University of Waterloo
Foundations of Data Science – Group Project
April 14, 2024

Investigation of Population Growth on Water Quality in Ontario

Group 13
Jaiden Nikolic, Parisa Rezaiy, Timothy McLarty, and Frank Miceli

Overview

Water is central to quality of life. Water quality can be affected by numerous factors such as urbanization, agriculture, and industry. As the Ontario economy grows, so does its population. As a population grows, so does the demand for natural resources such as water. But as a population grows, agriculture demands increase, productivity increases, and construction projects increase. If the increased activity is not well managed, pollution of natural resources, especially water, could also increase.

Our group project explores if water quality metrics are affected by population growth.

Data Source

There were numerous data sources that had to be retrieved and cleaned. (REF)

Population data was sourced from Statistics Canada Census datasets. In Ontario, a Provincial Water Quality Monitoring Network (PWQMN) captures and publishes water quality results.

Census Subdivision groups were utilized to filter data from PWQMN dataset. The census data is based on geo-spatial protocols where the PWQMN used lat/long co-ordinates. PWQMN data need to be transformed into a geo-spatial to map populations to monitoring stations and measurements.

The metric data was also vast and inconsistent. Some stations contained metrics for 1 or 2 years but then ceased. There were not many metrics that contained values for the entire period. Care was required to remove measurements from the analysis.

Finally, Ontario is a large and demographically diverse as its geography. Data was categorized into 3 categories: “Most Rural”, “Rural/Urban”, and “Most Urban” for the analysis.

Analysis

The number of metrics collected by Ontario, totaled 156 elements. For the scope of the project, we reduced it to a handful of elements:

Parameter	PARM Code	Common Sources
E. Coli	ECMF	Human and animal feces
Chromium	CRUT	Naturally occurring (erosion of minerals); releases or spills from industrial uses
Nitrates, total	NNOTUR	Naturally occurring; leaching or runoff from agricultural fertilizer use, manure and domestic sewage; may be produced from excess ammonia or nitirification in the distribution system
Lead	PBUT	Leaching from plumbing (lead service lines, lead solder and brass fittings)
Coliform, total	TCMF	Human and animal feces; naturally occurring in water, soil and vegetation
Arsenic	ASUT	Naturally occurring (erosion and weathering of soils, minerals, ores); releases from mining; industrial effluent
Mercury	HGUT	Releases or spills from industrial effluents; waste disposal; irrigation or drainage of areas where agricultural pesticides are used
pH	PH	N/A - pH is a measure of how acidic or alkaline (basic) the water is.

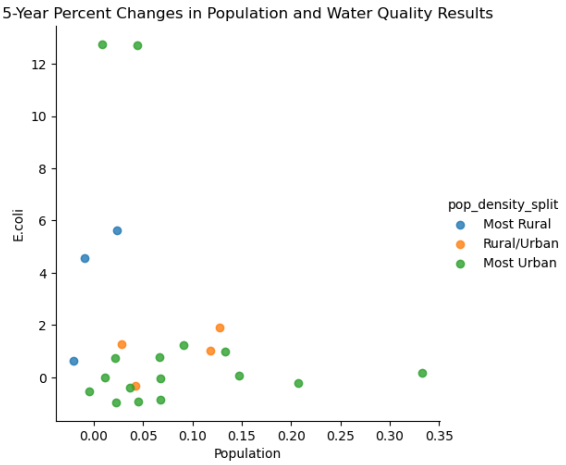
The number of observations for each different water quality parameter varied widely, ranging from 10,909 observations for nitrates to 298 observations for arsenic and 2 observations for Coliform which would ultimately be dropped.

