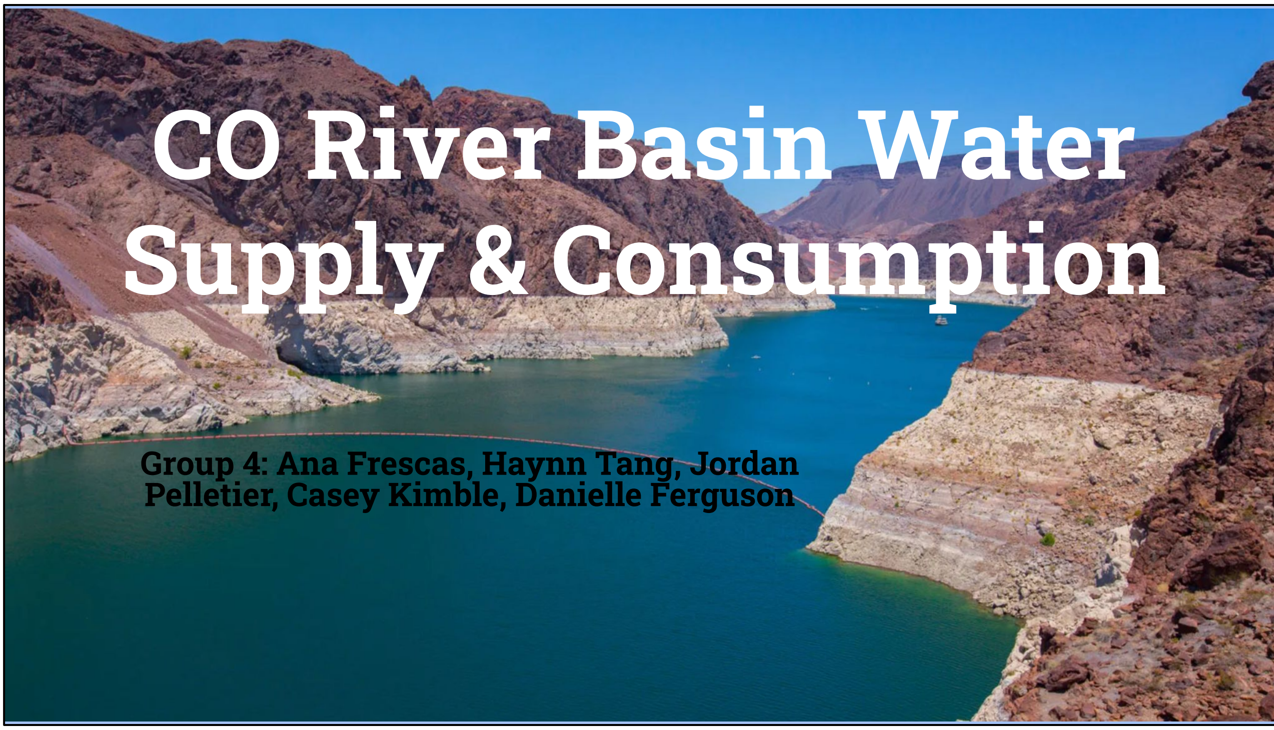
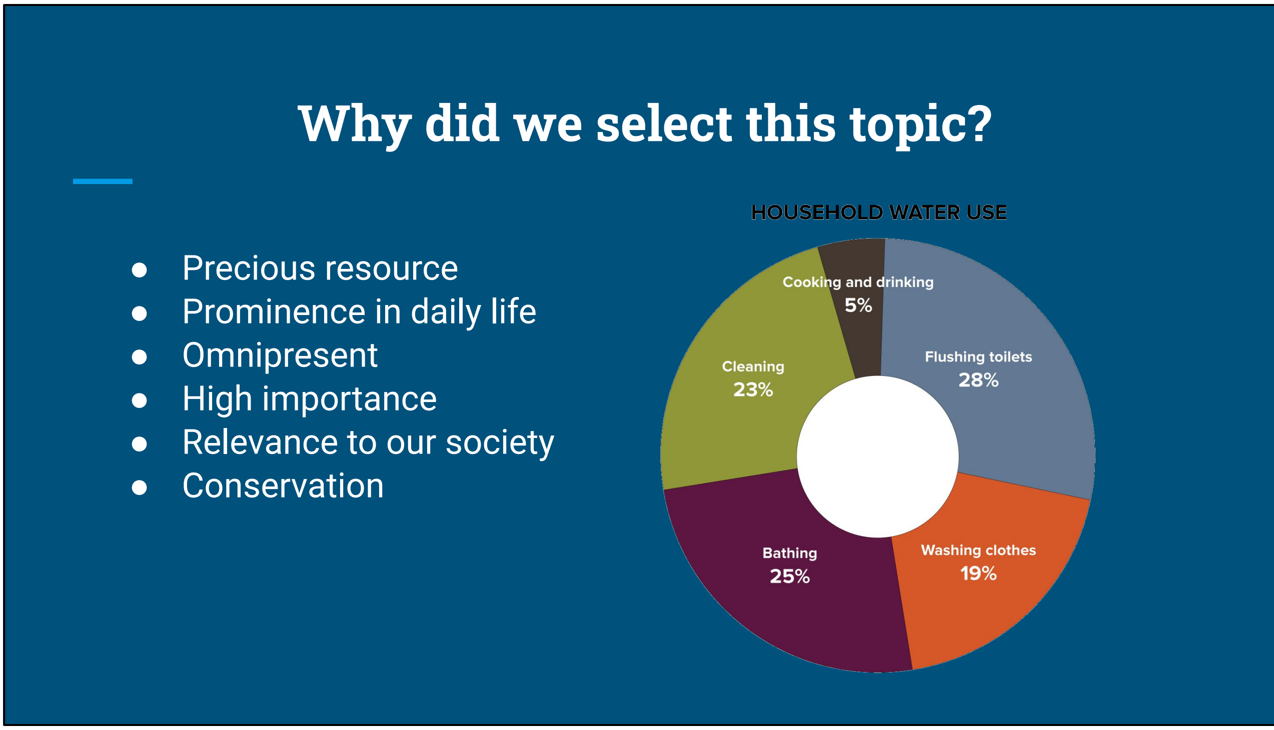

