**Semester Project Paper**

Group 3: Milena Hakobyan, Vsem Voskerchyan, Kristina Ghahramanyan, Elen Gabrielyan

American University of Armenia

ENV105 - Numbers Responsibility Environment

Mr. Sean Reynolds

December 10, 2021

**Essay**

Water pollution is a significant issue worldwide, and the goal of the project is to figure out how it progresses in some areas and how it affects people's health. The different pollutants that occur in the rivers, lakes, seas, oceans, and other water bodies are the reasons for water quality drop and pollution. According to the World Health Organization, at least 2 billion people drink polluted water, and half of the world's population possibly will live in water-stressed areas by 2025 (reference). Besides humankind, it can lead to wildlife poisoning and long-term environmental harm.

Three hypotheses of the research will help to understand the water pollution situation in the world, the change of the amount of Enterococci bacteria in the water bodies over time, and the difference in the pollution at the beach's water which is closer to the shore. The first hypothesis suggests that Sub-Saharan countries have the worst drinking water and sanitation conditions globally—the data from WHO and EPI helped find the answers to this question. Further, we collected the historical data provided by the US state of Virginia and New York to understand the issue of Enterococci bacteria in the water. Moreover, the datasets from the two states provided a massive amount of information, which showed the contrasts between the pollution amount at the different parts of the beaches.

The Environmental Performance Index (EPI) results from the collaboration of the Yale and Columbia universities, whose goal is to create a more sustainable future. The unsafe drinking water conditions come from water pollution; that's why it is essential to analyze the combined data of "unsafe drinking water and unsafe sanitation." The European countries are at the highest positions of the data, indicating the high quality of water and sanitation in those countries. In Excel, by using conditional formatting we can see the contrast of European and African countries. In the African continent, there is a considerable difference between Northern and Sub-Saharan Africa. The data of EPI shows that the countries of Northern Africa such as Tunisia, Algeria, or Egypt are much higher in the list than the Sub-Saharan countries. By sorting the countries in Excel, we found that 44 countries from the bottom half of the list belong to the Sub-Saharan region, which is somewhat equal to 48,8%. Moreover, in among the worst 50 countries, 40 are from this region which is 80%. In the end, it is obvious that the Sub-Saharan region has the worst drinking water and sanitation conditions in the world.

Another useful table we have examined was again on the WHO website, showing what percentage of each African country’s population uses safely managed drinking water. The last recorded data is from 2015; using the average function, we have found out that on average, only about 62% of Sub-Saharan countries’ population has normal drinking water. (WHO Regional Office for Africa, 2016). Using a pivot table, we put the low- to high percentages and created a corresponding bar chart to see the data better. As expected, all Sub-Saharan countries had very low percentages of safely managed water services. For instance, Eritrea was the first country on the graph with a percentage as low as 19.3%, then Uganda with 38.9%, and so many other countries with such low percentages.

After seeing the above results, we have decided to understand what adverse effects polluted water can have on human health. According to the World Health Organization, exposure to unsafe water, sanitation, and hygiene habits is directly linked to diarrheal diseases and intestinal nematode infections, and other diseases which can cause death (World Health Organization, 2016). We particularly examined a table found on the WHO website, summarizing data on the mortality rates for each country attributed to exposure to unsafe water. We have made a Pivot Table to make the data tidier - the original table was kind of hard to read. Then we put the numbers from high to low to see which countries have the highest mortality rates. There were 182 countries overall, and with a few exceptions, all the countries with very high results were Sub-Saharan countries. Even when we made a bar chart with the top 42 countries, all were from this region. Chad was on top of the list with a mortality rate of 100, then Somalia, Central African Republic, etc. All these countries had low percentages of safely managed drinking water services. So, the results of our small investigation did not contradict our expectations - Sub-Saharan countries have the highest mortality rates among all countries in the world since they lack normal water services.

When we were done with our analysis on water quality and related health issues all over the world, our group decided to look for two similar datasets of one specific type of water pollution to include some comparisons in our project. After looking at countless datasets, we managed to see that two US states: New York and Virginia have similar restrictions for the quality of beach water, more specifically, for the factor of the quantity of human-to-water transferred bacteria called Enterococci. Moreover, those two states had the same amount of beaches (approximately the same sizes), which was another key point for making an accurate comparison. To begin our analysis, we need a clean and nice-looking dataset, which was the first step of our work. Then, we started to make Pivot Tables based on our data so we could see any interesting trends or similarities. The Pivot Tables made it possible to create charts for a better representation of our data. Our first graph (bar chart) was about the overall average quantity of the bacteria in each beach, taking into consideration the data from 2015 to 2019. Doing the same procedure with both states resulted in some interesting findings. We saw that the average quantity of bacteria is much higher on beaches of New York, which can be caused by the larger number of tourists in the state.

Exploring our data further, we noticed that the dataset has a column for sample locations, with values differing from one another. That gave us an idea, that maybe sample location affects the resulting outcome. As Enterococci are mainly spread by humans, it can be inferred that the water taken straight from the shores will likely have a higher concentration of the bacteria due to the larger number of people. To prove our assumption, we decided to use a bigger dataset, consisting of 16 years of collected data from New York. We made a Pivot Table taking “sample locations” as columns and the “average MPN per 100ml” as values. Then, we made a pie chart to understand which sites of the beach water had the most bacteria. It turned out that the sample location did indeed matter when considering the bacteria quantities - water samples taken from the center had the least amount of bacteria per 100ml.

Throughout this whole project, our team has analyzed the data in the context of unsafe drinking water and unsafe sanitation together, keeping in mind that their combination gives more insights. Our first and main goal was to more clearly define the hypothesis for more quality data analysis results to achieve accurate conclusions. During our research, we had a goal to detect some activities which could have a correlation or explanation for our findings. Our first hypothesis, concerning African countries, and their drinking water and sanitation conditions was proven right based on our data. Also, we have found some correlation between the health data, mortality rate, and the quality of water in African countries. Next, our data showed that the locations of water sampling has a direct effect on the quantity of water bacteria. Lastly, our second hypothesis was that for the chosen two states (New York and Virginia), the ​​water pollution would get worse over time, but the data did not support our theory.

**Reference list**

The city of Virginia Beach. (2019). Beach Water Quality Test Results [Dataset]. <https://data.vbgov.com/dataset/beach-water-quality-test-results>

DOHMH Beach Water Quality Data | NYC Open Data. (2021, October 4). [Dataset]. <https://data.cityofnewyork.us/Health/DOHMH-Beach-Water-Quality-Data/2xir-kwzz>

Sanitation & drinking water. EPI. (2020). Retrieved November 30, 2021, from <https://epi.yale.edu/epi-results/2020/component/h2o>

WHO Regional Office for Africa. (2016). Population using safely managed drinking-water services [Dataset]. <https://aho.afro.who.int/ind/af?ind=292&dim=83&dom=Environmental%20risk%20factors&cc=af&ci=1&cn=Afro%20Region>

World Health Organization. (2016). Mortality rate attributed to exposure to unsafe WASH services (per 100 000 population) [Dataset]. <https://www.who.int/data/gho/data/indicators/indicator-details/GHO/mortality-rate-attributed-to-exposure-to-unsafe-wash-services-(per-100-000-population)-(sdg-3-9-2)>