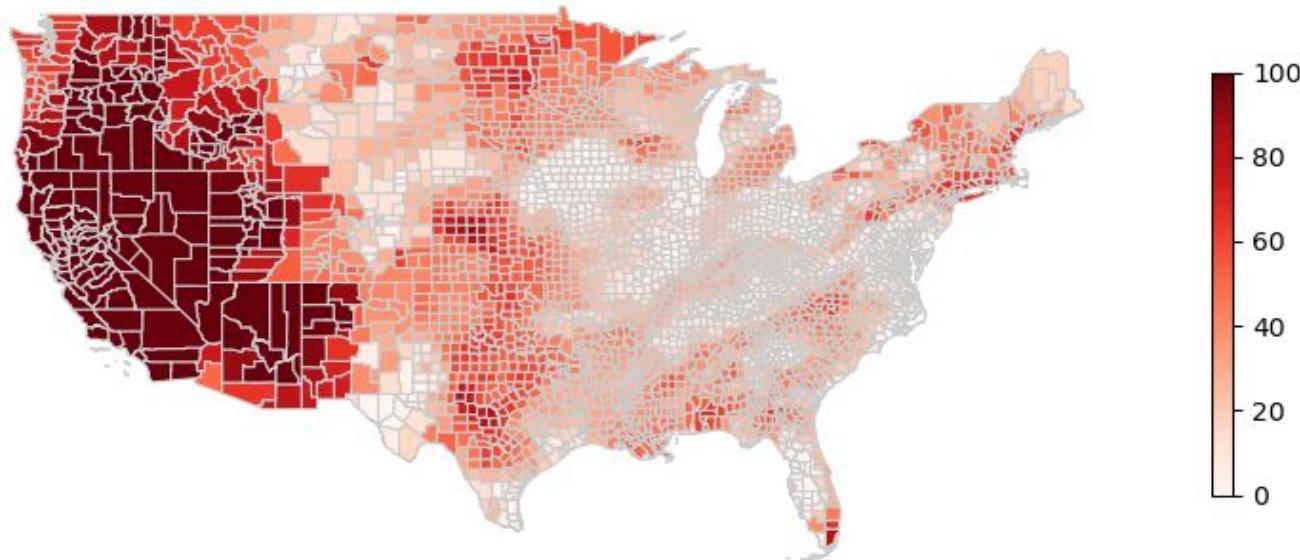
Total withdrawals, total (fresh+saline), in Mgal/d.



