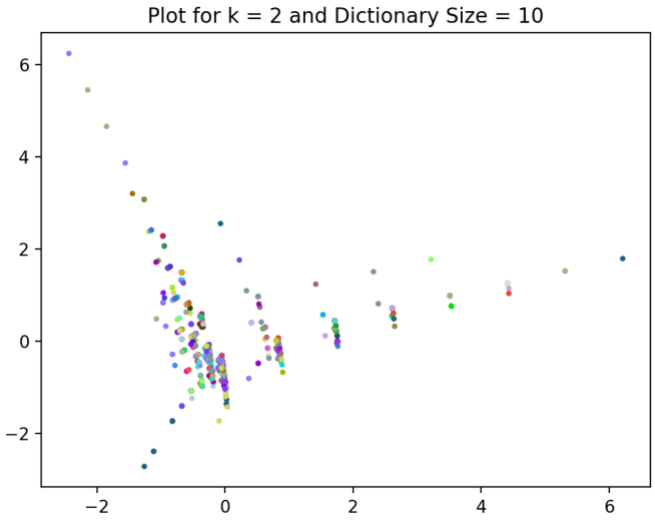
CS165B Final

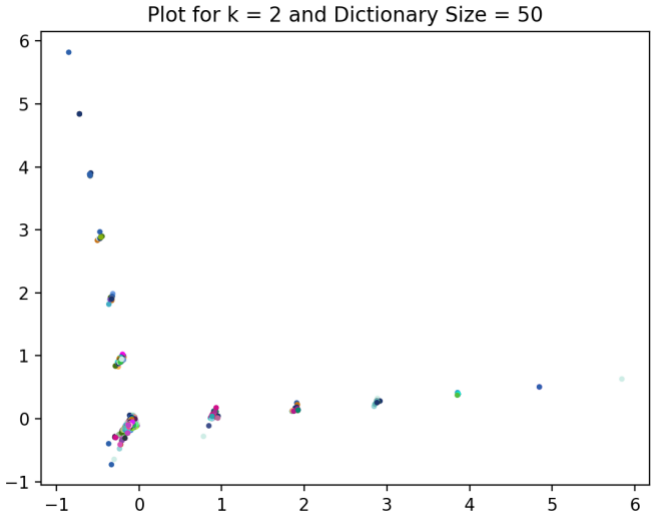
Image Recognition Report

By Amanda Serex

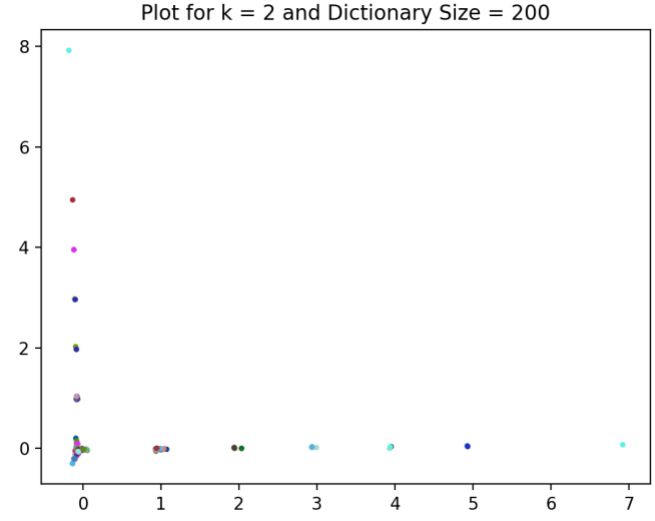
Plots When k = 2



Accuracy = 10.27%
Run Time: 160.4 sec

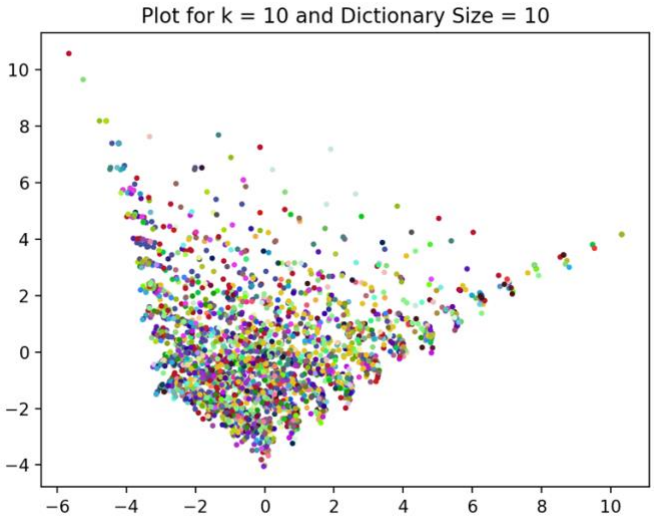


Accuracy = 12.02%
Run Time: 174.1 sec

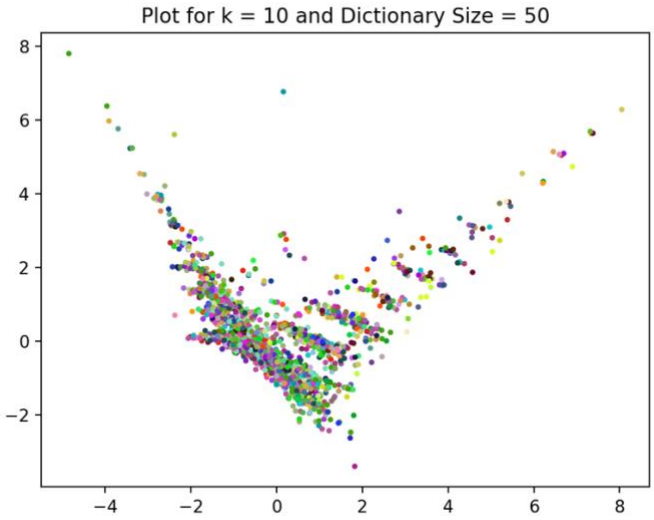


Accuracy = 12.70%
Run Time: 181.3 sec

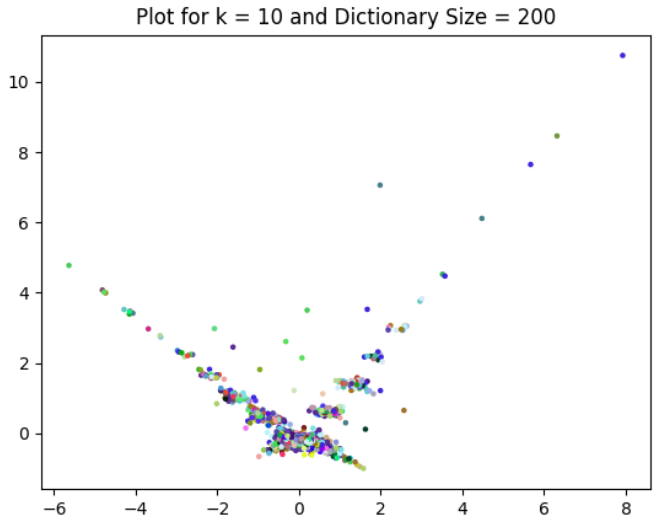
Plots When k = 10



Accuracy = 12.67%
Run Time: 146.3 sec

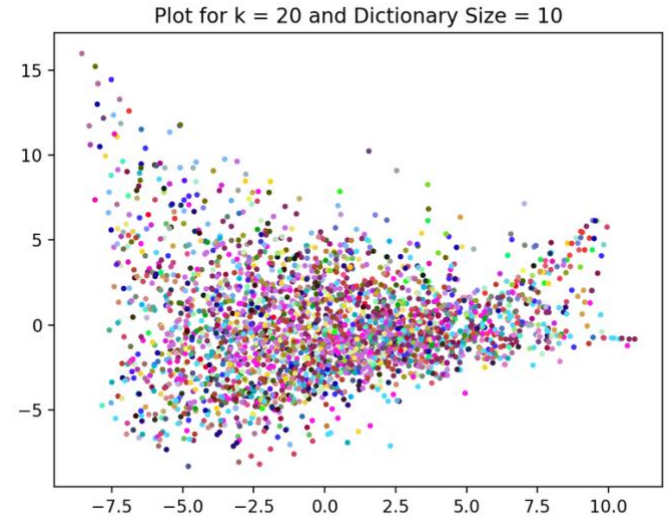


Accuracy = 15.78%
Run Time: 172.5 sec

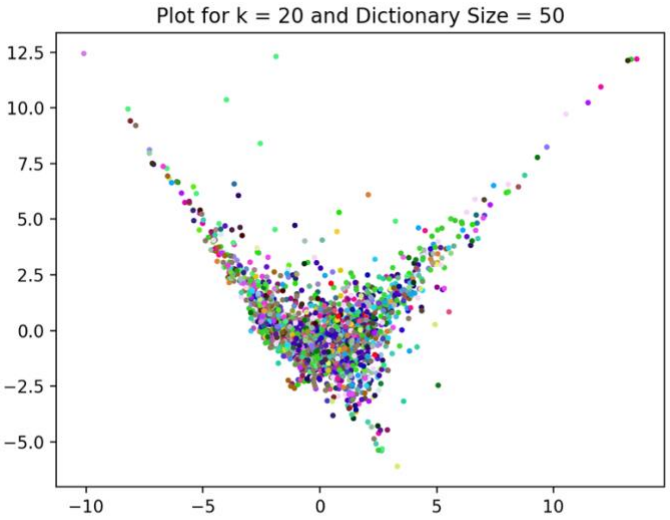


Accuracy = 16.69%
Run Time: 209.8 sec

