## 

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Санкт-Петербургский государственный технологический институт (технический университет)»

Дисциплина: «Проектирование систем интеллектуального анализа промышленных данных»

Отчёт по лабораторной работе № 1.

**Вариант 26**

**Выполнил: студент группы**

**№**4294м**:**

Пилецки А.Б.

**Проверил:** Соболевский В.А.

Санкт-Петербург

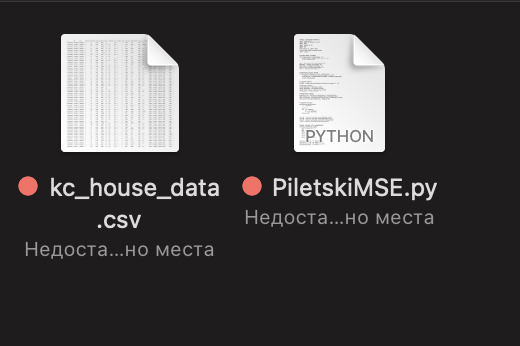
2023

**Цель:** разработка модели прогнозирования цен на недвижимость.

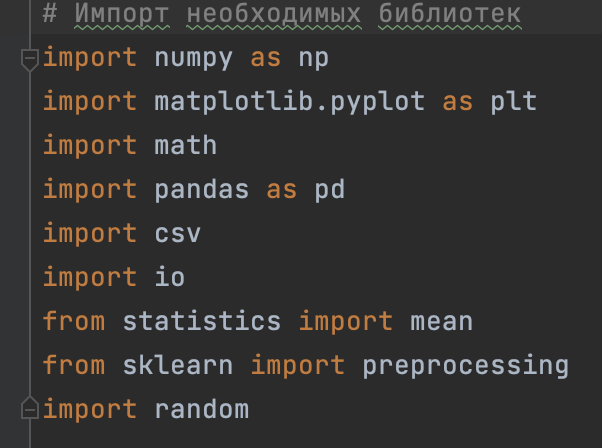
Пункт работы 1: Загружена выборка данных kc\_house\_data.csv

Загружена на локальный компьютер и разархивирована выборка

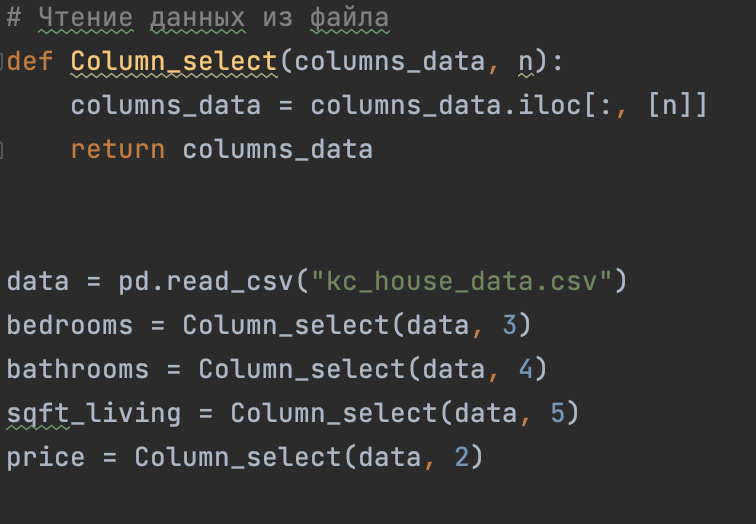
данных с сайта: (https://www.kaggle.com/code/djuuuu/general-linear-models/data)



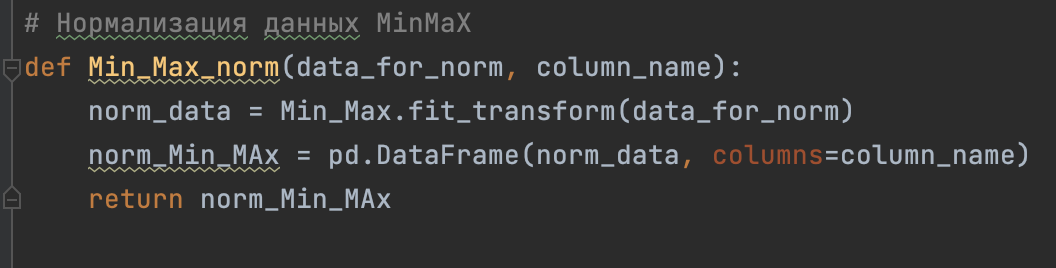
Также подключаем все необходимые библиотеки для реализации проекта.



