GLOBAL FRESHWATER USE & WITHDRAWALS (1901–2021)

FINAL PROJECT REPORT

DATA 5310 – Data Visualization

TEAM MEMBERS: PAVAN ARRAM, JEEVITHA BANDI, NIKHIL GHUGARE

1. INTRODUCTION

Freshwater availability is one of the most critical environmental concerns of the 21st century. Understanding how freshwater is used across time and across countries is essential for managing limited water resources, anticipating risks of scarcity, and designing sustainable policy interventions. In this project, we analyze historical global freshwater use from 1901–2014 and national freshwater withdrawal data from 1962–2021 to understand long-term trends, regional disparities, and reporting behavior.

Our datasets comes from KAGGLE, which compiles global and country-level annual water withdrawal statistics. The global dataset spans over a century, while the country-level dataset includes over 6,000 observations from nearly 200 countries. These data, while rich, also present challenges:

- Missing values are common in Africa, small island nations, and early years.
- ISO codes are inconsistent, requiring cleaning for geospatial matching.
- Withdrawals vary by orders of magnitude, making interpretation difficult without transformation.
- Reporting behavior has changed dramatically over time, affecting analysis.

Despite these nuances, the dataset provides a strong foundation for exploring meaningful questions about freshwater use, temporal trends, and global inequality in consumption.

2. MOTIVATION AND RESEARCH QUESTIONS

Our team selected five core questions that reflect diverse aspects of freshwater usage patterns, data reporting behavior, and global disparities. These questions are deliberately varied to ensure a comprehensive analysis:

1. How has global freshwater use changed from 1901–2014?
2. Which decades experienced the fastest and slowest growth in freshwater consumption?
3. Which countries withdraw the most freshwater, and how does this change over time?
4. How do the top water-using countries compare in their withdrawal trends?
5. What does the global distribution of freshwater withdrawals look like, and how does it evolve over time?

These questions are meaningful because they address both macro-level (global) and micro-level (country-specific) behaviors, identify inequalities, and highlight changes in environmental reporting quality.

3. DATA PREPARATION

Our data preparation process began by organizing the freshwater withdrawal dataset obtained from Kaggle, which contained multiple historical records compiled from international water-use databases. The raw file included several structural inconsistencies, such as duplicated fields (e.g., *Freshwater use.1*), which we removed to avoid redundancy. To maintain clarity throughout the analysis, we standardized all variable

names, renaming fields such as *Freshwater use* and *Annual freshwater withdrawals* into consistent identifiers like Global_Use_BCM and Withdrawal_BCM, ensuring uniform formatting across both global and country-level datasets.

We then addressed missing or incomplete values. Records without withdrawal amounts were excluded because they could not meaningfully contribute to the quantitative analysis. We also examined entries with missing ISO3 country codes, identifying roughly 96 cases with incomplete or non-standard identifiers. These gaps typically occurred in earlier decades or in regions where national water reporting systems have historically been weaker, particularly parts of Africa. Understanding these omissions was important for interpreting the temporal and spatial coverage of the dataset.

A substantial amount of effort went into correcting country name mismatches between the Kaggle dataset and the Natural Earth TopoJSON file used for our choropleth mapping. Several country labels in the Kaggle data did not align with the naming conventions of the map file. To prevent these countries from disappearing in the final visualization, we manually reconciled mismatches such as converting “United States” to “United States of America,” “Congo” to “Republic of the Congo,” “Micronesia” to “Federated States of Micronesia,” and “Eswatini” to “Swaziland.” These adjustments were essential to ensure full geographic coverage in the map representation.

Because annual withdrawals ranged from approximately 10^6 to 10^{12} cubic meters, the values spanned several orders of magnitude. Visualizing them directly would compress most countries into the lower range while exaggerating only the largest users. To create a meaningful distribution plot, we applied a $\log_{10}$ transformation, allowing us to compare small and large users on a consistent scale and reveal structural patterns in global water use.

Finally, we generated a separate dataset to measure the number of countries reporting water data each year from 1962 to 2021. By grouping the Kaggle records by year and counting non-missing entries, we produced a metric that reflects global reporting coverage over time. This helped us evaluate data reliability and understand when the dataset becomes sufficiently complete for meaningful global analysis.

These preparation steps collectively ensured that the dataset was clean, consistent, and well-structured, allowing our visualizations to accurately reflect true global freshwater use patterns.

