**Лабораторная работа 10. Кластерный анализ**

Методические рекомендации:

1. Рекомендуемый язык программирования – Python.
2. Рекомендуемые среды разработки:
   1. Пакет Anaconda (Jupiter Notebook) <https://www.anaconda.com/products/individual>
   2. <https://colab.research.google.com/notebooks/welcome.ipynb?hl=ru>

К лабораторной работе приложено 6 вариантов данных с описанием. Вам необходимо выбрать свой вариант.

Задание 1:

Вам предлагаются для анализа наборы данных в различных форматах таблиц.

1. Изучить описание данных, загрузить и провести предобработку данных. Аргументированно объяснить ваши действия.
2. Провести иерархический кластерный анализ, определить оптимальное число кластеров. Развернуто аргументировать в тексте ноутбука.
3. Добавить в таблицу столбец с номером кластера и сохранить полученную таблицу в файл.

Задание 2.

1. Повторить вышеописанные действия для кластерного анализа с использованием метода k-means.

Все ваши действия по исследованию таблицы, проведению преобразований и предобработке данных, кластерному анализу должны быть письменно описаны и аргументированы в тексте ноутбука в виде комментариев Markdown.

Важно: «голые» коды без развернутых текстовых объяснений не принимаются.

Полезные ссылки:

<https://habr.com/ru/post/585034/>

<https://pythobyte.com/hierarchical-clustering-with-python-and-scikit-learn-768953fe/>