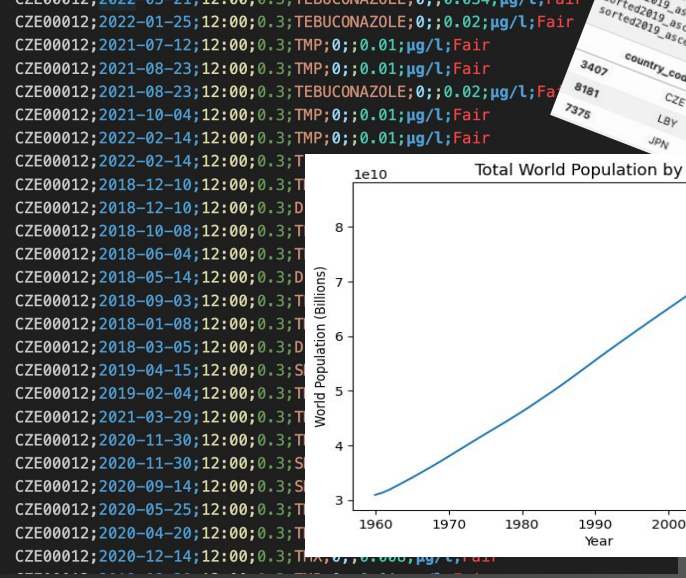
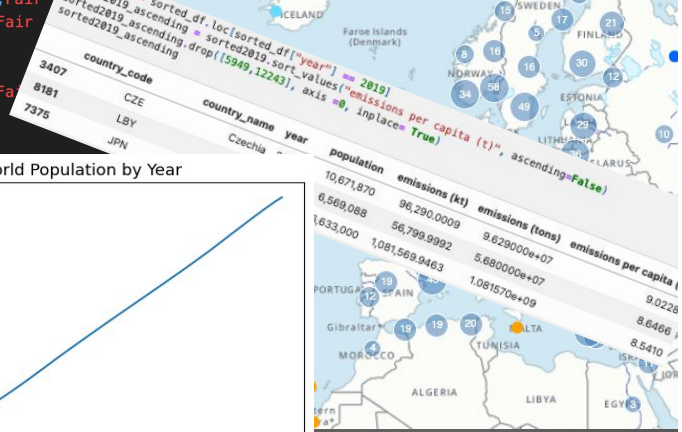
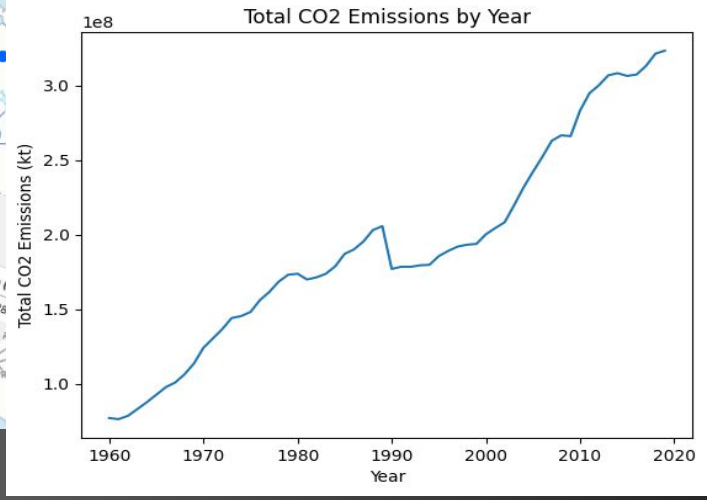






Correlation of CO_2 Emissions — and Water Toxins

By Jesse Rivera, Jerrica Raemer, Blake Singewald and James Montgomery

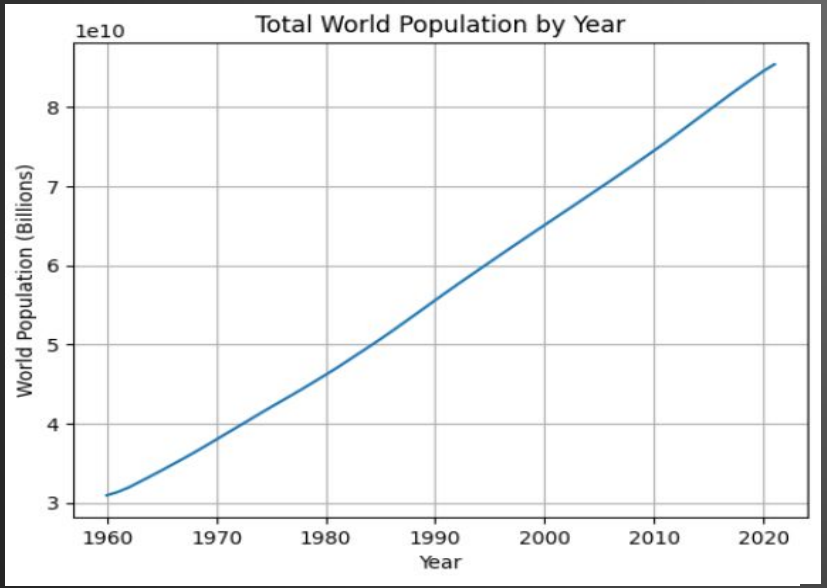
CO_2 - Water Toxins Global Impact

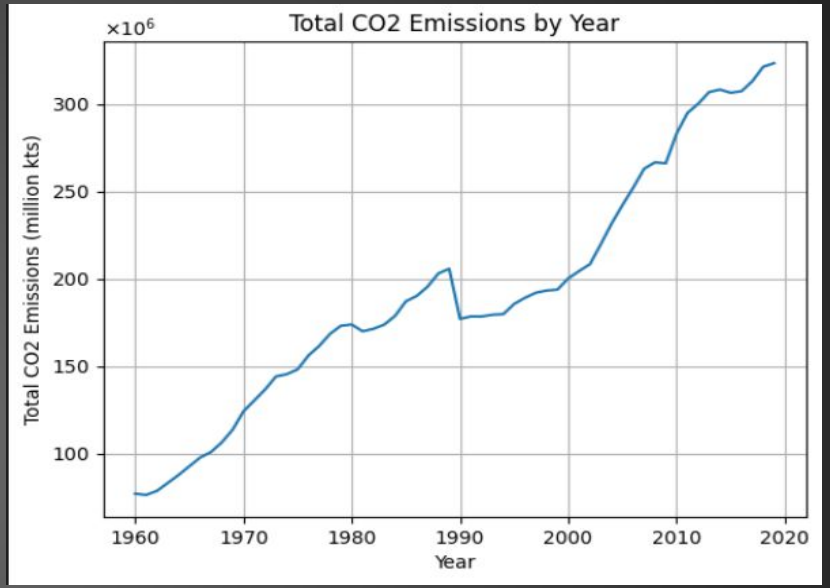
- What are the biggest factors contributing to global CO_2 emissions?
  - Fossil fuels - coal, oil and gas
  - Industry, transportation and electricity

- In 2019 the Global CO_2 emissions were at 33.2 gigatonnes (Gt)

- In 2022, globally, at least 1.7 billion people use a contaminated drinking water source

Carbon Dioxide is released from the combustion of hydrocarbons as a main source of energy.





