Addressing Water Inequality: An Analysis of California's Public Water Systems and Policy Solutions for Disadvantaged Communities

By

Neha Burri(11694929)

Akhila Reddy Kommiti(11705751)

Pooja Thella(11713291)

Simha Sai Ravi Kumar Marisa (11684252)

University of North Texas

Toulouse Graduate School of Business

ADTA 5940 – Capstone

Jingjing Tong

Spring 2025

Contents

CHAPTER 1: INTRODUCTION

Background

Being able to have safe drinking water stands as a fundamental right, yet different communities across the world, particularly in California, experience considerable gaps in their access to water quantity and quality. People from most parts of California benefit from sustainable water services, yet several marginalized communities in rural and low-income urban areas experience elevated risks from contaminated water supplies (Arche et al., 2024). Various social, economic, and environmental elements, together with racial discrimination, unstable economic markets, and outdated infrastructure systems, drive this inequality. The absence of equal clean water access leads to critical effects which cause health deterioration and economic difficulties and sustain negative poverty and social inequality patterns.

California's water system, despite its size and complexity, struggles with issues of equity in water distribution. The affluent parts of cities get steady, clean water supplies (Carchi et al., 2023). Still, poorer areas, mainly inhabited by Latino and African American residents, face insufficient water infrastructure due to poor quality water sources. Inspection results indicate higher exposure of residents in these regions to dangerous chemicals like arsenic and nitrates in their drinking water, which creates severe health concerns (Chen & Franklin, 2023). Historical, along with present social inequality patterns, combine with climate change issues to intensify these challenges so vulnerable groups remain with no sufficient access to vital water resources.

This study investigates the elements contributing to California's public water dispersion inequity, with particular attention paid to its effects on susceptible population groups. The research establishes connections between water quality standing, socioeconomic factors, and neighbourhood location to reveal how some neighbourhoods struggle more with clean water access (Fernandez-Bou et al., 2023). Existing policies, technological solutions, and potential new approaches will be examined in the research to determine ways of reducing discriminatory water distribution patterns across the state.

Addressing water inequality in California stands as an essential requirement because it defends both community health among disadvantaged residents along environmental equality between groups in society. Secure equal access to protected water supplies across California remains an emergency requirement since this western state faces escalating water scarcity during drought periods (Arche et al., 2024). The study adds to water justice discourse through systematic investigation of structural inequalities which maintain racial disparities and through the presentation of practical methods to attain water equity.

This investigation will answer three main questions by determining what social elements, together with economic conditions and environmental aspects, result in unequal clean water access throughout California. Which factors determine that marginalized communities face unequal water treatment, and how do racial, class and geographical locations influence water distribution (Carchi et al., 2023). The state needs new policies and technological innovations to guarantee fair water distribution, particularly for disadvantaged groups. This research investigates water inequality at multiple levels in California while creating strategies for securing water rights for all members of society.

Research Questions

1. Three dominant elements that create unequal access to clean drinking water in California include social aspects, economic conditions, and environmental factors?

2. Water quality in public systems reflects the interaction between residents' racial background, income level, and geographic positions?
3. The operational status of public water systems in disadvantaged communities depends to what extent on infrastructure quality combined with regulatory oversight?

CHAPTER 2: LITERATURE REVIEW

Historical Context: Overview of California's water infrastructure and inequality

The water systems of California deal with multiple complicated challenges which encompass both reliable water availability and water quality together with proper distribution among different populations. The state operates under extensive challenges because it exists across multiple geographic zones and demonstrates financial disparities between regions while dealing with climate change consequences. As the biggest and most intricate water system in the United States exists in California the state faces water shortage alongside pollution and unequal distribution of clean drinking water. Each distinct area within the system shows different water conditions because its dependable water resource supply levels change accordingly.

The state of California retrieves its water supply from Sierra Nevada snowmelt and maintains reservoirs as well as aqueducts to operate its water distribution network. Continuous pressure on these water supply systems continues during drought years because climate change has produced both longer and more intense dry spells (Agustí Pérez-Foguet, 2023). Water resource management in the state must balance sustainable water delivery to its 40 million residents together with broad agricultural operations through scarce resources (Agustí Pérez-Foguet, 2023) Water storage and distribution fails to meet increasing challenges because dry periods affect California with increasing regularity.

The principal water quality problems continue to threaten communities across the entire state of California. The residents throughout rural areas and low-income districts of Central Valley and southern California lack reliable access to water purity. Federal water requirement limits are exceeded by unsafe drinking water, which affects more than one million people across California, according to national reports. Public health risks plague communities that obtain water from systems contaminated with nitrates, arsenic, and lead contamination. The water systems cannot afford to upgrade their infrastructure because they do not have sufficient financial resources to maintain it.

Groundwater supply represents the primary water source for the state as it attempts to meet escalating water requirements, which creates a major problem. The excessive extraction of groundwater has led to surface land destruction and deteriorating water quality in defined areas. The contamination of groundwater has made it harder for communities to access clean water through these pollution problems.

California faces limitations in solving its water issues because water supply concentrations are unbalanced across the state. The Central California cities of Los Angeles, San Francisco, and San Diego operate secure, clean water supply systems, yet rural communities, especially those located in the Central Valley, encounter water contamination issues. Better funding for infrastructure combined with wider water management involvement will result in equal water accessibility for the people of California.

Environmental Racism & Disparities: Disproportionate effects on Latino and African American populations

The effects of polluted water on Latino and African American populations in California represent a major issue that recent research has started to analyze more intensively. The research conducted by Acquah & Allaire (2023) and Fernandez-Bou et al. (2023) shows that these populations typically reside in places where water pollution rates are high, thus negatively impacting their medical situation and life conditions. Communities with poor water quality mainly consist of residents from lower-income brackets who have little influence on political matters.

The populations seek water from outdated and maintained unsatisfactorily water supply systems. The rural Latino communities located in the Central Valley particularly depend on secondary water service systems that tend to harbour nitrates, arsenic or bacterial contaminants (Chen & Franklin, 2023). These pollutants present significant dangers to public health and affect three main vulnerable groups: children, elderly citizens, and women who are pregnant. Those who continually encounter hazardous substances in their water supply face increases in their potential to develop cancer alongside birth abnormalities and developmental delays.