CO River Basin Water Supply & Consumption
Group 4: Ana Frescas, Haynn Tang, Jordan Pelletier, Casey Kimble, Danielle Ferguson



The Colorado River is being overused, nearly 36 million people rely on the river. As the Colorado river starts to decrease in water levels. The population that relies on it, will start to suffer and the consequences are disastrous.

Water is one of our most precious resources because humanity couldn't exist without it. In addition to its key role in our survival, water is essential in nearly every facet of modern life from industrial processes to domestic pursuits. However, water is omnipresent and so we don't always take the time to consider how much we use, when and where we use it, and why we need it. We selected this topic due to its high importance and relevance to our society. We want to predict how our water usage and supply might look in the future to make recommendations for conservation and preparation.

The USGS, U.S. Geological Survey, website maintains national databases of water-use information. The data are collected and compiled every five years for each State, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. We have compiled the datasets from 1985-2015, 7 total, into a csv file and have sorted by the 7 states in the CO basin. The data contains water-use information that looks at how each of the 7 states uses water,the total amount of water withdrawals, and withdrawals broken down by type of water (fresh or saline) and withdrawal type (groundwater and surface water). The amount of withdrawals for water-use categories is also listed. All numbers are measured in million gallons per day (Mgal/d).

Questions:

Where is the water going and how do the 7 states compare?

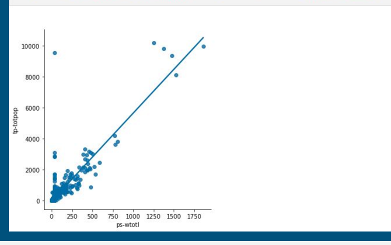
What is the predicted decrease of water levels?

Which factors have the most impact in water consumption reduction?

Data Exploration

- Finding the data
- Compiling the data
- Cleaning the data
- Analyzing the data

```
# visualizing how ps-wtotl(Public Supply - Total withdrawals in million gallons/day)
sns.lmplot(x="ps-wtotl", y="tp-totpop", data= temp_df, fit_reg= True, ci= None )
```




```
#1985 sort by state
test_1985 = pd.DataFrame()

State_val = ['WY', 'UT', 'NM', 'CO','AZ', 'CA', 'NV']

for val in State_val:
    test_1985=test_1985.append(DF_1985.loc[DF_1985['state'] == (val)])

#1985 rename state/year/area column/population total/public supply pop total/pub supply gw total
test_1985.rename(columns = {'state':'STATE', 'year': 'YEAR', 'area':'CountyCode', 'po-total':'TotalPop', 'ps-popto':'PS-TOPop','ps-popgw'

#Drop columns
test_1985 = test_1985.drop(columns = ['agency',
 'ps-wswfr',
'ps-wgwsa',
 'ps-wswsa',
'co-wgwfr',
```

Our group decided on water usage as a topic of general interest, and it took some time for us to find datasets that would work with our concept. We finally settled on the USGS water usage data.

The group then spent a good deal of time cleaning this data and merging the datasets by using python jupyter notebook.

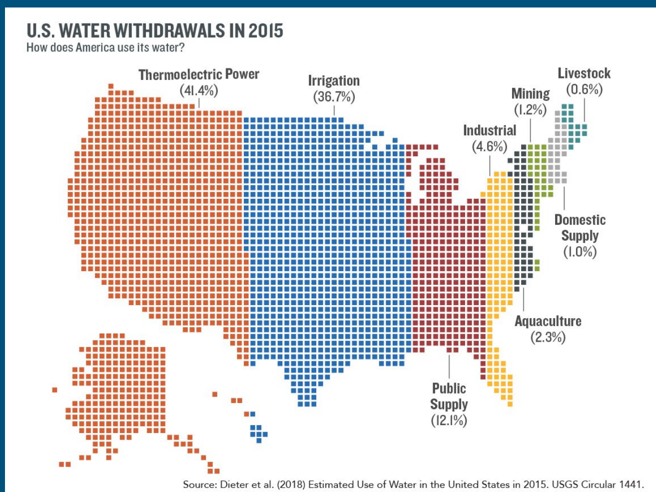
The mock database was created using Postgress/PGAdmin.