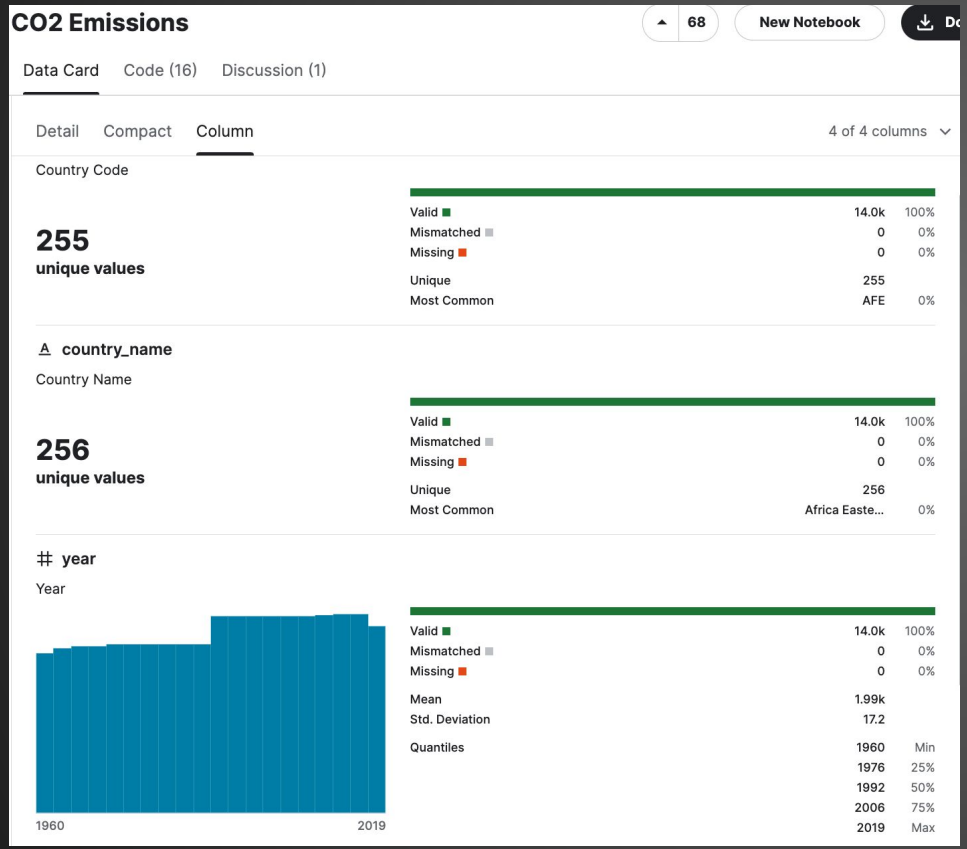
Is there a correlation between CO_2 emissions and toxins in waterways?

Hypothesis:

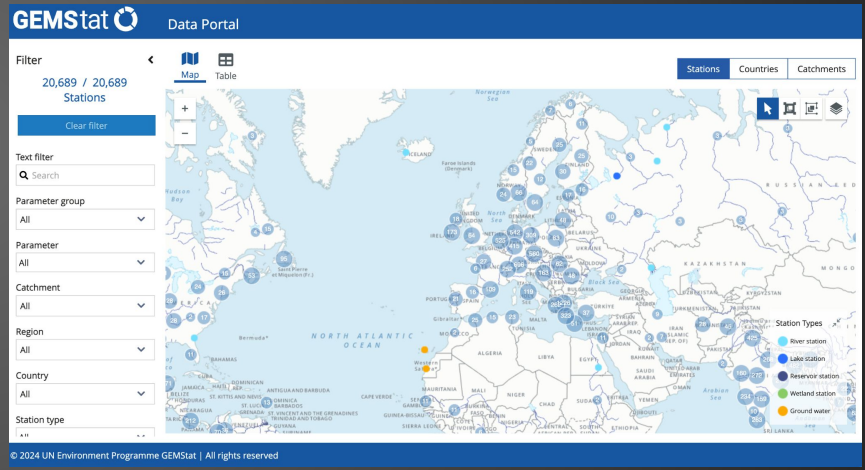
Within countries with high CO_2 emissions there will also be moderate to high levels of toxins in their water catchments.

Data Collection Sources:

https://www.kaggle.com/datasets/ulrikthygepedersen/co2-emissions-by-country



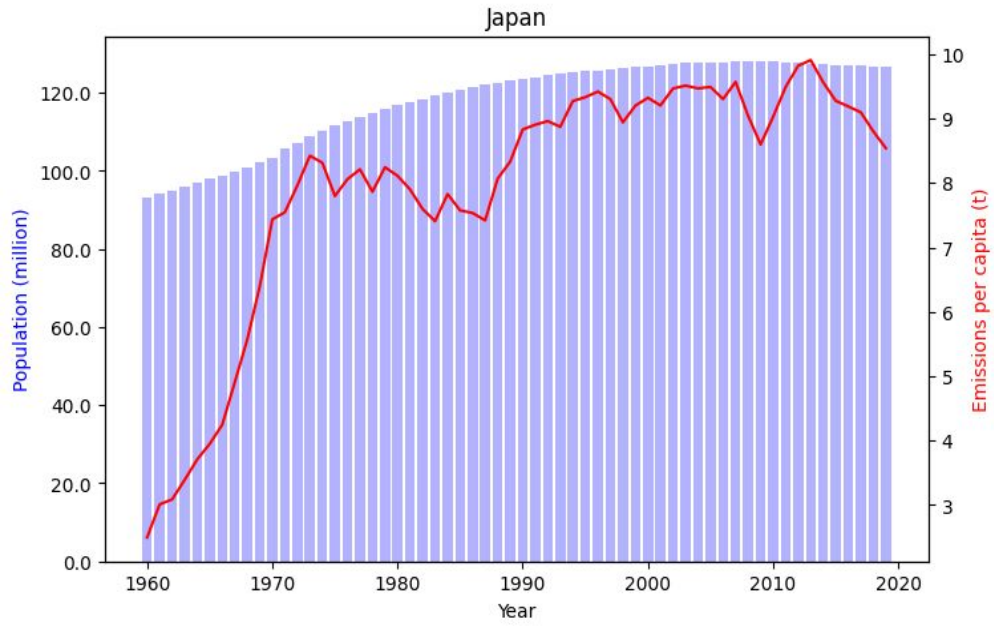
https://portal.gemstat.org/applications/

