

## Подготовка данных

`pd.read_csv()` — происходит чтение набора данных из файла CSV. Выбираются признаков и целевой переменной, удаляются колонки целевого показателя из набора данных признаков.

Исходные данные распределяются на тренировочный набор, предназначенный для обучения модели, и тестовый набор, используемый для оценки качества построенной модели.

## Создание и обучение модели:

Инициализируется линейная регрессия. Модель обучается на подготовленных заранее данных функцией `fit()`.

## Тестирование модели:

Значения прогнозируются на тестовом наборе. Затем качества модели оценивается с использованием метрик среднеквадратичной ошибки и коэффициента детерминации  $R^2$ , `r2_score`.

## Развёртывание модели:

Данный этап здесь отсутствует, потому что программа служит лишь демонстрацией. Обычно развёртывание включает интеграцию модели в рабочую среду приложения или сервера, настройку API-интерфейсов для приёма запросов и выдачу результатов прогнозирования.

Среднеквадратичная ошибка MSE и коэффициент детерминации  $R^2$  помогают оценить качество построенной модели:

**Среднеквадратичная ошибка** показывает среднюю величину отклонения между фактическими значениями и предсказанными моделью. Чем ниже эта величина, тем лучше работает модель. Если MSE близко к нулю, значит модель даёт практически точные прогнозы.

**Коэффициент детерминации  $R^2$**  оценивает, насколько хорошо модель объясняет вариацию данных. Чем ближе этот показатель к 1, тем лучше модель подходит данным. Если  $R^2$  равен примерно 1, значит почти вся вариация данных объясняется моделью. Если же  $R^2$  близок к нулю, это означает, что модель плохо описывает зависимость.