**Московский государственный технический**

**университет им. Н.Э. Баумана**

Факультет "Информатика и системы управления"

Кафедра ИУ5 "Системы обработки информации и управления"

Технологии машинного обучения

Отчет по лабораторной работе №1

Выполнила: Пасатюк Александра, группа ИУ5-63Б

Москва, 2020

**Задание**

Создать ноутбук, который содержит следующие разделы:

1. Текстовое описание выбранного Вами набора данных.
2. Основные характеристики датасета.
3. Визуальное исследование датасета.
4. Информация о корреляции признаков.

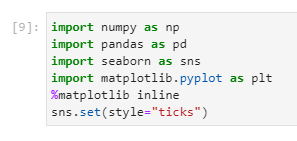
**Выполнение лабораторной работы**

# 1.Текстовое описание данных

Набор данных о подержанных автомобилях Toyota в Великобритании. Набор данных содержит следующие колонки:

* model - модель автомобиля
* year - год выпуска
* price - цена (в фунтах стерлингах)
* transmission - тип коробки передач
* mileage - пробег (в милях)
* fuelType - тип используемого топлива
* tax - налог на автомобиль
* mpg - расход топлива (миль на галон)
* engineSize - объем двигателя

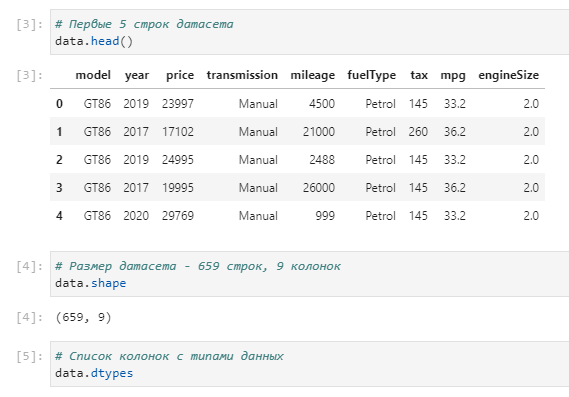
# Импорт библиотек

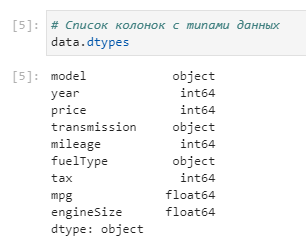


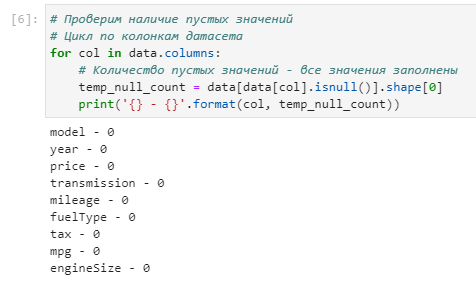
Загрузка данных



2.Основные характеристики датасета







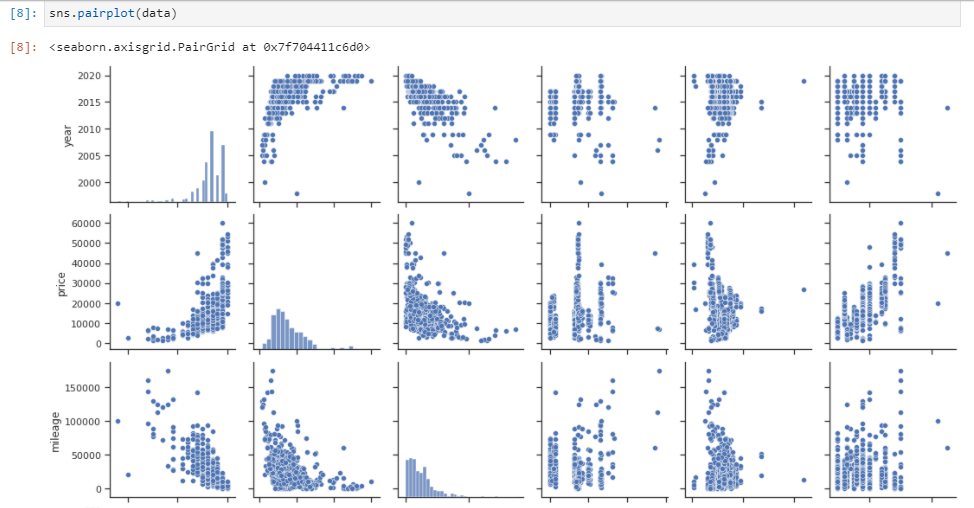


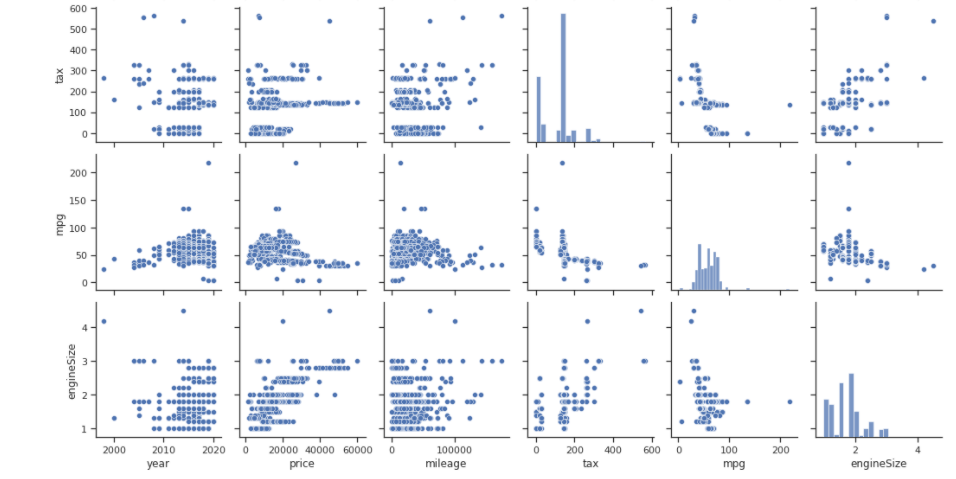
3.Визуальное исследование датасета

### "Парные диаграммы"

Комбинация гистограмм и диаграмм рассеивания для всего набора данных.

Выводится матрица графиков. На пересечении строки и столбца, которые соответстуют двум показателям, строится диаграмма рассеивания. В главной диагонали матрицы строятся гистограммы распределения соответствующих показателей.





4.Информация о корреляции признаков

