

Report:

Dataset Name	K=2				K=True Value				K=12			
	MI	AMI	RI	ARI	MI	AMI	RI	ARI	MI	AMI	RI	ARI
Seeds (K=3)	0.55	0.42	0.73	0.46	0.71	0.70	0.87	0.71	0.57	0.39	0.78	0.43
Iris (K=3)	0.67	0.51	0.76	0.53	0.75	0.74	0.88	0.74	0.57	0.41	0.78	0.47
Segmentation (K=7)	0.14	0.08	0.53	0.07	0.51	0.48	0.85	0.39	0.58	0.49	0.87	0.42
Vertebral (K=3)	0.42	0.33	0.64	0.29	0.42	0.41	0.67	0.31	0.39	0.24	0.68	0.20

Observations:

We plotted the clusters and made the above table. The screenshots for the True values are given below. Comparisons of the true values are made as compared to values of K away from true value. MI and AMI are compared. RI and ARI are compared.

Clearly, the values of MI and AMI are similar for K=True Value. However, so is not the case for RI and ARI. The difference between (MI and AMI) and (RI and ARI) increases as we move away from the truth value of K. It is also observed that smaller the number of clusters for the ground state, higher the chances of accurate plotting. Segmentation plot given below is most different from ground plot as compared to other datasets. Higher the number of clusters, higher the probability of a certain point to be wrongly allocated and hence higher the chances of a less accurate plot.



