

## Домашняя работа (максимум 30 баллов)

### 1. Методика оценки (максимум 30 баллов)

- Представлен код программы (20)
- Наличие развернутых комментариев (5)
- Предоставленная программа работает (5)

### 2. Предположительное время выполнения 2 часа

### 3. Задание

Проведите небольшое исследование «грязных» данных с целью обнаружения разного рода ошибок, а также устраните выявленные ошибки. В учебных и реальных проектах всегда есть соблазн сразу строить модели или проверять гипотезы, рекомендуется не поддаваться этому искушению и сделать предварительную очистку данных.

3.1 Выполните загрузку dataset, который Вы предполагаете использовать в проекте, используя библиотеки Python, C++ или в kaggle. Альтернативно можно загрузить default dataset, где вычислен индикатор momentum (прилагается к заданию).

3.2 Вычислите 3 статистики, которые помогут выявить ошибки: аномальные значения, дублирование и т.д. Прокомментируйте полученный результат.

3.3. На основании 3.2 устраните эти ошибки в данных. Прокомментируйте выбранный способ устранения ошибок.

3.4. Вышлите код или ссылку на для оценки задания.

Желаем успехов.