AI Prediction of Michigan Water Use

With the human population still increasing, the demand on industries to produce consumer goods is increasing as well. To keep up with demand, these industries must increasingly tax their limited supply of resources. This especially becomes a problem for shared resources like water. The Great Lakes are a huge source of shared water across multiple industries. It provides a source of revenue for recreational businesses, helps generate hydroelectric energy, and is used to irrigate crops and rear livestock.

In June 2024, Martusiuk aggregated the Department of Environment, Great Lakes, and Energy's water usage data for the Great Lakes Basin from 2013 to 2022. This decade's worth of data will be used to gain insight into the increasing demands on Michigan's water supply. The dataset includes a list of the various Michigan counties, industries that reside within those counties, and their annual water consumption. The data does not include population growth data which will be supplemented by the U.S. Census Bureau.

The primary objective of this project is predicting the future impact on Michigan's water supply based on archived industry and geographical water consumption levels. More specifically, this project will showcase how water consumption changes over time across industries and counties. Population growth data will also be included as an explanatory variable for the increased demand on the industries. With the annual population data, correlations may be identified across the industries to explain changes in demand. Additionally, highlighting the relationships between industries' water consumption will provide insight into any underlying effects on water use.

Although water use data is not available for 2023, the predicted results will be ready to compare with the actual water use results when they become publicly available. Data from the past decade is useful in developing prediction models for 2025 and forecasting future trends. Additionally, aggregate data on the total available water is currently inaccessible. This data would be crucial in identifying the threshold at which the water basin would be critically impacted; so future studies will be required to determine preventative measures. Despite the missing water level data, measuring the current impacts is an important step in predicting high-risk industries.

Policy reform is the biggest use-case for this project. Identifying the most impactful geographical or industrial segments, will enable policy managers to lobby for environmental protection, or penalize industries from consuming too much of the market. Additionally, competitive industries may use this data to identify market opportunities to develop more efficient processes to limit their consumption. Finally, environmental lobbyists will be interested in the impact of the various industries to better target their efforts.

By constructing machine learning models to forecast the future impacts of industries, Michigan's water use data can serve as a comparison to analyze the other seven states that border the Great Lakes.

References

Martusiuk, O. (n.d.). *Michigan water use data (2013 to 2022)* [Data set]. Kaggle. https://www.kaggle.com/datasets/oleksiimartusiuk/michigan-water-use-data-2013-to-2022/data

U.S. Census Bureau. (n.d.). *State population totals: 2010-2020* [Data set]. U.S. Census Bureau. https://www2.census.gov/programs-surveys/popest/datasets/2010-2020/state/totals/

U.S. Census Bureau. (n.d.). *State population totals: 2020-2023* [Data set]. U.S. Census Bureau. https://www2.census.gov/programs-surveys/popest/datasets/2020-2023/state/totals/