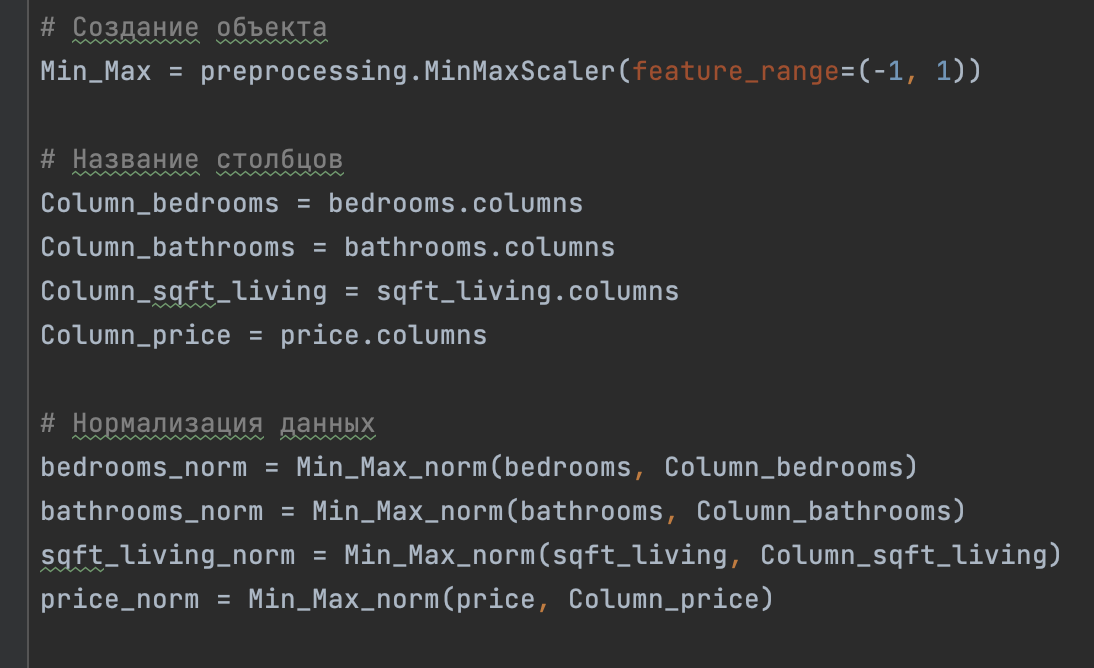
Пункт работы 2: Загрузка и выборка столбцов согласно варианту (столбцы bedrooms, bathrooms, sqft\_living)



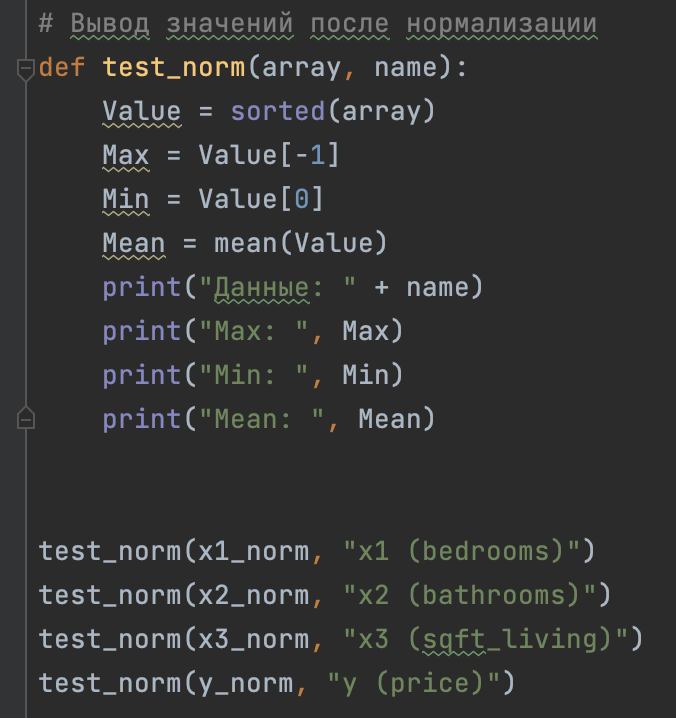
Пункт работы 3: Функция нормализации данных по алгоритму MinMax для каждого набора данных



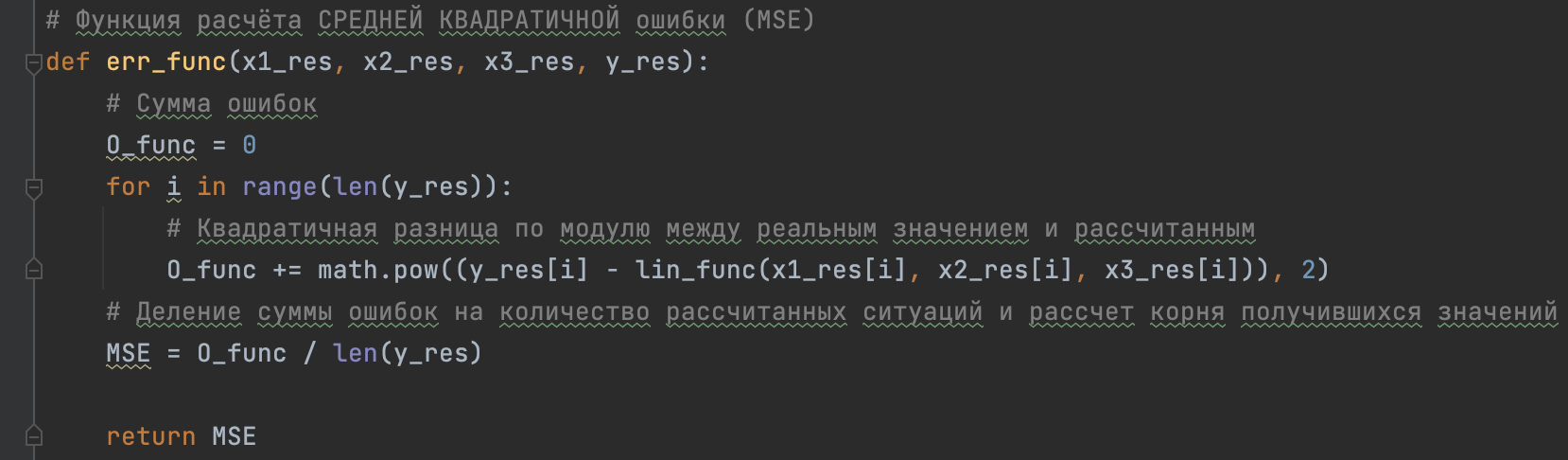
Пункт работы 4: Нормализация данных по алгоритму MinMax (согласно варианту) с использованием библиотеки Sklearn.