Agustí Pérez-Foguet, (2023) argues that historical practices alongside housing segregation and environmental racism together expose Latino and African American communities to unequal amounts of environmental hazards that include polluted water. His claims are seconded by Carch et al., 2023 and (Chen et al., 2023) who argues that Such communities struggle to get clean water because they face different barriers which include restricted funds along with weak political standing coupled with complicated administrative systems that overlook their needs.

Thanks to the environmental justice movement, the nation now recognizes the unequal distribution of water services among communities and seeks justice through improved water infrastructure development in these areas (Acquah & Allaire, 2023). The California Department of Water Resources, together with other state agencies, currently tackles these disparities through their implementation programs. California Directorates focus their water infrastructure investment funds on areas that have higher pollution rates (Fernandez-Bou et al., 2023). Many efforts must be enacted to deliver clean drinking water safety to all people throughout California without discrimination based on either race or economic level.

Governance and Policy Developments: SGMA, California Water Plan 2023

The development of water policies in California spans throughout history to respond to the fast growth rate and complex natural resource management needs across the state. Since the start of the state's water management system, California has built big infrastructure elements, including dams, reservoirs, and aqueducts (Goddard et al., 2021). The State Water Project and Central Valley Project development enabled northern California water allocation to transport water throughout southern regions, supporting agricultural growth and urban expansion.

The centralized water management system authorized to serve agricultural practices and urban development has received growing opposition (Hanak & Chappelle, 2025). California has started implementing more sustainable water management approaches because water scarcity and environmental preservation have become vital issues (Goddard et al., 2021). The Sustainable Groundwater Management Act (SGMA) stands as a major achievement, becoming law in 2014 to control excessive groundwater usage while avoiding additional damage to the environment. Under SGMA, local agencies must write operational plans for managing their groundwater source areas and work toward sustainability by 2040.

The California Water Plan undergoes its most recent restructuring in 2023 to serve as the guiding document for state water resource management. Through this plan, the state focuses on climate change adaptation, water system resilience enhancement, and water equality between all communities (London et al., 2021). The 2023 revision of the California Water Plan emphasizes forward progress that serves disadvantaged communities and native tribes because they typically face exclusion from water management procedures. Building upon previous top-down water management systems, California now adopts an inclusive strategy to serve the entire population's social and cultural requirements.

Water policies in California need to address intricate challenges regarding restoring natural environments and preserving habitats. The Sacramento-San Joaquin Delta, together with other critical ecosystems, makes California its home while supplying water to millions of residents (Siirila-Woodburn et al., 2021). Controlling water management between human supply requirements and environmental sustainability demands rigorous oversight because climate change impacts extend with increasing intensity.

Technology and Tools in Water Management: Predictive analytics, remote sensing, recycling/desalination

The water management issues across California require breakthrough technological solutions to find optimal repair measures. Additions of predictive analytics enable better water quality outcome predictions and water resource availability forecasts through its operation combined with data modelling and real-time monitoring systems for water distribution networks (Archer et al., 2024).

Water management systems become smarter and more efficient through implemented technological solutions because they reduce water scarcity during droughts and critical emergencies.

According to (Agustí Pérez-Foguet, 2023) these monitoring systems let water utilities recognize pollutants during their occurrence, resulting in the simultaneous protection of public health systems. Remote sensing allows for combined monitoring of underground aquifers and water surface elevations, enabling the detection of water levels that mostly exist in subsurface zones.

Water management effectiveness in California has grown significantly because of advanced predictive modelling technologies now spread throughout the field. Modern predictive modelling systems produce water supply outlooks by analysing recorded data, which combines actual climate measurements with projected climate conditions (Hanak & Chappelle, 2025). The data collection provides water executives with tools to choose suitable reservoir control systems and decide appropriate water allocations between agriculture, cities, and environmental protection needs.

Water recycling technology in California operates jointly with desalination methods to preserve water supplies. The wastewater facilities in Orange County operate their water recycling operation through state-financed capital and produce potable water from their treatment facilities (Agustí Pérez-Foguet, 2023). The desalination plant currently under development in Huntington Beach could supply water to California, but its implementation creates problems related to finances and environmental effects.

California must implement these modern technologies to increase water stability. The state should dedicate funding to infrastructure expansion through targeted budget allocations while simultaneously developing water technology policies which need appropriate financial support.

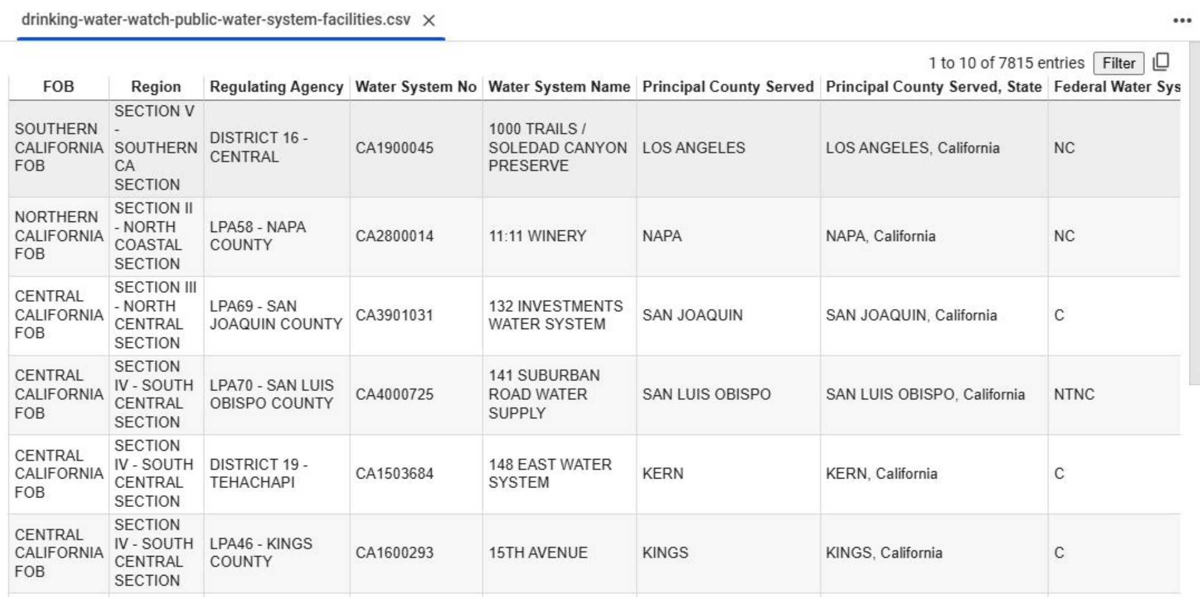
CHAPTER 3: METHODOLOGY – DATA PREPARATION

Software

This research is dependent on Python to execute its data analysis because this programming language provides exceptional capabilities in data manipulation and analysis features. The data manipulation and preparation tasks could be completed through the panda's library, which enables easy processing of extensive data, including reading, cleaning and summarizing data operations. The data visualization relied on Matplotlib and Seaborn tools to visualize patterns, distributions, and relationships in the data contents. Pattern recognition tasks were carried out using sci-kit-learn for regression analysis, clustering procedures, and cross-validation. These analytical instruments enabled detailed research of the public water systems data while enabling researchers to study trends between the different variables.

