Northeastern University

Microplastics Database

Anna Yang, Karen Cheng, Kathy Sha, Lauren Morris

TA: Jessica Van de ven

Professor Rachlin

June 24, 2021

Introduction

Microplastics are small plastic debris measuring less than 5 mm in diameter caused by the degradation of human plastic waste. These plastics are found in the stomachs of fish and pose potential threats to aquatic life (NOAA).

Additionally, Microplastics may have lasting impacts on human populations. Researchers from the Department of Obstetrics and Gynecology in Roma, Italy found that the microplastics entered women's bodies through ingestion and inhalation then translocated into the placenta. Researchers claim that the continual exposure to microplastics could cause long-term damage or upster a fetus's developing immune system due to the carried chemicals in the plastics (Ragusa).

Microplastic concentrations are collected across the United States by the United States Geological survey (USGS). The USGS Program supports the collection of streamflow and water-level information at approximately 8,500 sites across the United States. The colored dots shown in Figure 1 depict streamflow conditions as a percentile and they also represent the locations of all USGS sites across the United States.

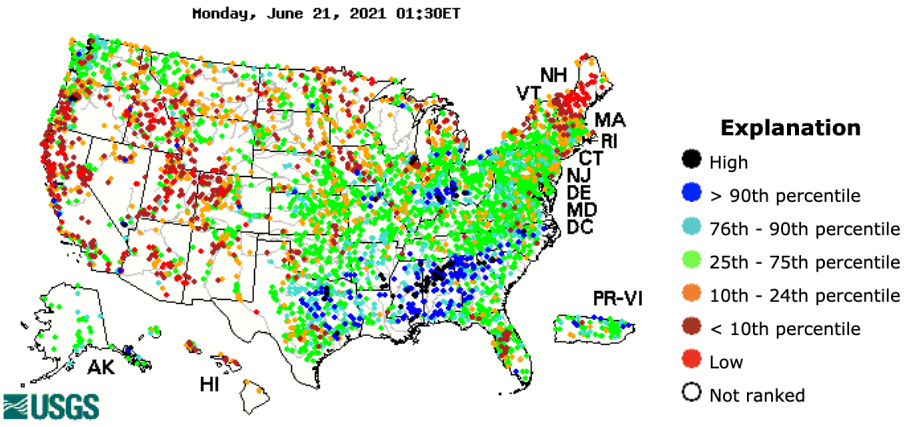


Figure 1. All USGS stations mapped in the United states. Percentile ranges in the legend represent percent stream flow in the lake as measured at each station. mage received from the USGS website. ("U.S. Geological Survey")

While microplastics are a developing worry amongst the National Ocean Service, very little accumulated data is available for analysis. Therefore, the goal of the project was to create a database for samples taken in lakes and oceans that could be accumulated for research purposes.

A database was created that can log microplastic samples across all 8,500 USGS stations around the United States. Currently data is only available for samples taken around the great lakes; however, for future directions, it is intended that scientists can log microplastic concentrations taken at any USGS station in the US. Data was then connected to city demographics as collected from the US Census Bureau for the 29 locations surrounding the collected USGS station's samples. Population size, poverty rate, and per capita income were individually recorded with 2019 information collected from the US census bureau. Analysis was conducted in MYSQL, MatLab and excel to find potential connections between human populations and microplastic concentrations.

