

A Visualization Tool for Global Wastewater Treatment Plants

Michael Yin
University of British Columbia
Vancouver, BC, Canada
jiyin@cs.ubc.ca

Madonna Huang
University of British Columbia
Vancouver, BC, Canada
hhuang99@student.ubc.ca

ABSTRACT

1 INTRODUCTION

Water pollution from industrial chemicals has become a rapidly pressing concern for the global society. Pollutants arising from chemical agents can have significant, long-term deleterious effects on human water consumption, marine life, and climate change - salvaging our water sources requires a massive effort on the part of environmental and economic policy [9]. One important facet in addressing the water pollution crisis is that of water purification. Pollutant-contaminated water is coined as wastewater, and the purifying process of removing the solids in wastewater is termed wastewater treatment [21]. Wastewater treatment is a complex and ever-evolving process that may involve a number of steps involving the physical removal of solids, decompositions of minerals through chemical interactions, membrane technology to separate based on molecular weights, etc [2, 11, 21]. Globally, wastewater treatment presently takes place on-site at highly specialized facilities called wastewater treatment plants (WWTPs).

Although WWTPs all share similar goals of water purification, they vary greatly in terms of design and treatment processes, which vary depending on regulation and requirements. For instance, country-level regulation and industrial processes can affect both the needs for wastewater treatment as well as specific requirements for technology and innovation [18]. The consequences of urbanization and population growth in a rapidly changing world have also affected the need for WWTPs to provide safe and sufficiently abundant drinking water [24]. As the needs of each community served by a WWTP can vary, WWTP themselves can vary by a number of different metrics, such as the number of people served, treatment efficiency, energy efficiency, etc [16, 20]. Certain WWTPs may also run into a variety of different challenges, mainly revolving around energy usage and process demands [8].

A recent dataset released by Ehalt Macedo et. al. provides global metrics regarding WWTPs, including population served, dilution factor, etc [5]. The researchers additionally provide an initial exploration of the dataset, highlighting the relationship between WWTP metrics and the river basins along which they are located. However, we find that the current analysis is rather esoteric, and that this dataset provides the potential for any user to understand and perform analysis on WWTP metrics. The recency of this published dataset, coupled with the importance of the application, motivates our present proposal. In terms of personal expertise and motivation, none of our group members have any particular research tie into the area, but we identify it as a pressing issue within our lives, and something that we share an interest in exploring.

Thus, in this project, we aim to develop an interactive tool to visualize and compare the efficiency metrics and degree of treatment for global wastewater treatment plants. Beginning with data and

task abstraction, we would first decompose the complex, high-level problem into a more workable, specific set of requirements. Based on our requirements, we would build out our visualization tool, which tentatively would primarily focus on aspects of browsing, searching, and comparison. More specifically, our tool would ideally allow users to understand wastewater treatment statistics at a per-country basis at a glance, allowing more fine-grain information through a zoom function and detailed juxtapositions between countries through a side-by-side comparison tool. We additionally provide a couple of exploratory use cases of our tool for summarizing WWTP metrics within a region of interest as well as through comparing and contrasting country-level metrics.

2 RELATED WORK

In order to contextualize the scope of our visualization design, we consider past research into ecological assessment analysis, WWTP visualizations, geovisualization, and interactive designs.

2.1 Existing Ecological Assessments

There have been significant efforts in both academia and governments to assess the impacts of wastewater effluent on ecosystems and biodiversity. Researchers examine factors such as the ratio between the discharge of WWTP effluent and the natural discharge of the nearby water bodies to provide better regulations of WWTP outflows [13]. Other studies require consistent and comparable geospatial data of WWTP in water quality modeling to identify the source of pollutants discharged into the water cycle [5]. Municipal wastewater treatment is also of great concern in environmental protection planning. The Canadian government reports the population served and discharge volume of WWTPs annually to analyze the level of wastewater treatment and risk of wastewater contaminates at a national scale [15]. However, WWTP-related visualization (vis) is lacking in academic papers and government publications except for statistical analysis results and simple graphics like static visualization of a single attribute.

