**So Far**

Clustering

* Meanshift
  + Estimate the number of clusters in the data
  + Main input is bandwidth
    - Sensitivity done on the same trials to determine the effect on certain bands and see effect on silhouette score
* Kmeans
  + Impose the number of clusters onto the data
  + Only input is number of clusters
    - Sensitivity done on the same trials to see the effect on silhouette score
* MST (Not sure if this is useful)
* DBSCAN
  + Similar to meanshift as it estimates number of clusters
    - Main inputs are eps and minimum samples to set the number of objects surrounding a point to be considered a “cluster”
  + Useful as it determines outliers in the data
  + Compared using silhouette score
* Other scikit clustering methods to see if they can be applied to the data

Plotting

* Wrote cmd line program to make different plots of data
  + Wavelength and Colour histograms
  + Colour Colour plots
  + Wave colour plots
* Results plots
  + Score vs number of clusters
  + Score vs bandwidth
  + Score vs fractional size
* Worked Noise/Signal ratio into data trimming

**Next Steps**

Data

* Preprocessing
  + Except for removing -99 and filtering with noise/signal no other “processing” has been done
  + Does PCA do this?
  + Should the data be fit using a Gaussian?
* Size of sample
  + Varied number of points used and found different results of clustering
  + What is the benefit of using small size?
* Training Set
  + Could a small sample of data be used as training set if they are identified before and then clustered?

Clustering

* I’ve been able to have the algorithms make clusters, but what is the next step of the analysis?
* Should I be trying to figure out the best clustering method/band combination?
  + What makes it the “best”?
* How do I transfer the results of clustering to an ID catalogue?

DS9&NED surveys

* I think the next step is to try and use DS9 to identify objects and compare them to the clusters
* Crash course on Ds9?

Papers

* Thresholds for colour-colour plots
* PCA analysis and Gaussian smoothing