Database design

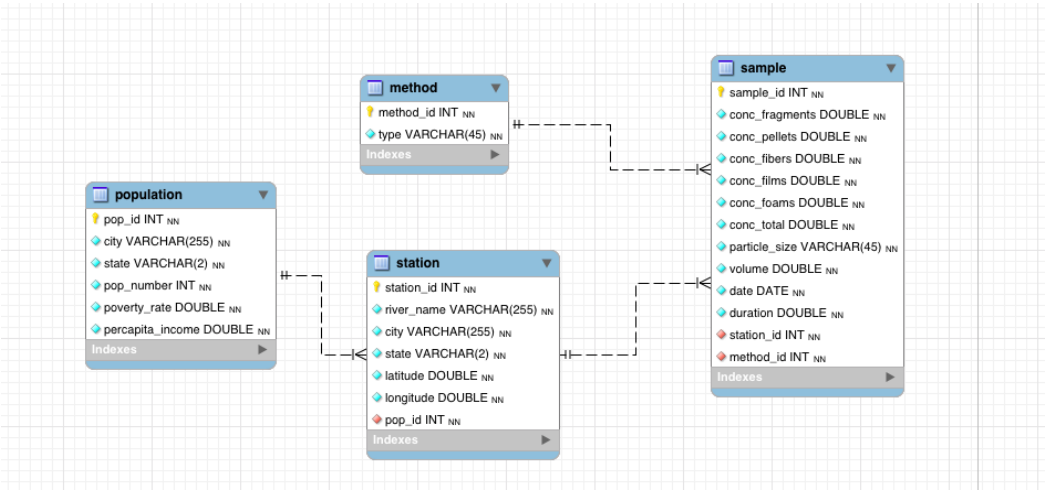


Figure 2. Model of database from MYSQL. Tables represent the samples taken, the method used to take samples, the USGS station, and the population/ demographic data of the surrounding town

In tackling the design of the microplastics database, an ER diagram was created representing the different entities, relationships, and attributes. Four separate entities were chosen based on the data: the population, USGS station, sampling method, and microplastics sample. As shown in the diagram, the sampling method and station entities have one or more microplastic sample entity, and a population entity has one or more station entity. The relationships between each entity gave insight into how each table connected to each other and which fields were necessary for data analysis.

Methods

Data accumulated for the database development included samples taken at the United States Geological Survey stations spanning the great lakes area. The data from these stations was used to log individual samples of microplastic concentrations in tributaries across the great lakes. There were 428 taken samples at 29 different USGS stations in the northern United States as shown in Figure 3. However, the data only represents 0.3% of all USGS stations in existence, so the conclusions drawn were only applicable to a small area.

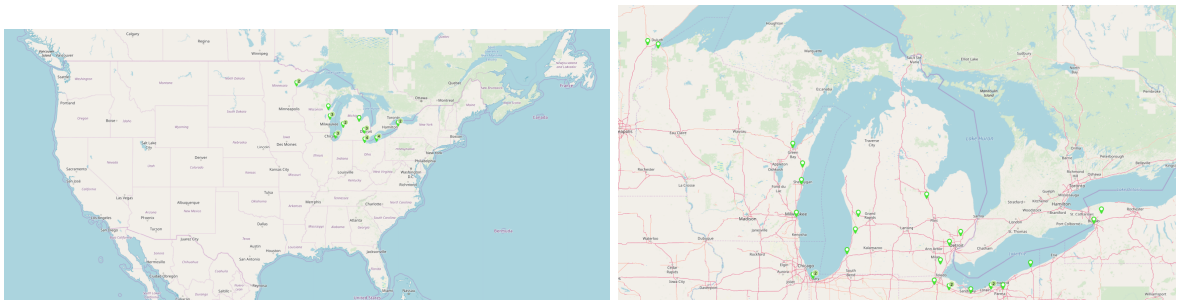


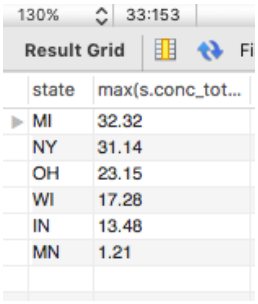
Figure 3. Map of USGS stations included in database. Image A shows the location of the great lakes and station in the united states. Image B shows a zoomed in version of the great lakes region. All points mapped using decimal GPS coordinates of the USGS station from the ID. All are surrounding the great lake regions in the Northern United States. Map created using MapLine

Raw data was manipulated in order to be suitable for the nature of the database. Columns were added into an excel sheet to represent the decimal gps coordinates of each station as collected from the USGS website. Strings were parsed in excel to obtain separated location, city and state data. Auto-Incremented foreign keys were manually created for the population tables and methods tables to identify each separate item in the tables.

