**QUESTION 4**

**Comparison with kmeans:** Now use the same dataset from (a), and using points with ids (210, 247, 265, 278, 288) as initial centers, identify the 5 clusters (use default settings for everything else). Compare the clusters you found here with the clusters you found in (a). Did you notice any difference? Compare in terms of cluster assignment and SSE.

Ans:

**Comparison Based on cluster sizes:**

Bisecting Kmeans: 78, 46, 45, 49, 82

Kmeans: 72, 54, 39, 10, 125

Bisecting Kmeans perform better by generating clusters with generally uniform sizes. As compared to Kmeans, which generates clusters of uneven sizes (e.g. 10 for C4 and 125 for C5)

**Comparison Based on SSE:**

Bisecting Kmeans: 54629.01

Kmeans: 59302.03

Bisecting Kmeans has slightly lower SSE as compared to Kmeans. It gives better results than Kmeans.