Московский Государственный Технический Университет

им. Н.Э. Баумана

Рубежный контроль №2

по курсу

Технологии Машинного Обучения

# Выполнила:

## Костян Алина

## ИУ5-53

\_\_\_\_\_\_\_\_\_\_\_\_\_

# Проверил:

## Гапанюк Ю.Е.

\_\_\_\_\_\_\_\_\_\_\_\_\_

Москва, 2019

# Задание

Данный вариант выполняется на основе материалов лекции.

Необходимо решить задачу кластеризации на основе любого выбранного Вами датасета.

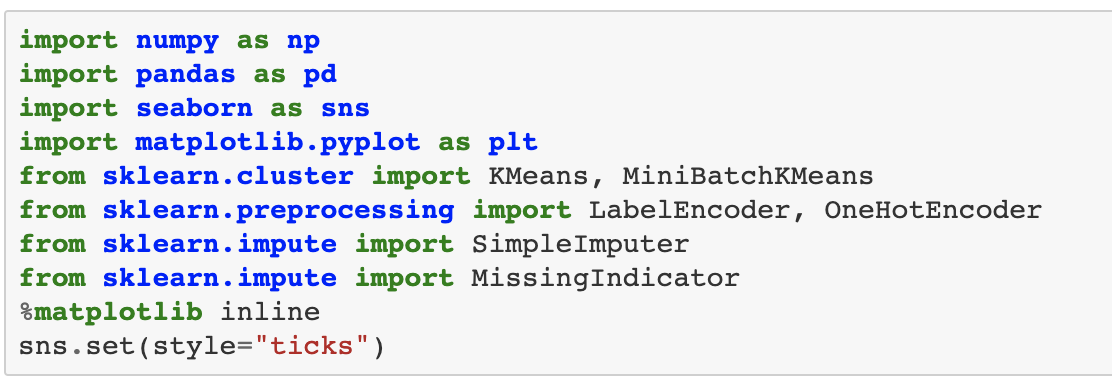
Кластеризуйте данные с помощью трех различных алгоритмов кластеризации. Алгоритмы выбираются произвольным образом, рекомендуется использовать алгоритмы из лекции.

Сравните качество кластеризации для трех алгоритмов с помощью следующих метрик качества кластеризации:

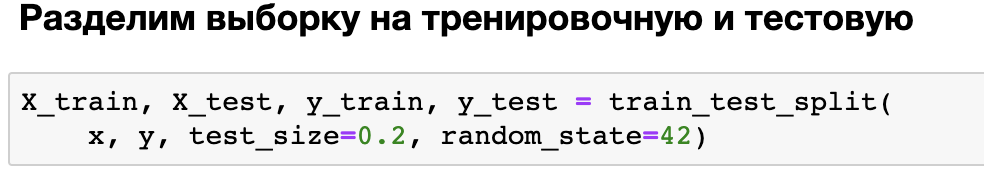
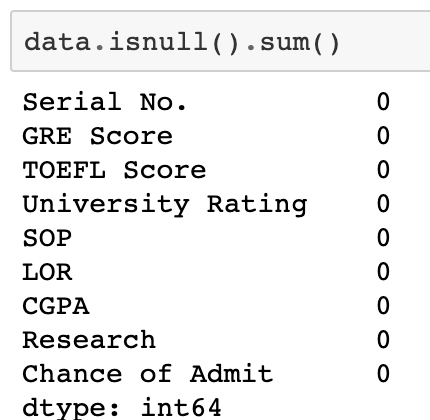
* Adjusted Rand index
* Adjusted Mutual Information
* Homogeneity, completeness, V-measure
* Коэффициент силуэта   
  Сделате выводы о том, какой алгоритм осуществляет более качественную кластеризацию на Вашем наборе данных