Data Collection

Information regarding California water systems came from the Drinking Water Watch Public Water System Facilities dataset. The available dataset consists of 7,815 entries containing 35 columns with essential variables, including Water Source Type, System Status, Residential Population and Regulating Agency. The database originates from publicly available information the California Department of Public Health distributes. The dataset contains numerical and categorical variables, permitting exploratory data analysis (EDA) and machine learning model applications. New data entries continue to update the information base to maintain the highest accuracy and current status inall analyses. fig.



drinking-water-watch-public-water-system-facilities.csv

1 to 10 of 7815 entries Filter

FOB	Region	Regulating Agency	Water System No	Water System Name	Principal County Served	Principal County Served, State	Federal Water Sys
SOUTHERN CALIFORNIA FOB	SECTION V - SOUTHERN CA SECTION	DISTRICT 16 - CENTRAL	CA1900045	1000 TRAILS / SOLEDAD CANYON PRESERVE	LOS ANGELES	LOS ANGELES, California	NC
NORTHERN CALIFORNIA FOB	SECTION II - NORTH COASTAL SECTION	LPA58 - NAPA COUNTY	CA2800014	11:11 WINERY	NAPA	NAPA, California	NC
CENTRAL CALIFORNIA FOB	SECTION III - NORTH CENTRAL SECTION	LPA69 - SAN JOAQUIN COUNTY	CA3901031	132 INVESTMENTS WATER SYSTEM	SAN JOAQUIN	SAN JOAQUIN, California	C
CENTRAL CALIFORNIA FOB	SECTION IV - SOUTH CENTRAL SECTION	LPA70 - SAN LUIS OBISPO COUNTY	CA4000725	141 SUBURBAN ROAD WATER SUPPLY	SAN LUIS OBISPO	SAN LUIS OBISPO, California	NTNC
CENTRAL CALIFORNIA FOB	SECTION IV - SOUTH CENTRAL SECTION	DISTRICT 19 - TEHACHAPI	CA1503684	148 EAST WATER SYSTEM	KERN	KERN, California	C
CENTRAL CALIFORNIA FOB	SECTION IV - SOUTH CENTRAL SECTION	LPA46 - KINGS COUNTY	CA1600293	15TH AVENUE	KINGS	KINGS, California	C

Figure 1 water information CSV file

CHAPTER 4: EXPLORATORY DATA ANALYSIS (EDA ANALYSIS)

Detailed Description of the Dataset

The dataset consists of 35 columns that provide information about 7,815 records recorded in the Drinking Water Watch Public Water System Facilities. The database incorporates whole data about California water systems to present essential data about different water distribution systems alongside their sources and service areas. The dataset delivers essential information about public water systems through details concerning regulatory control bodies and their water source types alongside community service connection distribution data.

Data Structure

The database contains 35 fields consisting of both numeric and categorical data points. The collection contains FOB (Field Operations Branch) and Region among several categorical variables together with Water System No, Water System Name, Pry Water Source Type, and, other elements. The categorical variables in the dataset explain essential aspects of water systems by describing geographical regions and determining authorities and water origin types. The database contains numerical variables that present information about the population served and service connection quantity and link to fee code frameworks.

Key Variables

The data separates water systems into different operational areas and geographic zones using FOB and Region variables, which include Southern California and Northern California. The FOB column displays information about field operations, while the Region column specifies particular areas of California that the systems serve.

The regulating Agency: This column mentions the agency with authority to monitor and control operations at each water system. The database uses Water System No alongside Water System Name to uniquely identify and name its water systems.The dataset contains multiple fields that classify water systems through their Federal Water System Type and State Water System Type formats. Each water system receives designation through this set of columns into Community status and Non-Transient Non-Community type or Transient Non-Community designations based on its service type.

System Status functions as an essential identifier displaying each water system's operational status, maintenance, or inactive conditions. This particular variable is essential in evaluating the water service dependability and availability across diverse areas.

The dataset contains population statistics described through four columns: Residential Population, Non-Transient Population, Transient Population, and Total Population. The columns deliver data that reveals the groups' demographics that depend on each water system. Several columns demonstrate data incompleteness based on the fact that Residential Population data exists for only 3,764 systems. This provides evidence of missing information in particular geographic regions.

The dataset contains multiple columns presenting total service connections among Agricultural, Commercial, Institutional and Residential divisions. The columns indicate the number of connectable customers and entities in each water supply system. The Number of Agricultural Service Connections

(AG) demonstrates the total number of agricultural users relying on that particular system. The dataset contains empty data points inside Combined Service Connections and Commercial Service Connections columns.

Fee Code and Fee Code Description provide information about the different fees associated with water service agreements. These codes establish the different methods used to fund systems and determine acceptable billing systems for users.The Treatment Plant Class and Distribution System Class act as columns to identify the plant treatments and distribution network classifications in each water system. The classification system reveals information about treatment quality standards and distribution system intricacy.

Sanitary Survey Date presents the latest inspection date regarding water systems in a different column. The regulatory checks and system maintenance activities appear chronologically in this document.

Missing Data

The information base includes numerous data columns which exhibit absent entries. The Principal County Served state column has 208 unrecorded entries, while Residential Population has a deficit of 4,051 values. Systematic data collection efforts are limited within obscure or rural locations based on the noticed gaps across the database. There is a substantial amount of missing data for the Non-Transient Population and the Transient Population among water systems across the country.