2.2 Existing WWTP Data Visualizations

Our work builds on HydroWASTE, a global database of 58,502 WWTPs and their characteristics published by researchers from McGill University in 2021 [5]. HydroWASTE is the first database to combine all regional data worldwide and include derived attributes missing in the source datasets [5]. Ehalt Macedo et al. focus on documenting how they manipulate and augment datasets to construct HydroWASTE in their paper [5]. The visualization published with the database therefore only includes limited information: the quality of attributes, ratios of treated wastewater discharge, and statistical analysis results using geographic maps like the topographic terrain map and other common visual encodings like scatter plots [5].

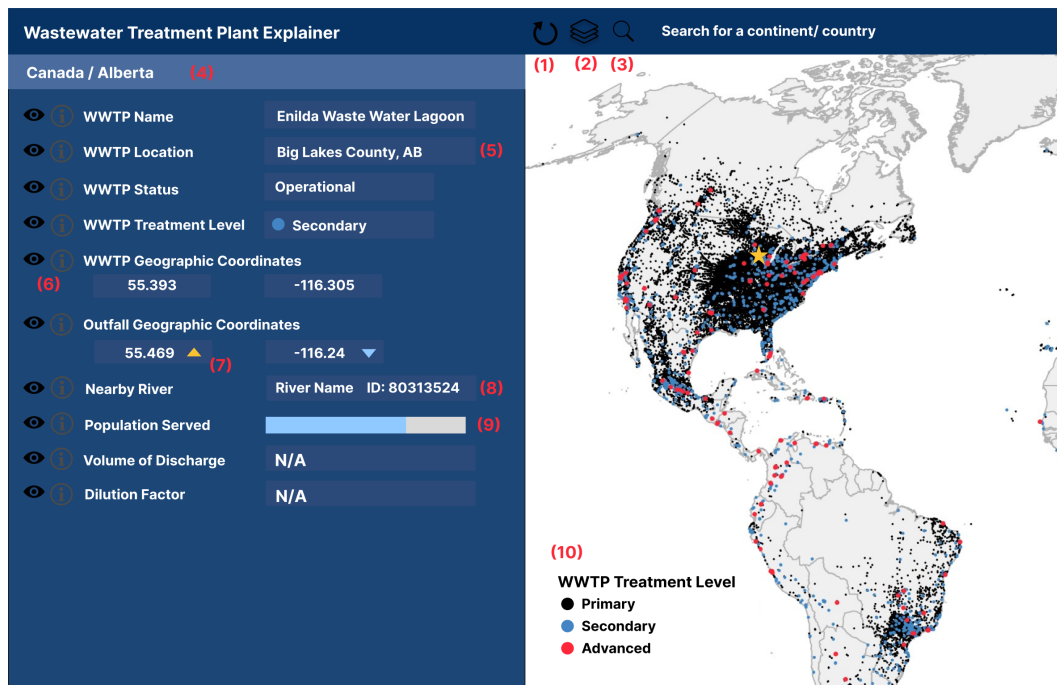


Figure 1: Mockup of a dot map. The color scheme and icons are only tentative due to technical limitations. (1) - (3) are buttons to restart a search, choose visualization levels, and search. The icons in (6) are used to disable displays and show tooltips.

Wongburi et al. highlight the importance of data visualization to help assess the operational status of WWTPs [26], but the visualization in their study is data analytic oriented rather than an information visualization approach. The tool most similar to our work is the interactive geovisualization of WWTP compliance status [25], however, its scale is limited to the dataset published by the U.S. Environmental Protection Agency (EPA) in 2020.

Other applications in the visualization of hydrographic data use the HydroSHEDS database, which incorporates HydroWASTE as one of its products [23]. For example, the freshwater ecosystems explorer developed by the United Nations Environment Program is an interactive tool with a high geographic resolution to show the change in surface water and wetlands at national, sub-national, and basin levels [17]. Users could quickly navigate through the parallel layout to select the region of interest, view the dynamic changes, and zoom into case studies shown as point marks. World Wide Fund for Nature also creates a Water Risk Filter that uses choropleth maps with attribute filtering to evaluate the risk of water quality for industry stewardship [6]. Our work differentiates from this kind of application by the target domain since no prior vis tool exists for WWTPs.

2.3 Geovisualization

Since our database consists mainly of geographic attributes of WWTPs, we investigate research on geovisualization in particular. Visualizing geographic data is tricky: mark representations, screen estates, and factors alike could all affect the usability of vis tools. For example, Klippel et al. show that adding values to the point marks on the map could slow down the decision-making process

[10]. Dillemath suggests that a larger map extent could lead to higher performance of navigation tasks [3].