Analysis

The goal of the database was to create a platform in which data can be gathered and stored in order to easily create conclusions from data that was previously not available or easily manipulable. Knowing which methods of sample collection yielded the most consistent or highest results would aid researchers in focusing efforts on those data gathering techniques for the database. This query could give implications for sampling errors in the given data and direction for further sampling method priority. Through results in MYSQL, boat and bridge yielded the most consistently high results; therefore, it may be beneficial to invest most funds and time in these methods for the most reliable data results in order to draw more conclusions and gain more information on microplastics.

The given state that a lake resides in may also influence the pollution due to human impact. To determine this, a query was written to determine which states had the highest total concentrations. Implications of results could give insight into which states need revised pollution laws or recycling policies. It was found that Michigan had the highest concentrations as shown in Table 1 with 32.32 m^3 of pollution as the maximum concentration total of the lakes. While the NCSL shows that Michigan has refund laws on bottles for customers, increased motivation or consequences may be necessary to decrease water pollution (Schultz).


130% 33:153

Result Grid Fi

state	max(s.conc_tot...
MI	32.32
NY	31.14
OH	23.15
WI	17.28
IN	13.48
MN	1.21

Table 1. Maximum concentration total in a given state

Population impact on microplastics may also be linked to the given area's economic bracket. To determine this, a query was written to determine the link between per capita income and the average concentration of microplastics in the area. The results were then graphed using excel in figure 4. The average concentration of microplastics for per capita income over 20k outputted an estimated 1.99 m^3, and 2.8 m^3 for per capita income less than 20k. It was found that in areas with per capita income over 20K, the majority of concentrations stayed below 4 m^3 with a few outliers. A linear regression model was run and found that with a linear correlation of y=2.4x, there was a r^ value of 0.0058 showing there was no statistical relationship between the income over 20k and the total concentration of microplastics in the lakes. For income under 20K the results appeared more heavenly distributed from 0 to 15

m^3;however, the linear regression found in matlab showed a linear correlation of y = 12.8 - 0.00055 with an r^2 value of 0.0328. This also shows that there was no definitive correlation between per capita income below 20K and the total concentration of microplastics in m^3. Therefore, per capita income of a population did not impact the surrounding pollution.

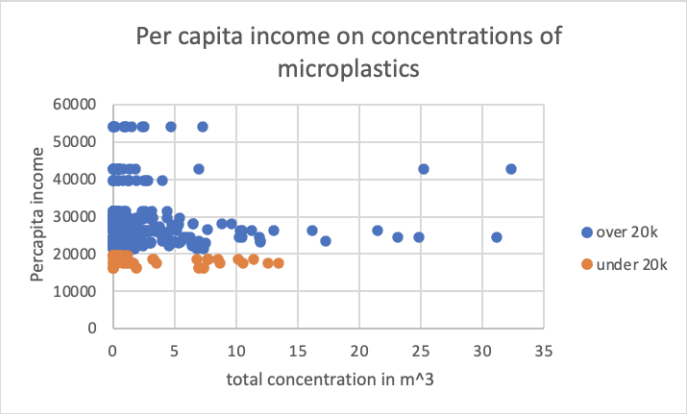


Figure 4. Effect of per capita income on microplastic concentration in surrounding states for high and low income areas.

Determining if there is a seasonal effect on microplastic concentrations can help see if there is a monthly change based on environmental factors of societal practices. For example, perhaps there is more plastic usage in colder months due to buying more packaged food in order to stay sheltered. Results from the query show that for at least 2 months in a row, one state would have the highest microplastic concentration as shown in figure 5. It can be concluded that for any given season, a specific state would dominate in microplastic concentration showing seasonal effects for those states. This is significant in giving reason for those specific states to look into their environmental policies for those months.

