**Задание №3**

**Определить наилучшую кластеризацию данных по индексу с помощью распределенных вычислений с использованием MPI**

Обучающая выборка (DS\_2019\_public.xlsx) загружается одним из процессов. Значения атрибутов привести к диапазону от 0 до 1. Далее этот процесс отправляет загруженные данные всем остальным процессам. Все процессы начинают запускают несколько раз алгоритм кластеризации данных алгоритмом K-Means (https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html). Каждый раз при этом вычисляется значение индекса качества. Среди решений выбирается лучшее (это решение имеет либо самый большой, либо самый маленький индекс – смысл выводится из уравнения индекса). Далее все лучшие индексы процессов сравниваются и процесс с лучшим результатом выводит свой результат.

Help: https://nyu-cds.github.io/python-mpi/

Задача по полученному варианту (выбор варианта осуществляется по последнему номеру в зачетной книжке).

Вариант 1 (последняя цифра в зачетке 0) SD индекс, в задании использовать функции MPI\_Isend, MPI\_Reduce

Вариант 2 (последняя цифра в зачетке 1) VNND индекс, в задании использовать функции MPI\_Bcast, MPI\_Gather

Вариант 3 (последняя цифра в зачетке 2 или 9) RS индекс в задании использовать функции MPI\_Isend, MPI\_Gather

Вариант 4 (последняя цифра в зачетке 3) RMSSTD-индекс в задании использовать функции MPI\_Bcast, MPI\_Reduce

Вариант 5 (последняя цифра в зачетке 4 или 8) Maulik-Bandoypadhyay – индекс в задании использовать функцию MPI\_Isend

Вариант 6 (последняя цифра в зачетке 5) Индекс Данна, в задании использовать функцию MPI\_Gather.

Вариант 7 (последняя цифра в зачетке 6) Индекс оценки Силуэта, в задании использовать MPI\_Send, MPI\_Gather

Вариант 8. (последняя цифра в зачетке 7) Индекс Дависа-Болдуина, в задании использовать MPI\_Bcast, MPI\_Reduce