Based on prior work [17], we adopt the dot map, which is good for visualizing absolute values in thematic cartography [7], to allow users to navigate through each individual WWTP and view its detailed statistics. However, a dot map is usually less precise and informative than other thematic maps that provide definite figures for users to refer to [27]. Therefore, we consider adding a choropleth map to address the vague distribution of WWTPs on a dot map. Katarzyna et al. have shown that users tend to have better performance in various map use tasks using the choropleth map, which is recommended for showing the overall geographical pattern of the mapped variable [22].

2.4 Interactive Designs

Interactivity is one of the most widely used concepts to provide users with different perspectives of spatial data and the capability to configure vis contents. Crampton provides a ranked typology of interactivity types in geographic visualization (GVis) [1]. Nöllenburg elaborates on two types of user interactions in geovisualization: focusing on individual views and linking multiple views [14]. While 3D visualization and animation are gaining momentum in GVis, Dorling's study shows that cartographic animations need more careful justification due to potential perceptual overload [4].

Another important facet that has been explored is the concept of linking views within geovisualizations. Roberts highlights the various architectures in which different views can "share" data to coordinate entire visualizations, contextualizing the discussion around prior geovisualization software such as ArcView and xGobi

[19]. Such systems extend the features found on the basis of an interactive map.

With these backgrounds in mind, we wish to build an interactive vis article to illustrate WWTPs distribution and crucial attributes that assist in detailed water quality assessments.

3 DATA AND TASK ABSTRACTION

In this section, we describe and provide abstractions for the data and tasks of the proposed vis tool based on the framework discussed in the textbook [12].

3.1 Data Abstraction

As highlighted in the previous section, we use the HydroWASTE database as the dataset in this project, which describes characteristics of WWTPs at a global scale - the full data can be downloaded in CSV format from the following website: https://figshare.com/articles/dataset/HydroWASTE_version_1_0/14847786/1. This dataset combines national and regional datasets based on their most recent updated versions as of writing to provide a global snapshot of WWTPs [5]. The dataset comes in the form of a simple flat table - each row encodes information about a single WWTP item, and columns represent its specific characteristics. There are a total number of 58502 items and 25 attributes in the dataset. To better understand our data, we start with an abstraction process to explain and analyze the properties of each attribute. We detail information regarding the semantic meaning, type, cardinality/range of each attribute in Table 1.

Since HydroWASTE does not include names of a WWTP's location and the nearby river, we plan to augment the current dataset by deriving these two categorical attributes shown in Figure 1 (5) and (8).

3.2 Task Abstraction

Our main goal is to provide an interactive display of WWTPs for policymakers to understand, evaluate, and adjust current wastewater treatment regulations. Under this study's context, Policymakers refer to government officers and intergovernmental organizations in the environment protection sector. Our database HydroWASTE includes key metrics in regulatory requirements of wastewater systems: the discharge volume of wastewater effluent, the effluent locations, the number of the population served, etc. Moreover, policymakers interested in regulations of other countries and regions can view information at different granularity levels: provincial, national, and global. Our tool will allow higher-level tasks such as aggregating data across locations and filtering out attributes of interest. Policymakers don't have to investigate individual WWTP and can easily find answers to questions like "which region/country has the highest effluent discharge?".

Ordinary water users can also glance at the distribution of WWTPs and the nearby river networks to better understand the surrounding water quality. Specialized analysts might take an overview of global WWTPs to compare their served capacities and discharge amounts relative to that of the receiving waterbodies (i.e., the dilution factor) to assess the risk of excessive effluent.

In general, our work allows users to *analyze*, *search*, and *query* the key metrics of WWTPs for water quality assessment. Different

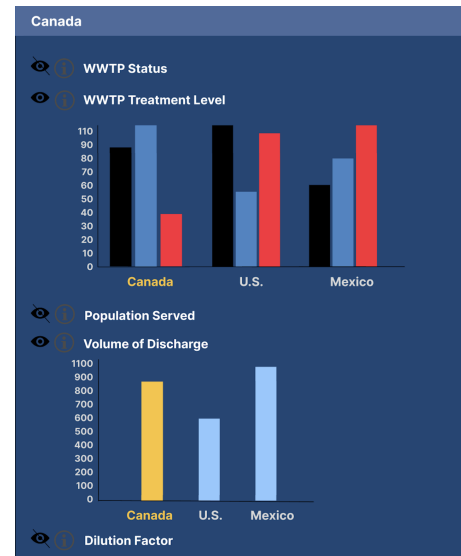


Figure 2: Mockup of bar charts. The axis annotations and legends are omitted for simplicity and the color scheme is tentative.