# Код и результаты выполнения

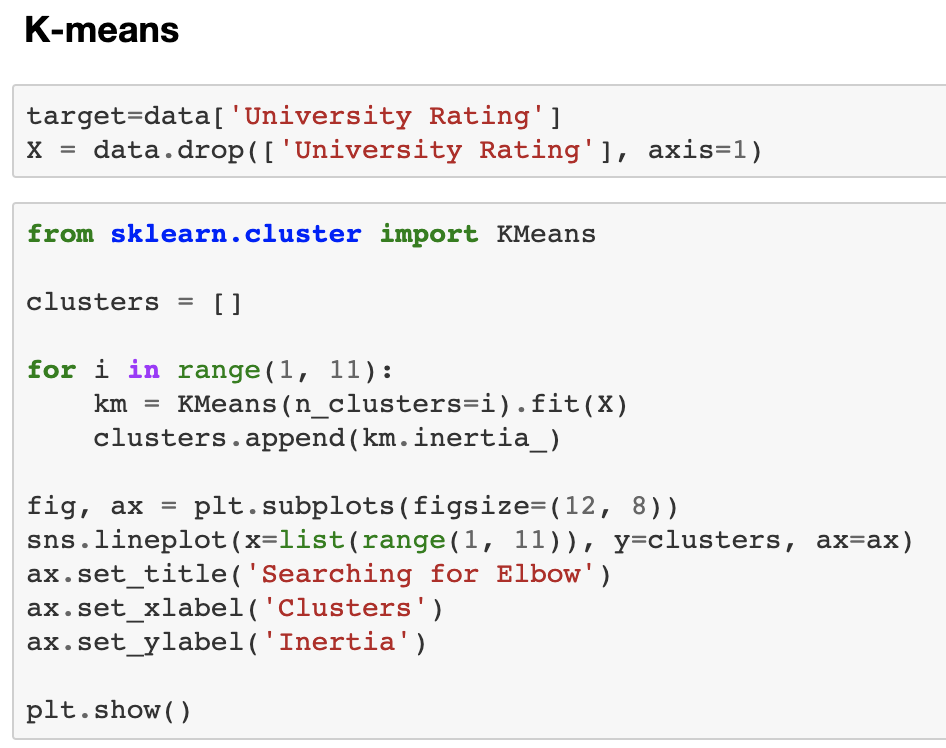
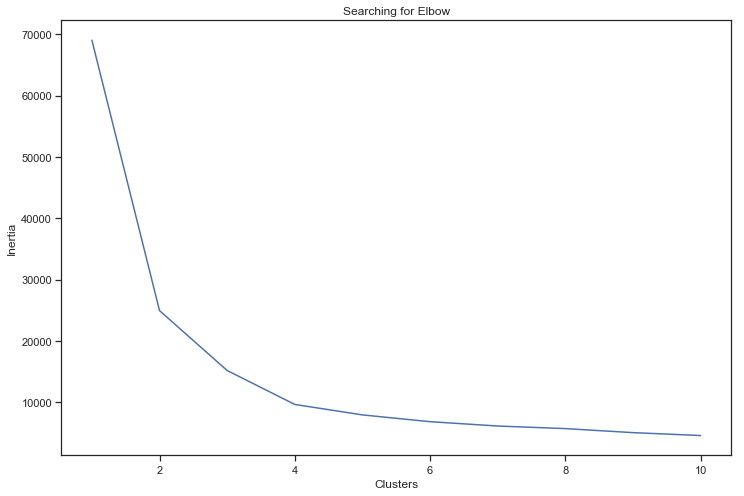
1. Подключим библиотеки:

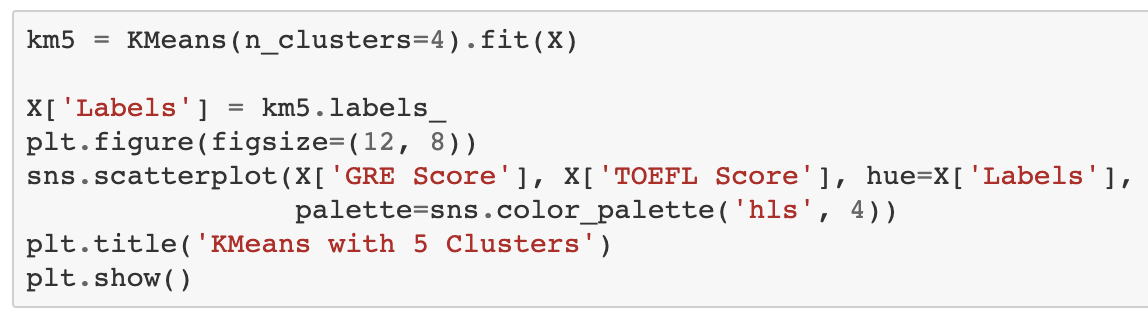
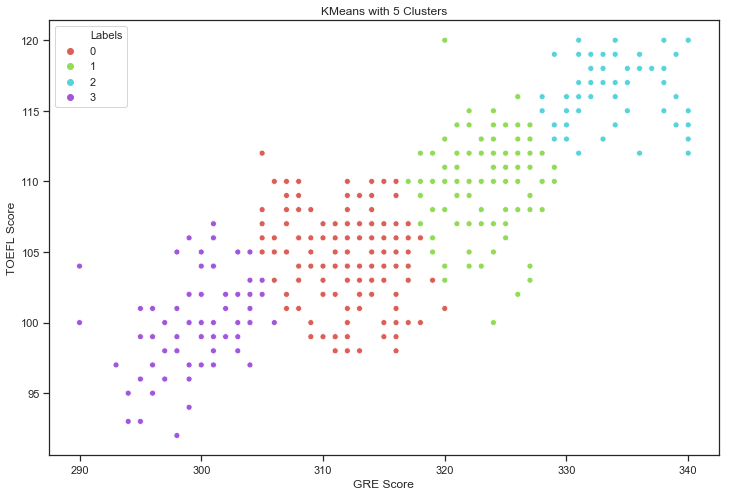


