Johnathan Sevick

1/23/26

CIVE 202

Project 1 SOW

Team members: Johnathan Sevick

Project Goals:

The main goal of this project is to provide the client (UNMC Water, Climate and Health working group) with meaningful statistical analysis of air quality data provided by the client on a CSV file. They are interested in knowing If the air quality in Nebraska meets National Air Quality Standards, due to "hotspots" and pollutants being a potential health hazard to those in sensitive groups. This statistical analysis includes determining if temperature, humidity, elevation and geographic location play a role in determining air quality. The client has requested that the following tasks be completed, and have requested deliverables by the end of the project.

The tasks requested to be completed by the client include the following specific goals:

- Task 1 – What are the 5 locations in Nebraska with the highest mean and median concentration of VOC, PM 2.5, PM 10.0
    - What are the coding steps needed to complete the task
        - Group by sensor
        - Create mean and median columns
        - Sort descending
        - Analyze data values for each pollution
    - What research is needed that involves no coding
        - Understanding what VOC, PM 2.5, and PM 10.0 are
        - Understanding statistic terms like mean and median
- Task 2 – On what days did the maximum values occur and where did this maximums occur?
- a. What are some potential reasons these maximums occurred on these days in these locations based on initial research?
    - What are the coding steps needed to complete the task
        - Group by date

        - Sort descending
        - Print with date, sensor name, and voc columns
        - Extract dates of interest and counties for investigation
    - What research is needed that involves no coding
        - Understanding what VOC, PM 2.5, and PM 10.0 are
        - Investigate what was happening in Nebraska on extracted dates in extracted counties. (weather, natural disasters, local industries, etc.)
- Task 3 – Does humidity and temperature have a noticeable effect on air quality?
    - What are the coding steps needed to complete the task
        - Create functions for humidity and temperature
        - Group by temperature and humidity
        - Create columns for mean air pollutants
        - Analyze average values of pollutants for different temperatures and humidity levels
    - What research is needed that involves no coding
        - Clearly lay out different humidity levels
        - Clearly lay out different temperature levels
- Task 4 – Have there been any Air Quality Index (AQI) health risks (unhealthy for sensitive populations) at any of the locations in the dataset for PM 2.5 and PM 10 based on the EPA's AQI ratings?
- a. When did they happen?
- b. Based on exploratory research, what could have been a potential cause?
    - What are the coding steps needed to complete the task
        - Use filter to set limit on values
        - Ensure data is grouped by location
        - Need PM 2.5 to be a float type in Python
        - Create a data frame so Python recognizes PM 2.5 as something that can be manipulated
        - Extract data from dataset to find out where issues occurred
    - What research is needed that involves no coding
        - What is AQI and what does PM 2.5 represent
        - What values of PM 2.5 are considered unsafe for sensitive groups
        - Using extracted dates, what happened in Nebraska on those days

Technical experince:

The client will be able recognize our company's technical expertise through a well documented analysis. This could include an annotated coding document, 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 deliverables, we ensure the client can verify our results.

Deliverables:

The client, UNMC, wants to receive several deliverables. Firstly, the raw Python Code file along with the data CSV file and a Python code file that documents the code used to generate the solution. Next, a link to GitHub populated with python code, CSV with data, and README file. Then a well written SOW defining the required tasks to be completed. Following this, an ACD that explains the use of every line of code in great enough detail to be understood. Finally, a final report will also be delivered containing the following: an introduction that gives background info and lays out the tasks, a detailed section describing how the task was completed, a results and discussion section presenting the results of the done analysis and possible reasons for those results, finally, A list of all references used within the report , including documents sent from UNMC.