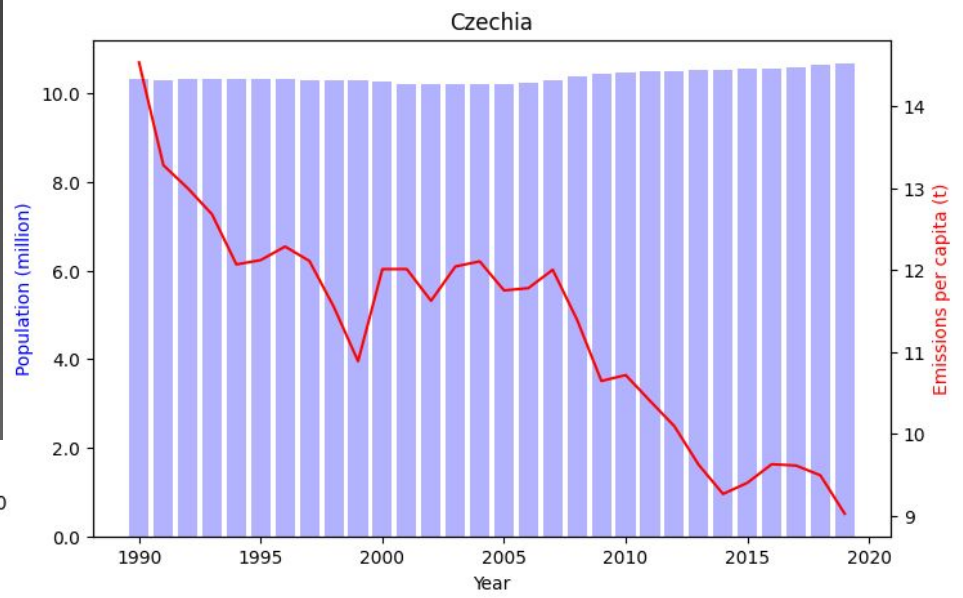


- Calculate emissions per capita
- Limit to 2019 data
- Highest 3 emitters*
- Lowest 3 emitters*

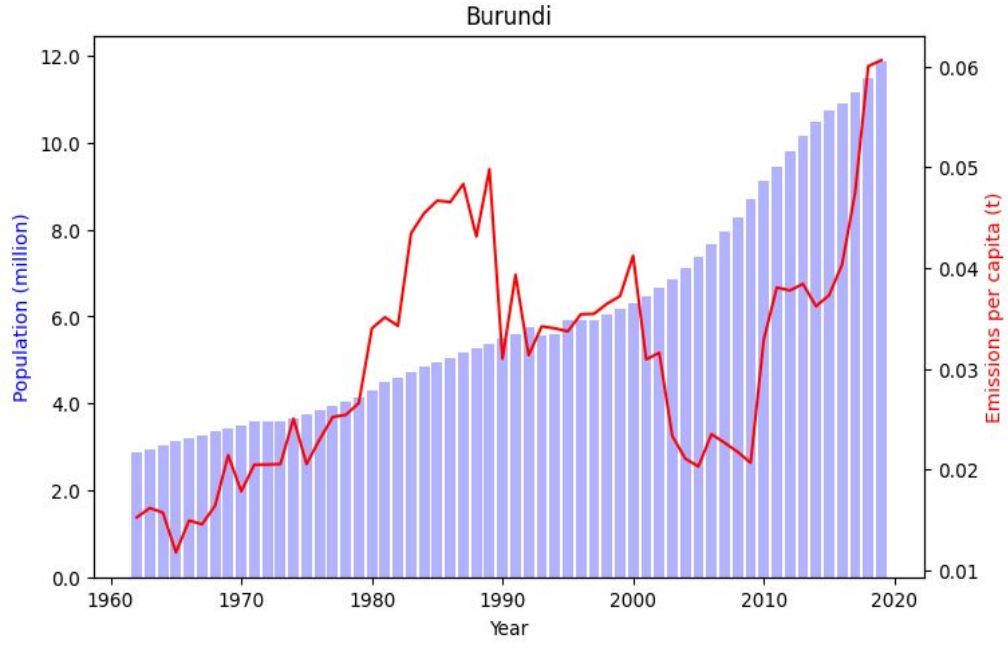
*water toxins data available

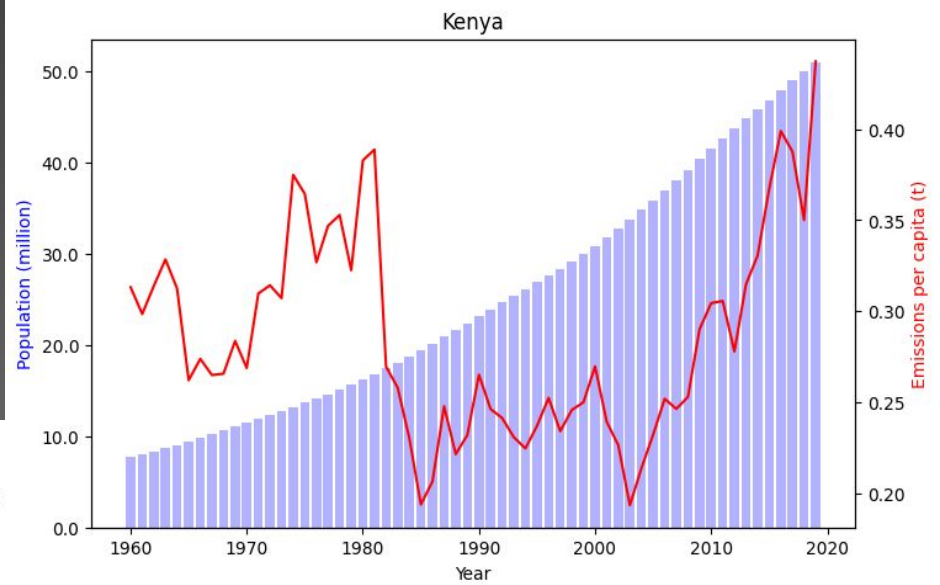
The Data:


Japan
Population (million)
Emissions per capita (t)
Year
0.0
20.0
40.0
60.0
80.0
100.0
120.0
3
4
5
6
7
8
9
10
1960
1970
1980
1990
2000
2010
2020

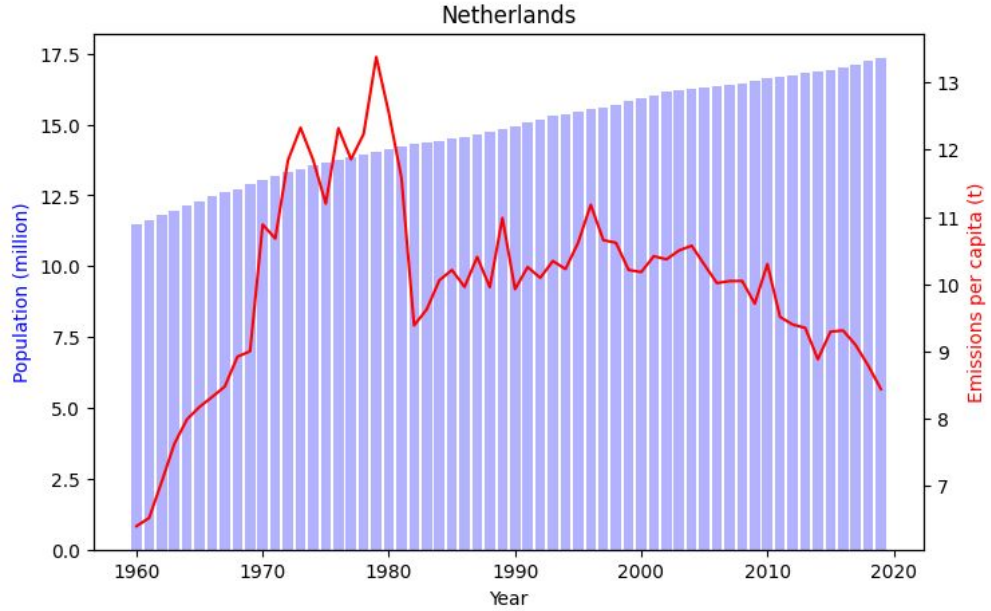

Czechia
Population (million)
Emissions per capita (t)
Year
0.0
2.0
4.0
6.0
8.0
10.0
9
10
11
12
13
14
1990
1995
2000
2005
2010
2015
2020