Evaluation of analysis goals will determine whether to impute or delete missing data from the Fee Code, Fee Code Description, and Treatment Plant Class columns, which have substantial missing entries. Before beginning the analysis, the analyst must address the missing values because they might affect the analysis results.

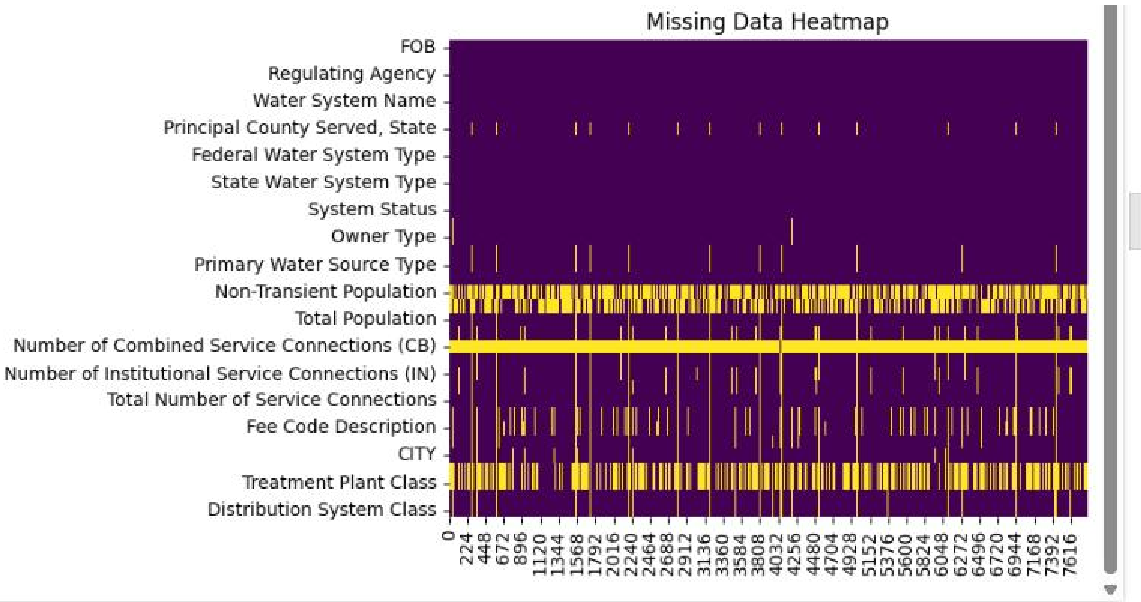


Figure 2 Missing data heat map

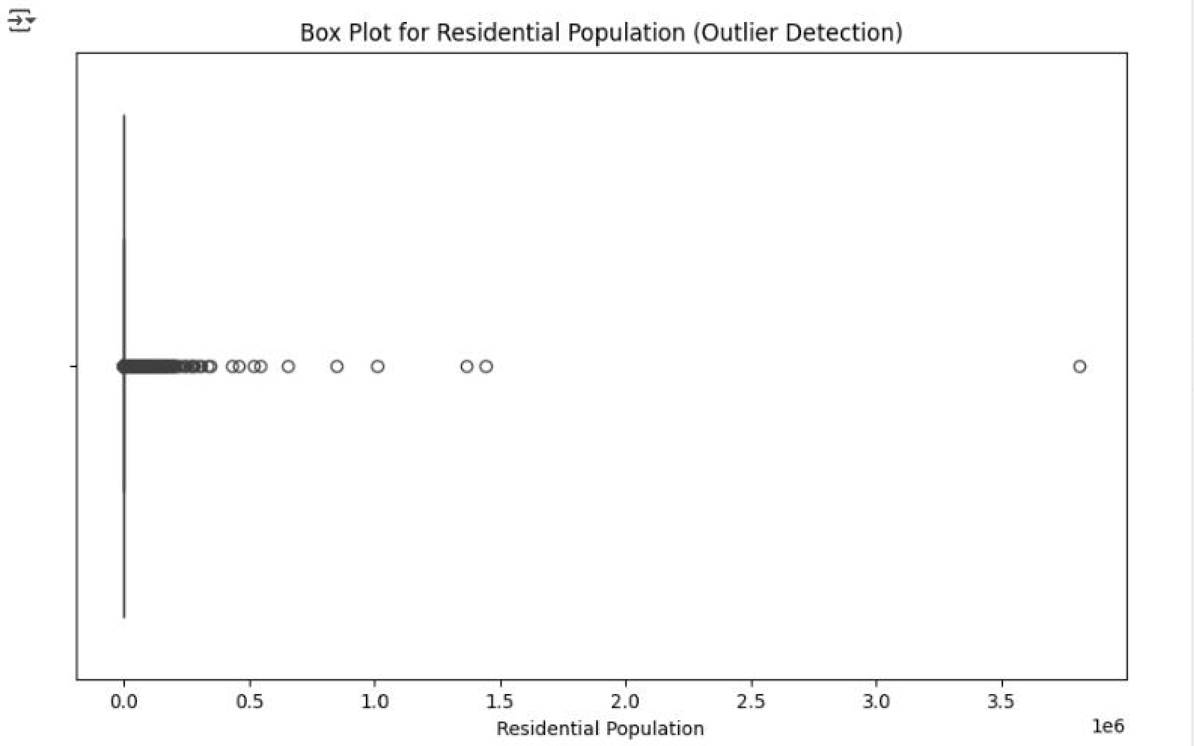
Data Summary

Multiple different values can be found throughout the dataset fields. In the Residential Population variable, the dataset spans from 1 to over 3.8 million and produces a mean value of about 10,400 people. Most water systems operate in small communities, whereas only a few systems provide water services to millions of residents. The Total Population shows a wide range of values that stretch from one single resident to more than eighteen million residents. The broad range of water system dimensions demonstrates how the water infrastructures differ in terms of their characteristics.

The measured values for Service Connections demonstrate extensive variation. The number of Agricultural Service Connections (AG) ranges between 0 and 2,462, which indicates that some systems focus on residential customers while others concentrate on agricultural needs.