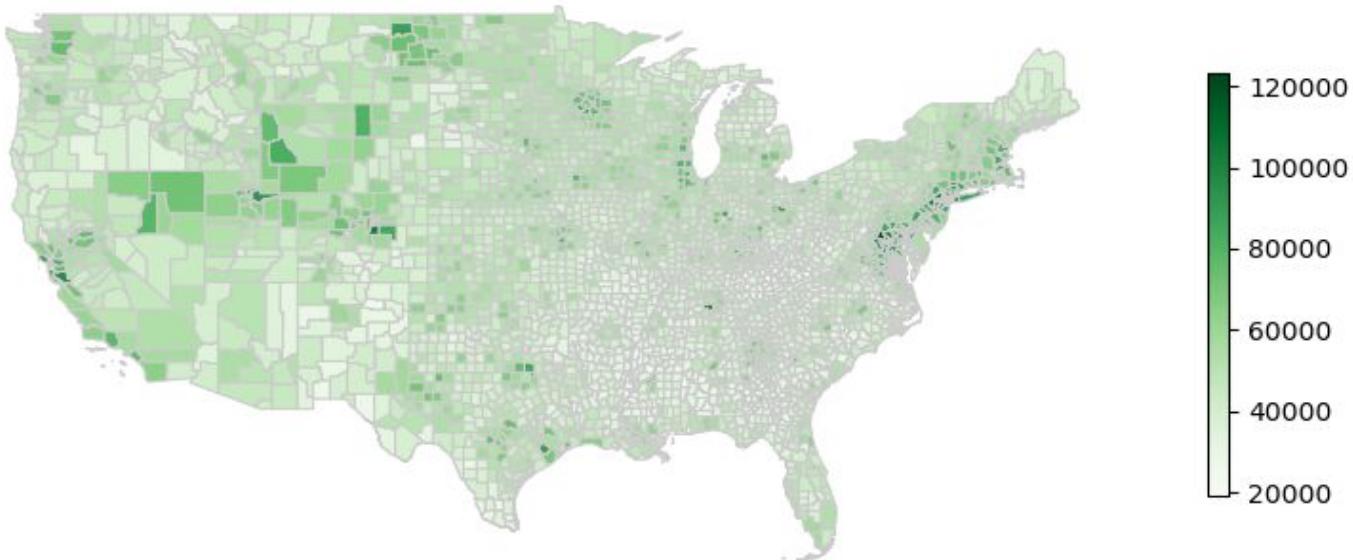
Moderate Drought Conditions



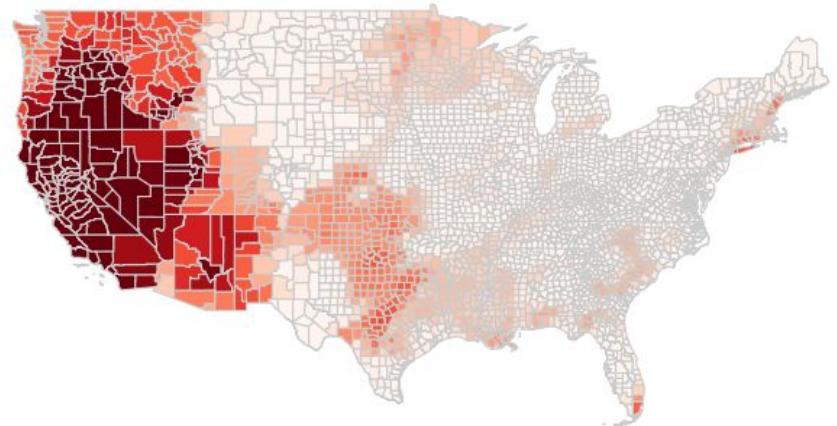
Abnormally Dry Conditions



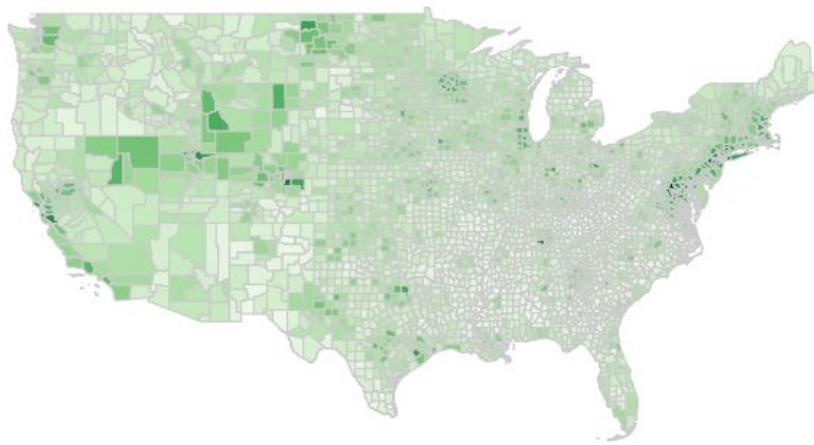
Median Household Income



Moderate Drought Conditions



Median Household Income



Expanded Context: Clustering

Rationale

K-Means and DBSCAN are clustering algorithms which group similar data points together based on their features, revealing underlying patterns and structures in the dataset

