

Контрольное мероприятие текущего контроля.

Обработка и анализ больших данных.

Необходимо выбрать большие данные с ресурса <https://www.kaggle.com>. Сформировать гипотезу добычи данных. Выбранные данные и гипотезу необходимо согласовать с преподавателем.

Написать программу на языке Python первичной обработки данных:

- 1) провести первичный анализ данных рядов;
- 2) провести очистку рядов;
- 3) исследовать ряды данных;
- 4) провести корреляционный анализ данных;
- 5) провести анализ автокорреляции данных;
- 6) провести кластеризацию данных;

В результате анализа данных необходимо написать выводы в подтверждение или опровержения выдвинутой гипотезы и заключения по каждому пункту.

Отчет должен содержать в себе: титульный лист, постановка задачи, код программы с комментариями, текст с пояснениями о результатах первичной обработки с графиками, выводы.

Работа может выполняться как индивидуально, так и в бригаде (не более 2 человек).