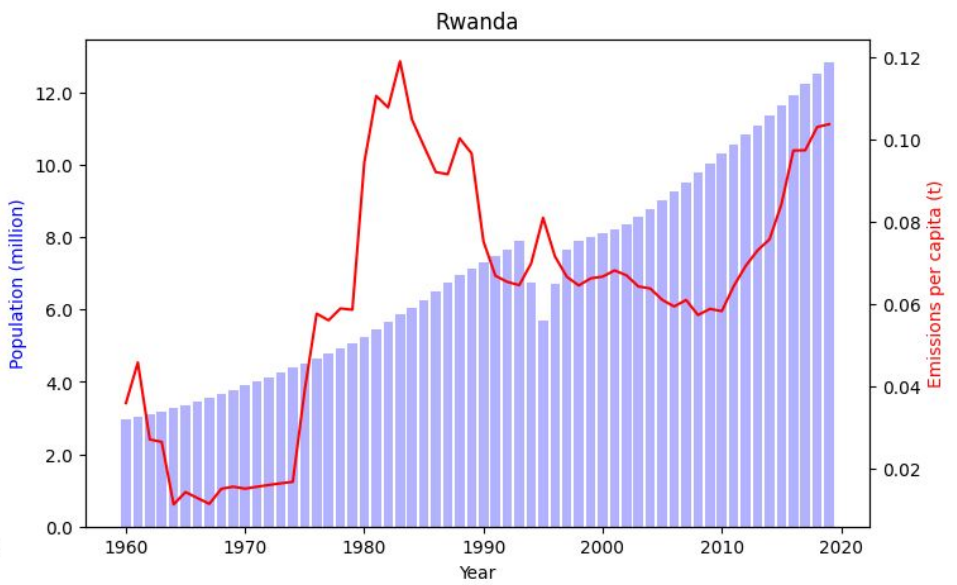
The Data:


Burundi
Population (million)
Emissions per capita (t)
Year
0.0
2.0
4.0
6.0
8.0
10.0
12.0
0.01
0.02
0.03
0.04
0.05
0.06
1960
1970
1980
1990
2000
2010
2020

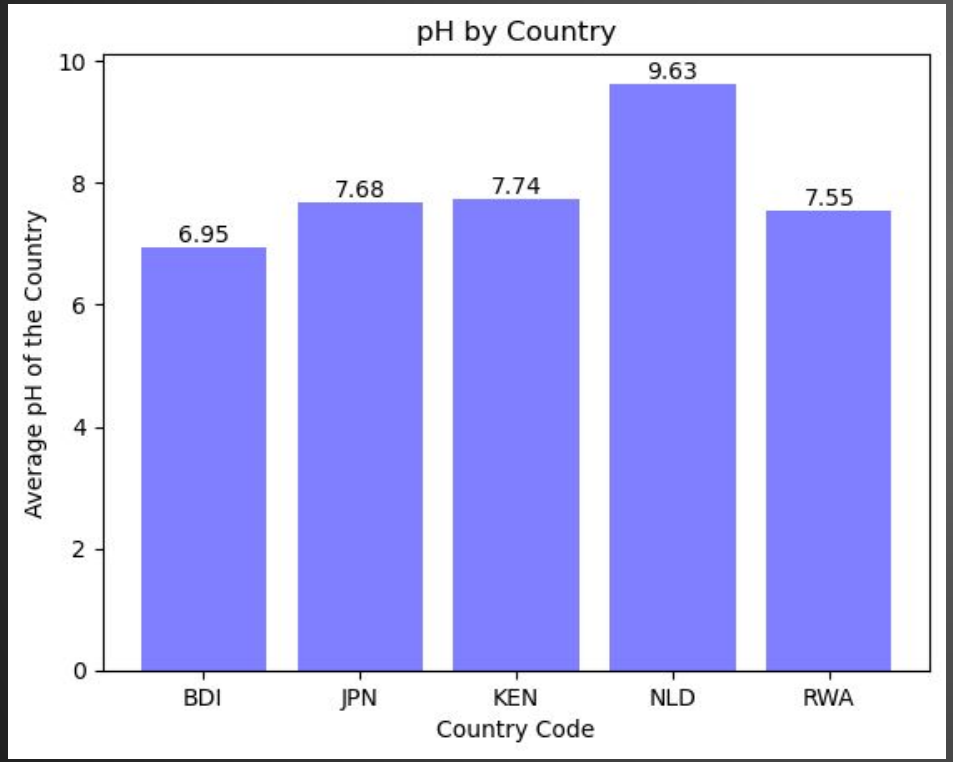

Kenya
Population (million)
Emissions per capita (t)
Year
0.0
10.0
20.0
30.0
40.0
50.0
0.20
0.25
0.30
0.35
0.40
1960
1970
1980
1990
2000
2010
2020

The Data:


Netherlands
Population (million)
Emissions per capita (t)
Year
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
7
8
9
10
11
12
13
1960
1970
1980
1990
2000
2010
2020


Rwanda
Population (million)
Emissions per capita (t)
Year
0.0
2.0
4.0
6.0
8.0
10.0
12.0
0.02
0.04
0.06
0.08
0.10
0.12
1960
1970
1980
1990
2000
2010
2020

The Data:



What can pH tell us?

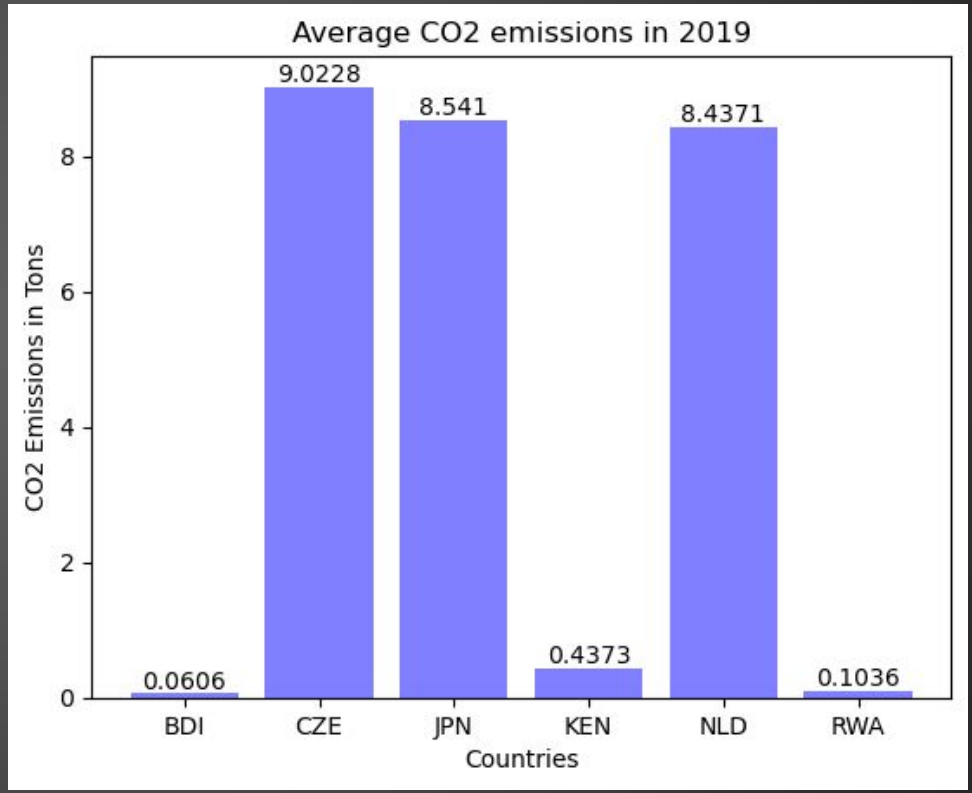
Czechia pH data was not available.
7.61 pH*