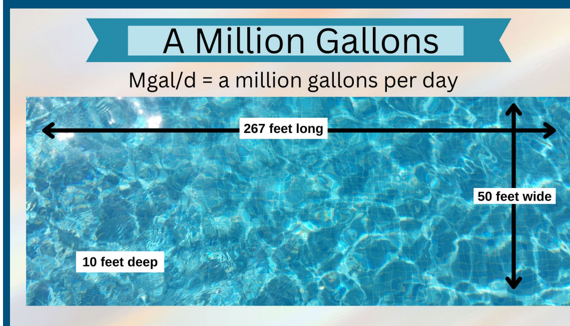
Our group struggled with making our data work with several different machine learning models in jupyter notebook,

so we turned to Tableau for our analysis to review the data through visualizations.

Our website dashboard was created by using HTML/Javascript.

Data Analysis





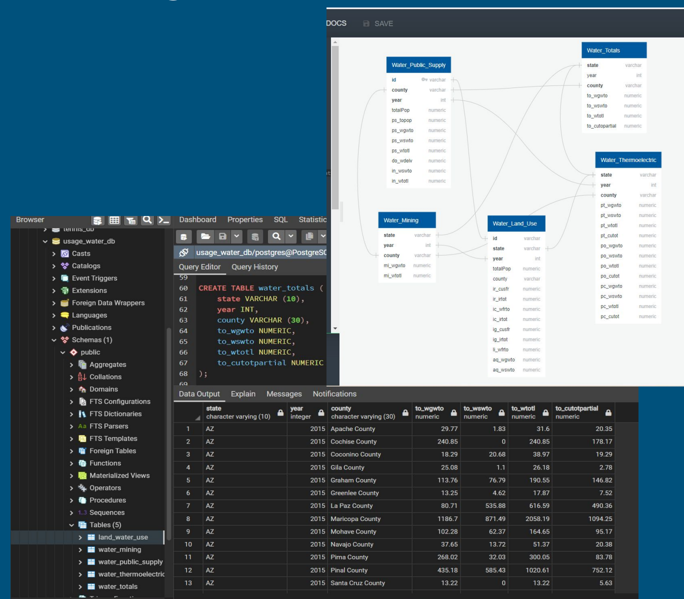
The 7 states we are analyzing make up the Colorado River Basin…..The Colorado River flows for approx. 1,450 miles and it provides water to each state....the 7 states are…. Colorado, California, Arizona, Utah, Wyoming, Nevada, and New Mexico..

The years we analyzed were 1985, 1990, 1995, 2000, 2005, 2010, and 2015.

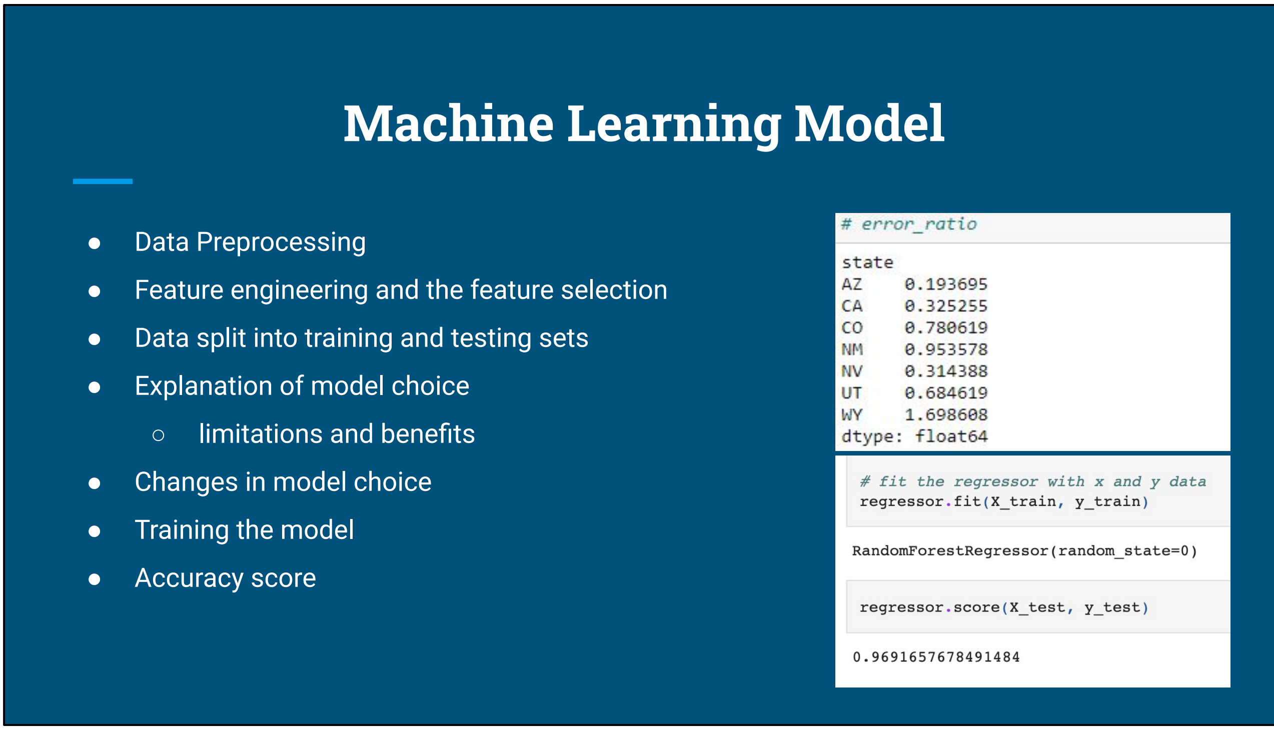
The analysis phase of the project involved analyzing the total withdrawals and total domestic water use in million gallons per day (Mgal/d) per state and per year. We looked at how each state uses water by looking at withdrawals and withdrawal type.

