**Cluster Validation Metrics from clValid R function.**

1. Internal measures
	1. “…take only the dataset and the clustering partition as input and use intrinsic information in the data to assess the quality of the clustering.”
	2. Connectivity
		1. Range: [0,∞)
		2. Goal: minimize…near zero is good.
	3. Silhouette
		1. Range: [-1,1]
		2. Goal: maximize…near one is good.
	4. **Dunn\***
		1. Range: [0,∞)
		2. Goal: maximize…the larger, the better.
2. Stability measures
	1. “…a special version of internal measures. They evaluate the consistency of a clustering result by comparing it with the clusters obtained after each column <variable> is removed, one at a time.”
	2. Average proportion of non-overlap (APN)
		1. Range: [0,1]
		2. Goal: minimize…near zero is good.
	3. **Average distance (AD)\*\***
		1. Range: [0,∞)
		2. Goal: minimize…near zero is good.
	4. Average distance between means (ADM)
		1. Range: [0,∞)
		2. Goal: minimize…near zero is good.
	5. **Figure of merit (FOM)\*\***
		1. Range: [0,∞)
		2. Goal: minimize…near zero is good.

**\*** The Dunn metric is all over the place. I would not trust its optimal score.

**\*\*** AD and FOM have an issue in that they will inherently get smaller as individual entities are put in their own, unique clusters. If you look at the dendrogram of the milk cluster analysis, entities split in to their own clusters quickly. I recommend limiting use of those metrics to include only the largest changes between cluster numbers. The milk data show a large drop between 2 and 3 clusters, so I would stick with 3 clusters.

Lastly, looking at the milk data results, my final interpretation would be to keep three clusters based on how consistent the validation metrics are between the hierarchical and k-means tests at cluster = 3. Even with that, look at the hierarchical dendrogram vs. the cluster assignments from k-means where k = 3. The two methods don’t always place the entities in the same clusters!