Assignment 3 Report
Students: skn003, tre081

How to run the program:
1. Run jupyter
2. Open "Assignment3_skn003_tre081.ipynb" in jupyter
3. Run the file

How did the clustering algorithms perform comparatively?
Overall the algorithms produced similar clusters with the main difference being that K-means had no clear overlap between the clusters, whereas Gaussian Mixture created some overlap between the yellow, bright green and dark red(?) clusters.

How many clusters did we get?
We got approximately 5 clusters with both algorithms when running based on both versions of the dataset.

Opinions about why the result turned out the way it did
We assume that the reason the results turned out the way they did is because K-means doesn't take variance into account, which Gaussian Mixture does. Because of this fact it would seem that the overlapping nodes result due to variance in what identifies the clusters and the value of the nodes.

2D scatter plot visualisations:
Engelsberger_Full.xlsx:

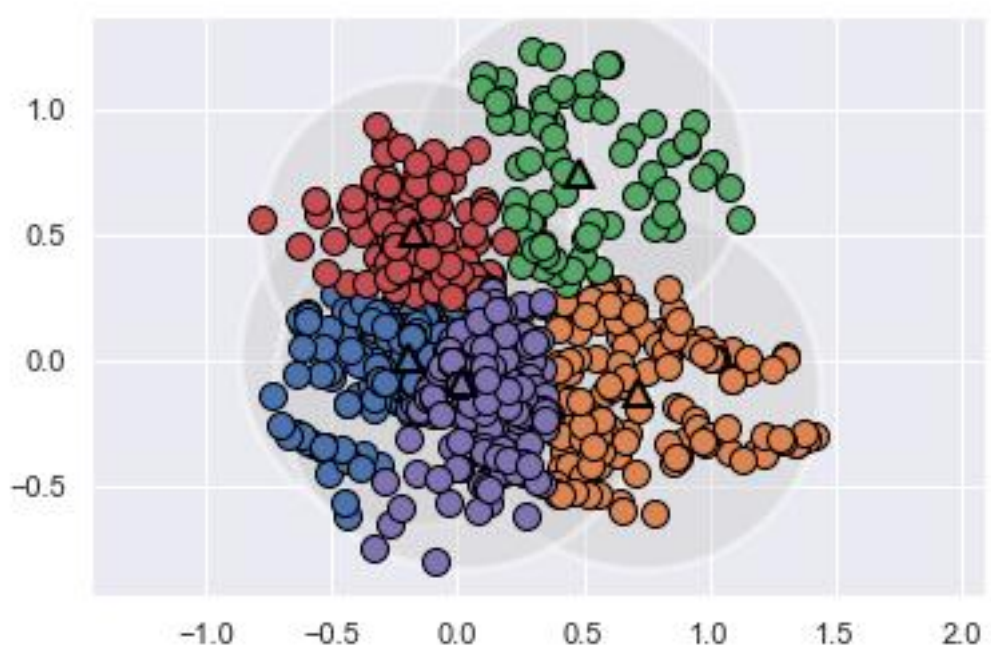

Figure 1: kmeans()

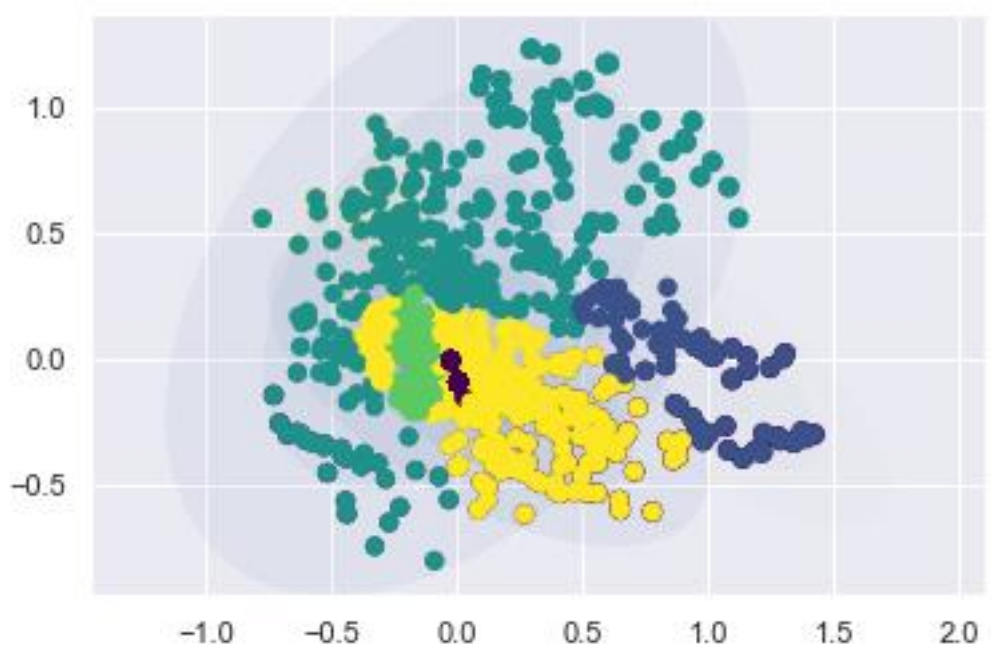

Figure 2: gaussian_mixture()