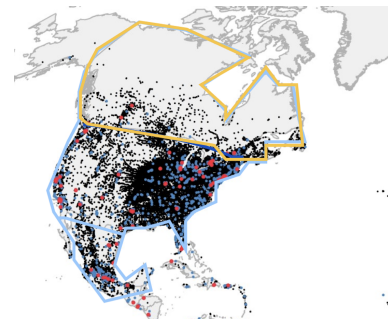


Figure 3: Mockup of region selection.

presentations of WWTPs can help end users in decision-making and planning. We will implement techniques to *explore*, *browse*, *locate*, and *lookup* WWTPs and their geographic locations. Users should be able to easily perform queries to *identify* each attribute, *summarize* quantitative attribute values of items, and *compare* across different levels of data.

4 SOLUTION

In this section, we propose the visual encoding and idioms for WWTP visualization, the implementation plan, and the resulting user scenarios.

4.1 Visual Encoding and Idioms

We want to build an interactive explainer to provide a fast exploration of individual WWTPs and flexible comparison in different scopes. The explainer format is well-suited for our database because we can provide a brief explanation and background information in this domain-specific study.

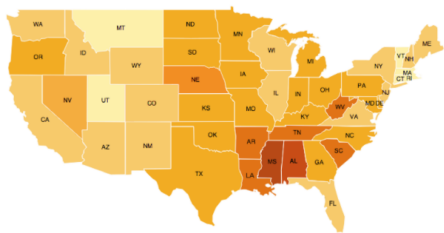


Figure 4: Mockup of a choropleth map with a single-hue color scheme.

4.1.1 Dot Map. Our first main page provides readers with an overview of WWTP distribution. On the right view, we present a dot map that uses the point mark on a spatial field to visualize the geographic distribution of WWTPs. To help users quickly identify the selected WWTP in a broader view, we choose the shape channel (the star shape in particular) to encode whether the WWTP is selected or not. The legend in Figure 1 (10) shows the color hue channel that encodes the level of treatment of WWTPs. This star shape provides the linking between the faceted views - the panel view on the left and the spatial field on the right of the webpage. This demonstrates overview-detail type linking - in which the left view provides more detailed information about a single selected data item (single WWTP) from the right. This is made easier through the ability to pan and zoom on the map, as described later. The synchronicity of the views is controlled through the user interaction of selecting (clicking on) a WWTP of interest from the map. Within this juxtaposition, both views are permanently visible side-by-side.

The left-hand side panel of the dot map displays the geographic and operational attributes of an individual WWTP. Since policy-makers are interested in the geography of WWTPs, we use arrow markings to show the change in geographic coordinates between the WWTP location and the outfall location. As shown in Figure 1 (7), the triangle pointing upwards represents an increase in latitude and vice versa. In addition, the color hue channel is used to double-encode the change in coordinates. The attribute of the population served is encoded by the length channel. Specifically, the length of the grey bar in Figure 1 (9) encodes the design capacity of the WWTP while the length of the blue bar encodes the actual amount of people served.

4.1.2 Choropleth Map. The second page displays a choropleth map (ex., Figure 4) to illustrate the general geographic distribution of WWTPs around the globe. The HydroWASTE dataset is preprocessed such that it aggregates the number of WWTPs in each geographic unit (i.e., country). The choropleth map uses a sequential, single-hue color scheme to represent different amounts of WWTPs: 0-10, 10-100, 100-500, >500. When the user moves the mouse over a certain country, a tool-tip window will show the statistical summary of all the WWTPs at the side.

