1. Read the file Q2.xml



* Remove all the ‘-’ in contact and change the column type to numeric
* Add ‘Mr. ’ in the beginning of all names
* Replace null with ‘UNK’ for strings and ‘-1’ for numerics

Get the output as a CSV file

1. Please write an application in python language that calls the USGS API and store the result in a relational database of your choice.

https://earthquake.usgs.gov/fdsnws/event/1/

* + 1. Please query all events that have occurred during year 2017
    2. Read the JSON response from the API
    3. Design the database objects (tables) required to store the result in a relational fashion.
    4. Flatten the response and store it in those objects.
    5. Add incremental fetch design to python script
    6. Provide query/analysis to give biggest earthquake of 2017
    7. Provide query/analysis to give most probable hour of the day for the earthquakes bucketed by the range of magnitude (0-1,1-2,2-3,3-4,4-5,5-6,>6 For border values in the bucket, include them in the bucket where the value is a lower limit so for 1 include it in 1-2 bucket)

For question (i-v) provide following

- Code for the integration

- DB schema

For question (vi and vii) provide following

- Queries for the analysis

- Any interesting visualization (using pandas/matplotlib) for these questions.

Refer:

Flatten/Normalize: <https://pandas.pydata.org/docs/reference/api/pandas.json_normalize.html>

Visualization: <https://www.w3schools.com/python/pandas/pandas_plotting.asp>

1. Push the code for the above 3 questions to Git. Provide the repository/project details and attach screenshots.