Database Integration

- Database stores static data
- Database tables
- Join using database language
- Database connection string



We used an ERD to group together categories of water in order to analyze the data further. According to the USGS website, Public Supply water usage is water that is withdrawn by public and/or private water suppliers in order to provide water for domestic, industrial and commercial purposes. This is why in the ERD we included the public supply, domestic and industrial water totals. We then grouped Irrigation, Irrigation Crop, Irrigation Golf, Livestock and Aquaculture together. We grouped these categories together because in some capacity the water withdrawals are used for land or animals. Aquaculture is also included within this group because before the year 2000 aquaculture was counted as a part of the livestock category. All of the thermoelectric categories are grouped together because before the year 2000, the subcategories thermoelectric once-through and thermoelectric recirculation were a part of the thermoelectric withdrawals as a whole. We also grouped all the mining categories together. Lastly, we grouped together overall water withdrawal totals, these totals included all freshwater and saline-water withdrawals. I then was able to create these tables within postgres and join a few of the tables together.



Our preprocessing involved a lot of cleaning of the data. This included selecting which columns or areas we wanted to test as well as making sure each column was the same throughout all the datasets as several either changed names, or were added or dropped throughout the years. A lot of the columns included a groundwater or a saline withdrawal calculation as well as a total calculation for each category. This led to a lot of repetitive or unneeded columns in our dataset. Therefore, we selected columns that displayed the total withdrawals to test against the population each year.

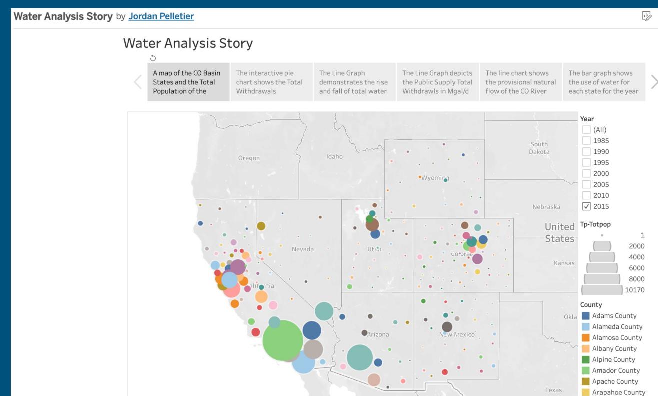
We originally wanted to test a linear regression model. To test this, we decided it would be best to split the training data into intervals from 1985 through 2005 for our training set and 2010 through 2015 as our testing set.

We started with a linear regression because it is one of the easiest ways of showing a direct relationship between the population and water usage. One of the limits for this, as we found out, was that the relationship between water consumption and population was not as linear as we expected and is affected by other aspects which could influence water consumption. We switched to a Random Forest Regression model as that requires less preprocessing and had more accurate predictions. However, the random forest regression model is vulnerable to overfitting, especially when multiple columns share some of the same data.

In our original testing of our Linear Regression model, the predicted outcomes were very inaccurate, as seen in the picture to the right. Notably, New Mexico and Wyoming were off by large amounts. Our Random Forest Regression model had a high amount of accuracy but several of our columns fed into each which led to probable overfitting.