Everything but the Netherlands is in a normal pH range:
6.5-8.5

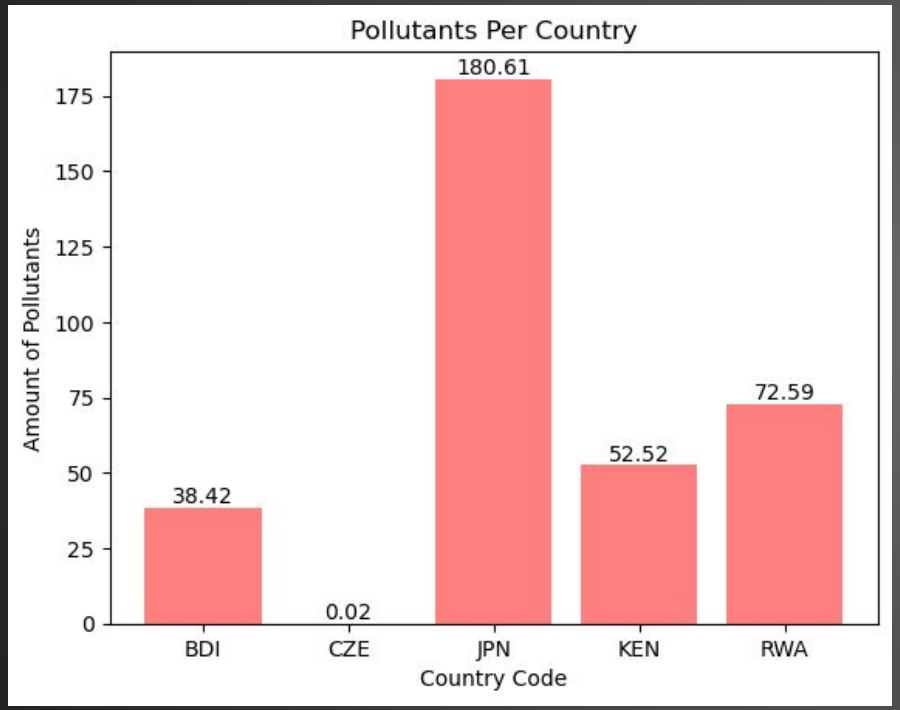
*https://intapi.sciendo.com/pdf/10.2478/johh-2022-0032

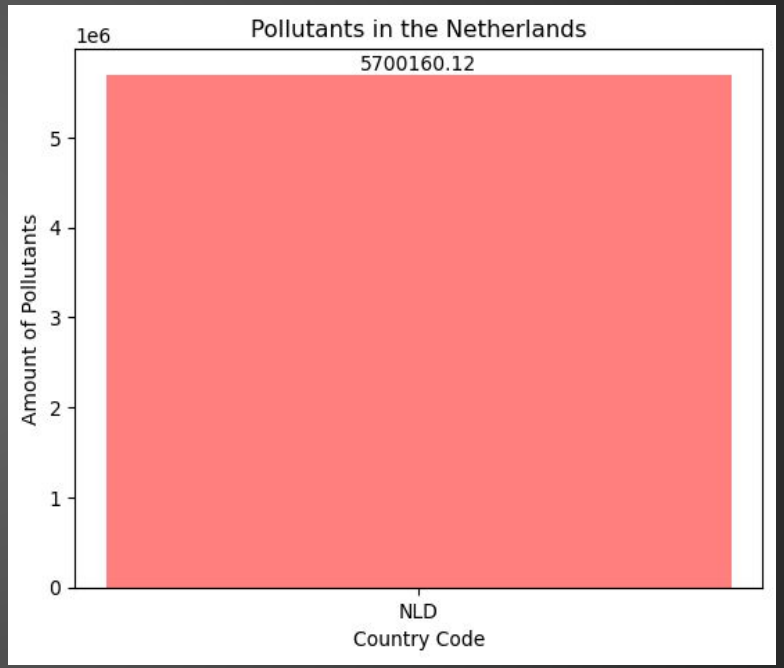
The Data:

Country Codes:
BDI: Burundi
CZE: Czechia
JPN: Japan
KEN: Kenya
NLD: Netherlands
RWA: Rwanda



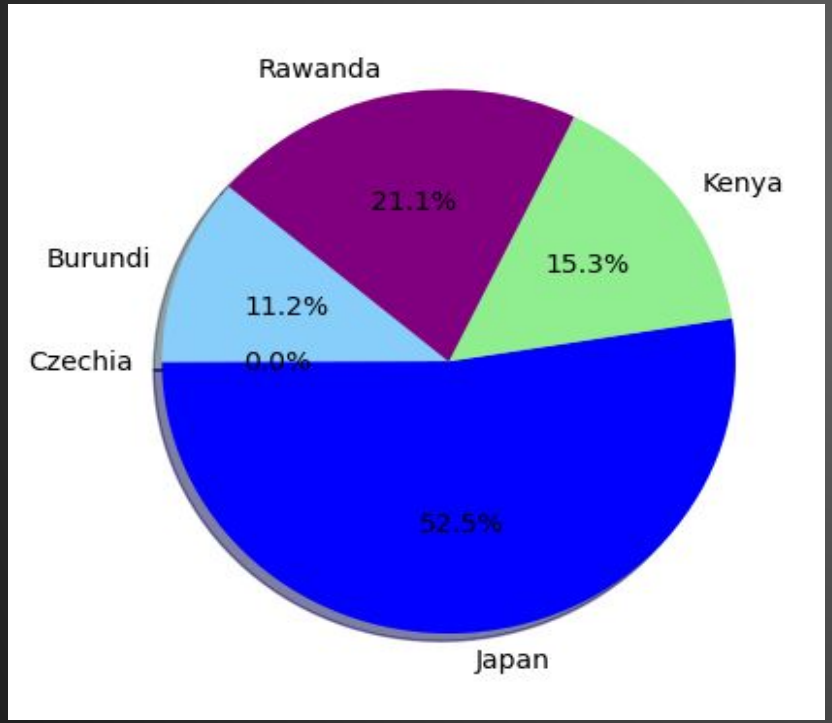
The Data:

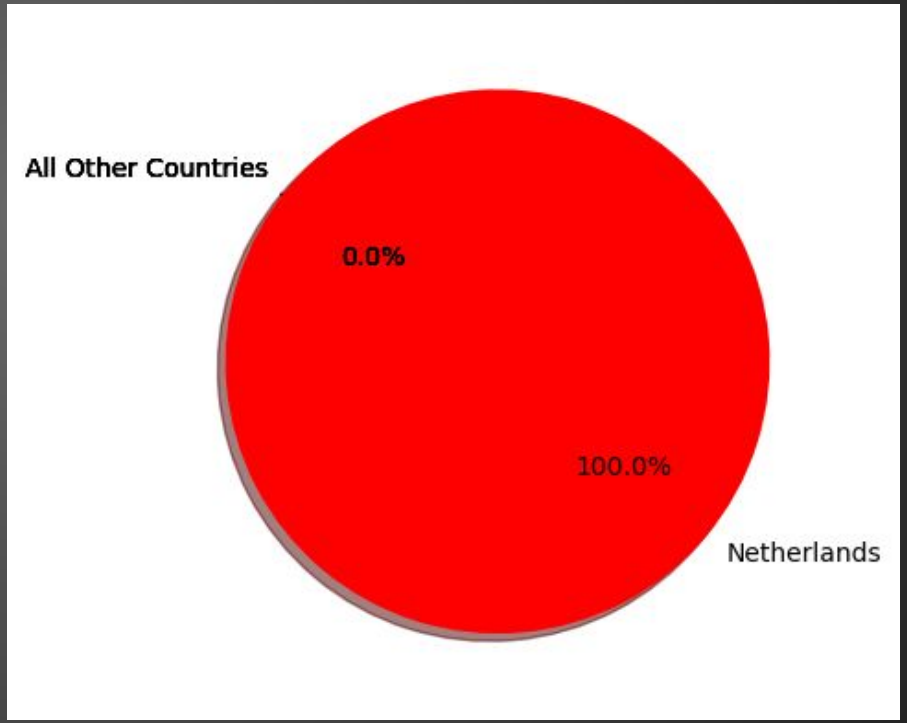

Pollutants Per Country
Amount of Pollutants
0
25
50
75
100
125
150
175
38.42
0.02
180.61
52.52
72.59
BDI
CZE
JPN
KEN
RWA
Country Code


Pollutants in the Netherlands
1e6
Amount of Pollutants
0
1
2
3
4
5
5700160.12
NLD
Country Code

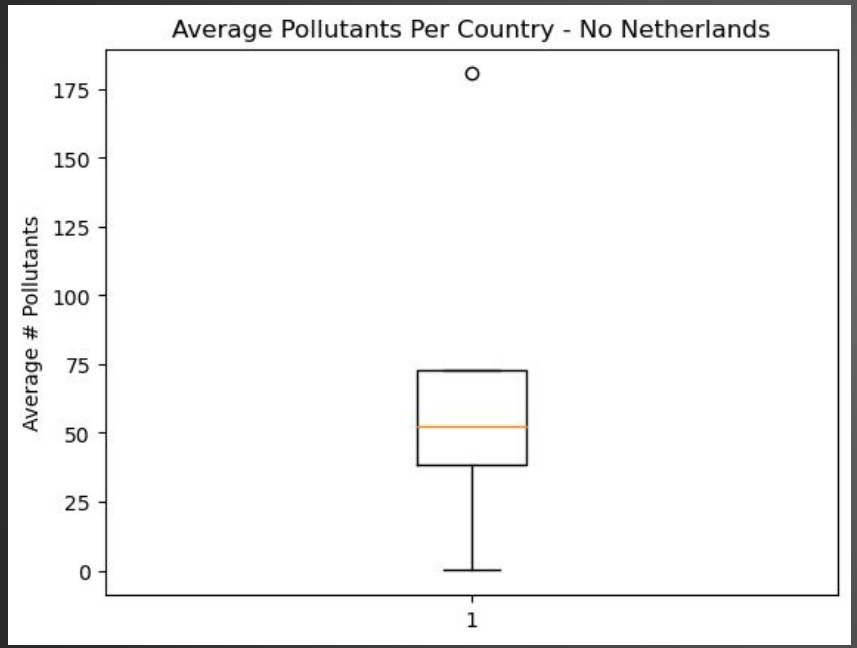
The Data:

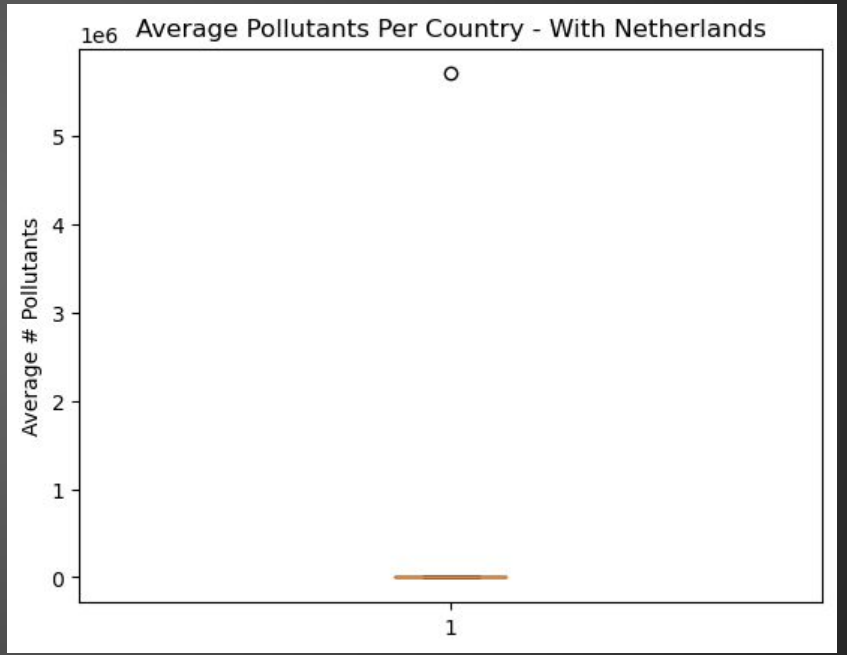
Average Water Pollutants by Country



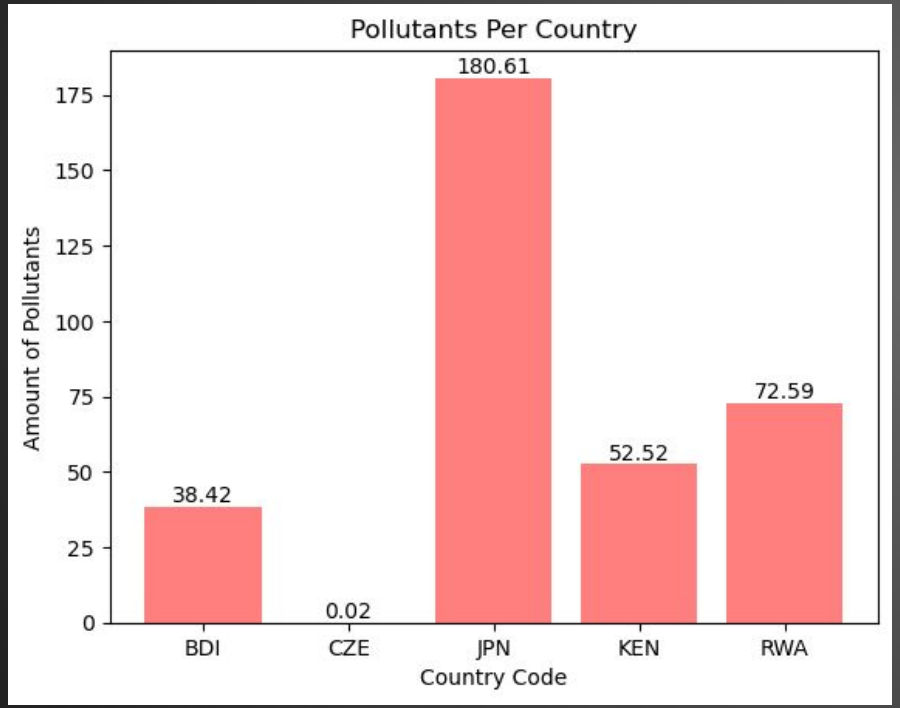


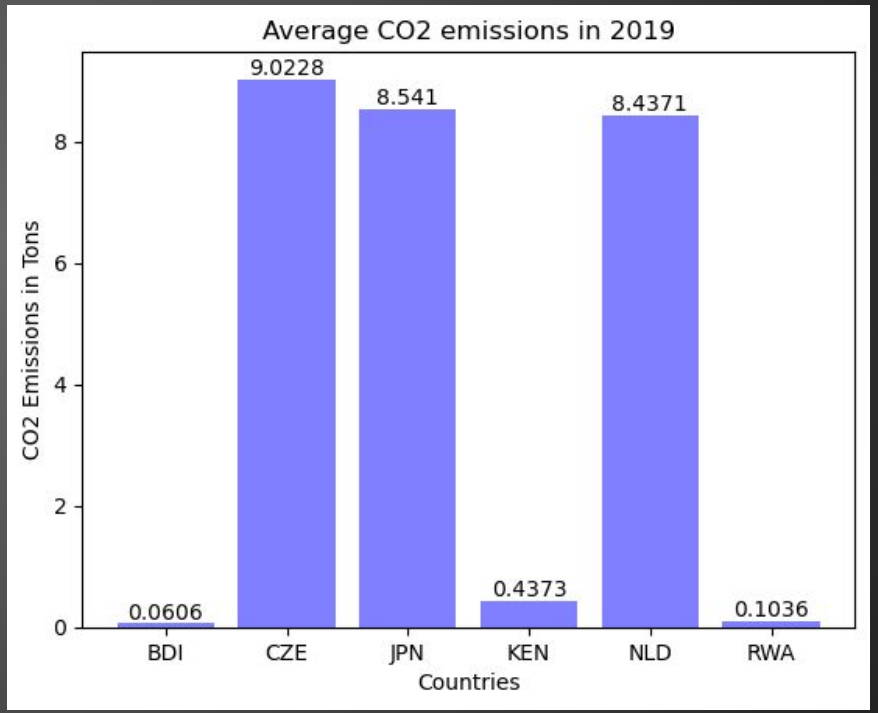
The Data:


Average Pollutants Per Country - No Netherlands
Average # Pollutants
175
150
125
100
75
50
25
0
1

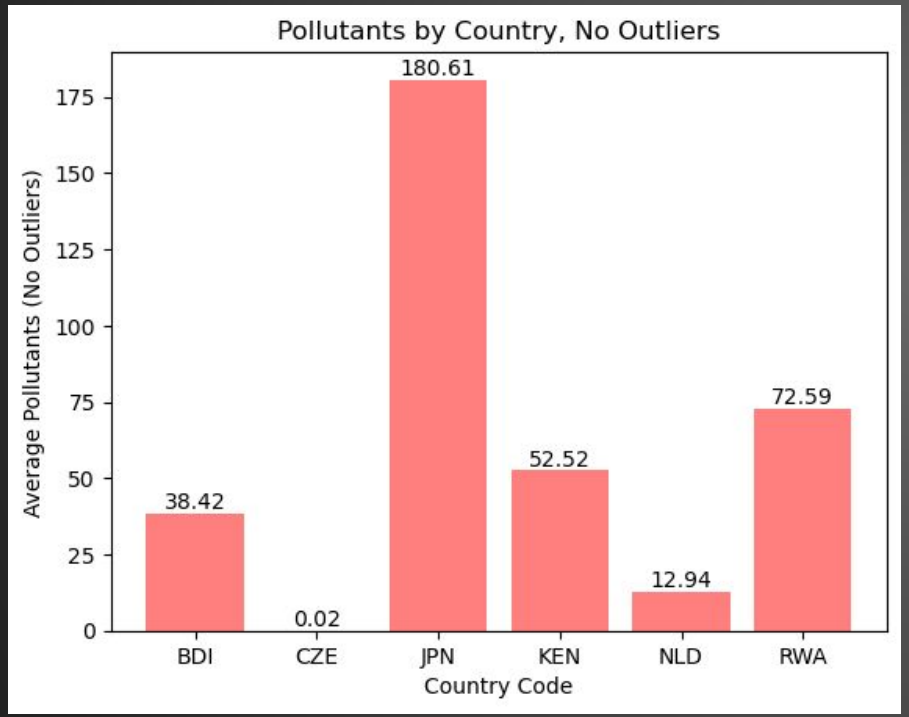

Average Pollutants Per Country - With Netherlands
1e6
Average # Pollutants
5
4
3
2
1
0
1