Models Chosen

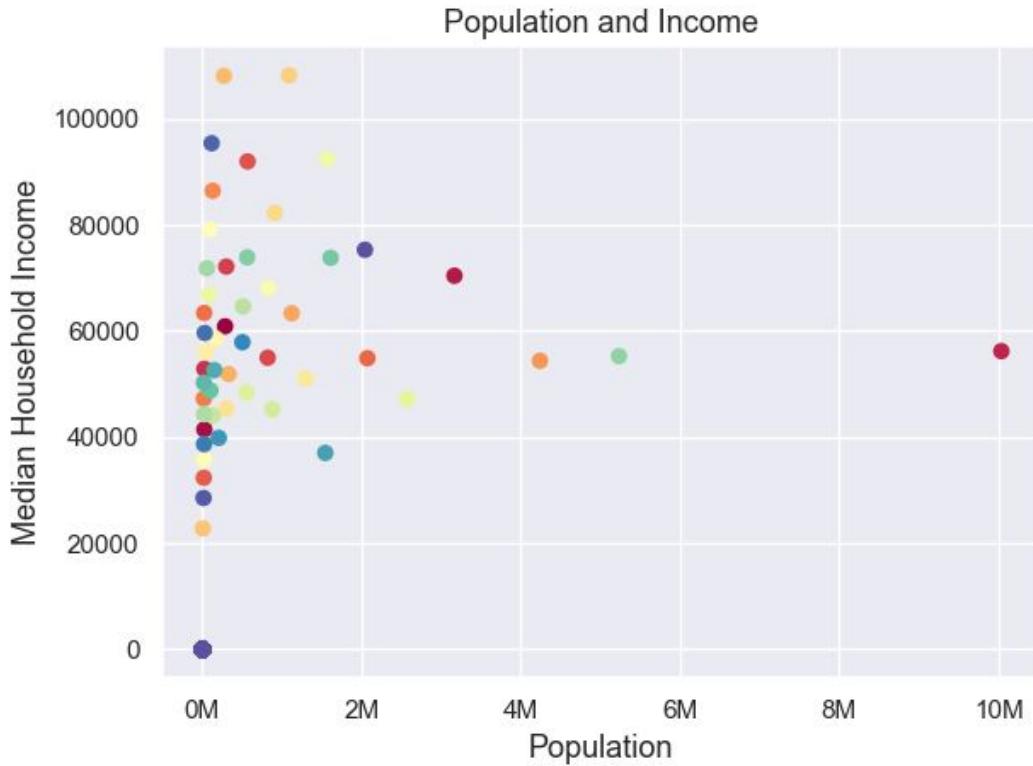
Individual Level

- **Demographic:** Population vs. Median Income
- **Public Supply:** Water Withdrawal vs. Domestic Use

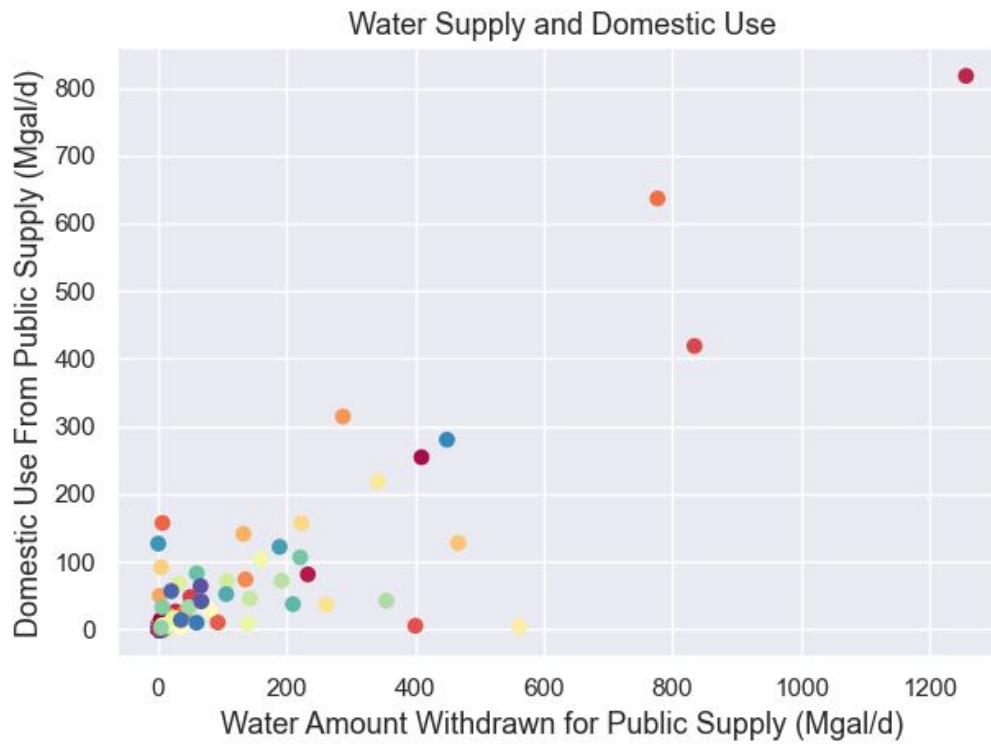
Policymaker Level

- **Irrigation:** Water Withdrawn vs. Reclaimed
- **Total Withdrawal:** Total vs. Public Supply Withdrawal