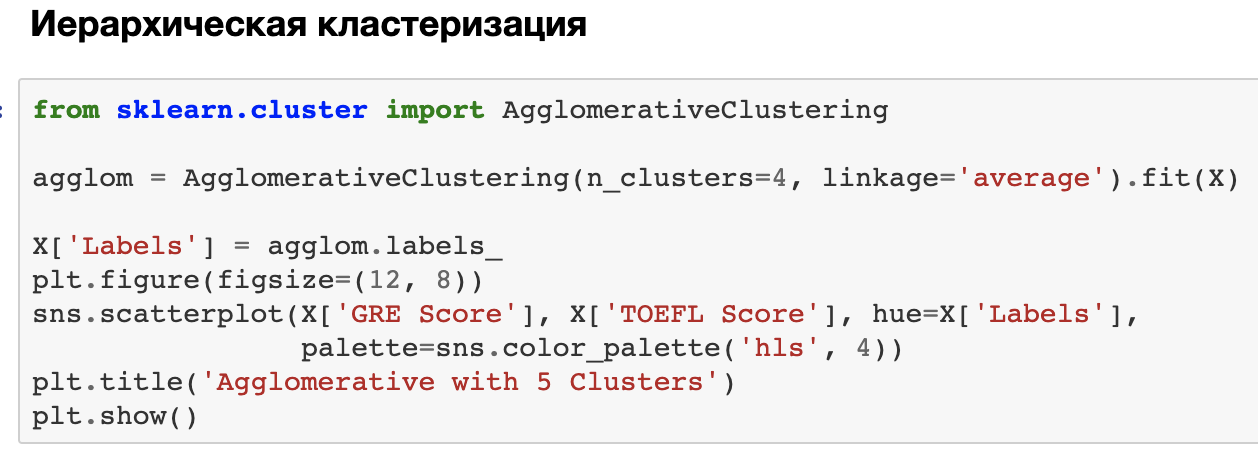
1. Подготовим данные

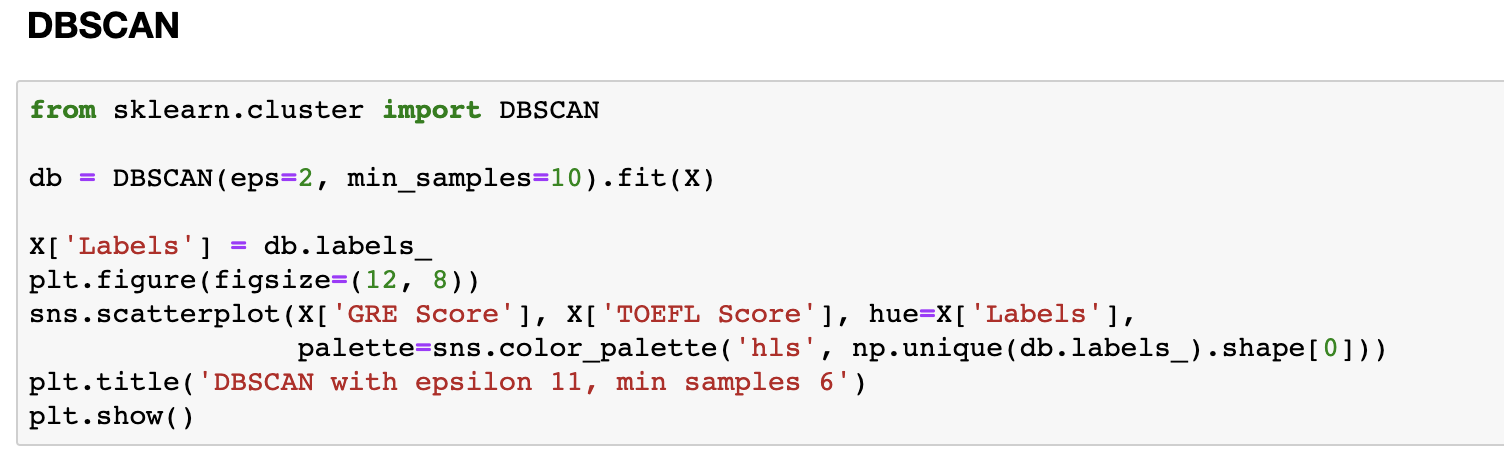


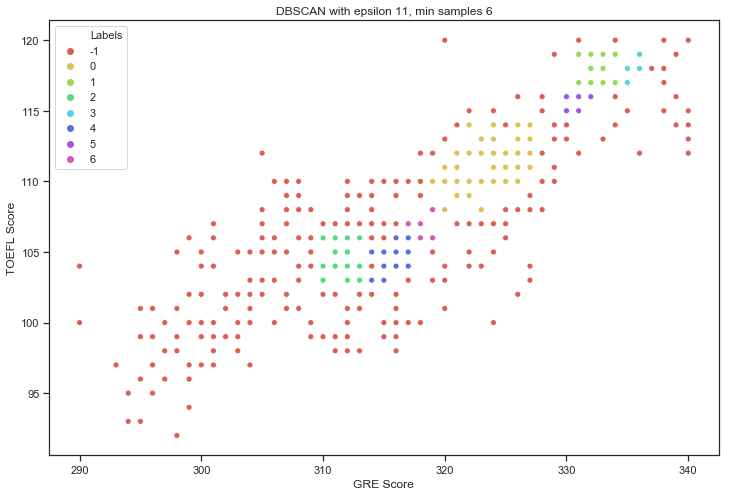
1. Кластеризация





1. Метрики качества