4.1.3 Bar Charts. Our third page provides the audience with the ability to compare and contrast WWTP metrics in different provinces and countries. Similar to the first page, users see a faceted view, with the right side displaying the spatial representation of the world,

and the left panel being used to display graphs and charts for comparison. These graphs and charts represent aggregate values for WWTPs that can be used to compare across countries, with the aggregation depending on the metric. For instance, the number of WWTPs of each type can be represented with a count, and the volume of the discharge can be represented by the sum within each country. In the grouped bar chart in Figure 2, the color hue channel encodes the country names while the length channel encodes the number of WWTPs. The color hue channel is further used to highlight the selected region of interest from others. In Figure 2, Canada is selected. In the interest of space and aesthetic constraints, we limit the number of country for comparisons up to 10. Bi-directional linking is present through interaction, hovering over the country-specific marks for each graph on the left panel highlights the country on the map on the right side of the page and vice versa.

4.1.4 Interactivity. On the dot map, users can either zoom into an individual WWTP's location or zoom out to observe the overall distribution of WWTPs. The search bar in Figure 1 (3) allows users to navigate to specific continents or countries on the map through a query. After zooming into a region, users can filter by either attributes or items. In the first case, users will click on the eye icon shown in Figure 1 (6) to enable or disable the display. In the second case, users can select the countries of interest to only view their WWTPs for comparison. Figure 3 shows an example of selecting three countries: Canada, the U.S., and Mexico to compare (borders highlighted in blue). Canada is again selected among the three countries (highlighted in orange) as the main subject of study. The corresponding bars for Canada in Figure 2 in the charts are also highlighted using the same orange color.

We note that, at broader levels of Zoom, the large number of dot marks create both visual clutter and cause high computation costs on the dot map. As such, we use aggregation techniques that are adaptive to different levels of Zoom to alleviate such issues. In particular, item aggregation is used - spatially close WWTPs of the same treatment level are grouped together as a singular point mark, and the size of the mark represents the number of aggregated WWTPs (a number is also presented that represents the same metric). When the viewer zooms in, the level of aggregation changes until each WWTP is once again represented as a singular point mark.

4.2 Implementation

The implementation of the proposed visualization will be done primarily using the visualization grammar of D3.js (<https://d3js.org/>), which allows for web-based displays through DOM manipulation. As follow, we plan to have a visualization hosted on the web. Users can navigate to a website, where they can read about WWTPs and view and play with our interactive visualizations. As such, we plan to integrate D3.js with web-based development frameworks and languages - React.js for the backend, and HTML and CSS for the frontend.

4.3 Result

4.3.1 Understanding WWTP Metrics near a Region of Interest. Emi is an environmental researcher working in water quality modeling

who is interested in understanding wastewater contamination in her province of British Columbia. As such, she wants to get a broad overview of how WWTPs are operated in this area of interest. To do so, she navigates to our visualization tool and makes a selection for her province (first selecting the country - Canada - by either locating it on the map or using the search bar, and then changing the layer to province and looking for British Columbia). The tool zooms in on the province and scans for WWTPs in it and displays a dashboard of graphs that summarize the WWTP metrics within the province. For example, if Emi is interested in the level of treatment of the wastewater treatment plants, she can refer to a bar chart that displays the number of WWTPs at each level. If she is interested in the population served, there will be a histogram that demonstrates this across the WWTPs in the province. If Emi specifically wants to get a sense of the metrics for a specific WWTP, she can hover over the WWTP mark on the map, which will highlight the associated mark on the charts (e.g. the bar on the bar chart) - letting her know the metrics at that specific WWTP. Overall, our visualization helps Emi in grasping the WWTP metrics in her province at a glance.

4.3.2 Comparing WWTPs Metrics across Countries. John is part of an environmental NGO in the U.S., and is concerned about wastewater treatment within his own country, especially how it compares and contrasts to treatment in other countries. To do so, he clicks his own country (USA) on the map, and then one that he is interested in comparing against (Canada). This brings up a series of graphs that compare the metrics against each other, for example, multiple-bar bar charts display the proportional distribution of operation for the two countries. There also exist graphs that compare the metrics against the world distribution, for example, John can see how the mean population served in Canada and the USA fall into a world distribution (shown as a histogram). Using this information, John can better understand how the metrics of WWTPs in his country compare to others, both in a direct comparison as well as in reference to global metrics.

5 MILESTONES

There are two deliverables of our project: a written report and a demo. We create a tentative timeline in Table 2 to keep track of project progress and delegate tasks. The timeline will be modified in an iterative manner to refine the project scope, tool requirements, and task assignments. We have completed all the tasks up-to-date and updated the task list according to the proposal feedback.