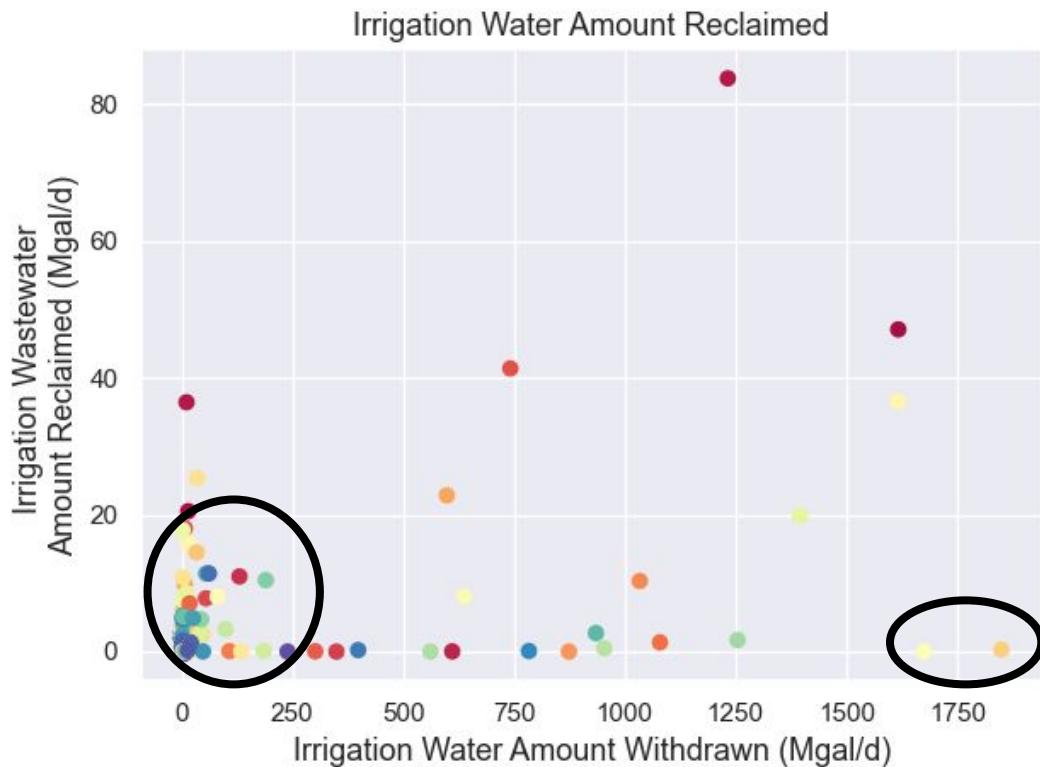
Population vs. Median Income



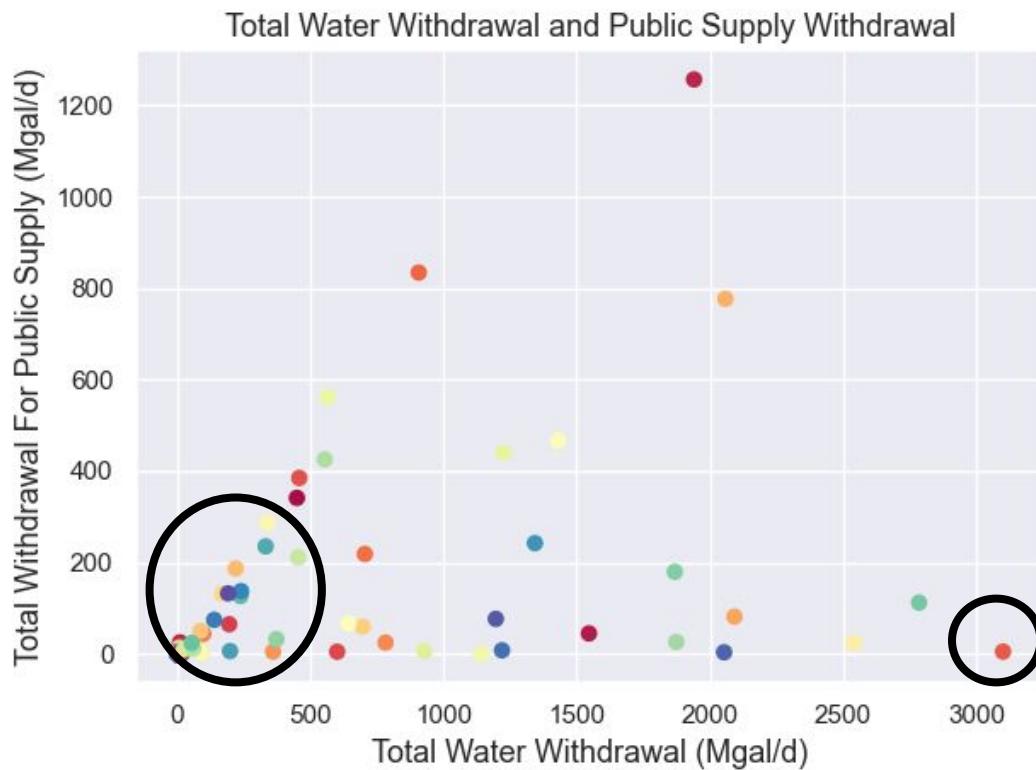
Public Supply: Water Withdrawal vs. Domestic Use



