Response Summary:

1. Student Information *

First Name	Ema
Last Name	Westerfeld
Major	Animation
Course ***(e.g. CGT 270-001)***	CGT 270-02
Term ***(e.g. F2019)***	S2022

2. Email Address *
(University Email Address is required.)
ewesterf@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	Weather.gov
Data source 2	Currentresults.com
Data source 3	Mapsofworld.com

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 structured into columns. Most of the data is temperatures, but there are also years, if Phil saw his shadow, and locations. It is secondary data.

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.)? *
The Phil column and locations are string data. The years are integers, but the temperatures are float points. The data is in a csv format 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? *
The year ranges from 1886 to 2016. The Punxatawny Phil Column records if he saw no, partial, or full shadow. The dependent variables are the overall average February temperature and the rankings.

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

I assume that the data is correct and complete. My audience is those who believe in the groundhog prediction.

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 reflects the belief that Punxatawny Phil can predict the weather.

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. *

Phil supposedly predicts whether winter will end in six weeks, but the data stops after February. There is a gap in temperatures from 1886 - 1894.

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

The focus of the data is the average temperatures.

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. *

I do not have access to the raw data, only the average temperatures. I can create my own categories.

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 data can be used for weather predictions and trends.