Задание по Теме 5

1. Dataset “blood.xlsx” (X1 = систолическое давление (зависимая переменная), X2 = возраст пациента в годах, X3 = вес в фунтах (две независимые переменные)):

* Постройте модель ридж регрессии для X1 как функции от X2 и X3. Используйте функции cv.plot и bias.plot для нахождения параметра k.
* Проверьте мультиколлинеарность данных.
* Проверьте значимость построенного уравнения регрессии.
* Проведите сравнительной анализ полученной модели с моделью линейной регрессии (когда оценки найдены по МНК).

1. Dataset “Kuiper.xls” (price – зависимая переменная, mileage, liter, cruise, sound, leather – независимые переменные):

* Постройте квантильную регрессию для различных значений параметров «тау», Y = price – зависимая переменная, mileage, liter, cruise, sound, leather leather – независимые переменные).
* Изобразите результаты на графиках.
* Постройте модель линейной регрессии и сравните эту модель с моделью медианной регрессии.

1. Dataset “cigarettes.txt”: (переменные по столбцам: carbon monoxide, tar, nicotine, weight):

* Постройте несколько моделей нелинейной (квадратичной и кубической) регрессии с зависимой переменной y=carbon monoxide, и независимыми переменными: x1=tar, x2=nicotine, x3=weight.
* Постройте модель линейной регрессии и сравните эту модель с моделями нелинейной регрессии.