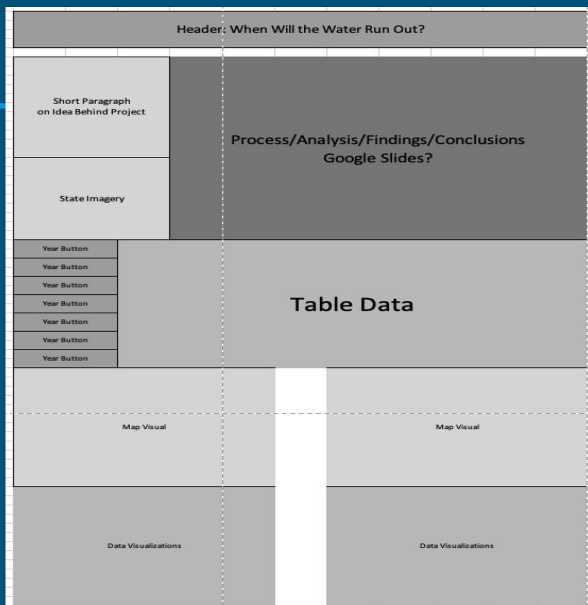
Dashboard Tools

- [Tableau]
- HTML/Javascript Website



The geographical nature of our data lends well to using maps created in Tableau to show analysis and predictions based on usage and availability. We also used Tableau to create charts relevant to our findings, and to pair with our analysis.

Storyboard Blueprint



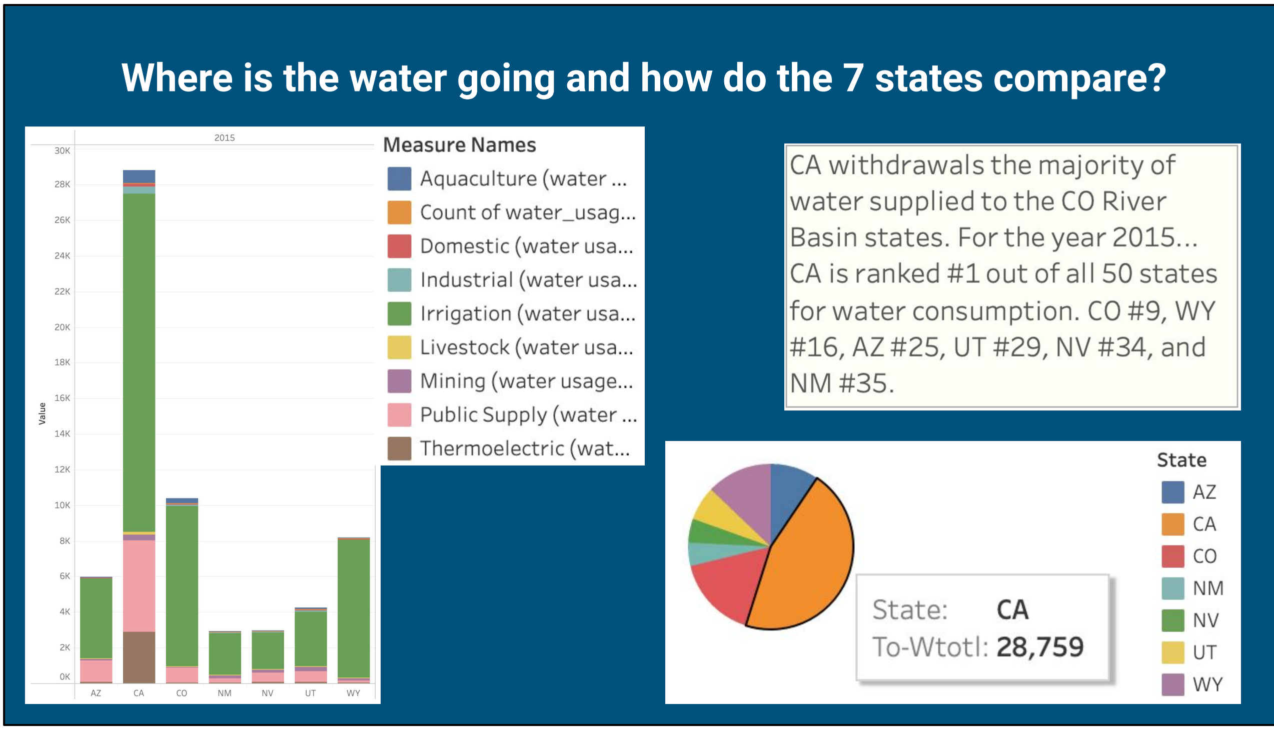
Website Rough Draft - building using HTML, Javascript and CSS. Incorporating Tableau Visualizations, Google Slides and Open Source Imagery.

```css
.jumbotron {
  color: antiquewhite;
}

.jumbotron {
  background:
  linear-gradient(
    rgba(197, 230, 250, 0.25),
    rgba(197, 250, 246, 0.25)
  ),
  url("../images/CO_River_Grand_Canyon.jpeg");
  background-repeat: no-repeat;
  background-attachment: fixed;
  /* color:white !important; */
  height:50vh;
  text-align: center;
  font-family: 'Verdana', sans-serif;
  font-size: 30px;
  font-weight: bold;
  }


  hr.rounded {
  border-top: 8px solid #1375b0;
  border-radius: 5px;
  }
```

Part of our group’s plan to present our data was to build a website to visualize our “why” and present some of our findings. Here is our original rough draft layout. We used HTML with Javascript and CSS to build and style our website, and embedded Tableau visualizations, google slides and open source imagery to add visualizations to our site to support our analysis.