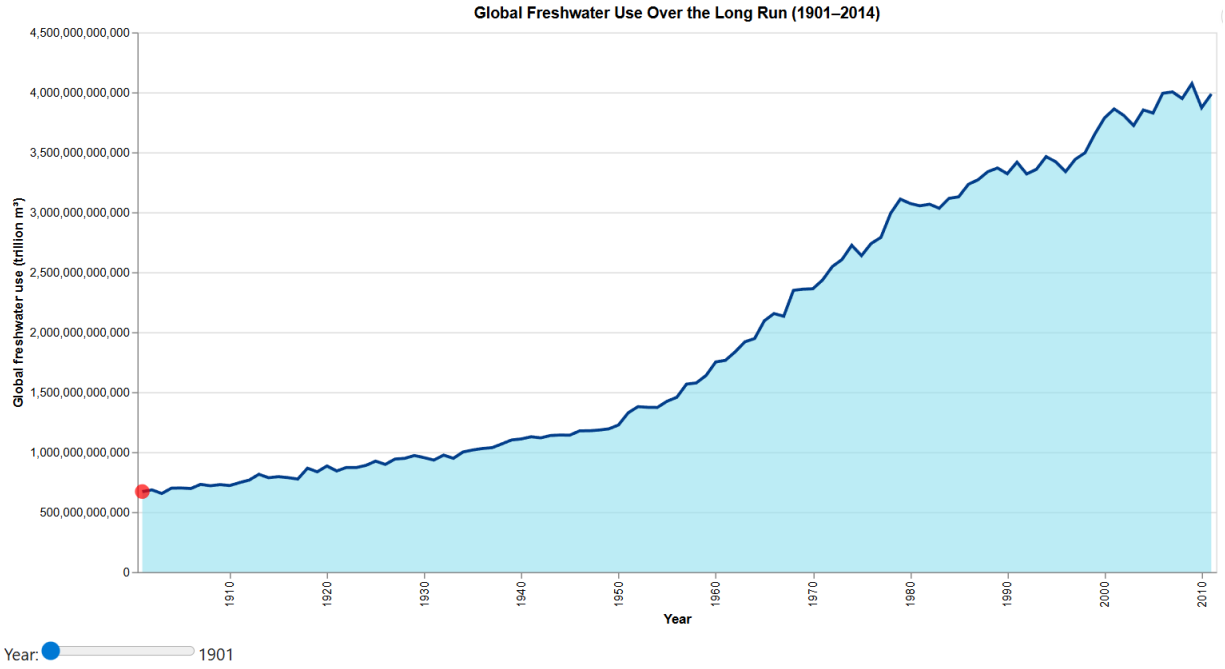
4. ANALYSIS AND RESULTS

4.1. Plot 1 – Global Freshwater Use (1901–2014)

Over the years, our use of freshwater has seen some significant changes. If we look back from 1901 to 2014, it’s clear that we’ve dramatically ramped up our consumption of this vital resource. From 1901 to around 1950, the increase was gradual but steady. It wasn't until after 1950 that we really started to see a sharp rise in freshwater withdrawals. This surge was largely fueled by industrialization, a booming

population, and the expansion of irrigated agriculture to feed everyone. The most remarkable growth happened between 1950 and 1990, a time marked by the Green Revolution and rapid economic development, especially in Asia. It was during these decades that our demand for freshwater really skyrocketed.

However, post-1990, the growth trend started to level off, indicating that in some areas, we might have hit a point where we're using water more efficiently or have stabilized our consumption levels. In summary, since the early 20th century, global freshwater use has increased more than fivefold, with the fastest growth occurring after the mid-20th century.

