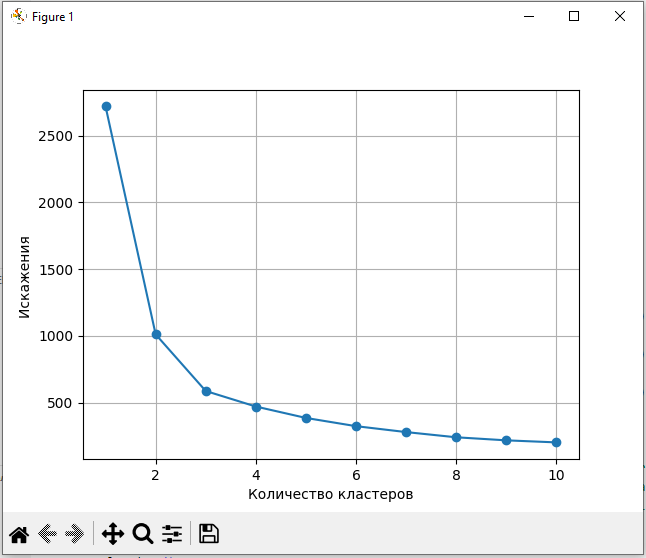
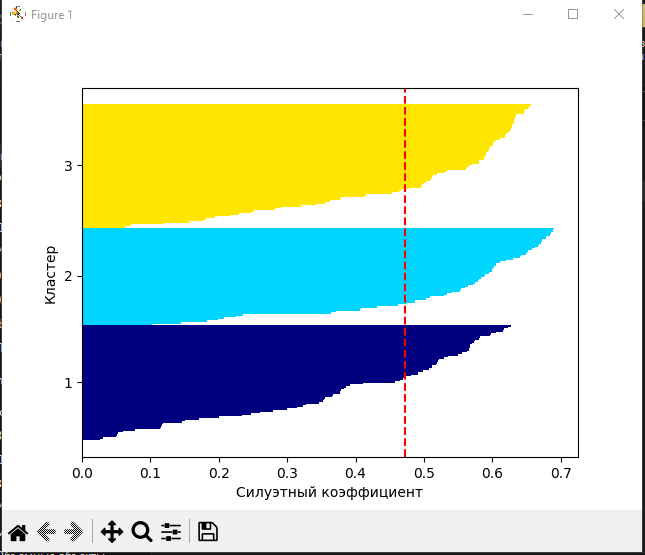


Исходя из полученного графика можно сделать вывод, что алгоритм достаточно хорошо показал себя на данном наборе данных. Особенно хорошо были разделен кластер 2



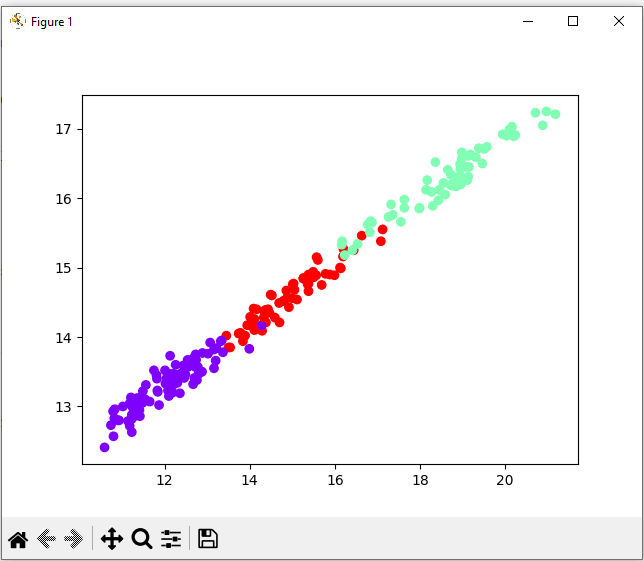
Этот график показывает, оптимальное количество кластеров. Отмечено на рисунке.



Визуальный анализ силуэтного графика позволяет быстро рассмотреть размеры разных кластеров и идентифицировать кластеры, которые содержат выбросы.

Как видно на приведенном выше силуэтном графике, наши силуэтные коэффициенты даже близко не находятся рядом с 0, что может служить индикатором хорошего объединения в кластеры.

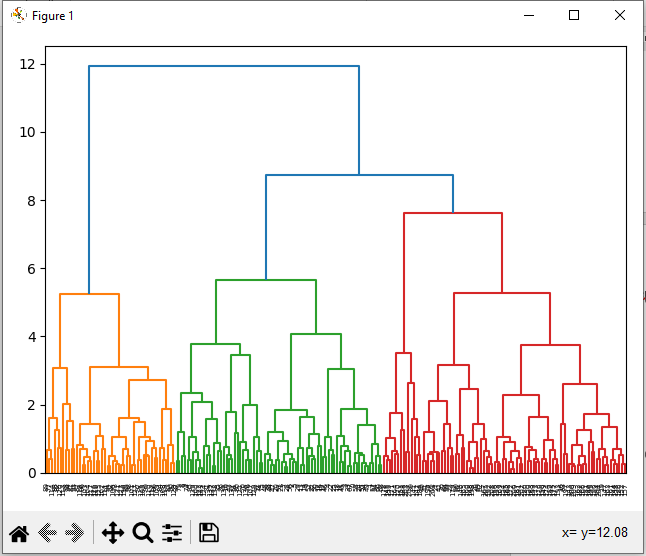
AgglomerativeClustering



Агломеративная кластеризация – это один из алгоритмов кластеризации, в котором процесс группировки похожих экземпляров начинается с создания нескольких групп, где каждая группа содержит один объект на начальном этапе, затем он находит две наиболее похожие группы, объединяет их, повторяет процесс до тех пор, пока не получит единую группу наиболее похожих экземпляров. Например, представьте пузырьки, плавающие на воде и прикрепляющиеся, и в конце процесса вы увидите большую группу пузырьков.

На рисунке так же видно, что алгоритм разделил данные на 3 кластера.

Дендограмма



Дендрограммы используются для разделения на множество различных кластеров. На изображении выше показано, что оптимальное количество кластеров должно быть 3 для данных, которые я использовал.