Engelsberger_Short.xlsx:

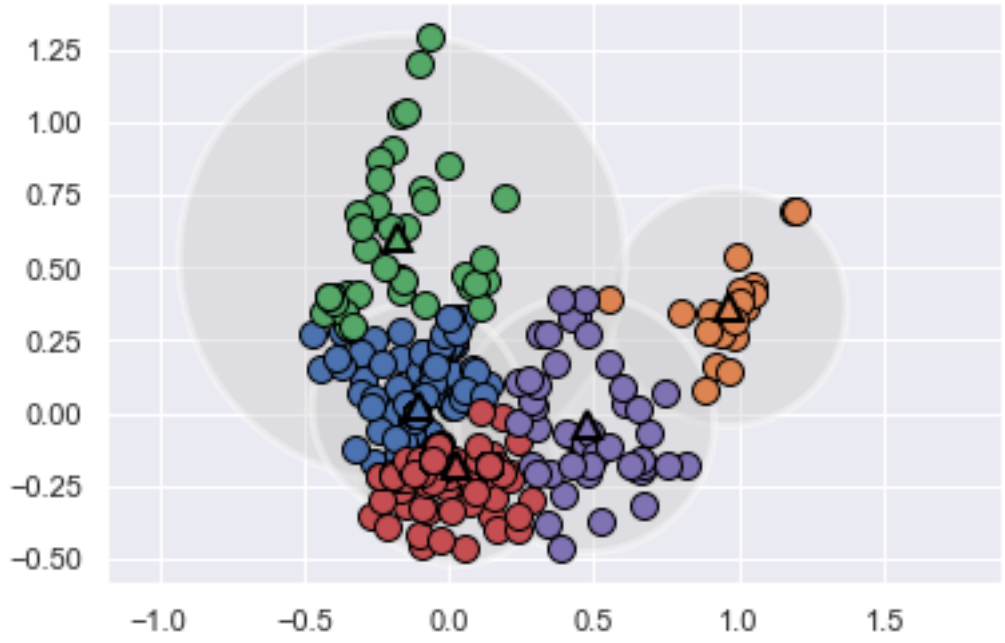

Figure 3; kmeans()

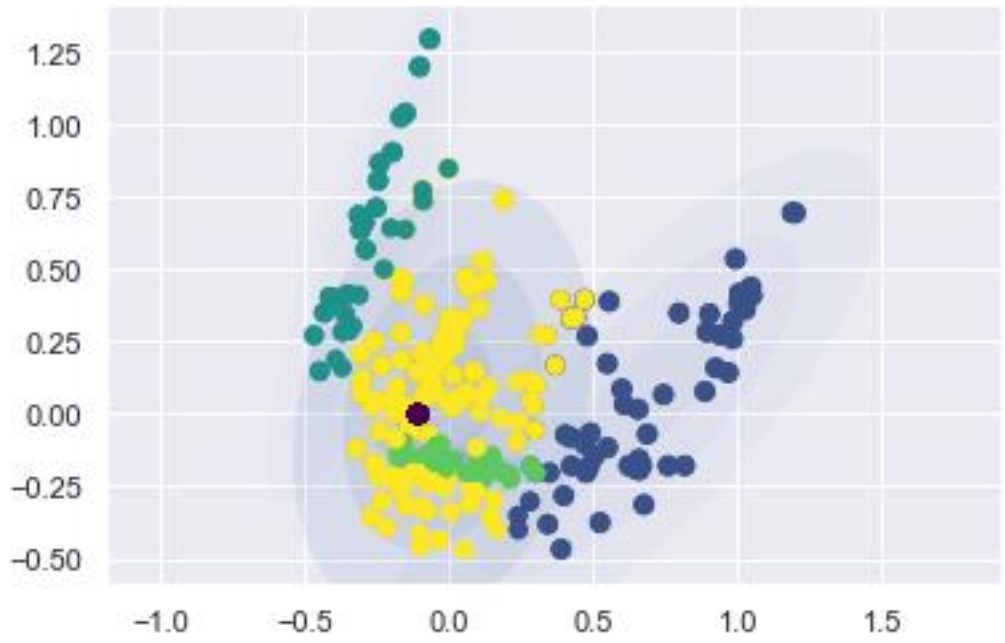

Figure 4: gaussian_mixture()