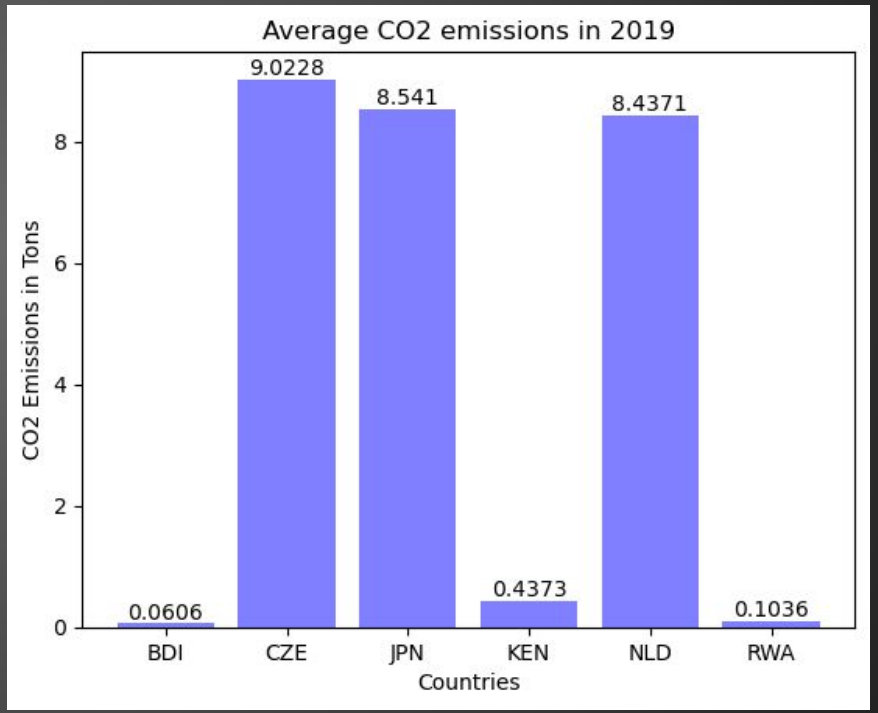
The Data: Side by Side



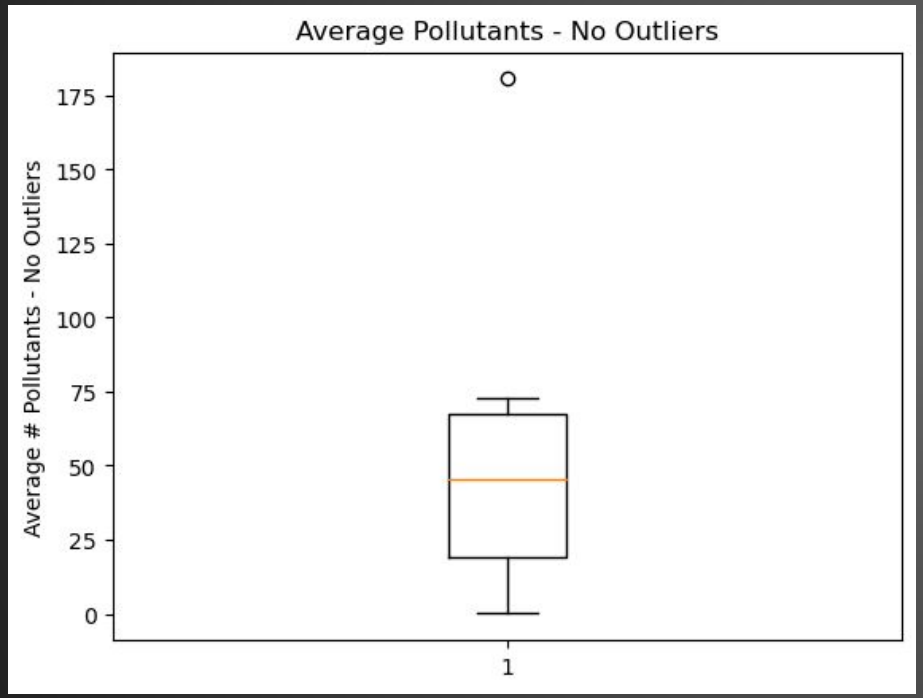


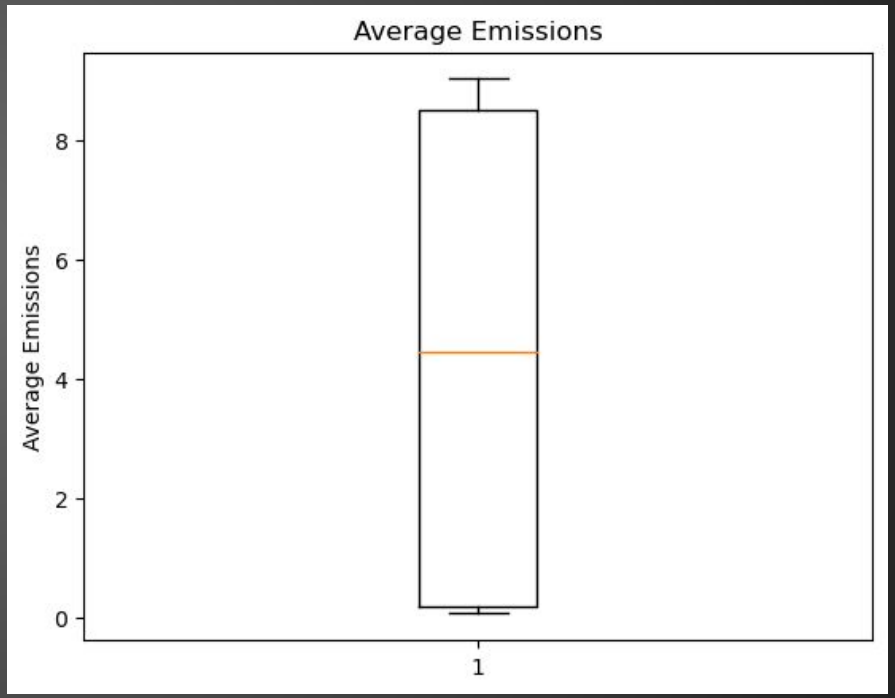
The Data: No Outliers





The Data: No Outliers





The Data: Quartiles

Average # of Pollutants:	
Lower Quartile:	19.31
Upper Quartile:	67.57
Interquartile Range:	48.26
Median:	45.47
Lower Bound:	-53.08
Upper Bound:	139.97

Emissions per capita(tons):	
Lower Quartile:	0.19
Upper Quartile:	8.52
Interquartile Range:	8.33
Median:	4.44
Lower Bound:	-12.31
Upper Bound:	21.01

The Data: Stats

Let X_1 represent the pollution and X_2 represent emissions. Then:

$\mu_1 = 59.52$, $\sigma_1 = 64.87$, and
$\mu_2 = 4.43$, $\sigma_2 = 4.64$.

Computing the two-sample t-test, we get:

$$t = \frac{59.52 - 4.43}{\sqrt{\frac{(64.87)^2}{6} - \frac{(4.64)^2}{6}}} = \frac{55.09}{697.77} \approx 0.085$$

Since the p-value is greater than 0.05, we cannot reject the null hypothesis.

Hence, we have not proven a correlation between CO2 emissions and water pollution.

Conclusion:

- The survey of pollutants (water toxins) within water catchments of top 3 CO_2 emitting countries compared to lowest 3 countries was **not statistically significant**.
- Therefore, our hypothesis of a correlation between CO_2 emissions and moderate to high levels of water toxins cannot be proven or disproven due to **inconclusive data**.
- Why?
  - Possible that high energy producing, industrialized nations have stricter laws against water pollution.
  - Data collection of pollutants across all countries may not be uniform.

Challenges:

- Decision to exclude high CO_2 emitting countries based on being oil rich with very small populations in 2019.

```
country_name
Bahrain                 22.259581
Burundi                  0.060632
Czechia                  9.022786
Japan                    8.540980
Kenya                    0.437279
Kuwait                  20.861949
Netherlands              8.437075
Oman                    16.518321
Qatar                   32.761775
Rwanda                   0.103623
United Arab Emirates    20.502283
Name: emissions per capita (t), dtype: float64
```

2019 Populations:
Qatar 2,760,385
Bahrain 1,477,469
Kuwait 4,360,444
Japan 125,244,761
Netherlands 17,434,557
Rwanda 13,146,362
Kenya 51,985,780

Challenges:

- Unable to use the true top 3 CO_2 emitting countries due to lack of water toxins data. Libya 8.65 tons of CO_2 emissions per capita but no data possibly due to geopolitical reasons.
  - Replaced with Netherlands (8.43) which has its own challenges.
  - Not every water pollutant recorded in every country.

- Netherlands “Bill Gates” water pollutant outlier skews data beyond recognition.

If We Had More Time:

- Deeper dive into specific pollutants that are found within all 6 countries.
  - Create scatter plots for each pollutant across top 3 and lowest 3 countries.
  - Calculate Pearson correlation coefficients, coefficients of determination and linear regression lines.

- Focus on water catchments only near major cities of the 6 countries. Retest hypothesis based on these specific data sets.

Questions?