

Практическая работа №7

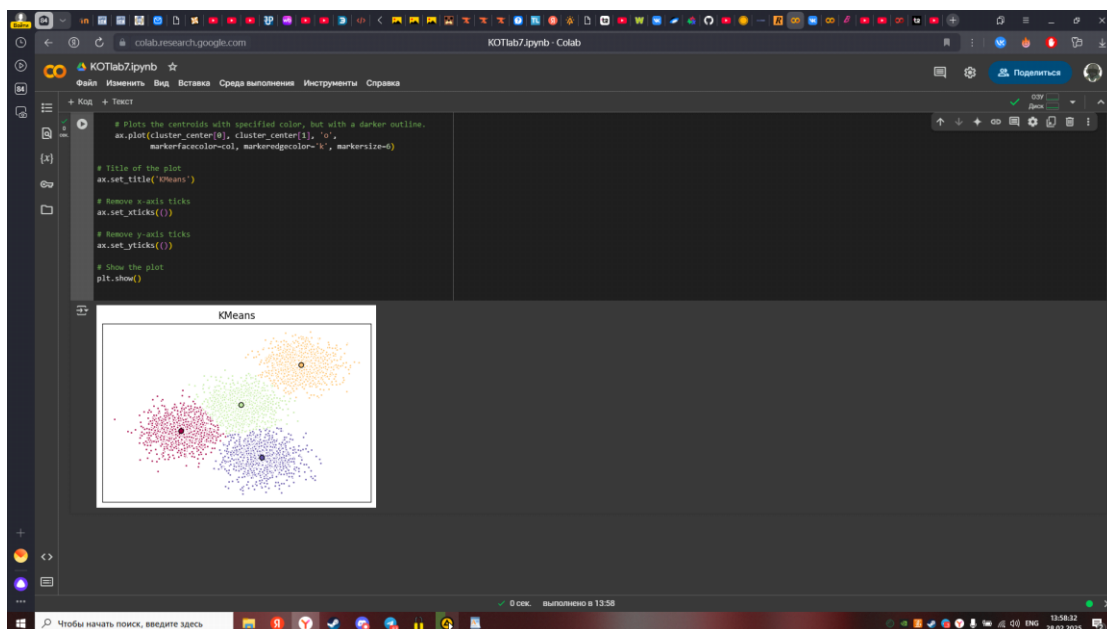
Машинное обучение. K-Means Clustering

Цель занятия: приобрести практические навыки в создании кластеров на методе К-средних на двух примерах.

Ход работы

В данной практической работе необходимо познакомиться с моделью К-средних. Такая кластеризация - самая простая, но при этом используется почти что повсеместно по обработке данных, потому научимся с ним работать. О том, как он работает, узнаем из методических указаний - на каждой итерации перевычисляется центр масс для каждого кластера, полученного на предыдущем шаге, затем векторы разбиваются на кластеры вновь в соответствии с тем, какой из новых центров оказался ближе по выбранной метрике.

Выполняя код из методических указаний, у нас выстраиваются следующие модели.



