**Air Quality in the World, the US, and Our Backyard**

**Kristen Lucas**

**Executive Summary**

How clean is the air we breathe? Is our air quality getting better or worse? Depending on where you live on the globe, the answers to these questions could vary greatly. I’m going to look at the air quality in the US compared to other countries, as well as the air quality in Nashville, compared to other large cities. I will be using an open data source from Google [here](https://docs.openaq.org/#api-Locations-GetV1Locations), as well as census info [here](https://www.census.gov/data.html) to determine which large cities at which to look. I’m interested to see how COVID-19 has affected air quality in 2020, so I will be using a yet-to-be determined dataset for that portion (perhaps some of the data we are using today could be helpful).

**Motivation**

There are not many things the entire world has in common, but breathing air is certainly one of them! I’m interested to see what trends there are in areas with good air quality vs bad air quality. I would like to see if there any correlations with the socioeconomics of an area and it’s air quality.

**Data Question(s)**

Some of the questions I would like to answer include:

1. Is air quality in the US improving overall, or getting worse?
2. Which large cities have the best air quality? Top 5?
3. Which large cities have the worst air quality? Bottom 5?
4. Which cities have the greatest air quality improvement over time?
5. Which cities have the worst air quality improvement over time?
6. How do the above questions relate to Nashville? What is Nashville’s ranking?
7. How has COVID-19 affected Nashville’s air quality, and to what extent?

**Minimally Viable Product (MVP)**

I believe I would like to have my finished product in Tableau in the form of a story. I envision starting with the largest scope and continuing to drill down. Start with looking at the US vs other countries, then look at just the US, then Nashville vs other cities, then look at the COVID-19 impact.

**Schedule (through 1/7/2021)**

1. Get the Data (11/17/20)
2. Clean & Explore the Data (12/8/20)
3. Create Presentation (12/12/20)
4. Internal Demos (12/19/2020)
5. Demo Day (1/7/2021)

**Data Sources**

[Google Air Quality Data](https://docs.openaq.org/#api-Locations-GetV1Locations)

[Census Data](https://www.census.gov/data.html)

[GDP Data](https://data.worldbank.org/indicator/NY.GDP.MKTP.CD)

**Known Issues and Challenges**

The main issues I believe I will have will be in some of the joins with this data. When taking the data from several different sources, figuring out how to put it all together will require a lot of figuring out, and likely some help! Another issue is not logistical, but rather within myself. Now that we are seeing the end of class in sight, I beginning to get nervous about this and my ability especially when thinking about finding a job. I’m ready for the challenge, and would love any help!

<https://openaq.org/#/why?_k=6narvj>

<https://openaq.medium.com/where-does-openaq-data-come-from-a5cf9f3a5c85>

<https://github.com/openaq/openaq-info/blob/master/FAQ.md#datakinds>

<https://www.airnow.gov/aqi/aqi-basics/#:~:text=Think%20of%20the%20AQI%20as,300%20represents%20hazardous%20air%20quality>.

<https://docs.airnowapi.org/>