6 DISCUSSION

7 FUTURE WORK

8 CONCLUSION

REFERENCES

- [1] Jeremy Crampton. 2002. Interactivity Types in Geographic Visualization. *Cartography and Geographic Information Science*. *Cartography and Geographic Information Science - CARTOGR GEOGR INF SCI* 29 (01 2002), 85–98. <https://doi.org/10.1559/152304002782053314>
- [2] Grégorio Crini and Eric Lichtfouse. 2019. Advantages and disadvantages of techniques used for wastewater treatment. *Environmental Chemistry Letters* 17, 1 (2019), 145–155.
- [3] Julie A. Dilleuth. 2009. Navigation Tasks with Small-Display Maps: The Sum of the Parts Does Not Equal the Whole. *Cartographica: The International Journal for Geographic Information and Geovisualization* 44, 3 (2009), 187–200. <https://doi.org/10.3138/carto.44.3.187> arXiv:<https://doi.org/10.3138/carto.44.3.187>
- [4] Daniel Dorling. 1992. Stretching Space and Splicing Time: From Cartographic Animation to Interactive Visualization. *Cartography and Geographic Information Systems* 19, 4 (1992), 215–227. <https://doi.org/10.1559/152304092783721259> arXiv:<https://doi.org/10.1559/152304092783721259>
- [5] Heloisa Ehalt Macedo, Bernhard Lehner, Jim Nicell, Günther Grill, Jing Li, Antonio Limtong, and Ranish Shakya. 2022. Distribution and characteristics of wastewater treatment plants within the global river network. *Earth System Science Data* 14, 2 (2022), 559–577.
- [6] World Wide Fund for Nature. 2021. *WWF Water Risk Filter*. Retrieved Oct 21, 2022 from <https://waterriskfilter.org/>
- [7] Eduardo Gomes. 2017. Creating a Dot Density Map: Resident Population in Mainland Portugal. *The Cartographic Journal* 54, 2 (April 2017), 157–162. <https://doi.org/10.1080/00087041.2016.1148106>
- [8] R Hamilton, Bill Braun, Ron Dare, Ben Koopman, and Spyros A Svoronos. 2006. Control issues and challenges in wastewater treatment plants. *IEEE control systems magazine* 26, 4 (2006), 63–69.
- [9] A Inyinbor Adejumo, O Adebisin Babatunde, P Oluyori Abimbola, A Adelani Akande Tabitha, O Dada Adewumi, A Oreofe Toyin, et al. 2018. Water pollution: effects, prevention, and climatic impact. *Water Challenges of an Urbanizing World* 33 (2018), 33–47.
- [10] Alexander Klippel, Frank Hardisty, Rui Li, and Chris Weaver. 2009. Colour-Enhanced Star Plot Glyphs: Can Salient Shape Characteristics Be Overcome? *Cartographica: The International Journal for Geographic Information and Geovisualization* 44, 3 (2009), 217–231. <https://doi.org/10.3138/carto.44.3.217> arXiv:<https://doi.org/10.3138/carto.44.3.217>
- [11] Yu Liu and Joo-Hwa Tay. 2004. State of the art of biogranulation technology for wastewater treatment. *Biotechnology advances* 22, 7 (2004), 533–563.
- [12] Tamara Munzner. 2014. *Visualization analysis and design*. CRC press.
- [13] Peta A Neale, Nicole A Munz, Selim Ait-Aissa, Rolf Altenburger, François Brion, Wibke Busch, Beate I Escher, et al. 2017. Integrating chemical analysis and bioanalysis to evaluate the contribution of wastewater effluent on the micropollutant burden in small streams. *Science of The Total Environment* 576 (2017), 785–795. <https://doi.org/10.1016/j.scitotenv.2016.10.141>
- [14] Martin Nöllenburg. 2007. *Geographic Visualization*. Springer Berlin Heidelberg, Berlin, Heidelberg, 257–294. https://doi.org/10.1007/978-3-540-71949-6_6
- [15] Government of Canada. 2020. *Municipal Wastewater Treatment*. Retrieved Oct 21, 2022 from <https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/municipal-wastewater-treatment.html>
- [16] Deborah Panepinto, Silvia Fiore, Mariantonia Zappone, Giuseppe Genon, and Lorenza Meucci. 2016. Evaluation of the energy efficiency of a large wastewater treatment plant in Italy. *Applied Energy* 161 (2016), 404–411.
- [17] United Nations Environment Programme. 2020. *Freshwater Ecosystems Explorer*. Retrieved Oct 21, 2022 from <https://www.sdg661.app/home>
- [18] Jiu-hui Qu and Maohong Fan. 2010. The current state of water quality and technology development for water pollution control in China. *Critical reviews in environmental science and technology* 40, 6 (2010), 519–560.
- [19] Jonathan C Roberts. 2005. Exploratory visualization with multiple linked views. In *Exploring geovisualization*. Elsevier, 159–180.
- [20] N Schwarzenbeck, W Pfeiffer, and E Bombal. 2008. Can a wastewater treatment plant be a powerplant? A case study. *Water science and technology* 57, 10 (2008), 1555–1561.
- [21] Amit Sonune and Rupali Ghate. 2004. Developments in wastewater treatment methods. *Desalination* 167 (2004), 55–63.
- [22] Katarzyna Słomska-Przech and Izabela Małgorzata Gołębiowska. 2021. Do Different Map Types Support Map Reading Equally? Comparing Choropleth, Graduated Symbols, and Isoline Maps for Map Use Tasks. *ISPRS International Journal of Geo-Information* 10, 2 (Feb. 2021), 69. <https://doi.org/10.3390/ijgi10020069>
- [23] World Wildlife Fund US. 2006. *Seamless hydrographic data for global and regional applications*. Retrieved Oct 21, 2022 from <https://www.hydrosheds.org/>
- [24] Hongtao Wang, Tao Wang, Bingru Zhang, Fengting Li, Brahim Toure, Isaiah Bosire Omosa, Thomas Chiramba, Mohamed Abdel-Monem, and Mahesh Pradhan. 2014. Water and wastewater treatment in Africa—current practices and challenges. *CLEAN—Soil, Air, Water* 42, 8 (2014), 1029–1035.
- [25] Research Watch. 2020. *Wastewater Treatment Plants*. Retrieved Oct 21, 2022 from <https://resourcewatch.org/data/explore/>
- [26] Praewa Wongburi and Jae K. Park. 2021. Big Data Analytics from a Wastewater Treatment Plant. *Sustainability* 13, 22 (2021). <https://doi.org/10.3390/su132212383>
- [27] John K. Wright. 1936. A Method of Mapping Densities of Population: With Cape Cod as an Example. *Geographical Review* 26, 1 (1936), 103–110. <http://www.jstor.org/stable/209467>