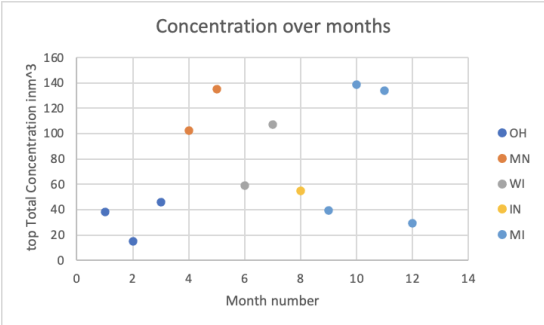


Figure 5. Maximum concentration of microplastics over the year. Each state in the legend was grouped to a season as determined through queries

Finally, population demographic impact on microplastic concentrations was analyzed by determining if poverty rate impacted the concentration in surrounding lakes. It was found that as poverty rate increased, as did the average concentration of microplastics in the surrounding areas as shown in figure 6. However, this linear trend was calculated using matlab and found to be 0.114x + 0.26 with an r^2 value of only 0.314 therefore, while there seemed to be a slight linear trend, it is not a strong enough correlation to definitively attribute the water pollution variable to only poverty in surrounding areas. Poverty rate may be only a small piece of the total picture.

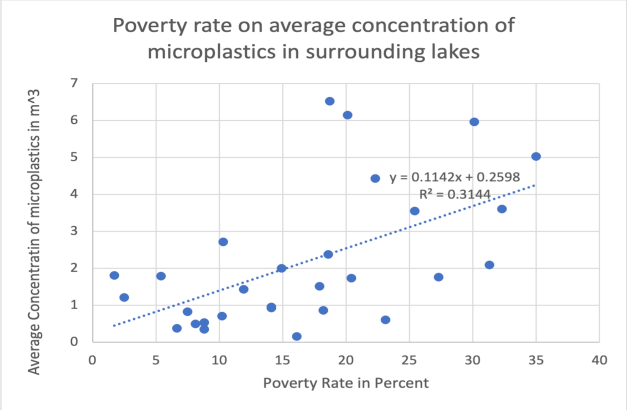


Figure 6. Effect of poverty rate on average concentration of microplastics .

Conclusions

A database was created in order to store sample information collected from USGS stations in the great lakes area as well as population data in the associated cities. The goal was that the database could be used by future scientists to log any USGS samples taken in a single database to be cross referenced with the surrounding areas populations to detect potential causes of water pollution.

It was found that certain states such as Michigan had higher total pollution levels despite established recycling laws showing that specific areas may contribute more to pollution due to state policies. It was determined through MYSQL, excel, and Matlab, that an area's per capita income had no linear impact on the water pollution; however, poverty percent had a potential, yet minimal, linear effect on the average concentration of microplastics. These findings could prove valuable in determining what causes increased levels of microplastics in our lakes such that new policies or preventative measures can be taken. Decreasing microplastics in the lakes and oceans may prove detrimental to human health in the long run.

For future directions, It would be advised to include tables relating to the fish populations to determine the impact that human waste has on the surrounding environment as well as agriculture relating to fishing. Furthermore it is advised to analyze any human health statistics alongside the concentration level to observe the effect that the waste has on populations especially developing fetuses.

Author Contributions

Anna wrote code for the queries and analyzed the results to draw out conclusions from the database. She arranged data for the methods to fit into sql.

Karen conceived of the presented idea along with the group and arranged data to fit into sql, specifically the sample table for the microplastics.

Kathy created and designed the ER diagram along with the help of the other team members. She also wrote the code for some of the queries, and analyzed and drew conclusions from the results.

Lauren wrote the final query as well as created the visualizations for the presentation and paper. She also wrote the analysis and conclusion pieces of the paper and added matlab statistical analysis.

All authors discussed the results and contributed to the final manuscript.

Works Cited

NOAA. “What Are Microplastics?” *NOAA's National Ocean Service*, 13 Apr. 2016, oceanservice.noaa.gov/facts/microplastics.html#:~:text=Microplastics are small plastic pieces,our ocean and aquatic life.&text=Microbeads are tiny pieces of,as some cleansers and toothpastes.

Ragusa, Antonio, et al. “Plasticenta: First Evidence of Microplastics in Human Placenta.” *Environment International*, vol. 146, 2021, p. 106274. *Science Direct*, doi:10.1016/j.envint.2020.106274.

Schultz, Jennifer. “State Beverage Container Deposit Laws.” *NCSL*, www.ncsl.org/research/environment-and-natural-resources/state-beverage-container-laws.aspx#mi.

U.S. Geological Survey, 2020, National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), accessed June 15, 2021, at URL https://waterdata.usgs.gov/nwis/