

Лабораторная работа №6

Кластеризация

Рассмотрим датасет с данными о клиентах супермаркета, включая их пол, возраст, доход и оценку трат ("**Customers.csv**"). Загрузить данные (read_csv())

- 1) Вывести первые и последние 10 строк таблицы (head(), tail()).
- 2) Вывести статистику по значениям признаков.
- 3) Вывести более подробное описание значений признаков (количество непустых значений, типов столбцов и объема занимаемой памяти).
- 4) Построить столбиковую диаграмму и круговую диаграмму, чтобы показать гендерное распределение по набору данных.
- 5) Построить гистограмму, для определения распределения возрастов клиентов. Также построить блочную диаграмму (ящик с усами).
- 6) Построить *гистограмму для анализа годового дохода клиентов, затем рассмотреть эти данные с помощью графика плотности.*
- 7) Проанализировать расходы клиентов, используя различные графики.
- 8) Используя *алгоритм K-средних*, изменяя различные его параметры, определить оптимальное количество кластеров
- 9) Осуществить визуализацию результатов кластеризации