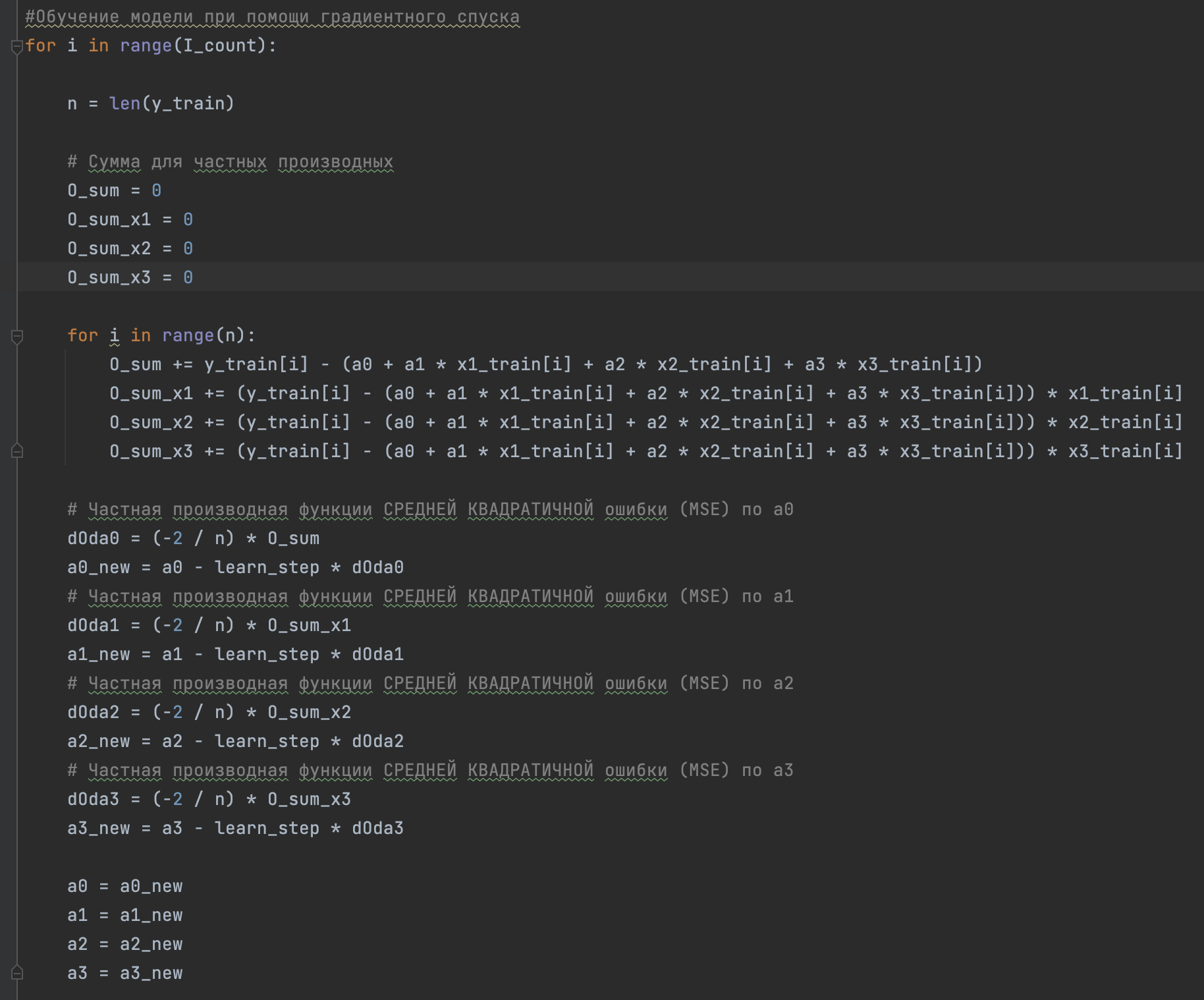
Функция для расчета максимального, минимального и среднего значения после нормализации:



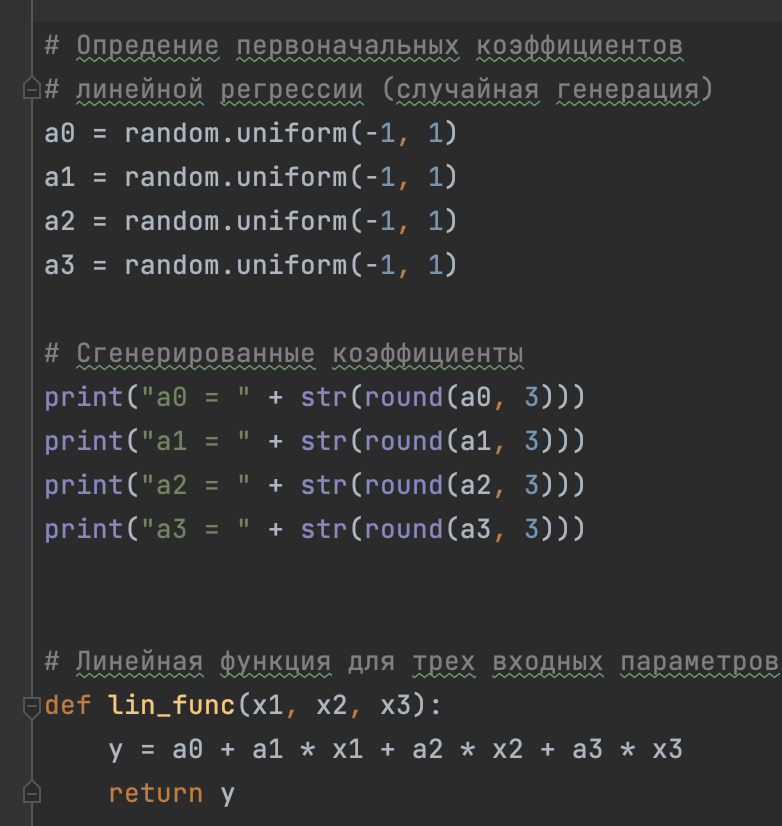
Пункт работы 5: Функция расчета ошибки и функция расчета производных



