Notes:

These plots are with k=4.

Top row is python, where I used HierarchicalAgglomeration assignments to initialize GMM EM. There are 40 possibilities total (manhattan/l1, euclidean/l2, cosine)metric x (ward, complete, average, single)linkage x (full,tied,diag,spherical)covariance, and ward linkage is only available with euclidean metric. Cosine affinity, which measures angles between points and single/complete linkage are not available in mclust.

Bottom row is mclust (14 total clustering possibilities), the VVV model is where each cluster has its own unconstrained variance.

Remember that this is a single dataset, and I iterated through many clustering modes.
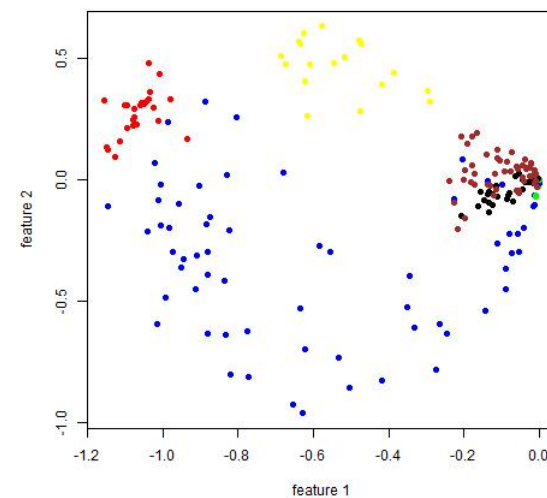
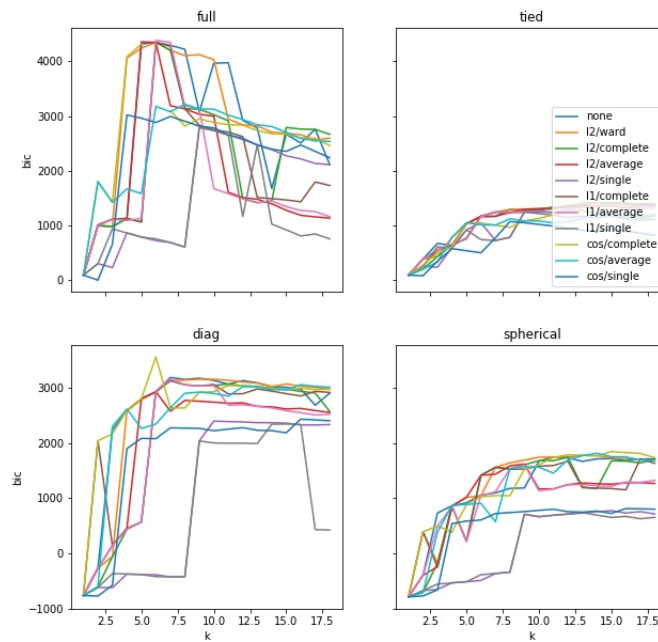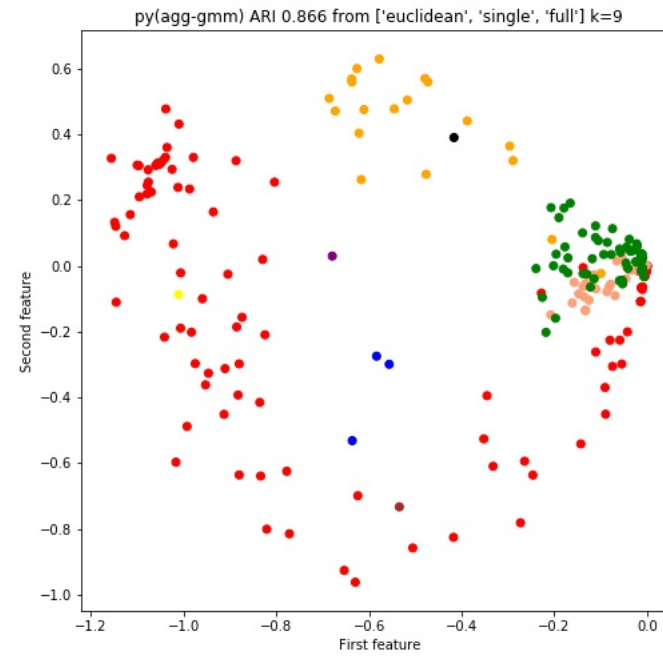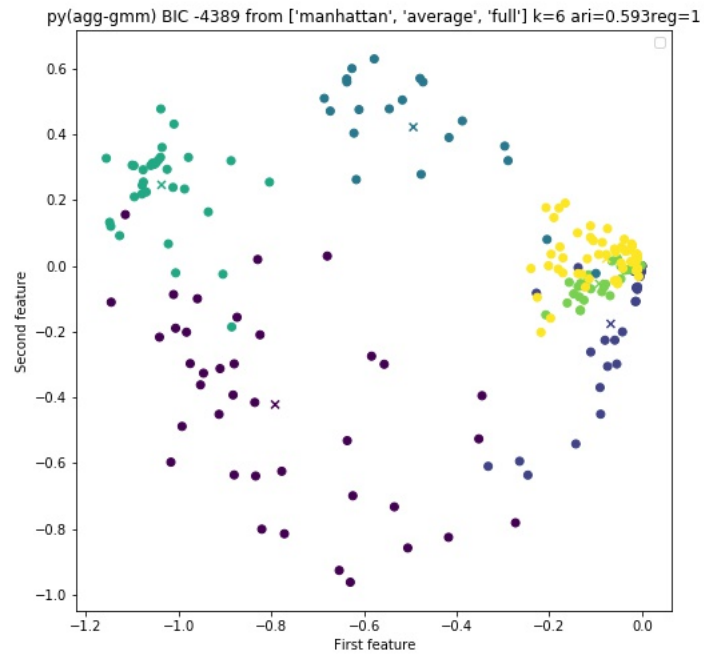py(agg-gmm) BIC -4389 from ['manhattan', 'average', 'full'] k=6 ari=0.593reg=1

