**Development Requirements**

**Project Title:** Exploring the correlation between lead levels in children and poverty rates (or insurance).

**Team Members:** Candy Tiu,Jessica Parrucci., Patrick Sande, Karim Ezzat, Emily Jacobs.

**Project Description/Outline:**The purpose of our project is to explore correlations of water quality and public health within Minnesota counties. We will identify and collect data on water quality, poverty levels, for children in the state of Minnesota.  After cleaning and analyzing our data, we will attempt to draw correlations while visualizing our findings.

**Research Questions to Answer:**

What is the correlation of water quality in Minnesota counties and relative public health in children?

**Data Sets to be Used:**

<https://data.web.health.state.mn.us/web/mndata/drinkingwater_query> (MN public Health)

<https://catalog.data.gov/dataset/lead-and-copper-data> (blomington,mn)

<https://www.census.gov/data-tools/demo/>  (US Census Bureau, Small Area Income and Poverty Estimates.):

**Breakdown of Tasks:**

1. Brainstorm topics
2. Research and identify available data
3. Select relevant data
4. Collect data files
5. Clean and validate data
6. Create master data sets
7. Visualize data
8. Create summary analysis
9. Create presentation

**Presentation Requirements**

**Brainstorming Session:**

The initial question our group set out to answer was, “Are we heading towards Flint, Michigan?”  The goal was to compare the water quality levels of Flint, Michigan to specific Minnesota counties, possibly Washington county. Washington county was a focus with the 3M suspicion of polluting ground water, however the lack of data directed the research to our current topic.  Our exploration focused on lead levels and its effects or correlation to poverty levels. We were able to find sufficient data by Minnesota county, blood lead levels in children, and poverty rates.

**Where and How We Located Data:**

We gathered data over and over again, refining and redirecting our research question as various data was either inaccessible or costly.  Once able to locate big data, (one of first data files held 64 lines, not quite big enough), we refocused our hypothesis. We googled, took suggestions from the TA’s and Paul for various sources, and we able to locate organizations with big data, specifically, the state of Minnesota and the US Census Bureau.

**Data Exploration and Clean Up Process:**

Exporting the data to csv files, the cleaning process involved searching for columns that were similar among data sets, using data from the same years, and merging files based on counties.

**Analysis Process:**

1. Defined our Questions
2. Set Our Priorities
3. Collected Data
4. Analyze Data
5. Interpret and Visualize the Results

**Conclusion:**

One county, Kandiyohi, has a higher percentage of blood levels of lead in children under the age of six, which is higher than Minnesota. The population is smaller in this county and they have a higher percentage of lead in their water. In Minnesota it's around 800 children with high levels of lead, most of these are in Ramsey and Hennepin.  It would be interesting to dig further into this, and do some sort of statistics. An interesting aspect of this study is that there is not a lot of data on lead levels in drinking water. The Federal government doesn’t have requirements or standards for schools to test lead level in the drinking water. Minnesota passed legislation in 2016 to start lead level testing in water in July of 2018.  As of the date of our study, October 2018, we could find the results of only one school. Karim called the state of Minnesota’s Health department, explained what we were trying to do, and what data we needed. They did not have any sources of data for us to use but seemed very interested in our study and the state hoped we would forward on any Findings.

**Implications of Study:**

Blood lead levels at least as low as 10 µg/dL are associated with adverse effects.

Lead affects virtually every system in the body.

Since a child’s brain is still developing, lead can lead to intellectual disability. Symptoms may include:

behavior problems

* low [IQ](https://www.healthline.com/health/iq-testing)
* poor grades at school
* problems with hearing
* short- and long-term learning difficulties
* growth delays

The role of lead in human health is still a fertile area of research.

This topic warrants further research.  We could spend months developing