Introduction to Data Science: Midterm Project

We've covered a lot of ground in the first half of class! By now, you have a solid introduction to Python, worked with the Pandas package, created visualizations, and written functions. You'll now combine those skills along with API calls to acquire your own dataset and complete your first data science project. Your goal is to come up with a question that could be answered using your current toolset. Pick a dataset that you can split into multiple groups to compare to one another, make observations and answer a larger question. See the rubric below for specific requirements.

	Below Standards	Meets Standards	Exceeds Standards
API Calls	I have made a few API calls but have been unable to do this systematically or programmatically in bulk.	I have used 1 API and encapsulated that in a loop to do so repeatedly.	I have employed multiple APIs or scraped multiple websites. I have written my calls or scraping logic as a function or class.
Functions	I have written few to no functions, or written either overly complex functions that should be simplified, or overly basic functions that provide little to no value.	I have written several functions to acquire and clean data in a systematic manner and then used them.	I have written complex functions, or functions on top of other functions to effectively perform complex tasks.
Visualization	I have made one or two visuals, but they are not labeled or clear.	I have incorporated at least 5 different visuals, using appropriate labels and	I have refined my colors and styles, and employed multiple types of visuals. They're stunning!
Coding Style	I have made little comments, used cryptic variable names, or rewritten the same lines of code repeatedly which should be encapsulated as functions.	I have appropriately broken down my code into functional chunks and commented my code appropriately.	I have cleaned and edited my code into a finalized product. I've followed PEP8 guidelines and my code could easily be reused or adapted by others.
SQL	I have not used SQL to store my data.	I have created a database and a table which I am populating using API calls.	I have created a database with multiple tables that are linked via primary and foreign keys.
Presentation	I have made some initial graphs or statistics, but have failed to present those as a cohesive story.	I have a defined problem and have at least a preliminary analysis of the question. I have summarized this as a blog post or presentation.	I have a defined problem and have analyzed the problem from multiple angles, synthesizing this into a convincing point of view.

Resources / Thought Starters

-Maryland Terrorism Dataset (GTD)
-Wikipedia
-Center for Diss Control and Prevention
-U.S. Census
-Yelp Api
-Google Maps API
-Twitter
-Kaggle Datasets
-MTA Turnstile Data
-Project Gutenberg