Irrigation: Water Withdrawal vs. Wastewater Reclaimed



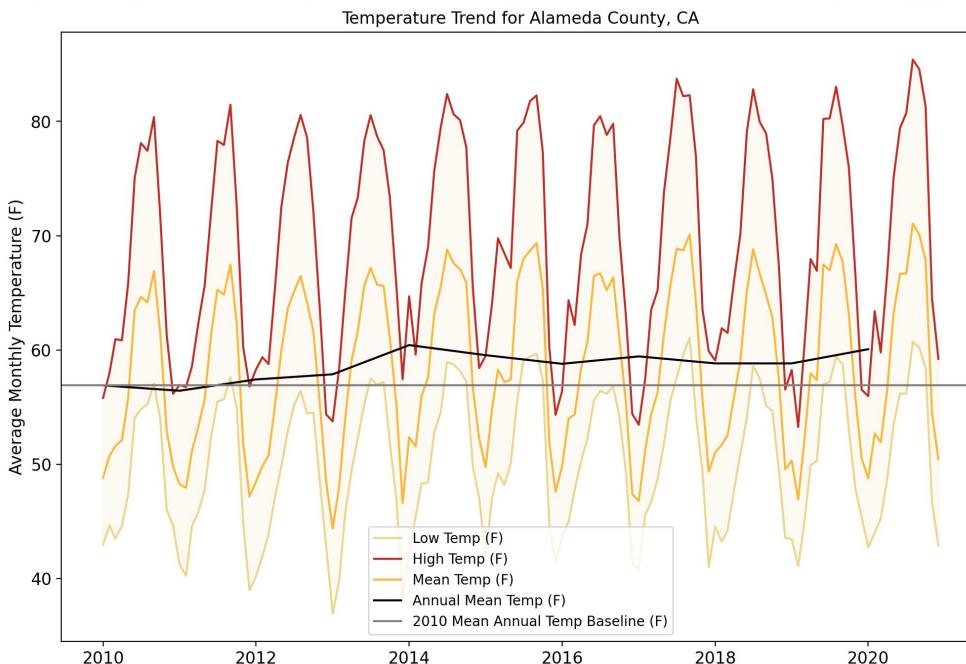
Total vs. Public Supply Water Withdrawal



So What? - Climate and Drought

June-July 2023, Worldwide Temperature Records were broken on four consecutive days^[1]
Increasing Temperatures yields increased water demand for irrigation, power generation, cooling...

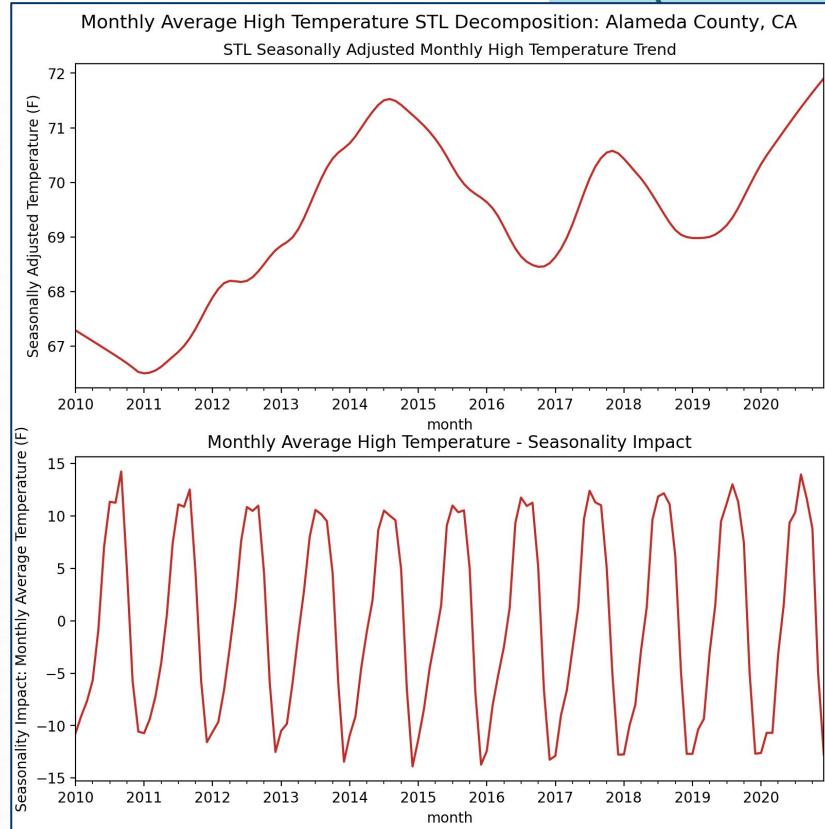
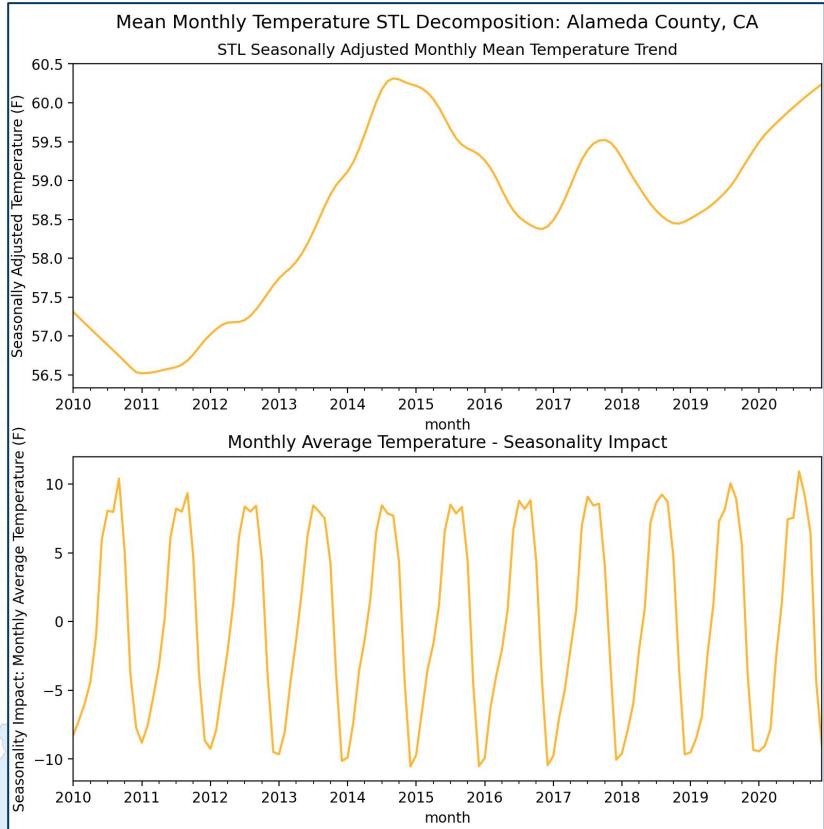
Temperature and Drought Conditions are likely good indicators of water usage trends - nationally and locally



