homework #3

1. what data have you gathered:
   * 10 years of National Data of the frequency of given names in the population of US births where the individuals have a Social Security number
2. how did you gather it
   * Obtained the 10 years of data sets from http://www.ssa.gov/OACT/babynames/limits.html
3. how have you explored it
   * opened the dataset in python
4. what have you cleaned
   * nothing, data is cleaned already
5. what still needs cleaning
   * nothing
6. what insights have you found
   * every year, the top 10 names do not repeat
7. what is your question and how does the data answer it
   * check the popularity and trend of names that my fiancé and I pick
8. do you need to gather more data or adjust your question
   * get names associated to crime
   * get names associated with celebrities
9. how might you use modeling to answer your question
   * classify names in categories