Baby Names

1. Please describe the format of the data files. Can you identify any
    limitations or distortions of the data?

 - The data is in csv format.
 - Names delinated to babies prior to 1935 are recorded after birth as social security
cards weren't issued till then. Therefore there are fewer names listed.

2. What is the most popular name of all time? (Of either gender.)
 - Top male name is James
 - Top Female name is Mary

3. What is the most gender ambiguous name in 2013? 1945?
Because gender equal names with count 5F/5M can score higher than gender equal
names with count 100F/100M we use a cutoff of count 100 when considering the
gender neutral names. Therefore in 2013 our top gender ambiguous names are Jael,
Milan, Lennon. In 1945 we have our top three as Frankie, Leslie, Jessie.

4. Of the names represented in the data, find the name that has had the largest
    percentage increase in popularity since 1980. Largest decrease?

99% change increase of Latoya from 1980 to 2014

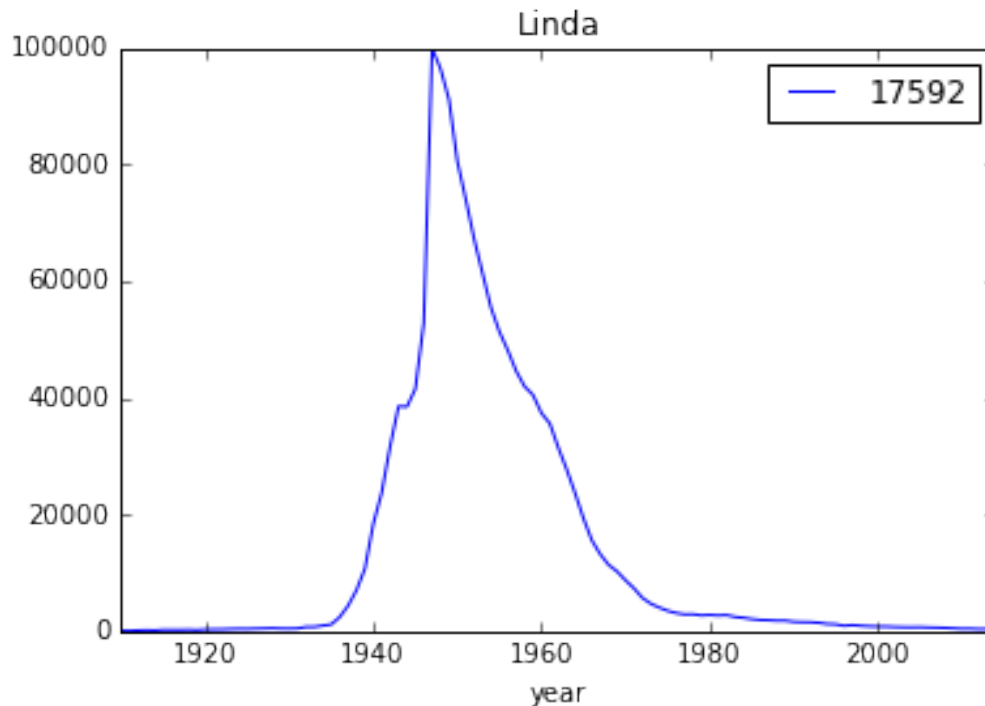| Latoya	5	2480	0.997984

99% change decrease of Colton from 1980 to 2014

Colton	6335	5	-0.999211

5. Can you identify names that may have had an even larger increase or decrease
    in popularity?
By pulling the Min and Max counts of all names in the data set and calculating the
difference, we can find the names with the biggest changes in popularity.

```
Linda           99757
Michael         92100
Robert          8621
```

Insight

A lot of names that sound similar are spelled different. Here we use the python fuzzy library to take a look at which names have the most diverse spelling and how they relate to their more common counterpart. This library has three algorithm measurements for word distance and we use all three to distinguish the names.

We graph the names using the Graphviz library. White boxed names are the most commonly used names of the group and the colored names are the next closest names. We use the Levenshtein distance to calculate how similarly spelled each name is from another and use this metric to space out the boxes and circles. The names in circles are the second and third most common form names of the group. As examples we have the names Ken and Karen.