Outliers

The box plot shows Residential Population values which surface important outlier measurements. The data contains specific points which extend beyond the boundaries of the interquartile range (IQR). Most water systems serve populations within moderate ranges according to the plot but several larger population systems stand apart from the rest as individual points in the extreme right. The extensive distribution of water system sizes causes potential analysis issues since it introduces unwelcome practical effects unless addressed appropriately. The plot highlights important points about investigating extreme data values since these points could represent both incorrect records and essential significant data points needing special treatment during analysis.



Categorical Variables

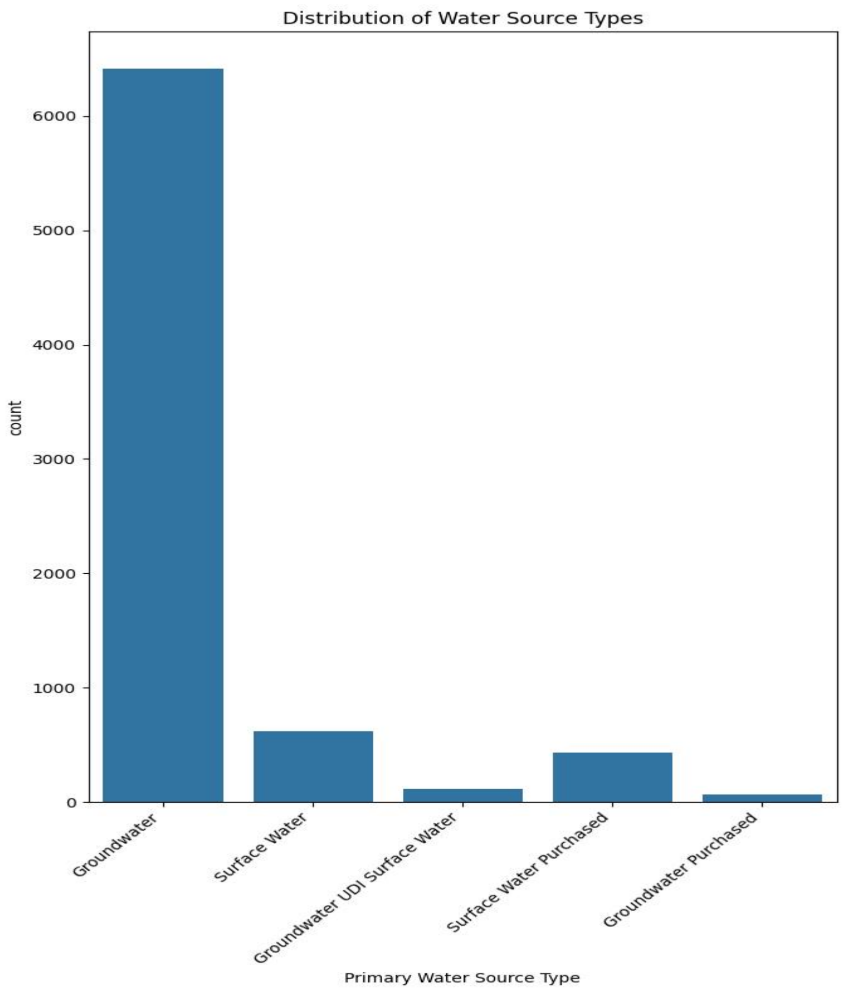
Categorical variables exist as variables that display categories and groups. They show characteristics through defined groups while remaining non-mathematical in nature. Such variables possess specific and predetermined value ranges that function as categories.

The dataset contains categorical variables which consist of two main categories:

The region serves as a classification variable which divides water systems between Southern California and Northern California regions.

The overseeing government organizations called "LPA58 - Napa County" or "District 16 - Central" act as regulating agencies for water systems.The Water System Type variable assigns classifications to water systems through distinct categories, including Community, Non-Transient Non-Community and Transient Non-Community.

This column provides information about the nature of fees by showing classifications such as "Small Community" and "Transient Non-Community Water System."The operational state of water systems is provided under this field which may apply as active or inactive.Researchers analyze categorical variables by counting their frequencies and providing percentages. They typically use bar plots, pie charts, and count plots for visualization.


Distribution of Water Source Types
count
Primary Water Source Type
Groundwater
Surface Water
Groundwater UDI Surface Water
Surface Water Purchased
Groundwater Purchased

