Response Summary:

Filter & Represent Worksheet

Goal: Produce visual representation of the data to address Question 1
Objectives: Students will extract data related to Question 1
Outcomes: Two visual representations of the same data showing different aspects of the data for Question 1

1. Student Information *

First Name	Joy (Chia-Hua)
Last Name	Lin
Course *(e.g. CGT 270-001)*	CGT 270-LC4
Term *(e.g. F2019)*	F2021

2. Email Address *
lin1424@purdue.edu

3. Visualization Assignment *
- Training Data

Q16. How many questions have visualizations?
- Two

Q17. Question 1
What were the top five most popular baby names in the US cumulatively from 1910 to 2012?

Q18. Question 2
How popular were biblical names in the state of California and how did this change over time?

Remember

Question 1: *

What were the top five most popular baby names in the US cumulatively from 1910 to 2012?

Apply

5. Filter the data: Remove any duplicate or any data unrelated to answering your question. Provide a description of the filtered data (what is needed to answer your question). *

I used the Occurrences field and the Top Name field. The Occurrences field is needed because I am trying to add all the number of occurrences together to figure out which ones were the most popular. The Top name field is the one I am trying to add together from all the states, so that is also necessary to keep to answer the question. I also included the Gender field for extra information with coloring the bubbles in my packed bubbles visualization, because I do not think it is distracting but also the extra information is nice to see, but it is not necessary to answer this question.

Evaluate

6. Next Step: Answer the following questions: *

Do you have enough data? Explain. If no, explain then revisit the Acquire Worksheet.	I have enough data.
Do you have the right data to answer Question 1? If yes, explain then proceed. If no, then revisit 'Filter the Data' question. Repeat until this answer is yes.	I have the right data to answer Question 1, because it has the amount of occurrences for each state for each top name. The five most popular baby names in the US are Mary with 2,510,913 occurrences, Michael with 2,274,105 occurrences, James with 1,566,897 occurrences, Robert with 1,213,750 occurrences, and Jennifer with 811,054 occurrences.

8. View 1 for Question 1 *

[Click here]

9. View 2 for Question 1 *

[Click here]

Remember

Question 2: *

How popular were biblical names in the state of California and how did this change over time?

Apply

Q41. Filter the data: Remove any duplicate or any data unrelated to answering your question. Provide a description of the filtered data (what is needed to answer your question). *

I filtered the data so only the data from the state of California remained. I kept the Year field because it was important to see I then created a calculated field in Tableau using Top Name as the parameter to label each entry's name as either true if it was a biblical name or false for non-biblical names. I did research to determine if a name was biblical or not, and then created this field based on my results.

Evaluate

Q43. Next Step: Answer the following questions: *

Do you have enough data? Explain. If no, explain then revisit the Acquire Worksheet.	I have enough data, after creating a field for labeling whether or not a name is biblical.
Do you have the right data to answer Question 2? If yes, explain then proceed. If no, then revisit 'Filter the Data' question. Repeat until this answer is yes.	I should have the right data to answer Question 2. Biblical names in the state of California had a spike in popularity around 1955 and also around 1985, while non-biblical names had a spike in popularity around 1946. Overall, biblical names were more popular than non-biblical names in California, with 551,247 occurrences of biblical names and 322,557 for non-biblical names.

Q44. View 1 for Question 2 *
[Click here]

Q45. View 2 for Question 2 *
[Click here]