Пункт работы 6: функция градиентного спуска

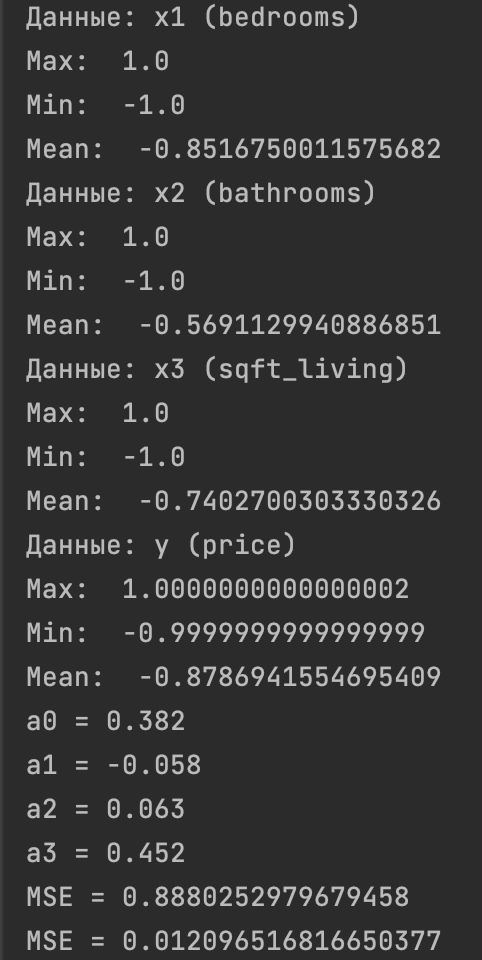


Пункт работы 7: Модель, инициализация коэффициентов рандомным образом и вызов функции из пункта 6, для обучения коэффициентов



