

**Практическое задание 7**  
**"Кластеризация"**  
**(Пример вопросов для подготовки к устному опросу)**

1. Объясните идею разделительной кластеризации и работу алгоритма k Means.
2. Напишите формулу меры для выявления кластеров в k Means (Sum of Squared Errors)
3. Назовите преимущества и недостатки алгоритма k Means.
4. Объясните один из способов (на выбор) подбора начальных центроидов в алгоритме k Means.
5. Объясните работу алгоритма k Medoids.
6. Объясните идею плотностной кластеризации и работу алгоритма DBSCAN.
7. Дайте определения основных понятий, используемых в алгоритме DBSCAN: окрестность точки, корневая точка, непосредственная достижимость, достижимость.
8. Назовите преимущества и недостатки алгоритма DBSCAN.
9. Объясните, почему алгоритм DBSCAN является нечувствительным выбросам и шумам в исходных данных.
10. Объясните, каким можно подбирать параметры *MinPts* и *Pts* алгоритма DBSCAN.
11. Объясните идею иерархической кластеризации и работу алгоритма агломеративной кластеризации.
12. Объясните идею иерархической кластеризации и работу алгоритма дивизимной кластеризации.
13. Объясните способ построения дендрограмм.
14. Назовите преимущества и недостатки иерархической кластеризации.
15. Дайте определения следующих мер схожести, используемых в иерархической кластеризации: Single linkage, Complete linkage, Average linkage.