The 7 states we analyzed are ranked from low to high out of all 50 states by the total amount of water withdrawals based on the 2015 dataset, which was our most recent data. You can see the pie chart here shows overall withdrawals by state, with California using almost half of the CO River Basin's water. The bar chart shows us a breakdown of withdrawals by type or industry.

While it may be obvious that states with a higher population like California and Colorado are going to lose a larger percentage of their water to the public supply, irrigation is actually the biggest user of water by percentage in all states, shown in green here on the bar chart, with California using almost 20,000 million gallons per day. We often hear about residents being asked to conserve water but having more regulation on irrigation would have a much larger effect on water usage and conservation. California also uses a good deal of water for thermoelectric energy, more than some states are using per day, across all categories. However, most of that water is saline. Most of the saline in Colorado River water actually comes from human activity such as irrigation, but some of it is simply due to the transfer of water and sediment/minerals from one basin to another.

#35. New Mexico

Total withdrawals: 2,900 million gallons per day (Mgal/d)

- Fresh water withdrawals: 2,810 Mgal/d
- Saline water withdrawals: 89 Mgal/d
- Surface water withdrawals: 1,460 Mgal/d
- Groundwater withdrawals: 1,440 Mgal/d

Total domestic water use: 170 Mgal/d (81 gal per capita, #20 highest for all states)

- Public supply: 262 Mgal/d (81 gal per capita, #21 highest for all states) --- 1.8 million people served (86% of population)
- Self-supplied: 25 Mgal/d (84 gal per capita, #18 highest for all states) --- 292,000 people served (14% of population)

Non-public total water use: 2,609 Mgal/d

- Irrigation: 2,370 Mgal/d
- Livestock: 32 Mgal/d
- Aquaculture: 24 Mgal/d
- Industrial: 3 Mgal/d (fresh), 0 Mgal/d (saline)
- Mining: 57 Mgal/d (fresh), 89 Mgal/d (saline)
- Thermoelectric power: 34 Mgal/d (fresh), 0 Mgal/d (saline)

#34. Nevada

Total withdrawals: 2,960 million gallons per day (Mgal/d)

- Fresh water withdrawals: 2,880 Mgal/d
- Saline water withdrawals: 82 Mgal/d
- Surface water withdrawals: 1,520 Mgal/d
- Groundwater withdrawals: 1,440 Mgal/d

Total domestic water use: 365 Mgal/d (126 gal per capita, #6 highest for all states)

- Public supply: 531 Mgal/d (122 gal per capita, #7 highest for all states) --- 2.7 million people served (93% of population)
- Self-supplied: 36 Mgal/d (186 gal per capita, #1 highest for all states) --- 193,000 people served (7% of population)

Non-public total water use: 2,401 Mgal/d

#29. Utah

Total withdrawals: 4,230 million gallons per day (Mgal/d)

- Fresh water withdrawals: 3,880 Mgal/d
- Saline water withdrawals: 350 Mgal/d
- Surface water withdrawals: 3,080 Mgal/d
- Groundwater withdrawals: 1,150 Mgal/d

Total domestic water use: 506 Mgal/d (169 gal per capita, #2 highest for all states)

- Public supply: 627 Mgal/d (169 gal per capita, #3 highest for all states) --- 2.9 million people served (98% of population)
- Self-supplied: 10 Mgal/d (169 gal per capita, #3 highest for all states) --- 61,400 people served (2% of population)

Non-public total water use: 3,593 Mgal/d

- Irrigation: 3,030 Mgal/d
- Livestock: 16 Mgal/d
- Aquaculture: 83 Mgal/d
- Industrial: 54 Mgal/d (fresh), 79 Mgal/d (saline)
- Mining: 3 Mgal/d (fresh), 258 Mgal/d (saline)
- Thermoelectric power: 61 Mgal/d (fresh), 8 Mgal/d (saline)

#25. Arizona

Total withdrawals: 5,980 million gallons per day (Mgal/d)

- Fresh water withdrawals: 5,980 Mgal/d
- Saline water withdrawals: 0 Mgal/d
- Surface water withdrawals: 3,220 Mgal/d
- Groundwater withdrawals: 2,760 Mgal/d

Total domestic water use: 987 Mgal/d (145 gal per capita, #4 highest for all states)

- Public supply: 1,200 Mgal/d (146 gal per capita, #4 highest for all states) --- 6.6 million people served (97% of population)
- Self-supplied: 24 Mgal/d (110 gal per capita, #10 highest for all states) --- 218,000 people served (3% of population)

- Thermoelectric power: 84 Mgal/d (fresh), 0 Mgal/d (saline)

#16. Wyoming

Total withdrawals: 8,140 million gallons per day (Mgal/d)

- Fresh water withdrawals: 8,050 Mgal/d
- Saline water withdrawals: 97 Mgal/d
- Surface water withdrawals: 7,400 Mgal/d
- Groundwater withdrawals: 748 Mgal/d