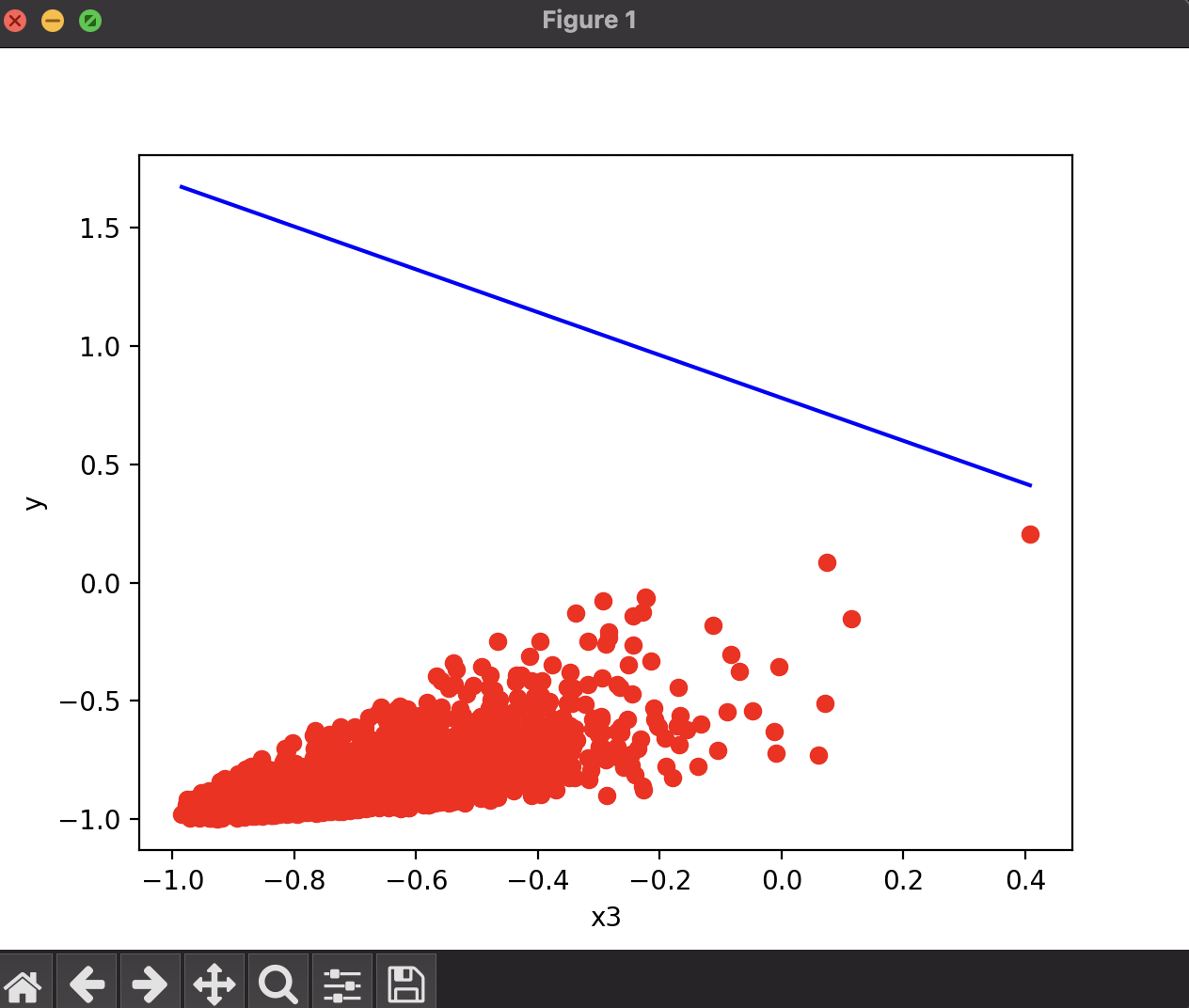
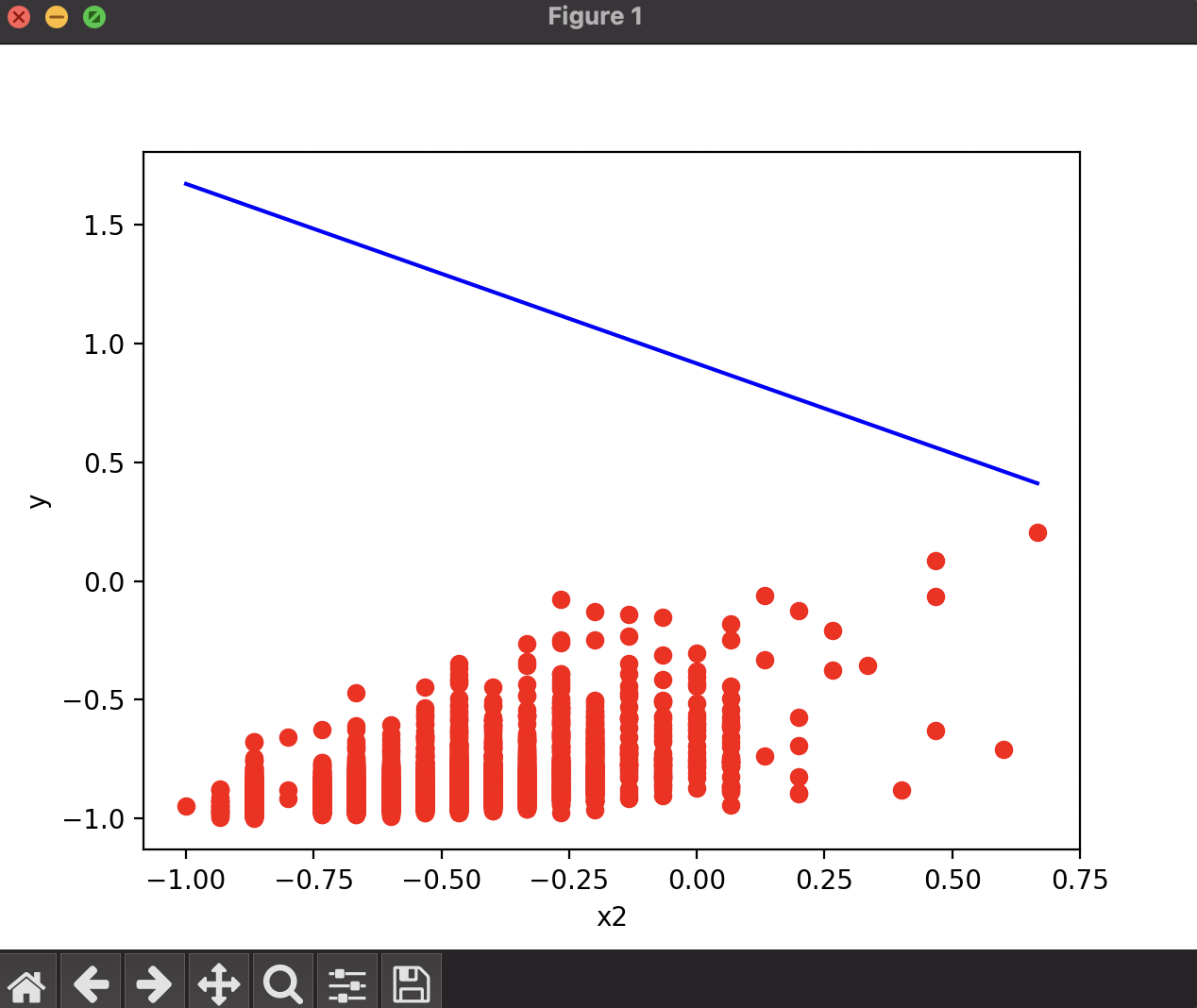
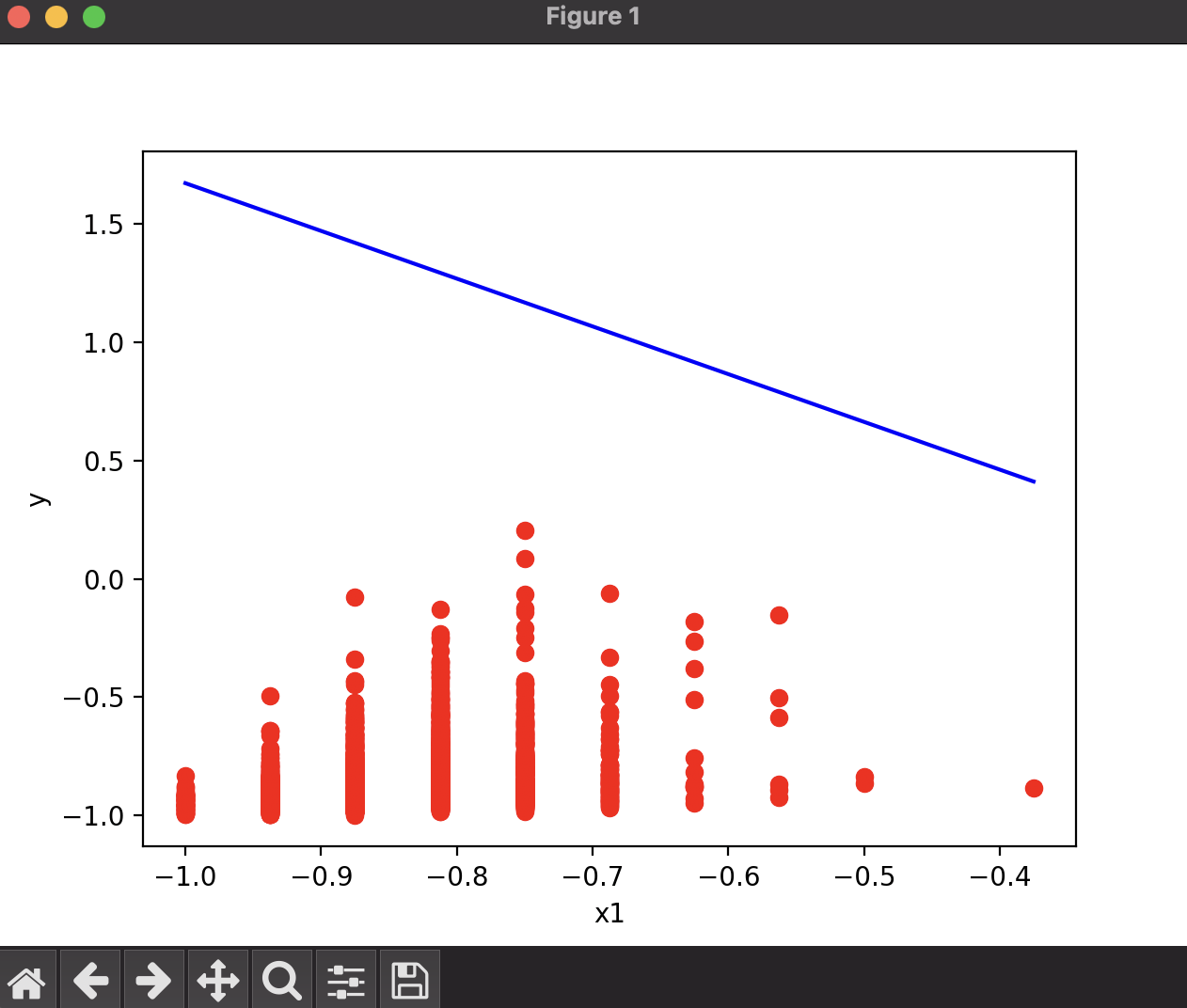
Пункт работы 8: в результате выполнения программы, были инициализированы коэффициенты модели случайным образов и обучены.

