Temperature and Crime Case Study Rubric

DS 4002 - Ahmed Ahmed
Due Date: Listed on Canvas
Submission format: Github link submitted on Canvas
Individual assignment:

Why am I doing this?
This case study gives you the chance to walk through a complete data science workflow using a real public safety question relevant to Charlottesville. You will practice retrieving data, cleaning and preparing it, producing visualizations, performing a statistical test, and interpreting the findings. The goal is to strengthen your ability to analyze real-world data and communicate your results clearly and confidently.

What am I going to do?
You will complete the case study using the provided GitHub repository (https://github.com/Ahmed-2026/DS4002-CS3) containing data files, starter code, and supporting materials. Your task is to work through the analysis steps in the Python notebook and generate your own results. By the end, you will calculate the Pearson correlation coefficient between temperature and crime for Charlottesville and summarize what the statistical evidence suggests. To complete the case study, you will:

- Use the notebook to process daily crime and temperature data
- Produce exploratory visualizations
- Compute Pearson correlations for both total crime and selected offense categories
- Summarize your findings in a results markdown file

Your final submission will be a github repository containing the completed notebook and written results.

How will I know I have succeeded?
You will meet expectations when your submitted repository when you follow the criteria in the rubric below.

Spec Category	Spec Details

Formatting	• Submit a link to your GitHub repository that contains all files from the starter template. Only two files should be modified by you, both located in the SCRIPTS folder:
	  ○ correlation.ipynb
	  ○ results.md
	• The repository must be organized so that another student or instructor can follow your workflow without confusion

Source Code File (correlation.ipynb)	● Complete all required code cells in the notebook ● Your notebook must generate ○ Two exploratory plots of your choosing ○ A Pearson correlation coefficient for temperature vs total crime ○ Pearson correlations for temperature vs specific offense categories ● Use meaningful variable names and clear comments so that a reader unfamiliar with your work can understand your process
Results (results.md)	● Write a clear summary of your findings using the numerical output from your analysis ● Include: ○ A short explanation of how Pearson correlation is interpreted ○ Numerical results from the notebook ○ A brief discussion of what the values suggest about temperature and crime in charlottesville ○ Possible factors that could explain the observed patterns