*Source: ABC News

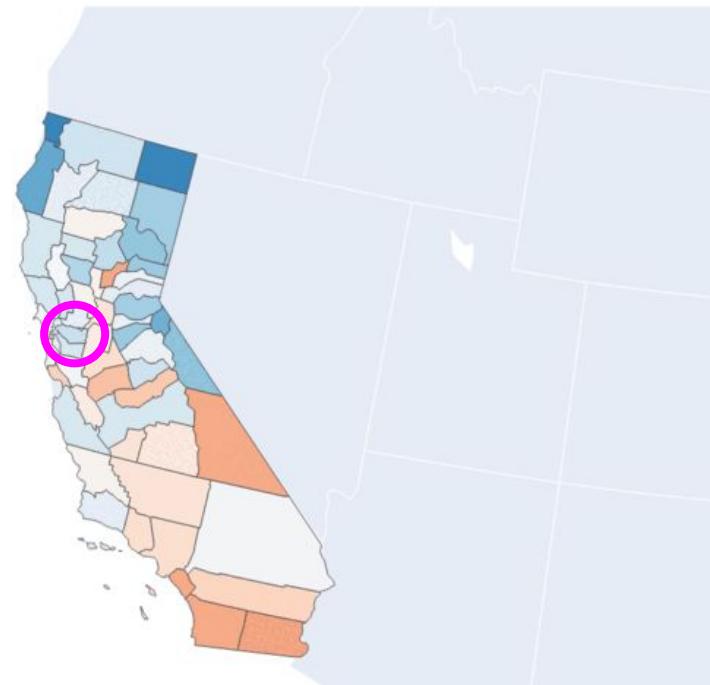
Isn't it Cyclical?

Time Series Decomposition helps identify longer term trends through seasonal swings



