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

2. Email Address *
myers436@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 *

ID: Integer - Mode, Median
Name: String - Mode
Age: Float - Mean, Median, Mode
sex: Boolean - Mode
breed: String - Mode
Date Found: Integer - Mean, Median, Mode
Adoptable: Integer - Mean, Median, Mode
Posted: Integer- Mean, Median, Mode
Color: String - Mode
Coat: String - Mode
Size: Boolean - Mode, Median
Neutered: Boolean - Mode, Median
Housebroken: Boolean - Mode, Median
Likes People: Boolean - Mode, Median
Likes Children: Boolean - Mode, Median
Gets along with males: Boolean - Mode, Median
Gets along with females: Boolean - Mode, Median
Gets along with cats: Boolean - Mode, Median
Keep in: Float - Mode

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. *

Nominal: Name, Sex, Breed, Color, Coat, Keep In
Ordinal: Size, Neutered, Housebroken, Likes People, Likes Children, Gets Along with Males/Females, Gets along with cats
Interval: ID
Ratio: Age, Date Found, Date Adoptable, Date Posted

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

No, It represents dogs posted from a variety of dates. It was updated multiple times to add new dogs that become adoptable. For example there are dogs that were posted from 2005-2019.

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

Data set is large with 2937 rows. The data is dense showing statistics from many different categories and info about adoptable dogs.

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 dog breed is most widely available?
Question 2	What dog breed is least widely available?
Question 3	Do most dogs get along with cats?
Assumptions	I can assume that there is a dog that is more widely available and less widely available. I will not know this until I calculate the data.