Total domestic water use: 91 Mgal/d (156 gal per capita, #3 highest for all states)

- Public supply: 101 Mgal/d (176 gal per capita, #2 highest for all states) --- 467,000 people served (80% of population)
- Self-supplied: 9 Mgal/d (75 gal per capita, #21 lowest for all states) --- 119,000 people served (20% of population)

Non-public total water use: 8,036 Mgal/d

- Irrigation: 7,790 Mgal/d
- Livestock: 16 Mgal/d
- Aquaculture: 29 Mgal/d
- Industrial: 8 Mgal/d (fresh), 0 Mgal/d (saline)
- Mining: 45 Mgal/d (fresh), 97 Mgal/d (saline)
- Thermoelectric power: 52 Mgal/d (fresh), 0 Mgal/d (saline)

#9. Colorado

Total withdrawals: 10,300 million gallons per day (Mgal/d)

- Fresh water withdrawals: 10,300 Mgal/d
- Saline water withdrawals: 24 Mgal/d
- Surface water withdrawals: 8,800 Mgal/d
- Groundwater withdrawals: 1,530 Mgal/d

Total domestic water use: 672 Mgal/d (123 gal per capita, #7 highest for all states)

- Public supply: 844 Mgal/d (123 gal per capita, #6 highest for all states) --- 5.2 million people served (95% of population)
- Self-supplied: 35 Mgal/d (123 gal per capita, #5 highest for all states) ---

- Mining: 8 Mgal/d (fresh), 24 Mgal/d (saline)
- Thermoelectric power: 37 Mgal/d (fresh), 0 Mgal/d (saline)

#1. California

Total withdrawals: 28,800 million gallons per day (Mgal/d)

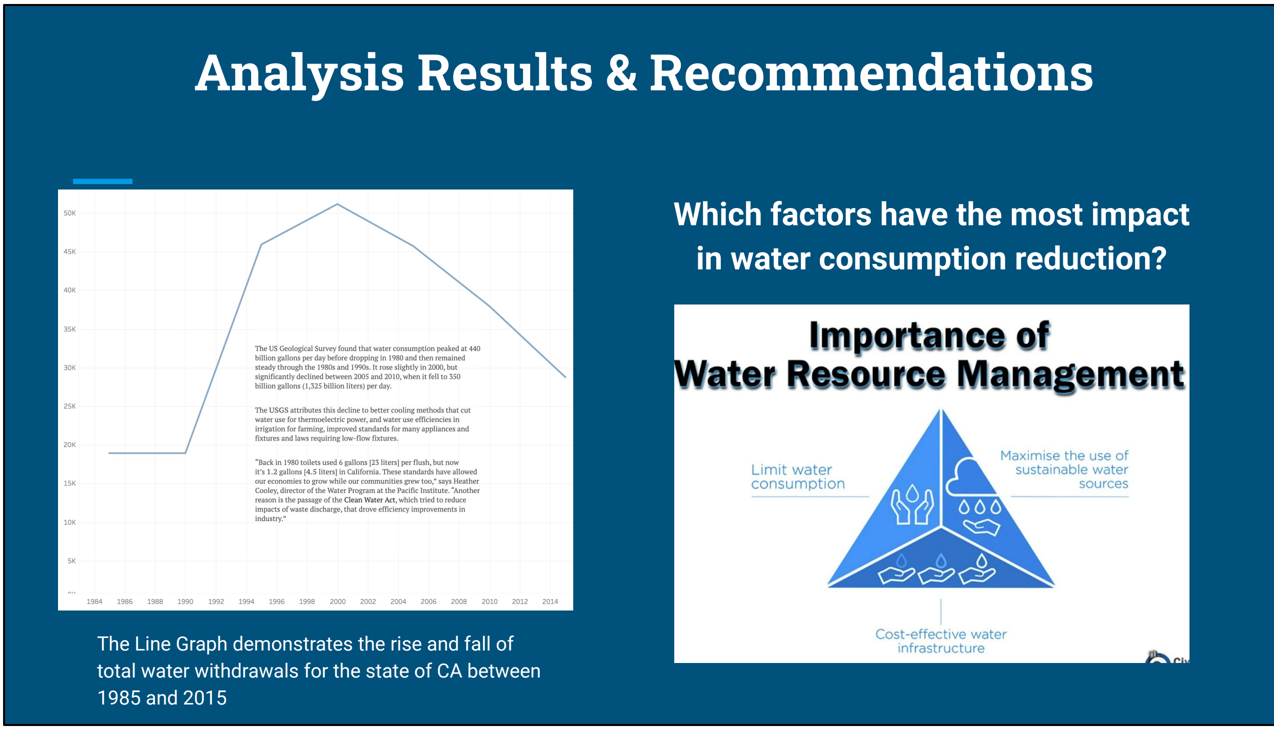
- Fresh water withdrawals: 25,600 Mgal/d
- Saline water withdrawals: 3,160 Mgal/d
- Surface water withdrawals: 11,300 Mgal/d
- Groundwater withdrawals: 17,400 Mgal/d