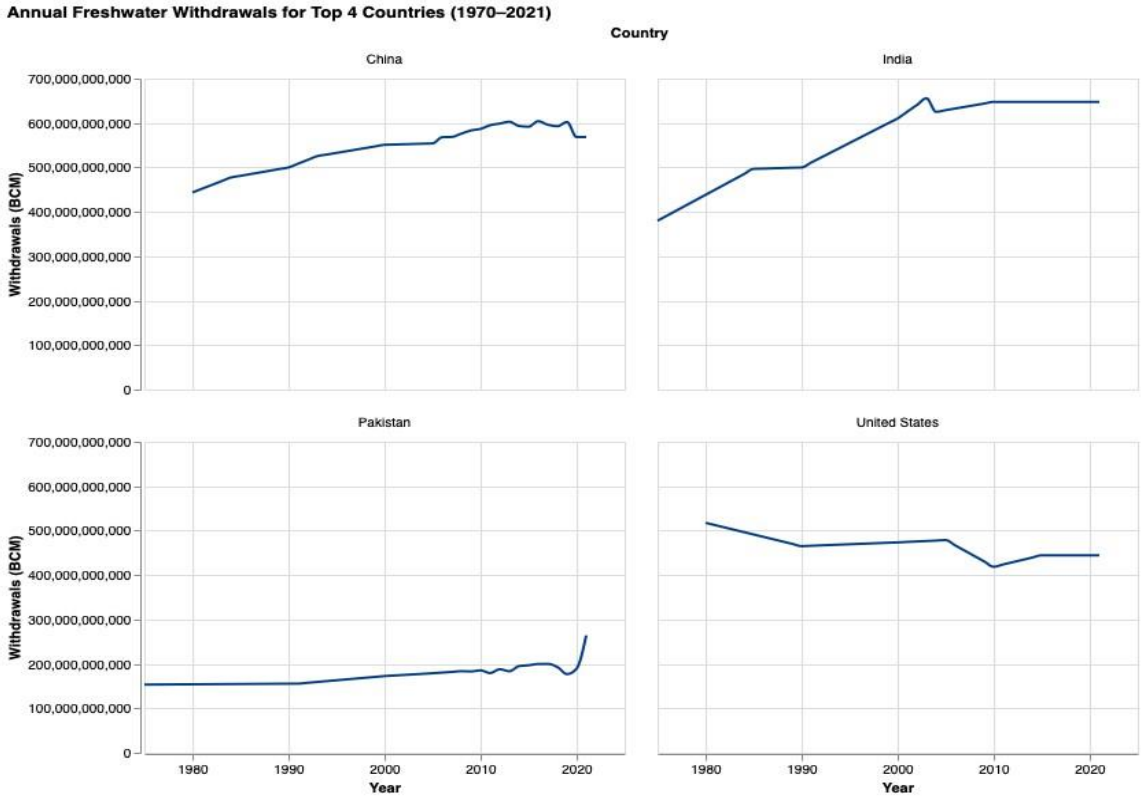


4.2.Plot 2 – Annual Withdrawals for Top 4 Countries (1970–2021)

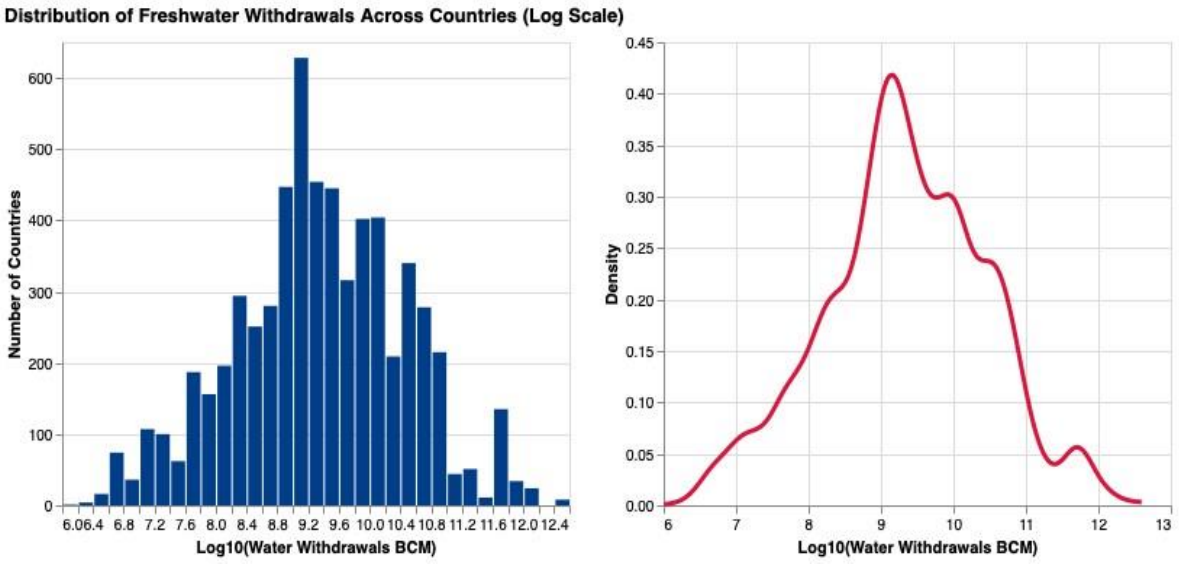
When we look at global water usage trends, it's clear that the fastest growth occurred between the 1950s and the 1980s. During this time, we saw a significant increase in water use, driven largely by agricultural expansion, the mechanization of farming, increased fertilizer use, and overall industrial development. Countries like China and India became major players in water consumption as their economies grew.
In contrast, the period from the early 1900s to the 1940s experienced much slower growth. During these decades, industrial and agricultural activities were on a smaller scale, and the systems for managing global water use were not as advanced.