True labels

py(agg-gmm) ARI 0.675 from ['cosine', 'single', 'diag'] k=6
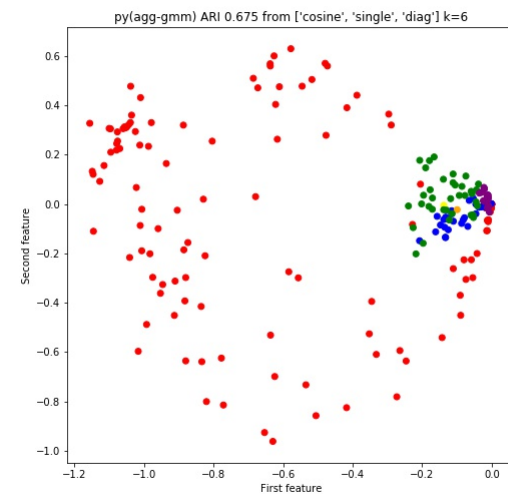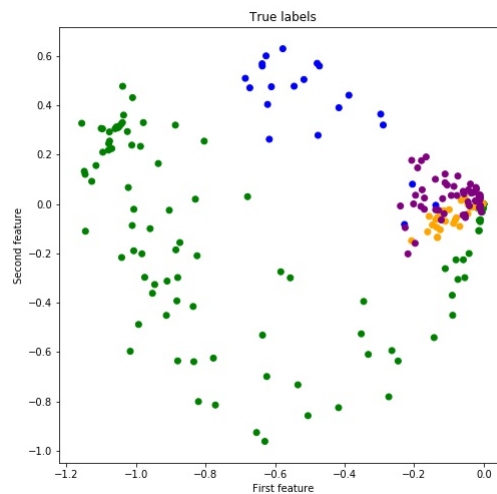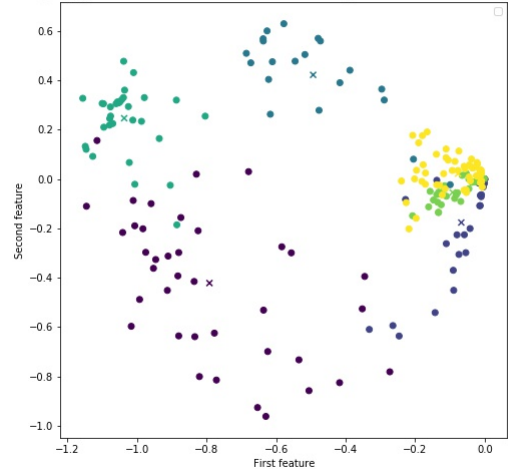
mclust Best BIC: 4362 from: VVV , k= 6 ari= 0.671

True labels
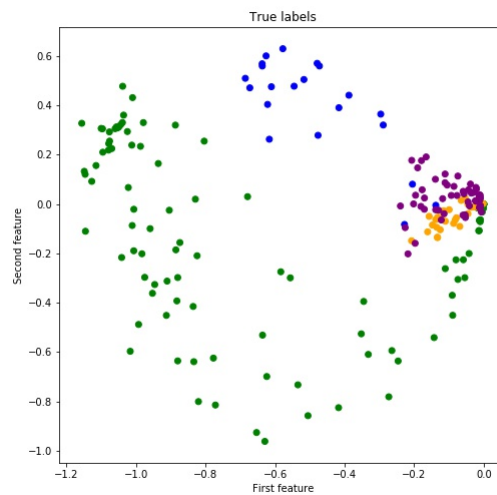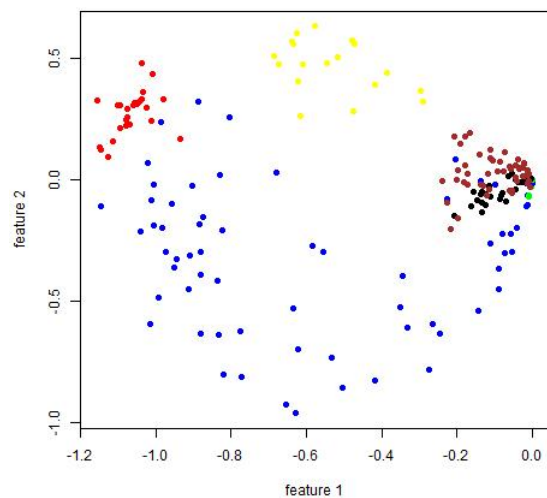
mclust Best ARI: 0.671 from: VVV , k= 6

k=6

(a) The subplots are divided by constraint used in GMM. The different
lines are the agglomeration methods for initialization.

mclust Best BIC: 4430 from: VVV , k= 7 ari= 0.574

mclust Best ARI: 0.671 from: VVV , k= 6