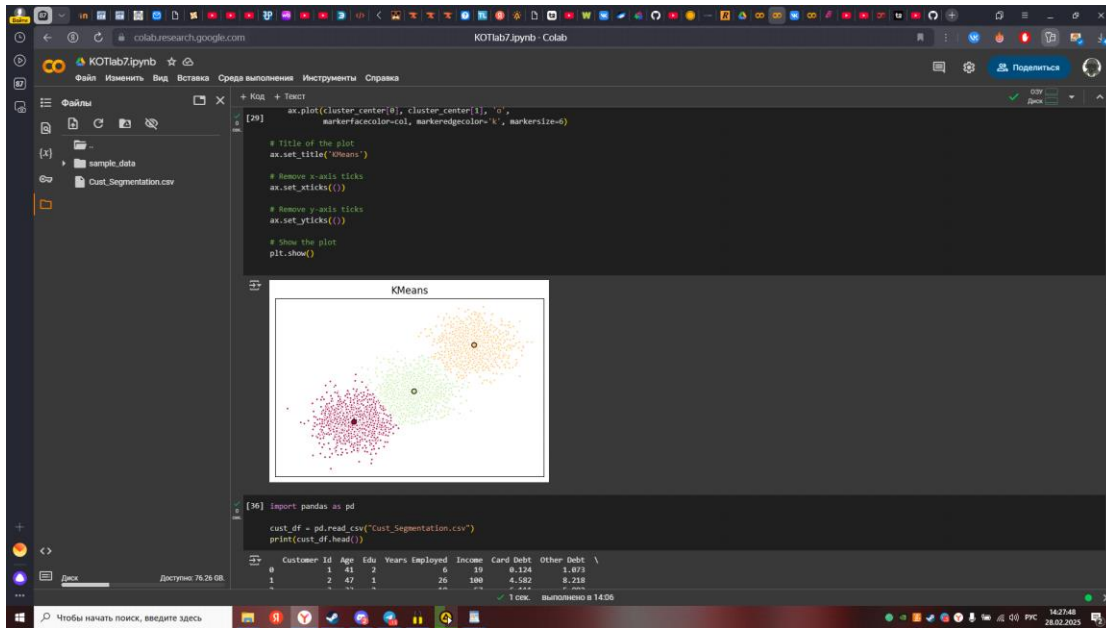


Рисунок 1, 2 – полученные кластеры, на второй сделали 3 кластера

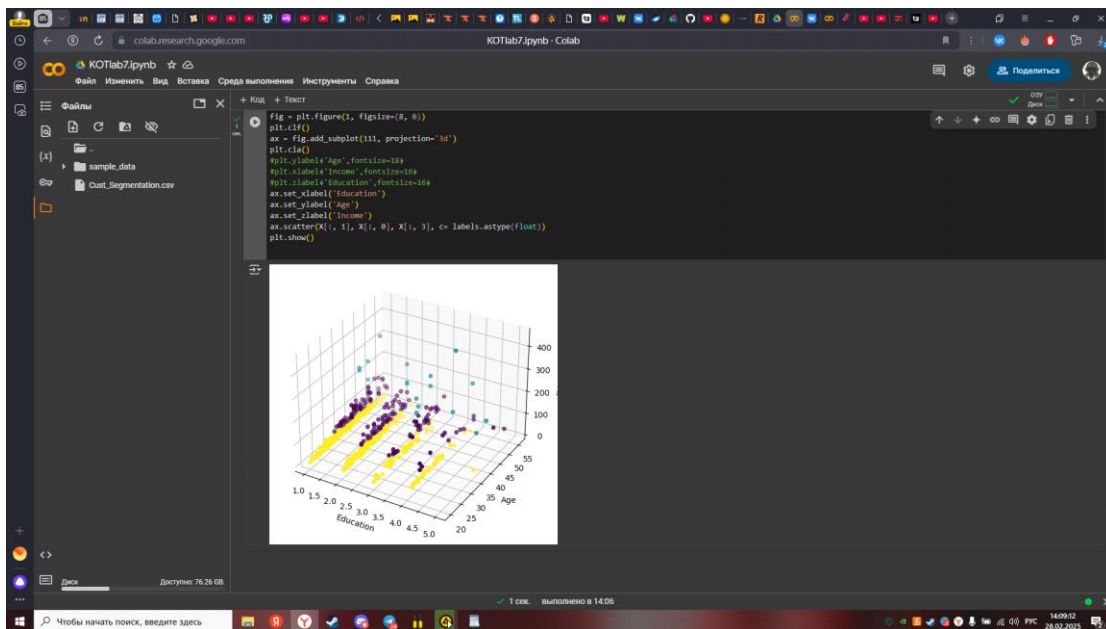


Рисунок 3 - граф распределения доходов по возрасту и образованию

Вывод: в результате практической работы мы узнали как работает метод кластеризации К-средних, а также научились строить графическое их отображение.