More recently, between 2000 and 2014, growth in water usage began to plateau. This slowdown is largely due to improved water-use efficiency and stricter policy regulations, as well as the saturation of demand in agricultural water use.

In summary, the decades from 1950 to 1980 saw the most rapid growth in water consumption, while the early 1900s and the period after 2000 had much slower increases in usage.



4.3. Plot 3 – Histogram + Density (Log Scale)

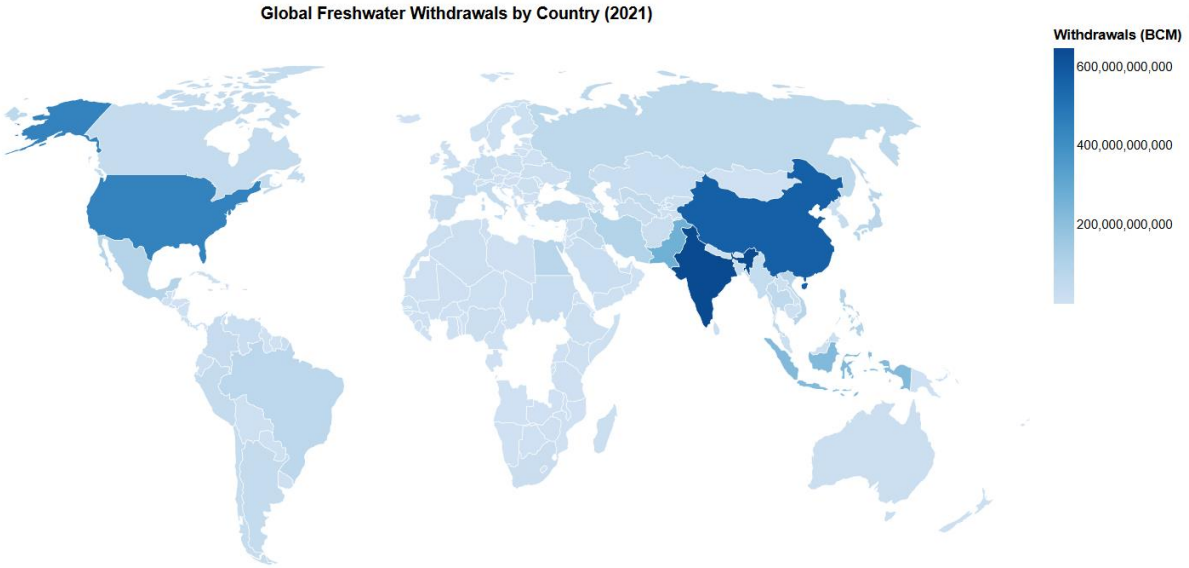


When looking at freshwater usage globally, a few countries stand out prominently. India leads the pack with the highest overall withdrawals, followed closely by China, which has seen its water usage grow rapidly. The United States comes next; although its water withdrawals are high, they have stabilized since around the 1980s, thanks to various efficiency improvements. Pakistan has significant agricultural water demands, while Brazil, Indonesia, and Bangladesh fluctuate in their rankings depending on the year.

If we consider changes over time, both India and China demonstrate strong, long-term growth in water usage. In contrast, the United States peaked in the 1980s and has since seen a decline in usage. Pakistan, on the other hand, has shown steady growth in its water demands. Most countries in Africa and Europe maintain low to moderate levels of freshwater withdrawals.

In conclusion, while India, China, and the United States are the dominant players in global freshwater use, their patterns and trends in water consumption reveal interesting differences over time.