Вывод пункта: в результате выполнения программы ошибка MSE уменьшилась. На основании этого можно сделать вывод, что точки данных находятся ближе к линии наилучшего соответствия, чем до обучения.

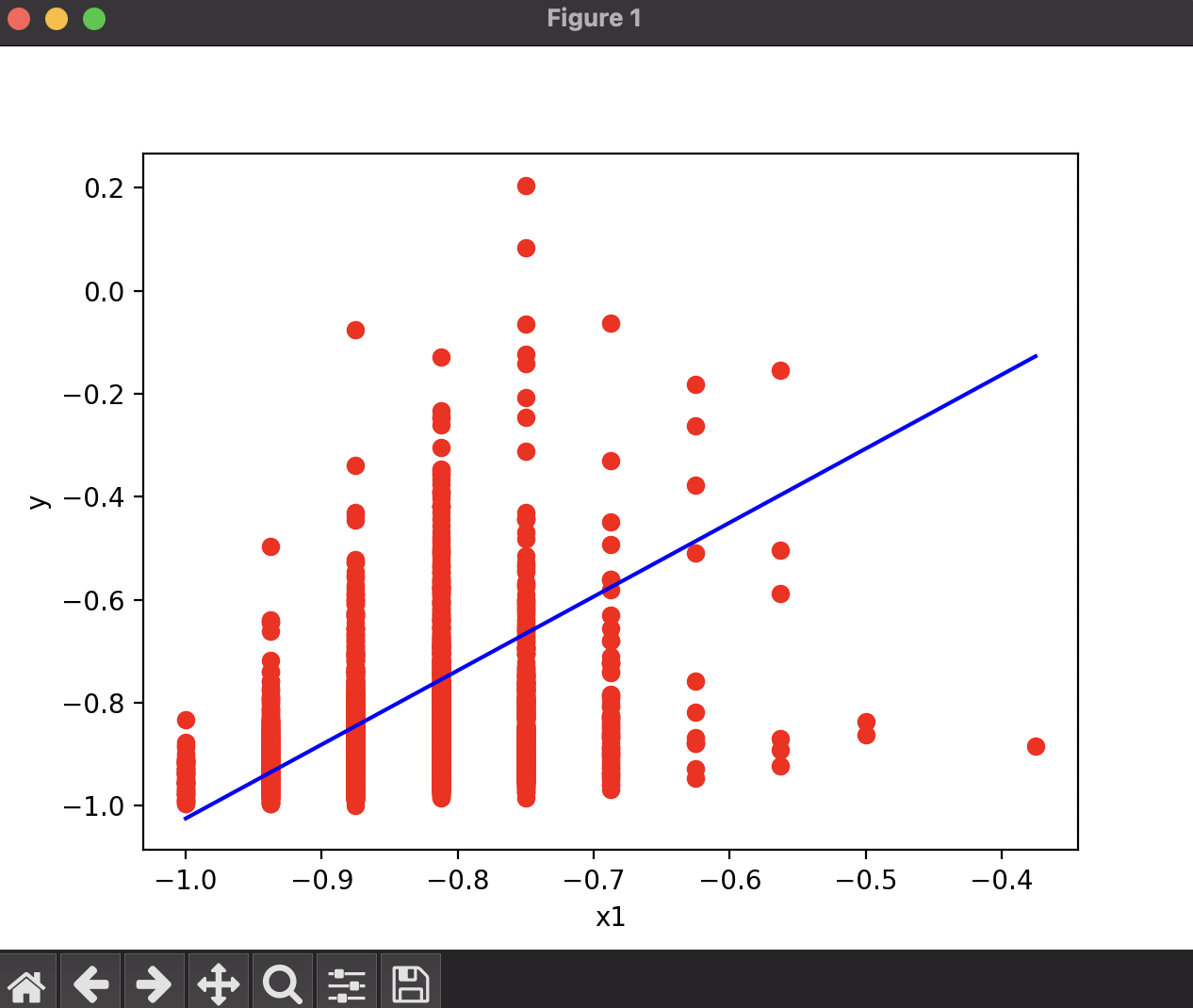


