**Project Goals:**

Our goal in this project is to provide a detailed analysis of public air data collected from PurpleAir sensors placed throughout different regions of Nebraska, this request was made by UNMC Water, Climate and Health Working Group. The client requested an evaluation of air quality conditions throughout Nebraska and whether or not pollution conditions are in line with NAAQS.

Specific goals Include:

1. Identify the top 5 locations in Nebraska with the highest mean and median concentrations of VOC, PM 2.5 and PM 10.0.

2. Determine the dates and locations where maximum pollutant values occur.

3. Evaluate if humidity, temperature, geographic location, and elevation have a noticeable influence on air quality.

4. Evaluation of potential Air Quality Index (AQI) health risks for PM2.5 and PM10 based on EPA AQI standards.

**Tasks:**

1. **Organize Air Quality Data:** Data will be collected from Purpleair monitors from different areas of Nebraska from Feb. 2024 to Mar. 2025, which will then be imported into Python. Cleaned CSV files will be included as final deliverables.

2. **Summary Statistics by Sensor Location:** Data sets will be grouped by sensor name and location to compute summary statistics. Summary statistics will include the mean, median, mode, and min/max for VOC, PM2.5, and PM10. Each location will be assigned by their correct pollutant type.

3. **Maximum Pollution Events:** Identify dates and locations that correspond to maximum observed pollution concentrations.

4. **Humidity and Temperature Analysis**: Humidity and temperature will be sorted to the client's liking. Air quality trends will be studied to determine if humidity and temperature affect them at all.

5. **AQI Risks**: PM10 and PM2.5 values will be compared to EPA values to clarify if any of the locations are categorized as "unhealthy" for sensitive air groups.

Deliverables:

- Clean CSV Files
- Annotated Python notebook
- Github repository
- written report

**Coding Tasks:**

Python will be used to complete the client's request, by analyzing and summarizing air quality data collected from PurpleAir from Nebraska. The raw CSV datasets will need to be imported, tables need to be formatted, and ensuring missing data is accounted for and corrected. The data sets will be grouped by location and sensor name to complete summary statistics including mean, median, mode, min/max values for PM2.5, VOC, and PM10 concentrations. Other coding tasks include identifying the locations of areas with maximum amounts of pollution. While also accounting for, humidity and temperature according to how the client needs/wants and according to the AQI standards. In order for complete transparency with the client, tables will be created for clients to visualize the project.

**Non-Coding Tasks:**

Research is the most important thing about this project and its data analysis. Reviewing the EPA National Ambient Air Quality Standards, will be used to determine the potential health risks. Researching the causes of air quality spikes, like agricultural activity, wildfire smoke, and vehicle traffic will also be conducted. Lastly, we could also research the potential influences on air quality, such as, humidity and temperature. A PurpleAir sensor documentation and Epa resources will be reviewed to make sure sensor data is being interpreted properly.

**Technical Expertise:**

The client will be able recognize our company's technical expertise through a well documented analysis. This could include annotated jupyter notebooks, which will be able to explain the steps of the data analysis. A clearly defined scope of work will include, tasks, project goals, and deliverables. With those variables, can ensure the client can verify our results.