Assignment – UNDP Crisis Bureau - Information Management and Data Science Internship

In this folder, you will find all the information you need to complete this technical test. The test questions are intentionally open to allow your creativity to shine through. There are technically no wrong answers - the idea is for you to mine both structured and unstructured data for insights and analysis using methods of your own choosing. You'll notice that at least one question will involve Natural Language Processing as several recent data-analysis tasks have involved NLP.

Please share all scripts you used and your responses to the questions below in the form of an .IPYNB notebook (preferred) or, in a .py file **AND** pdf/word/ppt (for the responses) by 18:00 on 10 September 2021 EDT.

Questions
1. What trends and patterns can you find when analyzing the ACLED data provided? How did you Analyze the Data? Provide a (very) brief description of the conflict dynamics for the time period / unit of analysis according to your findings. (40 Points)

2. What trends, patterns, and information can you find when analyzing the Newsfeed data-frame provided? How did you analyze the data? What was your methodology? What would you suggest pursuing if you had more time (e.g. a week or two weeks). (40 Points)

3. Can you identify any existing correlations or relationships between the two data frames? (20 Points)

Files:

- Nigeria2019_Newsfeed.csv
    o This file contains a collection of news articles covering events in Nigeria for the year 2019
- Nigeria2019_Newsfeed_Metadata.csv
    o Contains a definition for each column (variable) of interest in NigeriaNewsfeed2019.csv
- Nigeria2019_ACLED_Extract.csv
    o This file contains a catalog of conflict events that occurred in Nigeria in 2019.
- ACLED_Codebook_2019FINAL.docx.pdf
    o This is ACLED's methodology and includes definitions for the variables found in NigeriaACLED2019.csv on page 5.