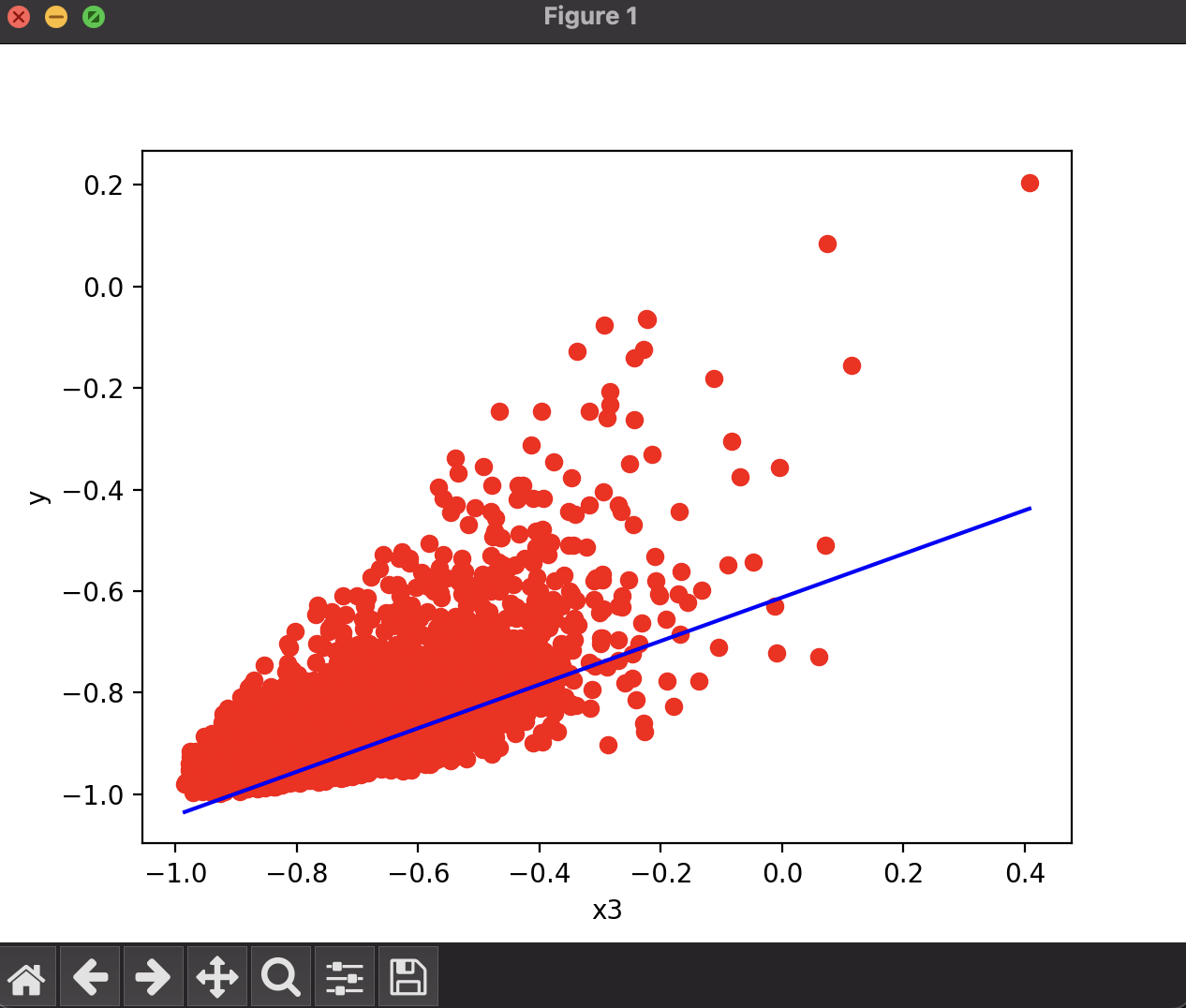
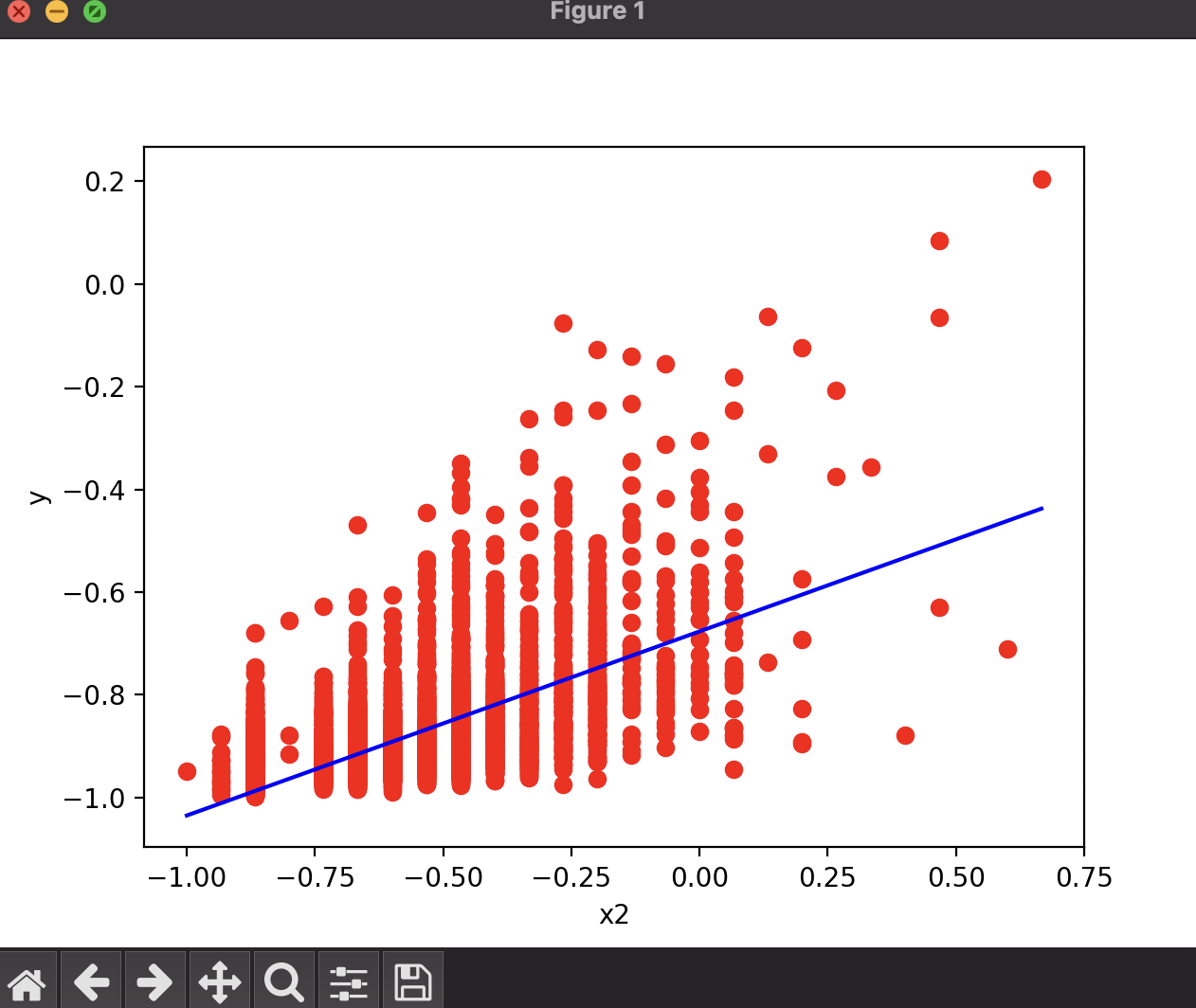
Пункт работы 9: 3 двухмерных графика обученной модели для каждого входного параметра относительно выходного параметра