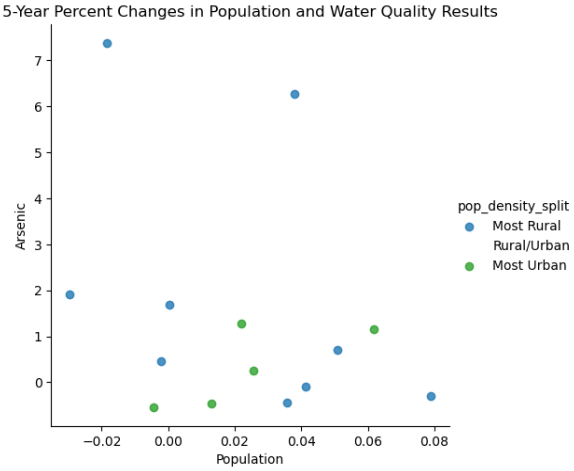
For each census year, the above parameters yielded the following correlation summary.

Water Quality Parameter	Most Rural CSDs	Rural/Urban CSDs	Most Urban CSDs
E.coli	0.826520	0.547852	-0.205104
Chromium	-0.084604	0.123818	0.018223
Nitrate	0.028008	-0.100048	-0.052025
Lead	-0.098779	0.010584	-0.022552
Coliform	*NaN*	*NaN*	*NaN*
Arsenic	-0.414951	*NaN*	0.747144
Mercury	0.086061	-0.261370	0.313155
pH	-0.155795	-0.150679	0.015458

Conclusions

There appears to be a strong positive correlation between population growth and E. coli levels in the "Most Rural" areas, indicating that as population increases in these areas, E. coli levels tend to increase as well. In rural/urban and more urban areas, the correlation weakens and becomes negative, suggesting a different relationship between population growth and E. coli levels in these areas. However, the sample size is very small suggesting that caution should be used in interpreting these outcomes.





In general, these findings do not substantially support a correlation between water quality and population growth. In scenarios where a strong correlation existed between population growth and changes in water quality, the sample size was too small to represent a robust conclusion. In scenarios with larger sample sizes, as in the case of pH level measurement, the correlation was much closer to zero. The significant variance in the observed results makes it challenging to determine any true correlation.

References (REF)

Census Data

1. 2021 Ontario Census Profile by Subdivision
    a. https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/details/download-telecharger/comp/GetFile.cfm?Lang=E&FILETYPE=CSV&GEONO=021
2. 2016 Ontario Census Profile by Subdivision
    a. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/download-telecharger/comp/GetFile.cfm?Lang=E&FILETYPE=CSV&GEONO=066
3. 2011 Ontario Census Profile by Subdivision
    a. https://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/details/download-telecharger/CSV.cfm?Lang=E&Geo1=PR&Code1=35&Geo2=PR&Code2=01&Data=Count&SearchText=Ontario&SearchType=Begins&SearchPR=01&B1=All&Custom=&TABID=1
4. 2006 Ontario Census Profile by Subdivision
    a. https://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/details/download-telecharger/comprehensive/comp-csv-tab-dwnld-tlchrgr.cfm?Lang=E#tabs2006
5. 2001 Ontario Census Profile by Subdivision¶
    a. https://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/details/download-telecharger/comprehensive/comp-csv-tab-dwnld-tlchrgr.cfm?Lang=E#tabs2001
6. 2021 Census Subdivision Boundary Files
    a. https://www12.statcan.gc.ca/census-recensement/2021/geo/sip-pis/boundary-limites/index2021-eng.cfm?year=21

Water Data

7. Ontario Water Station Coordinates
    a. https://files.ontario.ca/moe_mapping/downloads/2Water/PWQMN/PWQMN_Stations.csvs
8. Ontario Water Station Readings 2019-2021
    b. https://files.ontario.ca/moe_mapping/downloads/2Water/PWQMN/PWQMN_2019_2021Mar_update4.csv
9. Ontario Water Station Readings 2010-2018
    c. https://files.ontario.ca/moe_mapping/downloads/2Water/PWQMN/PWQMN_2010_2018_update5.csv
10. Ontario Water Station Readings 2000-2009
    d. https://files.ontario.ca/moe_mapping/downloads/2Water/PWQMN/PWQMN_2000_2009.csv

11. Census Reference
    e. https://uwg13foundati-ayr9249.slack.com/archives/C06JU03S55X/p1713114653391289