**Практическая работа 1**

Дедлайн: 25 сентября, 23:59

(за каждый просроченный день -10% от оценки)

1. Постройте как минимум 3 графика, описывающих данные. Какие предварительные выводы вы можете сделать по этим графикам? (2 балла)
2. Проведите предобработку данных (3 балла):
   1. Выявление аномалий
   2. Заполнение отсутствующих значений
   3. Кодирование категориальных переменных
3. Постройте модель линейной регрессии, и изобразите ее на графике. Рассчитайте MSE (mean squared error) (1 балл)
4. Напишите формулу получившейся линейной регрессии (1 балл)
5. Выведите summary модели. На основе полученных данных сделайте выводы о получившейся линейной регрессии. Какие из независимых переменных Х имеют большее влияние на Y, а какие Х не стоит использовать для прогноза Y? (3 балла)
6. Проверьте три допущения линейной регрессии и сделайте на их основе выводы (5 баллов)
   1. Линейная зависимость
   2. Постоянная дисперсия
   3. Нормальность остатков
7. Что-то от себя, что не обсуждалось на лекциях и семинарах (+5 бонусных баллов)

Полезные материалы:

1. Лекции
2. Семинары
3. <https://www.codecamp.ru/blog/linear-regression-assumptions/>
4. <https://www.simplilearn.com/tutorials/python-tutorial/data-visualization-in-python>
5. Все знания мира - в интернете