

IAT 355: Assignment 1 - Data Wrangling Report

Sophia Bhatti, 301206157 // Joanna Long, 301183187

Ideology

The question we are trying to investigate is the distribution changes of drinking water during the past decades, and how much the improvement in urban area differs from rural area of the same country. By studying this dataset, our hypothesis is the quality of water improvement is much slower in rural areas. Our aim is to revoke public attention on improving the water quality in rural areas globally.

Scope

Datasets are retrieved from the gapminder.org, and they are provided by *MDG indicators*. The datasets we chose allow us to differentiate between improved drinking water sources in urban and rural populations from the total population of each country. For the purpose of this assignment, we decided to only include the data from the year of 2000 and the year of 2010 of urban and total population of 165 countries. This is to show a significant comparison of how many water sources improved between those 10 years. Our first dataset is the percentage of clean water sources available to the total populations of each country, and we wanted to correlate this dataset with proportion of drinking water available to urban population of that country.

Methodology

We made use of the `d3.min`, `d3.max`, `d3.mean` and `d3.median` methods available in the `array` class to perform the operations within each column. In order to make the operation function more flexible and adaptable, we decided to store the key of each column into `key` array. A `for` loop is

used here in order to loop through the entire dataset. This way the operation function will not be limited to the chosen dataset.

Data Sources

Gapminder.org

- Proportion of the population using improved drinking water sources, total (%)
 - https://docs.google.com/spreadsheets/d/1I59aRsFJZjDQd7funiWgCQP6KV_s_MVEpTnB3WRUiSOk/pub?gid=0

- Proportion of the population using improved drinking water sources, urban (%)
 - https://docs.google.com/spreadsheets/d/1lGUfQUIIfQLJfHLCpuZS2MroWlVYtffQ_t9sOdYxk_s/pub?gid=0