Name	Meaning	Type	Cardinality/Range	Other Notes
WASTE_ID	Identifier for WWTP	Categorical	Range: [1, 58,502], consecutive integers; Cardinality: 58502	Key attribute
SOURCE	Dataset from which the data was compiled from (national/regional)	Categorical	Range: [1, 12], consecutive integers; Cardinality: 12	
ORD_ID	Identifier for the national/regional dataset	Categorical	Range: {1, ..., 78,000,001,002}; Cardinality: 47497	not unique
WWTP_NAME	Name of the WWTP	Categorical	Range: {set of WWTP names}; Cardinality: 49261	not unique
COUNTRY	Country in which the WWTP is located	Categorical	Range: {set of countries}; Cardinality: 188	
CNTRY_ISO	Standard ISO Defining Codes for Countries, Dependent Territories, Special Areas	Categorical	Range: {set of country ISOs}; Cardinality: 180	
LAT_WWTP	WWTP location latitude	Quantitative (Diverging)	Range: [-54.8, 71.6]	
LON_WWTP	WWTP location longitude	Quantitative (Diverging)	Range: [-175.3, 178.5]	
LAT_OUT	Outfall location latitude for the WWTP	Quantitative (Diverging)	Range: [-54.8, 71.6]	
LON_OUT	Outfall location longitude for the WWTP	Quantitative (Diverging)	Range: [-175.3, 178.4]	
STATUS	WWTP Status	Categorical	Range: {Closed, Construction Completed, ...}; Cardinality: 9	
LEVEL	Level of treatment at the WWTP	Ordinal (Sequential)	Range: [Primary, Secondary, Advanced]; Cardinality: 3	
DESIGN_CAP	WWTP design capacity (empty if not reported)	Quantitative (Sequential)	Range: [0, 11,200,000], integers; Cardinality: 7329	15835 empty values
POP_SERVED	Total population served by the WWTP	Quantitative (Sequential)	Range: [0, 10,100,000], integers	
WASTE_DIS	Volume of discharged treated wastewater	Quantitative (Sequential)	Range: [0, 3,070,000]	
HYRIV_ID	Identifier for outfall location river; empty if outfall is ocean or large sink	Categorical	Range: {10000009, ... ,80323236}, integers; Cardinality: 42822	379 empty values
RIVER_DIS	River discharge at WWTP outfall location; empty if outfall is ocean	Quantitative (Sequential)	Range: [0.001, 127,000]	10551 empty values
DF	Estimated dilution factor; empty if estimated outfall is ocean or large lake	Quantitative (Sequential)	Range: [1, 703,000,000]	11200 empty values
COAST_10KM	Outfall Location is within 10km of ocean or large lake?	Categorical	Range: {0, 1}; Cardinality: 2	
COAST_50KM	Outfall Location is within 50km of ocean or large lake?	Categorical	Range: {0, 1}; Cardinality: 2	
QUAL_LOC	Quality rating regarding location accuracy	Ordinal (Sequential)	Range: [1,4], consecutive integers; Cardinality: 4	
QUAL_POP	Quality rating regarding the POP_SERVED attribute	Ordinal (Sequential)	Range: [1,4], consecutive integers; Cardinality: 4	
QUAL_WASTE	Quality indicator regarding the WASTE_DIS attribute	Ordinal (Sequential)	Range: [1,4], consecutive integers; Cardinality: 4	
QUAL_LEVEL	Quality indicator regarding the LEVEL attribute	Ordinal (Sequential)	Range: [1,2], consecutive integers; Cardinality: 2	
QUAL_CAP	Quality rating regarding the DESIGN_CAP attribute	Ordinal (Sequential)	Range: [1,3], consecutive integers; Cardinality: 3	

