**GGIS224 Final Project Proposal**

My project is mainly centered around analyzing air quality in Jakarta, as it is one of the most polluted cities in Indonesia and globally. My hometown is Bandung, Indonesia; whenever I visit Jakarta on the weekends, I can feel the massive difference in temperature and air quality. In Bandung, the sky is very clear, and the average temperature is about 25˚C; in contrast, the sky is a lot more polluted in Jakarta, and the average temperature there is about 32˚C which is a massive difference considering that Bandung and Jakarta are just 150km apart (93 miles).

Aside from a lack of comfort and hygiene, the declining air quality poses a threat to the population’s health, especially respiratory diseases. Furthermore, due to the high percentage of the population smoking, 31%, according to Tobacco Atlas, it could be one of the main factors that contributes to the rising cancer rate in Indonesia. This project is really important to me as I often travel to Jakarta, and most of my friends live in Jakarta. Jakarta is the capital of Indonesia; therefore, it's really important to find the biggest factor that contributes to the city’s declining air quality.

The project will focus on the physical and natural environment, as I want to investigate the correlation between air quality and respiratory diseases. I will conduct spatial analysis to analyze the distribution of pollutants across different areas of the city and quantify its relationships with the number of respiratory illnesses in the city.

My research question is, “What is the correlation between particulate matter concentrations (PM2.5 and PM10) and the distribution of respiratory disease cases in Jakarta?” It’s important because we can now identify whether poor air quality plays a big role in causing respiratory disease among the Indonesian population. My hypothesis is that high levels of PM2.5 and PM10 are associated with increased cases/concentrations of respiratory diseases among the Jakarta population.

For the construction and modeling of this project, I will first focus on 3 main variables, which are particulate matter concentrations (PM2.5 and PM10), to quantify the levels of these pollutants in Jakarta’s air quality and compare it with health data to see the correlation. Second is the respiratory disease distribution/rates in areas across Jakarta to establish this correlation. I will use the number of respiratory illness cases per population in the given area to calculate the density of the cases. This will make the analysis more accurate as if we're only accounting for the number of cases, a given area with a larger population will likely have more cases. Dividing the cases with the population will make the data more consistent. Lastly, geographical location data of these 2 variables to conduct spatial analysis on the distribution of pollutants and respiratory diseases across Jakarta. For the data sources, I’ve found a dataset for particulate matter concentrations across multiple air quality stations located across Jakarta, which has the PM2.5 and PM10 readings. I will need to find a dataset of reported respiratory diseases across Jakarta to conduct this analysis.

I will conduct two analyses on my dataset. First, I will use spatial analysis to map the distribution of air pollutants and respiratory diseases using libraries like sf, sp, rgdal, rgeos, and tmap. Then, I will use correlation analysis to find the correlation between the levels of PM2.5 and PM10 and the concentration of respiratory diseases.

A diagram of data analysis

Description automatically generated

Works Cited

“Indonesia.” *Tobacco Atlas*, tobaccoatlas.org/factsheets/indonesia/#:~:text=Adult%20smoking%20prevalence%20in%20Indonesia%20is%2031%25.&text=Number%20of%20adult%20smokers%20in%20Indonesia%20is%2061%2C477%2C056.&text=Youth%20smoking%20prevalence%20in%20Indonesia%20is%207%25.&text=Adult%20smokeless%20tobacco%20use%20prevalence%20in%20Indonesia%20is%204%25. Accessed 9 Mar. 2024.