До обучения:



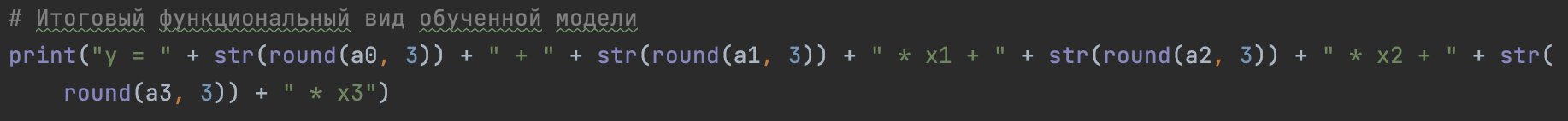
После обучения:





Вывод по пункту: в результате анализа графиков можно сделать вывод, что после обучения срез обученной регрессии находится в более оптимальном положении, относительно тестовых данных, чем перед обучением.

Пункт работы 10: Функциональный вид обученной регрессии





Вывод: Значительное влияние имеет параметр x2 (bathrooms),т.к имеет больше разброс значений, влияние других переменных x1(bedrooms) и x3(sqft\_living) значительно меньше.

Пункт 11: Итоговый вывод: в ходе лабораторной работы была разработана модель прогнозирования цен на недвижимость.

Были написаны функции линейной регрессии, основанной на трёх

переменных (характеристик дома).

В результате выполнения лабораторной работы была достигнута цель: разработана модель прогнозирования цен на недвижимость. Результаты указывают на то, что ошибка MSE уменьшилась, коэффициенты стали более оптимальными и срез регрессии находится в более оптимальном положении относительно входных данных.

