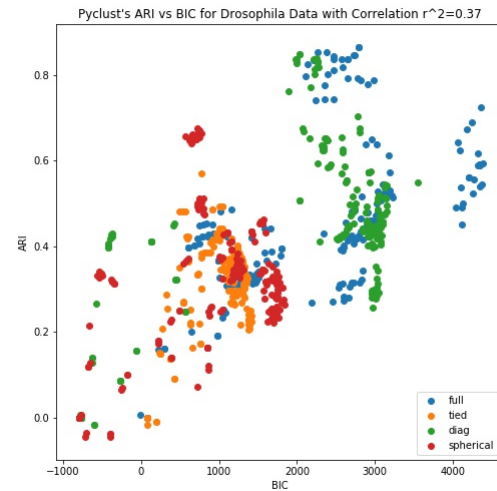
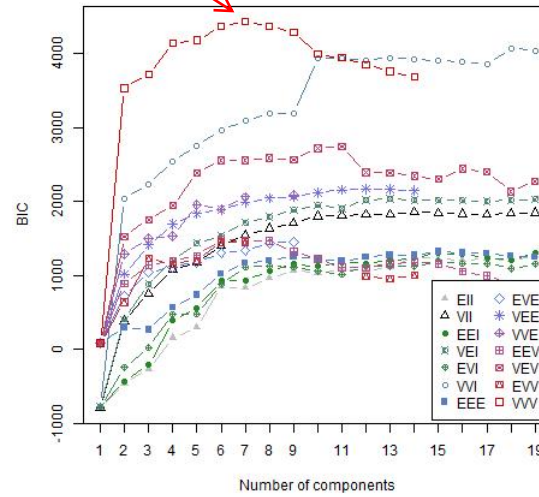
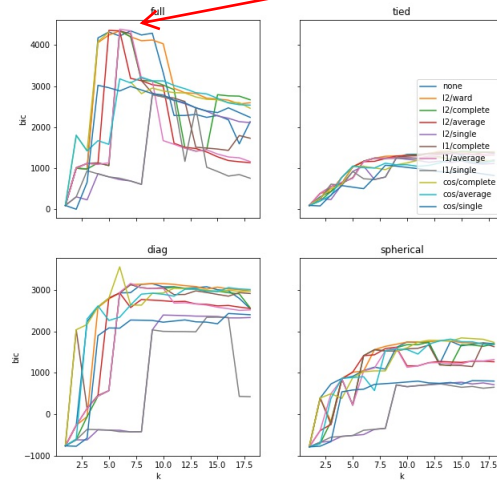
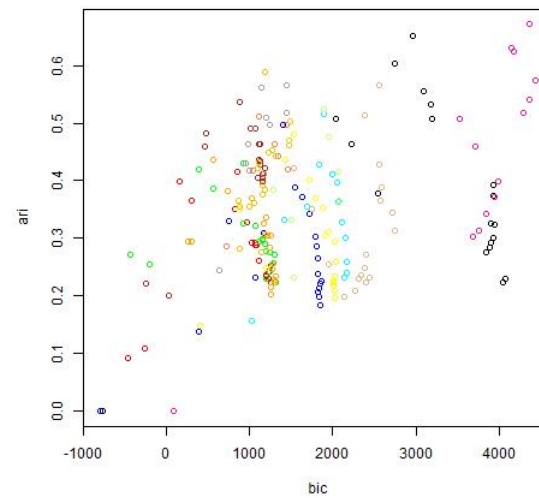


"best" clustering according to BIC



Mclust's ARI vs BIC on Drosophila Data with Correlation  $r^2=0.18$



Drosophila Data

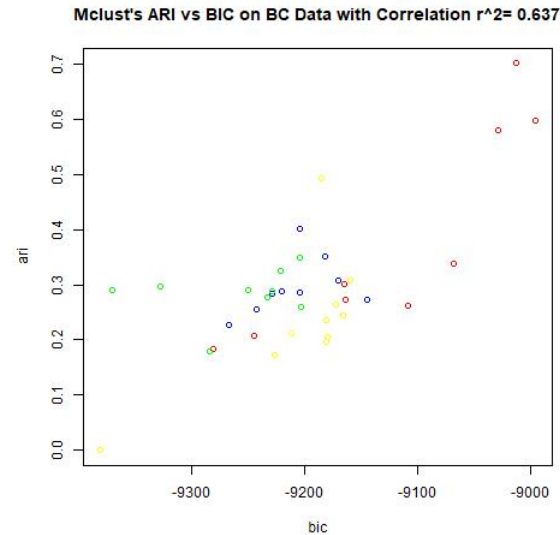
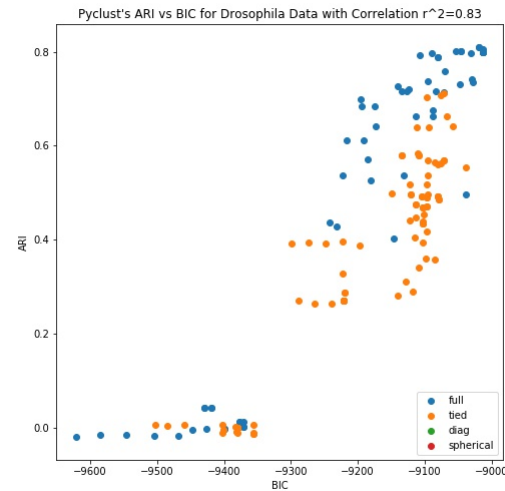
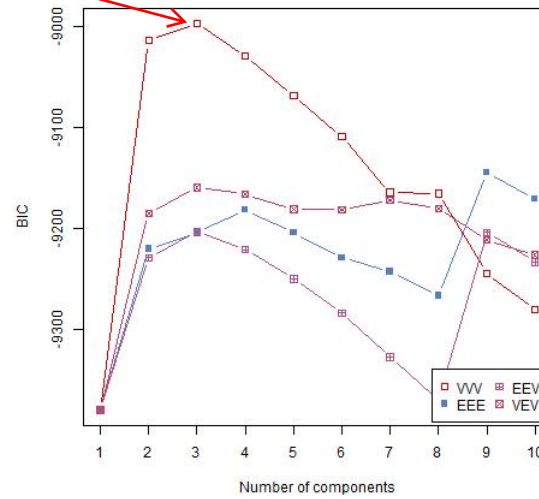
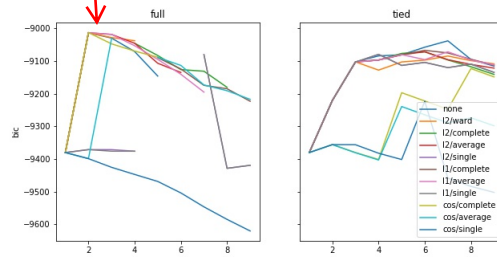
The left column is using Python's clustering methods - "pyclust,"

The right column is using mclust.

The top row shows the BIC results for all combinations of clustering techniques.

The bottom row shows the relationship between BIC and ARI for all the clustering techniques.

"best" clustering, according to BIC



Wisconsin Breast Diagnostics Dataset (<http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>)

As you can see, I limited the number of clustering combinations for both mclust and pyclust to replicate what was done in the original paper Fraley, Raftery Model-Based Clustering, Discriminant Analysis, and Density Estimation