

## Проверочная работа

1. Загрузите датасет `weight-height.csv`. Проанализируйте все столбцы записей представленных в этом датасете. Определите количество уникальных значений.
2. Переведите дюймы и фунты в более привычную для нас системы измерений.
3. Разделите датасет на 2 датафрейма по гендерному признаку. Напишите функции для определения количества записей в каждой коллекции, математического ожидания, стандартного отклонения, `min` и `max` значения по каждому из столбцов, разбиения на квантили.
4. Визуализируйте на гистограммах распределение роста, веса у обоих полов.
5. Воспользуйтесь методом, определенным в библиотеке `pandas` и сравните этот результат сданными, полученными в задании 3.
6. Кластеризируйте данные датасета `weight-height.csv` методом `k-средних` и иерархической кластеризации. Полученные кластеры отобразите на графике.
7. В функцию предсказания передайте свои данные о росте и весе и определите – какого вы пола.
8. С помощью множественной линейной регрессии найдите функцию определения пола.
9. Определите коэффициенты детерминации модели, полученной в задании 7.
10. Передайте в функцию предсказания вашей модели ваши данные о росте и весе и определите какого вы пола.
11. Сделайте вывод о точности данной функции относительно использованных вами методов кластеризации.
12. Определите в какие диапазоны попадают значения веса и роста у 68%, у 95%, и у 99,7% мужчин и женщин.

Примечание: категориальный признак привести в количественный, с помощью метода `LabelEncoder`, определенного в классе `preprocessing` библиотеки `sklearn`.