Visualizing Year over Year Change

Year over Year Monthly Average Temperature Change (F): CA

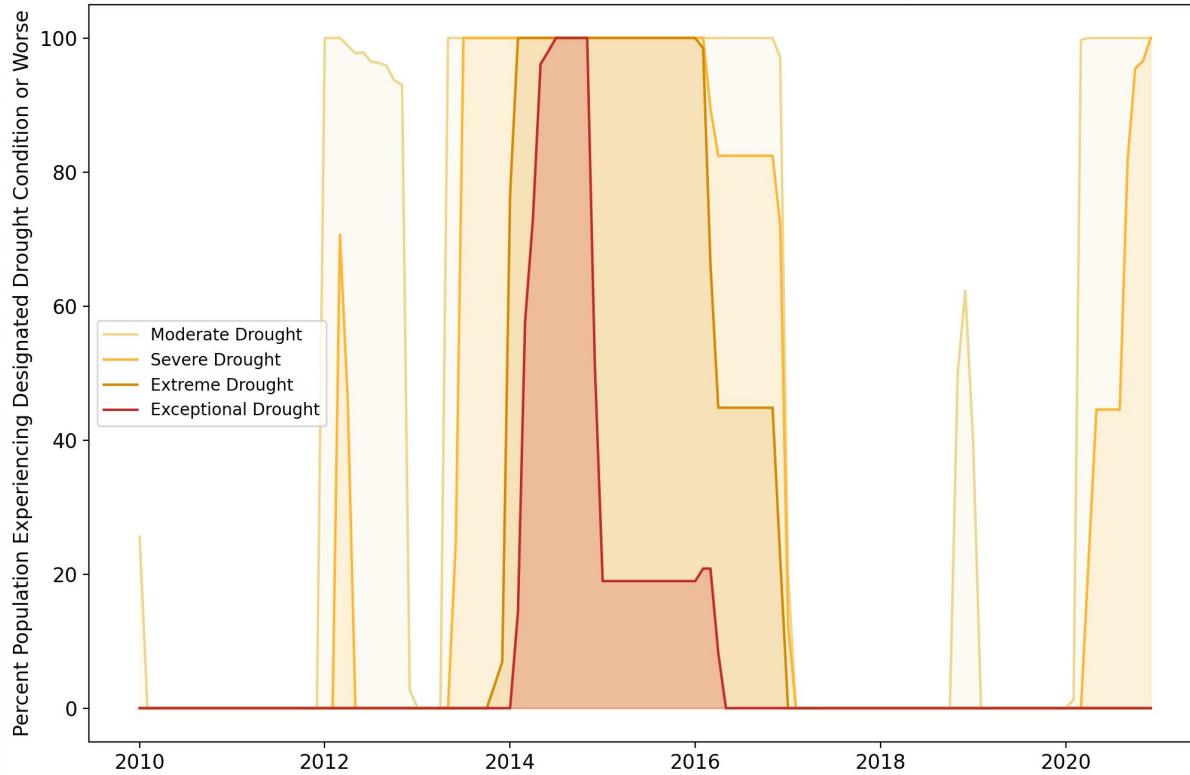


month=2012-10-01



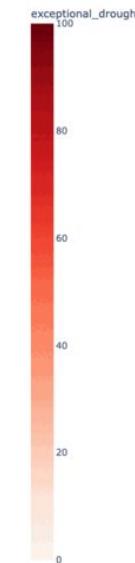
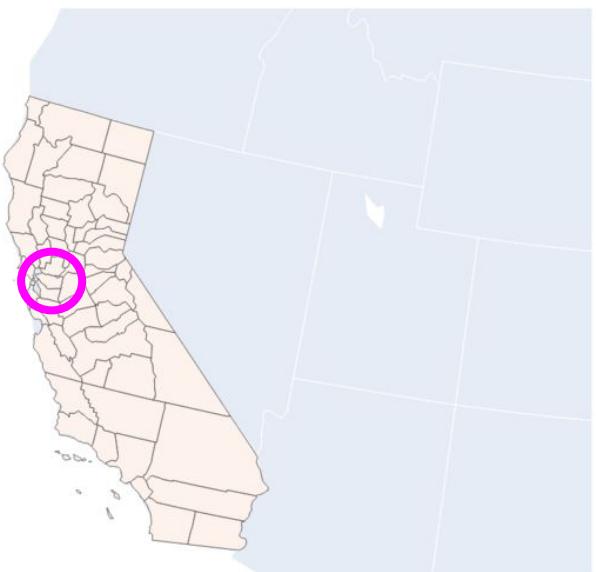
Climate and Drought

Average Minimum Drought Condition for Alameda County, CA



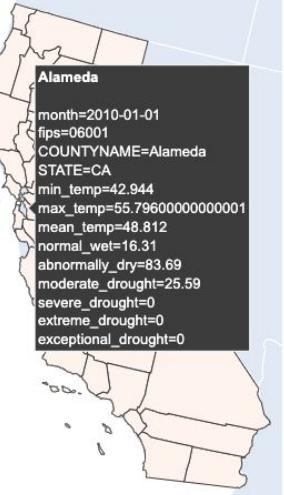
Visualizing Drought Conditions

Percent Population Experiencing Exceptional Drought: CA



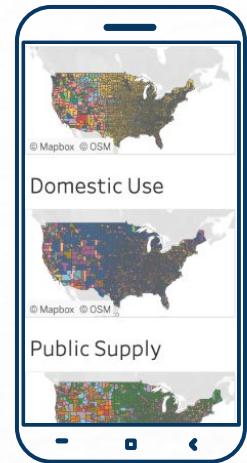
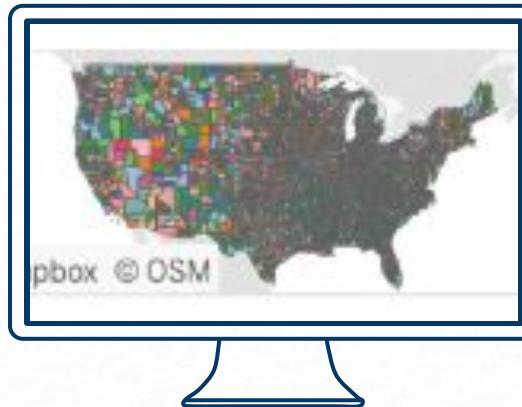
month=2012-12-01