Total domestic water use: 3,350 Mgal/d (86 gal per capita, #17 highest for all states)

- Public supply: 5,150 Mgal/d (86 gal per capita, #17 highest for all states) --- 37.7 million people served (96% of population)
- Self-supplied: 127 Mgal/d (88 gal per capita, #16 highest for all states) --- 1.4 million people served (4% of population)

Non-public total water use: 23,503 Mgal/d

- Irrigation: 19,000 Mgal/d
- Livestock: 183 Mgal/d
- Aquaculture: 727 Mgal/d
- Industrial: 399 Mgal/d (fresh), 0 Mgal/d (saline)
- Mining: 46 Mgal/d (fresh), 272 Mgal/d (saline)
- Thermoelectric power: 36 Mgal/d (fresh), 2,840 Mgal/d (saline)

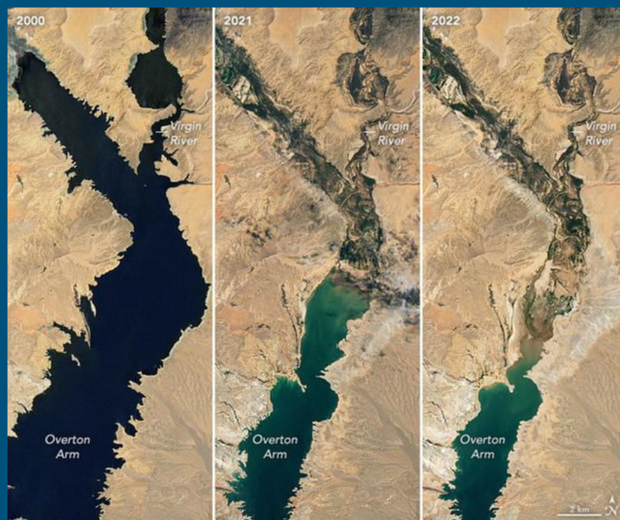


As mentioned during the last slide, the burden of reducing water consumption is often put on the public and individual consumers. Things like low flow toilets, water conservation while brushing teeth and taking showers.

“Back in 1980 toilets used 6 gallons [23 liters] per flush, but now it’s 1.2 gallons [4.5 liters] in California. These standards have allowed our economies to grow while our communities grew too,” says Heather Cooley, director of the Water Program at the Pacific Institute. “Another reason is the passage of the Clean Water Act, which tried to reduce impacts of waste discharge, that drove efficiency improvements in industry.”

But as we have seen from the data, industrial regulation would have a much greater impact on water conservation in the united states.

What is the predicted decrease of water levels?



NASA EARTH OBSERVATORY

Beyond what we know about where the water is going, climate change is playing its own role in the lack of replenishment of said water, with record low water levels in areas like Lake Mead and Lake Powell.

Some data from our more in-depth analysis:

“A new study predicts that a hot and dry climate may lead to a 29% decline in the

Upper CO River basin “baseflow” at the basin outlet by 2050s”

https://www.usgs.gov/news/state-news-release/groundwater-flow-colorado-river-may-decline-third-over-next-30-years

“The average flow of the Colorado River has already declined nearly 20% since 2000, with half of that attributable to rising temperatures. Temperatures in the Basin are predicted to rise another 2-5 degrees Fahrenheit by 2050, which could reduce river flows by another 10 to 40%.”

https://www.nature.org/en-us/about-us/where-we-work/priority-landscapes/colorado-river/colorado-river-in-crisis/

Recommendations for Decreasing Water Consumption

- Starting Water Banks
- Using Nature Based Solutions
- Rethinking Agriculture
- Upgrading Old Infrastructure

- Starting water banks. In Utah, these are voluntary agreements among users to share a water resource. In exchange for not using their share of water, users receive compensation and legal protection for their water rights. Water banks create more flexibility than the “use it or lose it” rule that has traditionally governed water rights in the West.
- Using nature-based solutions. About 90% of the water in the Colorado River Basin starts as snow or precipitation

- in forests, making these landscapes our largest natural reservoirs. The Nature Conservancy is working throughout the Basin to advance forest management that mitigates wildfire danger in critical watersheds. They're also working with landowners on the Upper Green in Wyoming to restore wet meadows, improving the land's ability to retain water through low-tech infrastructure.
- Rethinking agriculture. In the Verde Valley in Arizona, crop switching is used to reuse water and more efficient irrigation to keep water flowing.
- Upgrading old infrastructure. Obsolete irrigation systems can waste a lot of water.

What would we have done differently?

Next Steps...

What would have helped answer our original predictive questions would be to get to know our datasets more before investing so much time cleaning and organizing them, with the purpose of being able to determine more firm predictions for a machine learning model. Our group also attempted to find additional datasets for deeper analysis of our original info but were unsuccessful in doing so. We believe this could have helped answer our original question which was looking into when the water would run out.