Download as PDF

URL to view Results	[Click Here]

Response Summary:

Mine Worksheet

Goal: to identify patterns, extreme and subtle features about the data
Objectives: Students will identify basic descriptors for the data, and categorize the data according to the specifications from the Parse Worksheet
Outcomes: Three (3) specific questions to be answered using the data

1. Student Information *

First Name	Samruddhi
Last Name	Tawade
Course ***(e.g. CGT 270-001)***	CGT270
Term ***(e.g. F2019)***	F2021

2. Email Address *
stawade@purdue.edu

3. Visualization Assignment *
- Lab Assignment

Analyze

4. Basic Descriptors: for each data component from the Parse Worksheet, identify basic descriptors (basic statistics). Explain *
The variables Time, Date, country, city, id, area, and direction are all either characters or strings, and can therefore cannot be quantified. However, the Latitude, Longitude, Depth, distance, xm, md, and richter are all floats and can be quantified.

5. Categorize: consider what is similar and what is different? Categorize the

data. Are the variables categorical (normal, ordinal, or rank). Are they quantitative (discrete or continuous)? Show categories. Explain. *

All of these variables are quantitative because their values are a result of counting or measuring something. They are also discrete variables because they are being measured in a finite amount of time. The categories "date" and "time" are direct proof of this.

6. Temporal: is the data streaming data? How is it stored (all at one time, over several years in years, days, minutes, seconds)? Explain. *

The data is stored over the span of many years. The spread is not organized by year so it's difficult to say exactly how many years it was spread across. This is not streaming data because the earthquakes do not happen continuously.

7. Range and Distribution: what is the distribution of the data? Few values, small size, evenly spread, sparse or dense? Explain. *

The distribution has 24,007 values, so it is extremely dense. We can't tell whether or not it is evenly spread because the spread is not organized by year.

Evaluate

8. Questions and Assumptions: list at least 3 questions you plan to answer with the data or list the questions if they were provided. Must be complete sentences and end in a question mark. What assumptions are you making? *

Question 1	In which region did the most earthquakes occur?
Question 2	What is the average magnitude of these earthquakes?
Question 3	Do any of the earthquakes in this spread show up in the other spreads?
Assumptions	I am assuming the the hashtags are there because there wasn't any data to record for that category.