2010-01-01 2010-09-01 2011-05-01 2012-01-01 2012-09-01 2013-05-01 2014-01-01 2014-09-01 2015-05-01 2016-01-01 2016-09-01 2017-05-01 2018-01-01 2018-09-01 2019-05-01 2020-01-01 2020-09-01



Public Benefit, Public Access

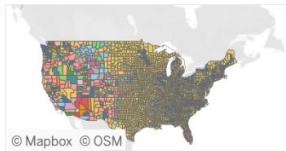
Surfacing insights and local trends on a publicly accessible microsite via Streamlit opens the door to discovery and conversation.



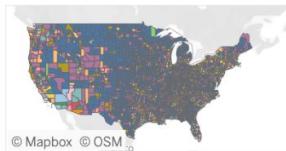
Population vs Median Income



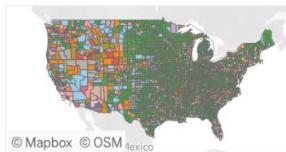
Irrigation



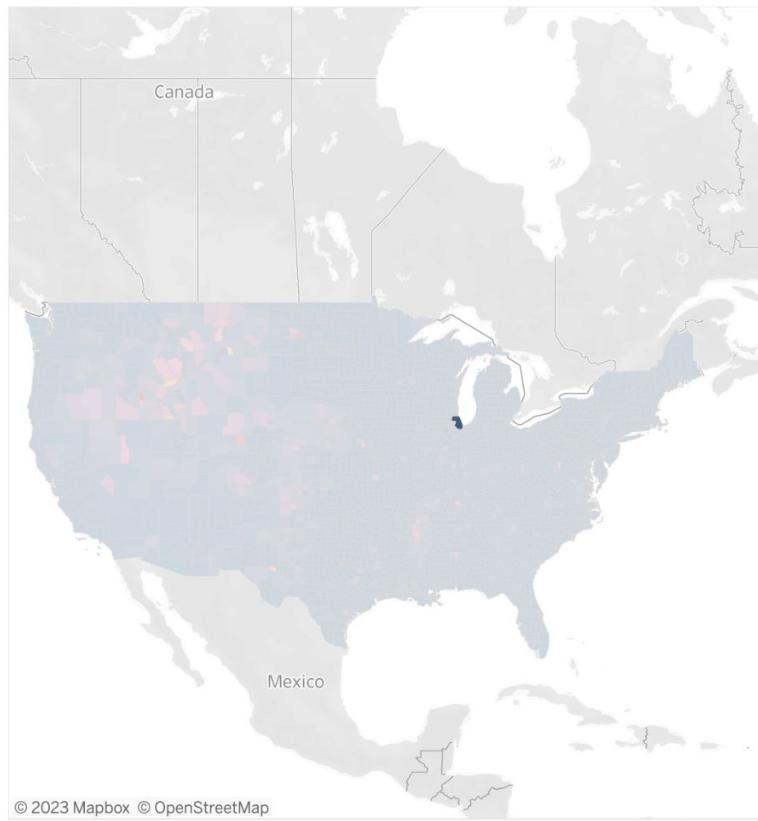
Domestic Use



Public Supply



Total



State

(All)

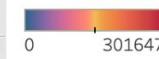
Countyname

(All)

Switch Map View

Total Per Pers...

Select Map



Page

Cluster Charts

Cluster Chart Selector

Select your state

AL

Select your county

Autauga

Model type

- Public Supply Water Withdrawal vs.
Domestic Use
- Irrigation Water Withdrawn vs.
Wastewater Reclaimed
- Total Water Withdrawal vs. Water
Withdrawn for Public Supply
- Population vs. Median Income

Page

Time Series

Time Series Trend Selector

Select the charts accordingly:

Select your state

AL

Select your county

Autauga

Select a time series chart

- Temperature Trends by County
- Drought Trends by County

Recommendations

Iteration & Testing

- Update for 2016-2022 data
- Conduct end user panel testing
- Iterate on visuals to ensure easy access to the full story

Production and Publication

- Consider higher capacity platforms for visualizations to minimize load time performance
- Source more recent climate data via API & Schedule Model Updates
- Publish findings to build public awareness

THANKS

Do you have any questions?

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, infographics & images by Freepik

Please keep this slide for attribution