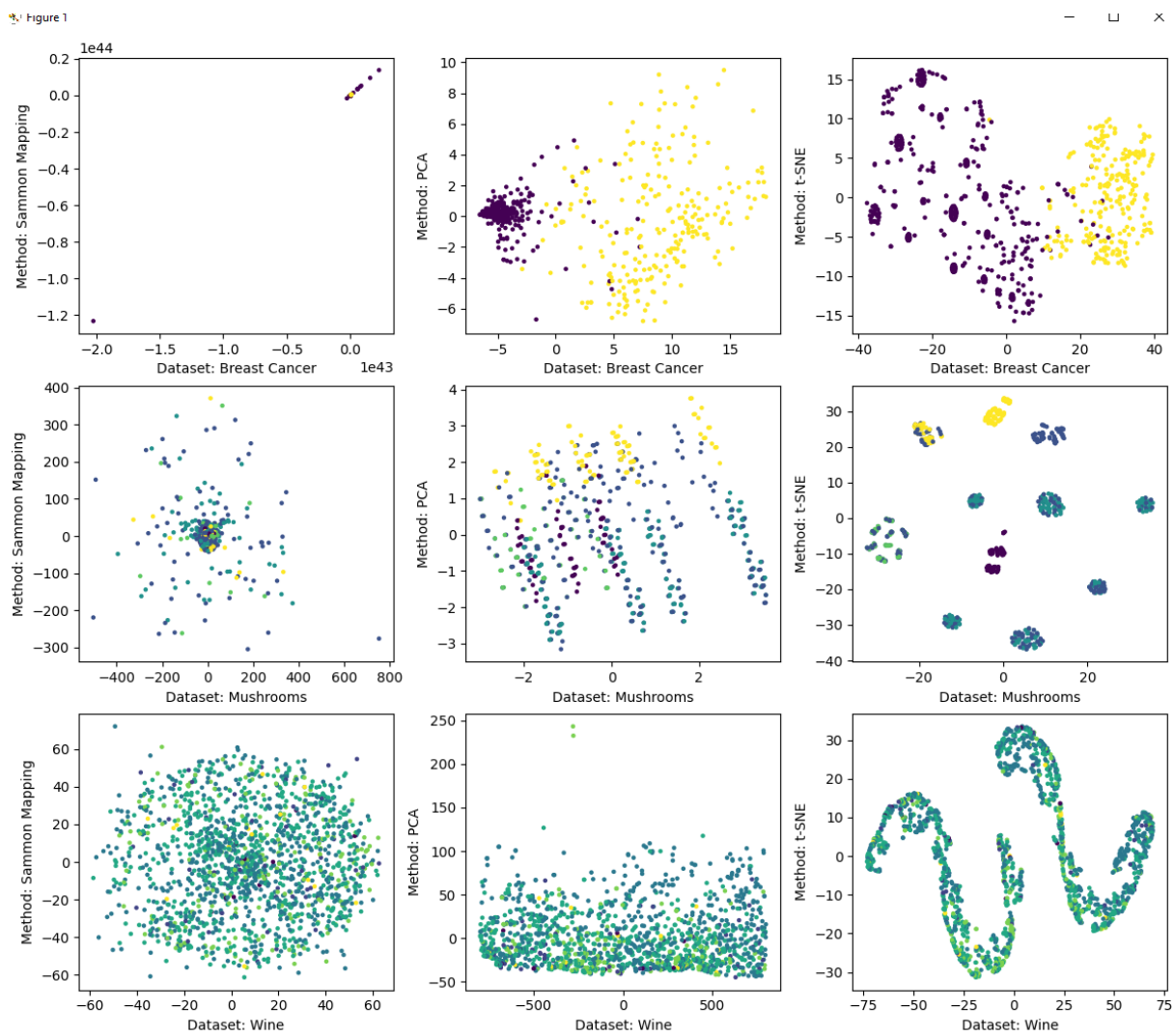


3.1. Comparison of DR Techniques:

Then answer this shortly (in a couple of paragraphs): In your opinion, which technique performed the best for each data set, regarding the separation of the classes? How are the classes in the data sets separated? Are some classes easier to separate than others?

Answer: Starting with the breast cancer dataset. We can clearly see that PCA and t-SNE performed best in separating the classes. The classes are separated in a way that they are mostly grouped together, much like clusters. For the mushroom dataset we can see that t-SNE performed best in separating the classes. The classes are grouped up like clusters and contains some errors. We can see in the plot that the yellow, purple and dark blue classes are best separated. For the wine dataset it is really messy. Here I would say sammon since it shows more of the labels spread out rather than clumped and getting lost behind each other. But I do not think sammon really is that good either.



3.2. Comparison of Clustering Techniques:

Then answer this shortly (in a couple of paragraphs): In your opinion, which clustering technique performed the best for each data set? How are the clusters in the data sets separated? Are some clusters easier to separate than others?

Answer: In my opinions the bkmeans and kmeans performed the best for all datasets. It is clearly visible the division between the clusters. The clusters are seperated in a way that they are clumped together and seperated by color. We have a yellow and a purple cluster that are neatly seperated.

