## **Response Summary:**

#### 1. Student Information \*

First Name	Adam
Last Name	Abounnaaim
Major	Computer Graphics Technology
Course (e.g. CGT 270-001)	CGT 27000-007
<b>Term</b> (e.g. F2019)	S2022

#### 2. Email Address \*

(University Email Address is required.) aabounna@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 integer components of the data such as the Total Yes, Total No, and Total Abstain columns all have a tendency to be primarily a majority of Total Yes in each row, with occasional exceptions. The integers in the count column also increase with value as it descends. Alphanumeric string components such as Vote Date, Question, and Bill Number also change accordingly throughout the data of course, and it appears that a majority of resolutions made it only to the House of Representatives since it comprises the majority of both Total Abstain and Total No.

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. \*

The data throughout the set is certainly discrete data, since there are no precise measurements taken over a long period of time to extrapolate patterns or ranges from that, nor are the values present in the document fractions or decimals; the data in the set is entirely either whole number integers or alphanumeric strings, and no continuous relationship can be established between each and every value. The variables present throughout the document are mainly numerical, such as Vote Count or Record, but others are categorical and specifically ordinal since columns such as Vote Result denote different categories that are ordered and ranked between each other. (Passed, Agreed to, Failed)

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 was compiled and stored at one time, since the source and context of the data has concluded, the 113th Congress. It certainly is not streaming data since the source of the data will have no more updates nor changes that will be done.

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

The distribution of the data is very spread; for Vote Result, the highest values concentrate around the values of "Passed" and "Failed," and a plateau of numbers emerges in between them, with only a smaller peak for the value of "Agreed To." The data is remarkably sparse in its patterns, though when it comes to the amount of bills a sponsor was involved in, the distribution changes a bit; the Republican party dominates and has much larger values than the Democrat party.

# **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	Are there any connections between bills of a certain category?
Question 2	Are any shifts in trend or pattern noticeable over the timeframe of the Congress?
Question 3	How many bills only made it to the House of Representatives and not the Senate?
Assumptions	I assume that the data is accurate, and has not somehow been colored or biased by any particular source towards any particular idea or narrative. I also assume that the data is clean fairly enough already to begin deducing patterns and observations, though some further tidying may be required.