Project #1: Organizing and Summarizing Air Quality Data

Client: UNMC Water, Climate and Health Group

Background Information:

UNMC has a working group composed of health professionals, environmental scientists and affiliated engineers who work on issues related to how water, air and soil pollution can have negative impacts on human health. UNMC recently started an initiative with AirPurple to monitor air quality conditions across Nebraska. Air quality is not a specialty of any of the scientists currently affiliated with UNMC, so the UNMC group has opened a request for proposals (RFP) from engineering firms and independent consultants to help them to analyze the air quality data set from the AirPurple monitors from February 2024 to March 2025 and answer a series of questions about air quality in Nebraska. Your engineering firm wrote and was awarded the request for proposal and is now tasked with completing the data analysis of the AirPurple dataset.

Client Requests:

The UNMC Water, Climate and Health working group is interested in understanding potential air quality in Nebraska, if air quality is meeting National Ambient Air Quality Standards (NAAQS: https://www.epa.gov/criteria-air-pollutants/naaqs-table) and where potential "hotspots" exist that may be related to public health concerns. The UNMC group wants to understand how temperature, humidity, elevation and geographic location all have an impact on air quality in Nebraska. Specifically, the UNMC group has asked that the following information be included in the final report from your engineering firm:

1. What are the 5 locations in Nebraska with the highest mean and median concentrations of VOC, PM 2.5 and PM 10.0?
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?
3. Does humidity and temperature have a noticeable effect on air quality?
   a. Please use the following categories for humidity:
      i. Low humidity occurs at a relative humidity less than 50%, high humidity occurs at a relative humidity between 50 and 80%, very high humidity occurs at a relative humidity above 80%
   b. Please use the following categories for temperature:

Temperature °F	Category
< 32	Below freezing
32-50	Cool
51-70	Warm
>70	Hot

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?
5. **BONUS:** Does the sensor altitude have an impact on the air quality values?

All results should be grouped by sensor name where possible throughout the data analysis.

Resources Provided by the Client:

1. AirPurple Map: https://map.purpleair.com/air-quality-standards-us-epa-aqi?opt=%2F1%2Flp%2Fa10%2Fp604800%2FcC0#5.21/40.002/-97.008
2. Purple Air Community: https://community.purpleair.com/c/data/api/18
3. EPA Air Quality Tables: EPA Air Quality Index Document.pdf
4. Air Now website: https://www.airnow.gov/

Deliverables:

The following items must be sent to UNMC (the client) by the end of the project. Please submit the following items to Canvas to receive full credit for Project #1.

1. Raw data files used in the Python Code file
    a. This allows your professor to run your code and provide you with detailed feedback. This should include the CSV of the data used to run your code.
2. Python code file that documents your code used to generate your solution. This must be turned in as a .ipynb file. No .py files or .txt files or html links are allowed.
3. ~~A link to your GitHub, populated with Project #1 data:~~
    a. ~~Python code~~
    b. ~~CSV with Data~~
    c. ~~README file that describes the project (you can use your SOW to help with this)~~
4. A well-written scope of work (SOW) related to the client prompt that defines the tasks undertaken by your group to complete the project
5. An annotated code document (ACD) that explains each line of code in detail (https://uofnelincoln-my.sharepoint.com/:w:/g/personal/klane11_unl_edu/EWQY3dadECBAppqDxlIi2UwByWjgxcJOOCsRPVudJPMq6Q?e=RDILxC)
6. A written report that summarizes the findings contained in the code, meets the requests of the client and that can be submitted to UNMC to provide proof of completion of the project.
    a. Report must have the following elements:
        i. Introduction/background information explaining the project, scope of work and objectives of the work completed

      ii.   Methods section detailing how data analysis was completed. Methods should be replicable (another analyst can complete the same analysis and get the same results) and detailed.

     iii.   Results and discussion section that presents the results of the data analysis and explains observations and possible reasons for observations where applicable.

     iv.   A reference list and in-text citations for all resources used in the report, including references to documentation provided by the client.

All files must be submitted to Canvas by Tuesday February 3rd , 2026 at 5:00pm CT.

Skills Students will learn:

- Review basic Python functions
- Write a scope of work
- How to outline functions/steps in code
- Identify sources of data
- Identify types of data
- Extract and load data into Python
- Basic operations in Python to compute statistics
- How to organize raw data

Grading:

Component	Description	Points Possible	Points earned	Instructor Feedback
Raw data files formatted and submitted	Raw data files used in the Python Code file are formatted and submitted to Canvas	5		
Python code file	Python code file (.ipynb file) is submitted to Canvas, can easily be opened and accessed by the instructor and contains all of the code used to generate the solutions in other deliverables	25		
GitHub Link	GitHub link contains raw data, python notebook (.ipynb) and README file	10		
Scope of Work	See scope of work rubric for more details	50		
Annotated Code Document	Please see the rubric for the annotated code document on Canvas	20		
Written summary	a. Report must have the following elements: v. Introduction/background information explaining the project, scope of work and objectives of the work completed	50		

	vi. Methods section detailing how data analysis was completed. Methods should be replicable (another analyst can complete the same analysis and get the same results) and detailed. vii. Results and discussion section that presents the results of the data analysis and explains observations and possible reasons for observations where applicable. viii. A reference list and in-text citations for all resources used in the report, including references to documentation provided by the client.			
TOTAL		160		