Response Summary:

1. Student Information *

First Name	Keegan
Last Name	Palonis
Major	Data Visualization
Course (e.g. CGT 270-001)	CGT-003
Term (e.g. F2019)	S2022

2. Email Address *
(University Email Address is required.)
 kpalonis@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 *
 time: From 1/5/1900 at 19:00:00.000 to 3/17/2014 at 5:11:34.980
 Date: From 1/5/1900 to 3/17/2014
 Time: Various Times Depending on earthquakes
 Latitude: Range [-71, 84.964]
 Longitude: Range [-179.99, 180]
 Depth: Range [0, 720]
 Magnitude: Range [6, 9.6] with one null value Avg: 6.56
 MagType: 12 Categories including null
 Nst: Range [0, 999] as well as nulls
 Gap: Range [8.7, 230.4] as well as nulls
 Dmin: Range [0.0180, 26.0630] as well as nulls
 Rms: Range [0, 1.68] as well as nulls
 Net: 6 categories
 Id: First part of the id matches the net category
 Updated: Range [2013-02-08T21:36:30.656Z (2/8/2013 at 21:36:30.656), 2014-03-19T16:40:31.884Z (3/19/2014 at 16:40:31.884)]
 Place: Landmark location relationship (101 km SW of Atka Alaska) 8304 values
 Type: 1 Category - Earthquake

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. *

Time: continuous, could be any date/time between the ranges
Date: Continuous, same as above
Time: Continuous
Latitude: Continuous
Longitude: Continuous
Depth: Continuous
Magnitude: Continuous
MagType: Nominal
Nst: Continuous
Gap: Continuous
Dmin: Continuous
Rms: Continuous
Net: Nominal
Id: Nominal
Updated: Nominal
Place: Nominal
Type: Nominal

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

This is not streaming data, as the last earthquake listed in the data is from March of 2014. There have been magnitude 6 earthquakes since then, and the data was not updated. It is stored all at one time, and cannot be edited by users.

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

The magnitude values have a small range (only taking into account magnitude 6 earthquakes and up) but is dense, but clusters at the lower values and not evenly spread, though there are some higher values.

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	Where do the most category six earthquakes occur? Where are they rare?
Question 2	Is there a pattern between magnitude and frequency? Do more frequent earthquakes cause smaller or larger earthquakes?
Question 3	Is there any trend of frequency and size since 1900? Are earthquakes in places more or less frequent? More powerful?
Assumptions	The earthquakes that occur in the ocean will be related to the country listed in the place category, or the closest country given by latitude and longitude. Fault lines are not present everywhere, and since they are what cause earthquakes, places without these will not be present.