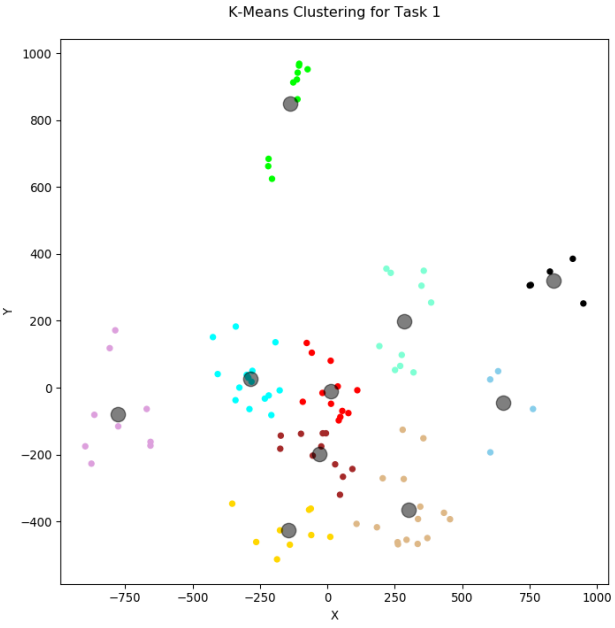
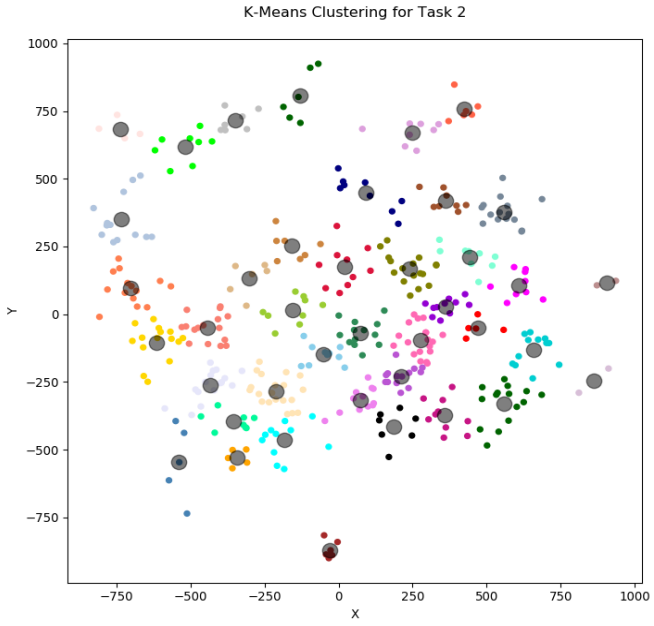
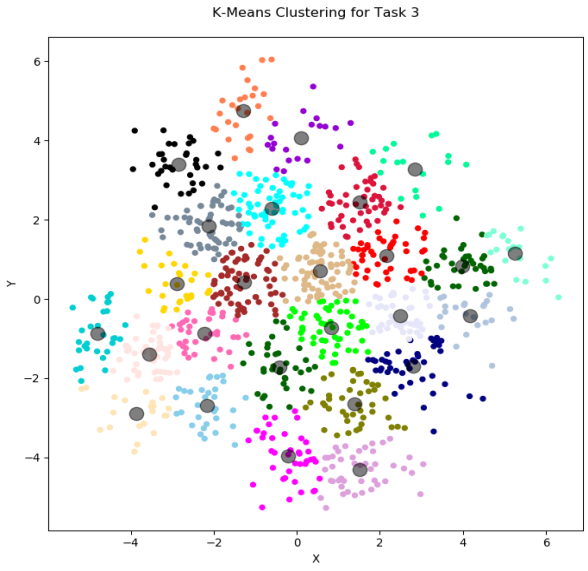
Project 2 Report

Plots were generated for each task using the **matplotlib** library. This was done to help visualize the resulting clusters of the K Means algorithm for each task. In order to properly plot each sample into a two-dimensional graph, the PCA dimensionality reduction algorithm was used to convert each feature vector into an equivalent x and y coordinate. These coordinates were then fed into the plot() function which mapped each class label to a unique color. The center of each cluster is visualized by a black circle.The resulting plots are below:







In order to obtain the accuracy of each task, I ran the K Means algorithm 5 times and computed the average accuracy. This was because I noticed the accuracy was slightly different every time I ran the algorithm and computed the confusion matrix. The accuracies are given below:

Task 1 – 94.2%

Task 2 – 74.75%

Task 3 – 48.56%