4.4.Plot 4 – Choropleth Map (2021 Withdrawals)



When we look at the faceted time-series plots of water consumption across different countries, several distinct trends emerge that help answer the question of how water usage varies globally.

Starting with **India**, we see a steep and consistent increase in water consumption over the past 60 years. By 2021, India had become the largest consumer of water, a trend that reflects its reliance on intensive agriculture and significant groundwater resources.

In **China**, water consumption began to experience major growth in the 1970s, coinciding with its rapid industrial boom. However, in recent years, this growth has started to stabilize, suggesting a shift in demand patterns.

Turning to the **United States**, the picture is somewhat different. Here, we saw high levels of water withdrawals between 1960 and 1980. After that period, due to regulations, the adoption of more efficient irrigation methods, and shifts in industrial practices, water withdrawals have declined significantly.

Finally, in **Pakistan**, there's been a steady upward trend in water consumption, largely driven by agriculture that relies heavily on water resources.

In conclusion, it's clear that different countries exhibit distinct patterns in water usage. Rapidly growing economies like India and China tend to see quick increases, whereas developed nations like the United States tend to stabilize or even decrease their consumption over time.

4.5. Plot 5 – Countries Reporting Freshwater Data (1962–2021)

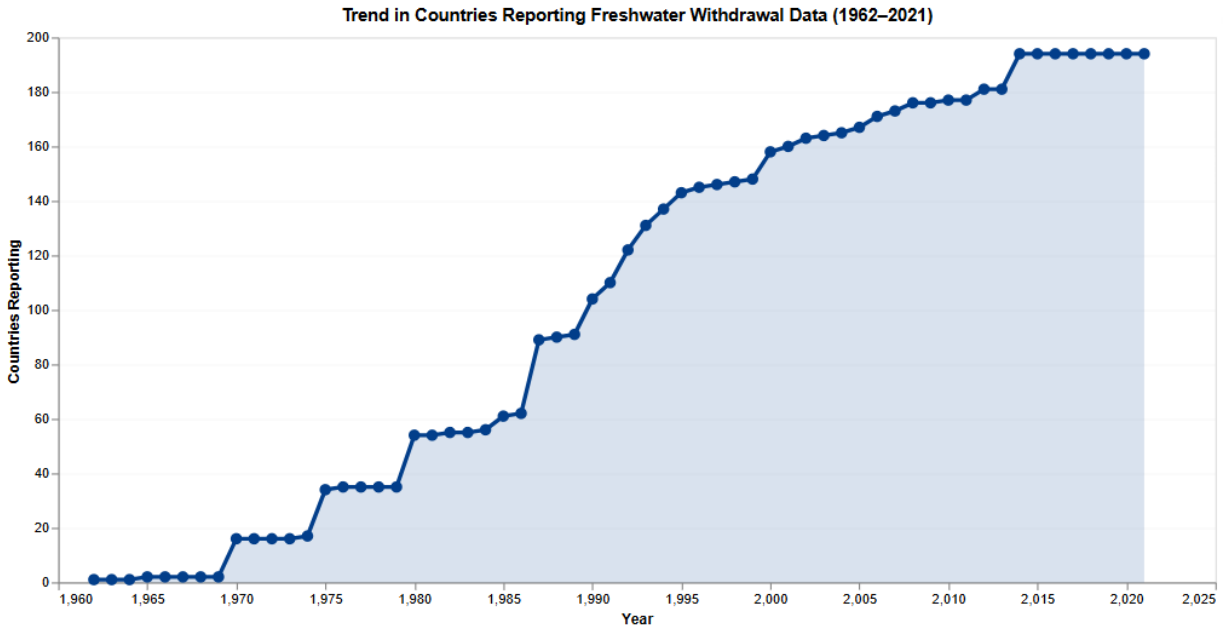
When we take a closer look at the histogram and the density on a log scale, along with the choropleth map, we can see some interesting patterns regarding water usage across different countries.

Firstly, it's clear that the distribution of water usage is quite skewed. Most countries are in the lower range, using very little water—typically between 10^7 and 10^9 cubic meters per year. However, there's a small group of countries that stand out because they consume enormous amounts, with figures reaching into the 10^{11} to 10^{12} cubic meters per year. This significant inequality in water withdrawal volumes has remained consistent over time.