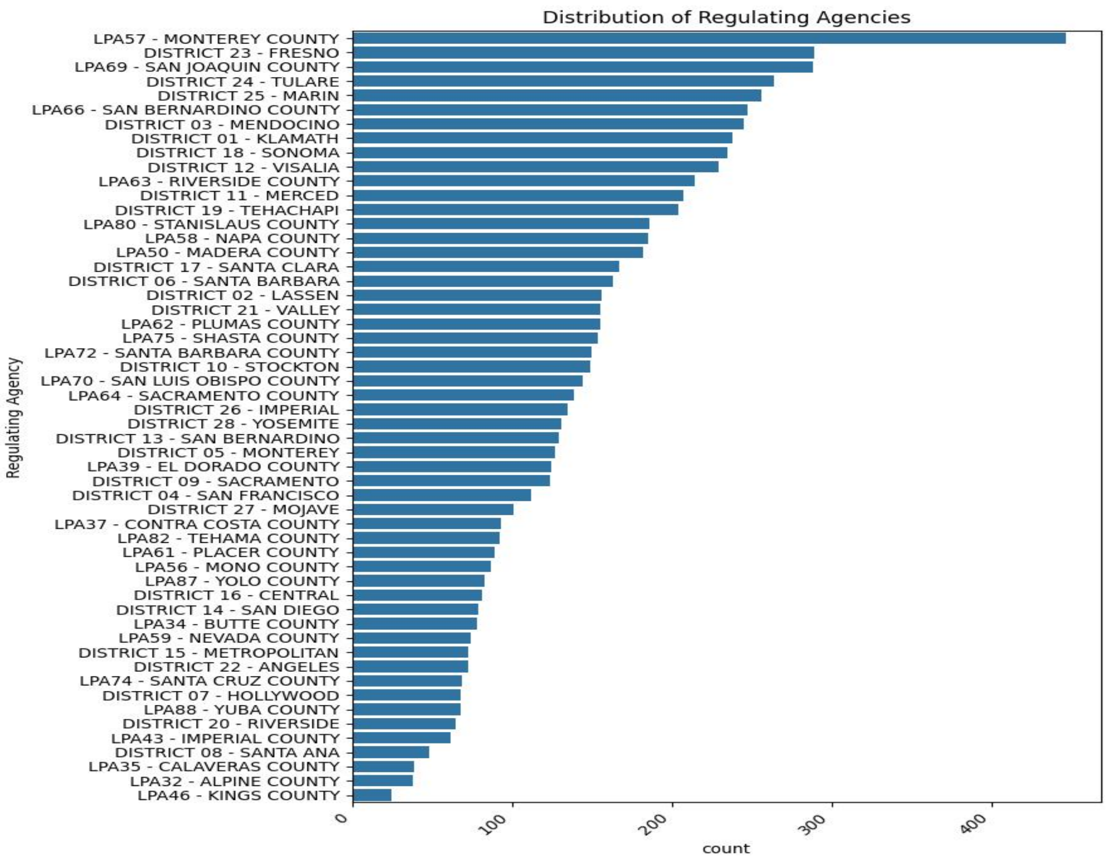

Distribution of Regulating Agencies
Regulating Agency
count
LPA57 - MONTEREY COUNTY
DISTRICT 23 - FRESNO
LPA69 - SAN JOAQUIN COUNTY
DISTRICT 24 - TULARE
DISTRICT 25 - MARIN
LPA66 - SAN BERNARDINO COUNTY
DISTRICT 03 - MENDOCINO
DISTRICT 01 - KLAMATH
DISTRICT 18 - SONOMA
DISTRICT 12 - VISALIA
LPA63 - RIVERSIDE COUNTY
DISTRICT 11 - MERCED
DISTRICT 19 - TEHACHAPI
LPA80 - STANISLAUS COUNTY
LPA58 - NAPA COUNTY
LPA50 - MADERA COUNTY
DISTRICT 17 - SANTA CLARA
DISTRICT 06 - SANTA BARBARA
DISTRICT 02 - LASSEN
DISTRICT 21 - VALLEY
LPA62 - PLUMAS COUNTY
LPA75 - SHASTA COUNTY
LPA72 - SANTA BARBARA COUNTY
DISTRICT 10 - STOCKTON
LPA70 - SAN LUIS OBISPO COUNTY
LPA64 - SACRAMENTO COUNTY
DISTRICT 26 - IMPERIAL
DISTRICT 28 - YOSEMITE
DISTRICT 13 - SAN BERNARDINO
DISTRICT 05 - MONTEREY
LPA39 - EL DORADO COUNTY
DISTRICT 09 - SACRAMENTO
DISTRICT 04 - SAN FRANCISCO
DISTRICT 27 - MOJAVE
LPA37 - CONTRA COSTA COUNTY
LPA82 - TEHAMA COUNTY
LPA61 - PLACER COUNTY
LPA56 - MONO COUNTY
LPA87 - YOLO COUNTY
DISTRICT 16 - CENTRAL
DISTRICT 14 - SAN DIEGO
LPA34 - BUTTE COUNTY
LPA59 - NEVADA COUNTY
DISTRICT 15 - METROPOLITAN
DISTRICT 22 - ANGELES
LPA74 - SANTA CRUZ COUNTY
DISTRICT 07 - HOLLYWOOD
LPA88 - YUBA COUNTY
DISTRICT 20 - RIVERSIDE
LPA43 - IMPERIAL COUNTY
DISTRICT 08 - SANTA ANA
LPA35 - CALAVERAS COUNTY
LPA32 - ALPINE COUNTY
LPA46 - KINGS COUNTY

Figure 3 Visualization of categorical variables

Numerical Variables

Categorical variables exist as variables that display categories and groups. They show characteristics through defined groups while remaining non-mathematical in nature. Such variables possess specific and predetermined value ranges that function as categories.

The dataset contains categorical variables which consist of two main categories:

- The region serves as a classification variable which divides water systems between Southern California and Northern California regions.
- The overseeing government organizations called "LPA58 - Napa County" or "District 16 - Central" act as regulating agencies for water systems.
- The Water System Type variable assigns classifications to water systems through distinct categories, including Community, Non-Transient Non-Community and Transient Non-Community.
- This column provides information about the nature of fees by showing classifications such as "Small Community" and "Transient Non-Community Water System."
- The operational state of water systems is provided under this field which may apply as active or inactive.
- Researchers analyze categorical variables by counting their frequencies and providing percentages. They typically use bar plots, pie charts, and count plots for visualization.