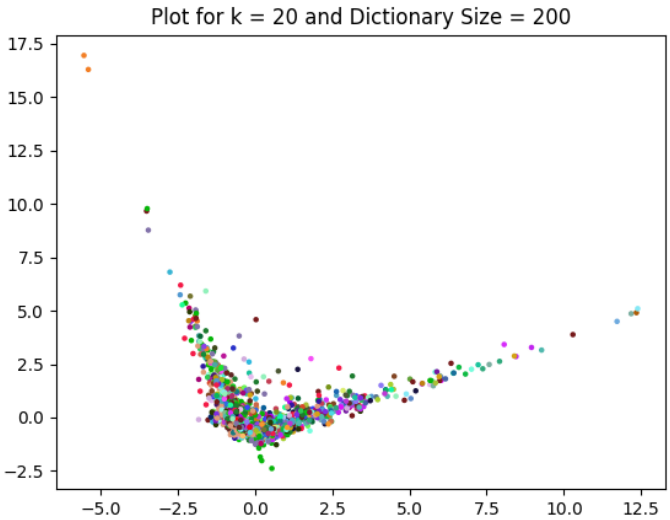
Plots When K = 20



Accuracy = 14.14%
Run Time: 169.1 sec



Accuracy = 18.78%
Run Time: 178.1 sec



Accuracy = 19.93%
Run Time: 189.8

Analysis:

The graphs above were created with the Background data set so the accuracies would likely be better if it had been excluded. My data shows that as k, the number of means, increases the accuracy increases as well regardless of dictionary size. Similarly, for all levels of k, as dictionary size increases the accuracy also increases. It can be seen in each set of charts with the same k value, as the dictionary size increases, the scatter plot becomes more and more refined. For example, the plot with inputs k = 10 all have same general shape, but the points become far less spread out in relation to each other as dictionary size increases. This is true for all three different k values and the corresponding dictionary sizes. Therefore, the highest k value and highest dictionary size I tested resulted in the best accuracy. I found that while this is the general trend it is not always true if poor kmeans are selected at the beginning. The times above are how long it takes to run the program once with those inputs. As k increased, the time increased. The time was pretty similar for all of the same size dictionaries though.