Geographically, Asia really stands out as the leader in high water withdrawals, particularly countries like India, China, and Pakistan. North America also plays a crucial role, mainly due to the massive consumption by the U.S. On the other hand, Africa shows a different trend with generally low water usage, but there is a noticeable increase in participation over the years. Europe, meanwhile, tends to maintain moderate and stable levels of water use.

It's also worth noting the trend in reporting. Back in the 1960s, only a handful of countries reported their water usage data, but now nearly all UN nations are involved. This increase in reporting has significantly improved our understanding of global water distribution patterns in the last few decades.

In conclusion, while global water distribution has always been skewed, with just a few countries using the majority of the world's freshwater—we're now in a much better position to analyze and understand these patterns thanks to improved data coverage.



5. DISCUSSION

The results of our analysis strongly reinforce the research questions we set out to explore. Taken together, the plots reveal meaningful long-term patterns in freshwater use, deeper structural inequalities across countries, and a clear evolution in global reporting practices. Below, we expand on each major finding.

5.1.Global Demand is Increasing and Uneven

Freshwater withdrawals have risen more than fivefold since 1901, with the sharpest increase between 1950–1990. This period aligns with industrial expansion, population growth, and the rise of irrigation-driven agriculture. The growth is not uniform across countries, highlighting major differences in development and water dependency.

5.2. A Few Countries Dominate Global Use

India, China, and the United States consistently withdraw the largest volumes of water. Their combined impact shapes global freshwater demand and raises concerns related to groundwater depletion, agricultural sustainability, and water-sharing across national borders. These countries are central to any policy changes or conservation efforts.

5.3. Strong Inequality in the Distribution

The histogram and density plot show an extremely skewed distribution: most countries withdraw modest amounts, while a small group consumes orders of magnitude more. This structural imbalance suggests that global water stress is driven by a limited set of high-usage nations.

5.4. Data Coverage Has Improved

The number of reporting countries increased sharply after 1980, improving the reliability of more recent estimates. Earlier decades remain sparse, which introduces some uncertainty in long-term comparisons.

5.5. Limitations

- Missing or inconsistent data, especially in parts of Africa
- Combined reporting of sectors (agriculture, industry, domestic), limiting detailed interpretation
- Apparent plateaus may reflect reporting gaps rather than real behavior

Despite these limitations, the visualizations collectively offer a clear and informative picture of how freshwater withdrawals have changed over time and where the major pressures on global resources lie.

6. CONCLUSION

This project combined interactive time-series graphics, faceted country comparisons, geospatial mapping, and distribution analysis to provide a multidimensional view of how global freshwater withdrawals have changed over the past century. Across these visualizations, a consistent message emerges: global freshwater demand has risen steadily, and this rise is driven disproportionately by a small number of countries with intensive agricultural and industrial activity.

The century-long trend highlights sustained growth in withdrawals, particularly during the mid-20th century when population expansion and agricultural modernization accelerated water use worldwide. Our faceted analyses reinforce that while many countries withdraw modest amounts, nations such as India, China, and the United States dominate overall global usage, shaping the trajectory of water resource pressures.

At the same time, the reporting-trend visualization reveals that global participation in freshwater data collection has improved significantly since the 1980s. This strengthens confidence in modern estimates but also emphasizes the need for caution when interpreting early-century patterns, where incomplete reporting may mask variability. Regional gaps—especially across parts of Africa—underscore the importance of continued investment in international data cooperation and uniform reporting standards.

Taken together, these findings point to an urgent need for sustainable water management practices, particularly in high-consumption regions, and for ongoing global monitoring to ensure reliable assessment of our shared freshwater resources. The project demonstrates how thoughtful visualization can help uncover structural patterns that are difficult to see in raw numerical form and can guide policy discussions around long-term water security.

7. REFERENCES

- Natural Earth. TopoJSON geographic boundary files used for world map projections. https://www.naturalearthdata.com/
- Altair Visualization Library. Grammar-of-graphics statistical visualization tool used for interactive chart construction. https://altair-viz.github.io/
- Vega-Lite. High-level visualization grammar for building scalable charts and interactive elements. https://vega.github.io/vega-lite/
- World Bank Open Data. Indicators and methodological references on freshwater withdrawals and environmental reporting.
https://data.worldbank.org/indicator/ER.H2O.FWTL.K3
- ChatGPT by OpenAI. Used as an assistive tool for clarifying concepts, identifying inconsistencies, refining explanations, and correcting errors in written analysis and code documentation. https://openai.com/chatgpt