



## Регрессоры

Самостоятельная работа.

Провести анализ данных из файлов ниже и решить задачу регрессии, выбрав в качестве целевой переменной параметр `p3`. Функция, описывающая данные из файлов ниже нелинейная (степенная, логарифмическая и т.д.), поэтому следует заранее предугадать аналитическую формулу искомой функции.

1. При помощи метода `describe` библиотеки `pandas` провести предварительный анализ данных, установить минимальные и максимальные значения для последующей визуализации.
2. Для одномерного случая при помощи методов `figure` и `plot` библиотеки `matplotlib` визуализировать данные. Использовать ограничения по осям `xlim` и `ylim`, включить сетку методом `grid`. Для многомерных случаев визуализировать проекции данных по всем осям.
3. Создать не менее 5 различных моделей регрессии, причем:
  - (a) две из моделей должны быть построены на нормированные данные (используйте библиотеку `sklearn`, сначала с помощью `StandardScaler()`, затем `MinMaxScaler()`);
  - (b) для каждой модели вывести полную формулу полученной функции в том формате, в котором у вас она задана, например, для синусоидальной регрессии на нескольких признаках формула может быть следующей  $y = k_1 \sin(a_1 x_1 + b_1) + k_2 \sin(a_2 x_2 + b_2) + \dots$ , причем вместо коэффициентов должны быть числа, полученные при обучении (используйте методы `coef_` и `intercept_`);
  - (c) одна из моделей должна быть протестирована при различных оптимизаторах (по умолчанию обычно используется метод наименьших квадратов, но в реализации, например, регрессии с  $l_2$ -регуляризацией на `python` есть возможность сменить метод оптимизации, он же `solver`);
  - (d) должно быть не менее двух метрик определение качества у каждой модели;
  - (e) результаты всех задач для каждой модели помимо метрик следует представить графически.
4. Сделать вывод о том, как модель качественнее описывает данные.

ВАРИАНТ 1. data1.csv  
ВАРИАНТ 2. data2.csv  
ВАРИАНТ 3. data3.csv  
ВАРИАНТ 4. data4.csv  
ВАРИАНТ 5. data5.csv  
ВАРИАНТ 6. data6.csv  
ВАРИАНТ 7. data7.csv  
ВАРИАНТ 8. data8.csv