Data Scientist Consultant Technical Test

In the zip file in the attachment, you will find a number of datasets of different formats that contain pseudo-historical information on variables related to **food security, economic shocks**, **extreme weather events** and **conflict** for several European countries (*). The time period at which the datasets refer is undefined and is identical in all provided data. Moreover, all datasets contain information on the same countries which are identified either through an ID, or a geometry. All geometries are expressed in longitude, latitude (**epsg=4326**), and you may find them in the **Well-Known Text** format (**WKT**).

Here is a list of the datasets:

1. **geometries.shp** contains the geometries of the countries.
2. **economy_index.json:** contains the time series of the national economical index of every country. This is to be interpreted as an indicator of the economical wealth of the country.
3. **rainfall.json** contains the rainfall time series for every country expressed in millimetres.
4. **conflict.json** provides a normalised conflict score for every country, where 1 means high levels of conflict while 0 means no conflict at all.
5. **food_security.csv**: provides a food security score for some of the countries where 1 means high levels of food security while 0 refers to a situation of extreme food insecurity. The formula on which the food security indicator is built can be assumed constant for all countries

The objective of this test is to read the data, understand and describe the story behind it. To do so you will have to merge all the datasets into one *"unique source of truth"* and by triangulating the different layers of information understand the relations that are present. During your analysis and modelling note that variables that relate to extreme weather events, economic shocks and conflict are to be considered drivers of food insecurity.

As guidance to your analysis, you can use the following set of tasks. It is not mandatory to complete all of them, and you may also attempt to go beyond what is written if you wish. They are meant to be used as a suggestion to shape your analysis.

The suggested tasks are:

1. Merge all datasets into one unique file exploiting the information on the provided geometries or the indexes.
2. Analyse and describe the events that occurred according to the data and the relationship between the provided food security indicator and the input data.
3. For the countries with no food security indicator infer its value from the input variables for the entire time window. The entire time series can be reconstructed
4. Analyse the performance of the inference model used.

The expected output is a brief report (2 page max) where your analysis will be outlined and described, and a python repository where all the code used for the analysis is contained. Overall code tidiness and good structure will be valued positively, and quality will be preferred to quantity in the report which should be clear and concise.

(*) The datasets contain synthetic data which has absolutely no relation with any real historical event.