# Report

## Pros/Cons of Clustering Methods

K Means Clustering Cons:

* a negligent edge of each cluster, because the priorities are set on the center of the cluster, not on its borders;
* an inability to create a structure of a dataset with objects that can be classified to multiple clusters in equal measure;
* a need to guess the optimal *k*number, or a need to make preliminary calculations to specify this gauge.

Unlike the centroid-based models, the EM algorithm allows the points to classify for two or more clusters – it simply presents you the possibility of each event, using which you’re able to conduct further analysis. More to that, the borders of each cluster compose ellipsoids of different measures unlike k-means, where the cluster is visually represented as a circle. However, the algorithm simply would not work for datasets where objects do not follow the Gaussian distribution. That is the main disadvantage of the method: it is more applicable to theoretical problems rather than the actual measurements or observations.