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	Thomas
Last Name	Cluff
Course *(e.g. CGT 270-001)*	CGT 270-009
Term *(e.g. F2019)*	F2021

2. Email Address *
tcluff@purdue.edu

3. Visualization Assignment *
- Training Data

Analyze

4. Basic Descriptors: for each data component from the Parse Worksheet, identify basic descriptors (basic statistics). Explain *
ID - float: no descriptors because this is essentially the primary key, Min and max could be calculated but its just 1 and the total number of Pokémon.
Name - String: Name is a string so length and mode (mode is useless because each Pokémon has a unique name) can be calculated.
Type - String: Type is a string so length and mode can be calculated.
Total, HP, Attack, Defense, Special Attack, Special Defense, and Speed are all integers so min, max, average, median, and mode could be calculated.

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. *

ID is technically Ratio as there is a meaningful zero. But you cannot calculate things such as median, average, etc. because it is a primary key. Because it is a primary key system, ID is discrete as some ID's can have decimals to them (but not all of them have decimals and it isn't an actual decimal. More of a sub-ID).
Name is a discrete Nominal variable because it has no clear ranking based on just the name. The ID is what gives ranking to the name of the Pokémon.
Type is also a discrete Nominal variable because there is no clear order for arranging/listing types. There are community standards for listing the types, but they are arbitrary and have no statistical/mathematic/data-vis purpose.
Total, HP, Attack, Defense, Special Attack, Special Defense, and Speed are all ratio data that is continuous for an integer. A zero is the lowest value you can have in a stat and all stats are built upon from there.

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

None of this data is dependent on time. It exists all the time in the newest generation of Pokémon games. It does not change while playing the game nor does it change the longer the game has been released

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

Total, HP, Attack, Defense, Special Attack, Special Defense, and Speed are evenly distributed with no real outliers. Obviously ID is extremely even distributed and so is Name. For the different types, Water has the most and Fairy the least. Every other Type is around 60 Pokémon with that type.

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	What is the average HP for each elemental type of Pokémon?
Question 2	Which elemental type of Pokémon has the highest average attack stat?
Question 3	Which elemental type of Pokémon has the highest average Defense and Special Defense stat?
Assumptions	I am assuming that when calculating stat type statistics that we are only considering Generation 8's stats, As these are the newest so older generations are deprecated.