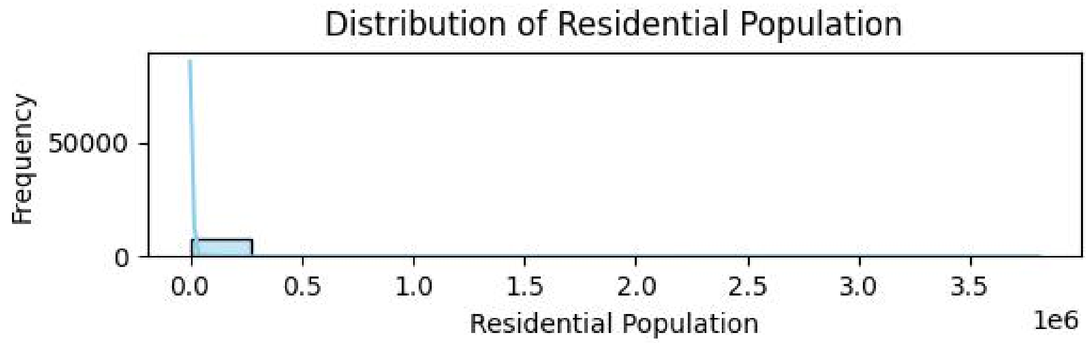


Figure 4 Distribution of population

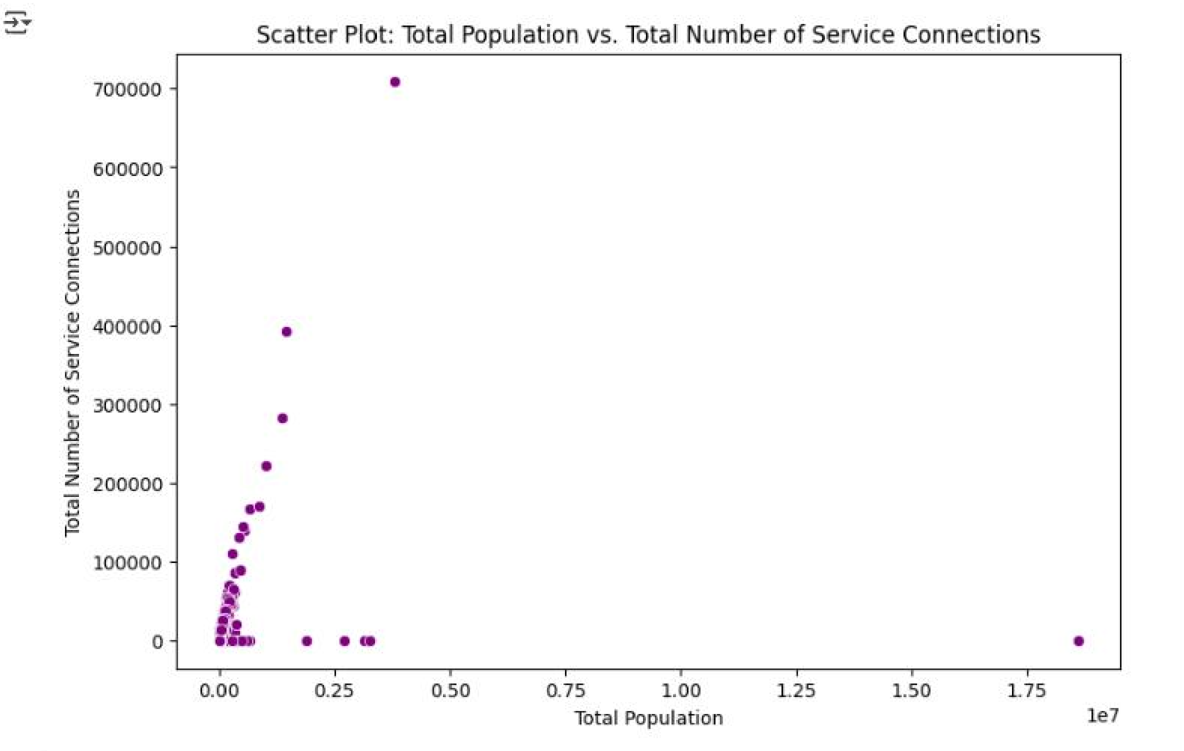


Figure 5 Plot of total number of services connection vs total population

CHAPTER 5: PRELIMINARY RESULTS AND DISCUSSION

The Drinking Water Watch Public Water System Facilities dataset includes 7,815 records that detail the characteristics of 35 variables for water systems throughout California. The main purpose of this research is to analyze water resource availability and distribution quality throughout the state by evaluating regional differences and system types while examining the social and economic effects of water availability. The preliminary findings derived from the data set receive analysis regarding water disparities and drinking water security in California.

Descriptive Statistics and Data Overview

The dataset includes essential variables, which include FOB (region), Water System Name, Regulating Agency, Water System No, Principal County Served, and Primary Water Source Type. These variables let users identify each system by assigning classifications that capture location and management details. The dataset shows that water systems exist in three major California areas known as Southern California, Northern California, and Central California, and each of these parts maintains unique characteristics for its managed water systems.

Research shows that groundwater stands as the main water source for many systems inspected through the dataset while matching California's overall dependence on underground water supplies in rural and urban transitional areas of the San Joaquin Valley and Southern California regions. The crucial water supply role of groundwater exposes these sources to greater vulnerability to pollution that develops safety and quality concerns in specific regions. The discovered preliminary data underscores the necessity for additional research about water supply systems that maintain their operations using groundwater as well as their approaches to managing water quality.

System Status flags down whether individual water supply systems operate or not. Preliminary data reveals that a substantial number of water supplies function actively leading to positive indications about state-wide water availability. The assessment of inactive systems requires closer inspection to determine both causes and effects because they present potential access barriers for specific communities found in rural areas.

```
Descriptive Statistics of Numeric Variables:
                                              count          mean  \
Residential Population                       3764.0  10399.722901
Non-Transient Population                     2314.0    267.784788
Transient Population                         3710.0    581.405391
Total Population                             7642.0  10565.251897
Number of Agricultural Service Connections (AG)  7206.0      3.543714
Number of Combined Service Connections (CB)    22.0     11.045455
Number of Commercial Service Connections (CM)  7507.0    114.879446
Number of Institutional Service Connections (IN) 7242.0      4.543911
Number of Residential Service Connections (RS)  7385.0   1229.964523
Total Number of Service Connections          7636.0   1310.158984

                                                    std  min   25%  \
Residential Population                       80314.024444  1.0  23.0
Non-Transient Population                      1556.268576  1.0  15.0
Transient Population                          4528.555683  1.0  26.0
Total Population                            230200.460994  1.0  45.0
Number of Agricultural Service Connections (AG)     52.946507  0.0   0.0
Number of Combined Service Connections (CB)      34.465272  1.0   1.0
Number of Commercial Service Connections (CM)    1182.584399  0.0   0.0
Number of Institutional Service Connections (IN)    75.835808  0.0   0.0
Number of Residential Service Connections (RS)  10975.441689  0.0   0.0
Total Number of Service Connections          11992.512266  0.0   2.0

                                               50%     75%         max
Residential Population                       109.5  812.25   3806272.0
Non-Transient Population                      43.0  132.00     41686.0
Transient Population                          75.0  200.00    141700.0
Total Population                             113.0  401.75  18606473.0
Number of Agricultural Service Connections (AG)   0.0    0.00      2462.0
Number of Combined Service Connections (CB)     2.0    4.00       164.0
Number of Commercial Service Connections (CM)   1.0    7.00     79294.0
Number of Institutional Service Connections (IN)  0.0    0.00      5415.0
Number of Residential Service Connections (RS)   2.0   52.00    623898.0
Total Number of Service Connections           11.0   62.00    708607.0
```

Regional Disparities in Water Access

The initial dataset analysis reveals an intense and significant difference in how water systems are available and maintained based on geographical location. The data shows Northern California as a region containing numerous community-based water systems which maintain stable and long-term access to water. The areas that have metropolitan structures possess extensive development systems whi which enable them to maintain large human settlements efficiently. The water systems in Central California and Southern California, along with their wide expanses of rural land and farming areas, present a combination of water sources that extend from transient non-community systems to systems that draw water from groundwater supplies.

