Proposed topic:
It is interesting to note how many people in the world have the same name. Sometimes it seems like a certain name is more common in a certain state or country.
My project topic is inspired by the above fact and would like to analyze the data sets of US Baby names and answer the below question:
1.Which is the most popular name at the national level?
2.Which is the most popular name at the state level?

The two different data formats available for answering the above questions are
1.SQLite database that contains two tables NationalNames and StateNames (130.4MB)
2.NationalNames.csv (11.54MB)

Source :From data.gov

These data sets would have the following fields:

- Id

- Name

- Year

- Gender

- State (only for StateNames)

- Count

I would need to convert the SQL data formats tables to CSV, then combine them using python packages (import sqlite3) to get a file with the all the names, id, year, gender, state and count at national level and state level.
I would be able to generate the frequency of the names as per state and national level.

One interesting visualization can be to plot a graph showing the trend of a certain name for the time period of last ten years . That could be X-Y (Year vs Count) and highlight the the trend as per gender.