

Response Summary:

1. Student Information *

First Name	Keegan
Last Name	Palonis
Major	Data Visualization
Course (e.g. CGT 270-001)	CGT 270
Term (e.g. F2019)	S2022

2. Email Address *

(University Email Address is required.)

kpalonis@purdue.edu

3. Visualization Assignment *

- Lab Assignment

Generate

4. Identify appropriate data sources: is the data publicly available? What search methods were used? *

Data source 1	Seismographic recordings published
Data source 2	New Reports after the earthquake
Data source 3	Books analyzing the earthquakes in different regions of the world

5. Data format: what format is the data in? Structured vs instructed? All text, a combination, multiple sources? Is it primary or secondary data? *

The data is in a plain text format, is structured, and is all text, with both letters and numbers. This is primary data, since it is listing events that occurred and when.

6. Data types: what types of data are in the data? How are they stored? What is the access to the data (API, JSON, txt, csv, etc.)? What structure holds the data (data base, spreadsheet, etc.)? *

There are both numbers and words in the data, in a file, and is an xlsx file. The data is held in a spreadsheet

Evaluate

7. Variables: list the data variables? What are the parameters? Give them names. What are the dependent variables and independent variables? *

Date and Time, Location (latitude and longitude), depth, magnitude, type of magnitude, number of reports, the gap, distance from the earthquake that data was recorded, travel time of the earthquake, the identification of the earthquake, when it was updated, and where the earthquake was located. The Dependent variables are everything about the earthquake itself, and the independent variables are the identification of the earthquake

8. Audience & Assumptions: list any assumptions you have about the data. Who is your audience? *

The assumptions about the data is that the data is accurate, that all earthquakes that were of magnitude 6 or higher were recorded, and that no data points were removed. The audience is anyone that is interested in earthquakes and where and when they were located.

Generate

9. What real life behavior does the data reflect? Does it show patterns of activity, regularity of events, a timeline, population data, etc? Explain. *

The data shows a natural disaster, where most earthquakes of magnitude 6 occur in similar areas and occur fairly regularly, as well as a timeline.

11. What are the weaknesses of the data source? Is it likely that the source will be available in the future? Is the data complete? What is the quality of the data? Is it specific to your needs for the current project? Is the data in the format you need? Are there missing data? Explain. *

The weakness is only that it has not been updated since 2014, and there is some data that is missing from the document, as certain earthquakes do not have the number of reports listed, or the gap.

12. What information is emphasized? What is the central focus of the data? Explain. *

The focus of the data is where and when magnitude 6 earthquakes occur, as the data collected is all of the magnitude 6 earthquakes recorded

13. At what level of granularity is the data provided? Is the data summarized, or do you have access to the raw data? Is the data categorized or is the data in a format that allows you to create your own categories, etc. Explain. *

The raw data is present, and is categorized by earthquake, and I would be able to modify the categories that are present. It is built in a spreadsheet, and has labels in the columns.

14. What is the scope of the data? What topics can be covered using the data? Is there a time range/frame? Is the data for a specific area/discipline/demographic etc.? Explain. *

The topics that can be covered are where and when earthquakes of a magnitude of 6 occur, how far away facilities to record the data are, and is mostly for academic use, as the layman wouldn't have much use for the data.