The county data in Central California demonstrates that Fresno Kern, and Kings tend to operate systems supplying water to mobile populations. The water infrastructure of agricultural regions specifically sections itself toward labour workers and short-term residents instead of long-term residence communities. The urban territories of Los Angeles County, along with San Francisco, maintain established water infrastructure networks to supply their permanent residential bases, while rural counties operate smaller water systems. The contrast shows a substantial economic inequality, which leads rural farmers to struggle more to secure trustworthy and uncontaminated water supplies.

Water System Types and Source Reliability

Both Federal Water System Type and State Water System Type fields analysed the water system types across California. The database uses three classification columns which distinguish community systems from non-transient non-community systems and transient systems because each group operates under unique regulatory frameworks. Community systems represent the most commonly found water systems in urban areas, as these systems receive extensive support for maintaining quality infrastructure. The regulatory frameworks for non-transient, non-community water systems match their compact dimensions as they provide water services for locations including educational institutions and workplaces. Water quality monitoring takes place less consistently among transient water facilities that serve temporary populations such as travellers or seasonal employees in rural locations and tourism zones.

The Primary Water Source Type data shows numerous California water systems utilize groundwater as their main source, even though this presents significant concerns because of contamination problems and excessive groundwater extraction issues throughout the state's aquifer system. Groundwater, which remains accessible in rural regions, exposes itself to high risks from pollutants like nitrates and arsenic, along with other dangerous chemicals that stem from agricultural runoff and industrial pollution. Thorough testing alongside proper water management becomes essential because the Central Valley agricultural regions heavily depend on groundwater resources, which present continuing water contamination risks.

Population Served and Water System Capacity

Each water system within the dataset presents data about Residential populations together with Non-Transient Population, Transient Population, and Total Population information. The designed columns enable an initial examination of water system dimensions to evaluate their serving capability for local populations. The data shows that city-based community water systems provide service to large permanent populations, but rural and touristic regions depend mostly on non-community systems for their transient populations.

A major problem exists regarding the empty fields in the Residential Population and Primary Water Source Type columns because both contain 172 missing data points each. The absence of data reduces our ability to conduct thorough evaluations of water accessibility in different communities. The evaluation requires complete data to correctly portray water accessibility conditions for both urban and rural areas, which face the most significant disparities.

Policy Implications and Water Management Challenges

Preliminary findings about water access and equity conditions in California bring forth major conclusions to the analysis. Active water systems constitute most supply systems across California but certain sections of the state retain substantial differences based on water system infrastructure and reliability alongside service population numbers. The communities of Central and Southern California that operate in rural and farming regions struggle with too much groundwater dependence but metropolitan areas get better and consistent water services.

The study analysis reveals the importance of specific infrastructure funding coupled with solid regulatory monitoring systems in critical water system areas. Defining water equity as an absolute priority of state policymakers remains essential because climate change creates prolonged droughts and rising water scarcity across California and serves overpopulation communities.

Limitations and Next Steps

The presented preliminary findings have useful value, but future stages of analysis need to resolve existing problematic aspects. Additional investigation must be conducted on the data gaps in the Residential Population and Primary Water Source Type fields because they could adversely influence the research outcomes. The assessment of water system quality, particularly aimed at groundwater-dependent systems, requires additional examination to incorporate both contamination levels and treatment standards data.

The research should extend to provide increased depth about regulatory standards which control water systems throughout California, especially within the farmlands and remote areas that demonstrate heightened contamination risks. The complete understanding of water equity in the state depends on adding water quality testing results along with regulatory compliance data.

References

Acquah, S., & Allaire, M. (2023). Disparities in drinking water quality: evidence from California. Water Policy, 25(2). https://doi.org/10.2166/wp.2023.068

Agustí Pérez-Foguet. (2023). Broadening the water affordability approach to monitor the human right to water. Cities, 143, 104573–104573. https://doi.org/10.1016/j.cities.2023.104573

Archer, H., Morello Frosch, R., Pace, C., Baehner, L., & Cushing, L. (2024, August). Drinking water arsenic contamination and COVID-19 outcomes in California, USA. In ISEE Conference Abstracts (Vol. 2024, No. 1). https://ehp.niehs.nih.gov/doi/abs/10.1289/isee.2024.0141\

Carchi, D., Orellana, M., Martínez, A., & Segovia, J. (2023). Affordability and sustainability in the human right to water. Journal of Agribusiness in Developing and Emerging Economies. https://doi.org/10.1108/jadee-06-2023-0151

Chen, H., & Franklin, M. (2023, November 21). Spatio-Temporal Modeling of Surface Water Quality Distribution in California (1956-2023). ArXiv.org. https://doi.org/10.48550/arXiv.2311.12736

Fernandez-Bou, A. S., Rodríguez-Flores, J. M., Guzman, A., Ortiz-Partida, J. P., Classen-Rodriguez, L. M., Sánchez-Pérez, P. A., ... & Medellín-Azuara, J. (2023). Water, environment, and socioeconomic justice in California: A multi-benefit cropland repurposing framework. Science of the Total Environment, 858, 159963. https://www.sciencedirect.com/science/article/pii/S0048969722070632

Goddard, J. J., Ray, I., & Balazs, C. (2021). Water affordability and human right to water implications in California. PLOS ONE, 16(1), e0245237. https://doi.org/10.1371/journal.pone.0245237

Hanak, E., & Chappelle, C. (2025). California's Water Quality Challenges - Public Policy Institute of California. Public Policy Institute of California. https://www.ppic.org/publication/californias-water-quality-challenges/

London, J. K., Fencl, A. L., Watterson, S., Choueiri, Y., Seaton, P., Jarin, J., ... & Bailey, C. (2021). Disadvantaged unincorporated communities and the struggle for water justice in California. Water Alternatives, 14(2), 520-545. https://www.water-alternatives.org/index.php/alldoc/articles/vol14/v14issue2/626-a14-2-4/file

Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., ... & Kaatz, L. (2021). A low-to-no snow future and its impacts on water resources in the western United States. Nature Reviews Earth & Environment, 2(11), 800-819. https://www.nature.com/articles/s43017-021-00219-y

Xenarios, S., Edwards, E. Y., & Buurman, J. (2025). Water Tariffs and Affordability in Urban Water Supply and Wastewater Systems. Water Economics and Policy. https://doi.org/10.1142/s2382624x24500164