Table 1: Attributes in the HydroWASTE dataset. Within this table, longitude and latitude values are rounded off to one decimal place, and the ranges of other qualitative attributes are rounded down to at most 3 significant digits.

A TIMELINE - TABLE

Task	Time	Deadline	Description	Assignee
Pitch	2min	Sept. 28	Present to the class	Madonna
Pre-proposal meeting	4h	Oct. 11	Data exploration and abstraction	Michael, Madonna
Pre-Proposal meeting with Tamara	15min	Oct. 12	Discuss project ideas and data/task abstraction	Michael, Madonna
Proposal meeting	4h	Oct. 20	Finalize the proposal	Michael, Madonna
Proposal	14h	Oct. 21st	Define the project scope and potential solutions	Michael, Madonna
Implementation	87h	Dec. 10	Implement the vis tool	Michael, Madonna
- Tool requirements	2h	Nov. 02	Decide visual encoding and corresponding implementation details	Michael, Madonna
- D3 basics and setup	15h	Nov. 08	Learn D3 and implement the explainer framework	Michael, Madonna
- Dot Map	5h	Nov. 08	Implement the dot map	Michael
- Graduated symbol map (abandoned)	5h	Nov. 11	Implement the graduated symbol map	Madonna
- Choropleth map	20h	Nov. 19	Implement the second page	Madonna
- Main page	20h	Nov. 26	Implement the main page for the dot map	Michael, Madonna
- Bar chart and bidirectional linking	20h	Dec. 03	Implement the page for comparison	Michael
Updated Project Writeup	20h	Nov. 15	Finish the related work, and update the solutions and milestones	Michael, Madonna
Project Peer Reviews	4h	Nov. 16	Provide and receive project feedback	Michael, Madonna
Post-Update Meeting with Tamara	15min	Nov. 23	Discuss the vis tool and final paper	Michael, Madonna
Draft the final paper	14h	Dec. 04	Finish the final paper draft	Michael, Madonna
Pre-Presentation Setup	15h	Dec. 11	Prepare slides, rehearse, and make the presentation video	Michael, Madonna
Final Presentation	10min	Dec. 14	Present to the class	Michael, Madonna
Final Paper	15h	Dec. 16	Finalize each section and submit	Michael, Madonna

Table 2: Tentative Timeline