****Московский государственный технический университет им. Н.Э. Баумана  
Кафедра «Системы обработки информации и управления»

Лабораторная работа №2  
по дисциплине  
«Методы машинного обучения»  
на тему

«Обработка признаков»

Выполнил: Ван Пэй  
Группа: ИУ5-21М

Москва — 2021 г.

**1. Цель лабораторной работы**

Изучение продвинутых способов предварительной обработки данных для дальнейшего формирования моделей.

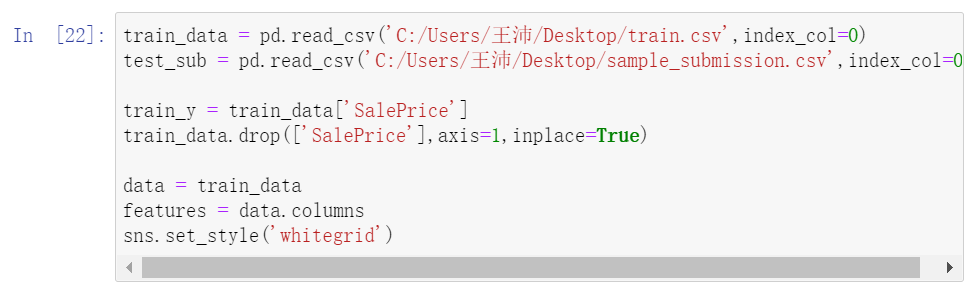
**2. Задание**

* устранение пропусков в данных;
* кодирование категориальных признаков;
* нормализацию числовых признаков.

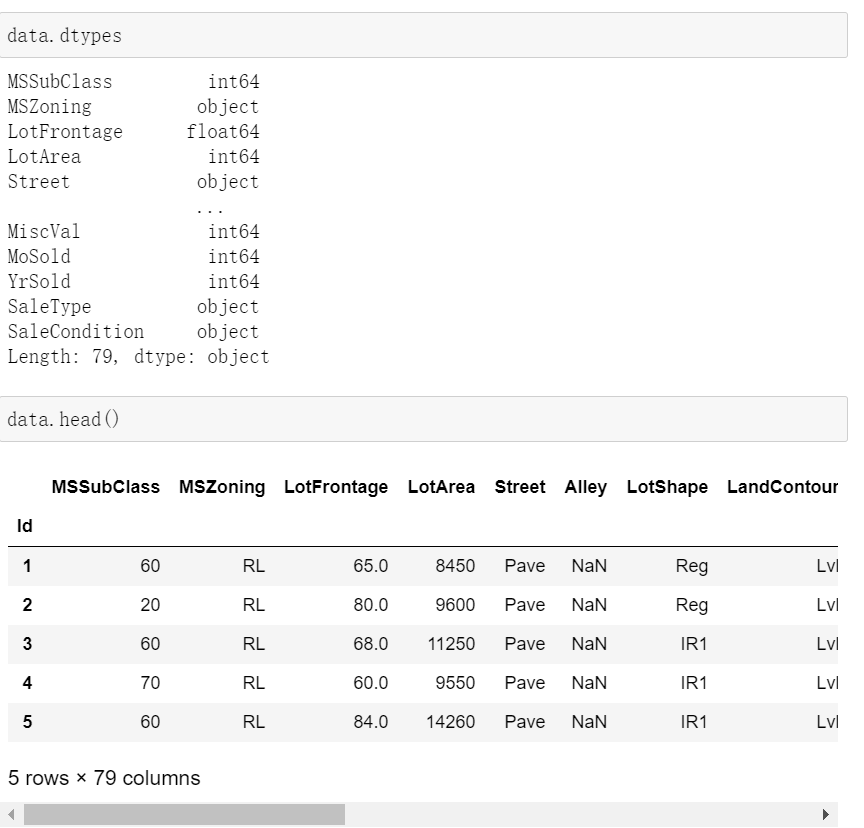
**3. Ход выполнения работы**Подключим необходимые библиотеки и настроим отображение графиков



Возьмём набор данных：



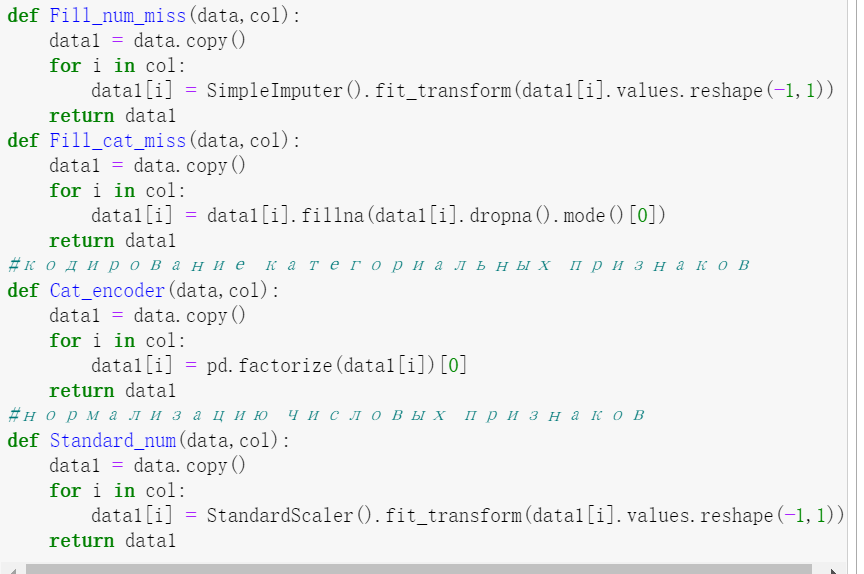
Посмотрим на эти наборы данных:



Функции:

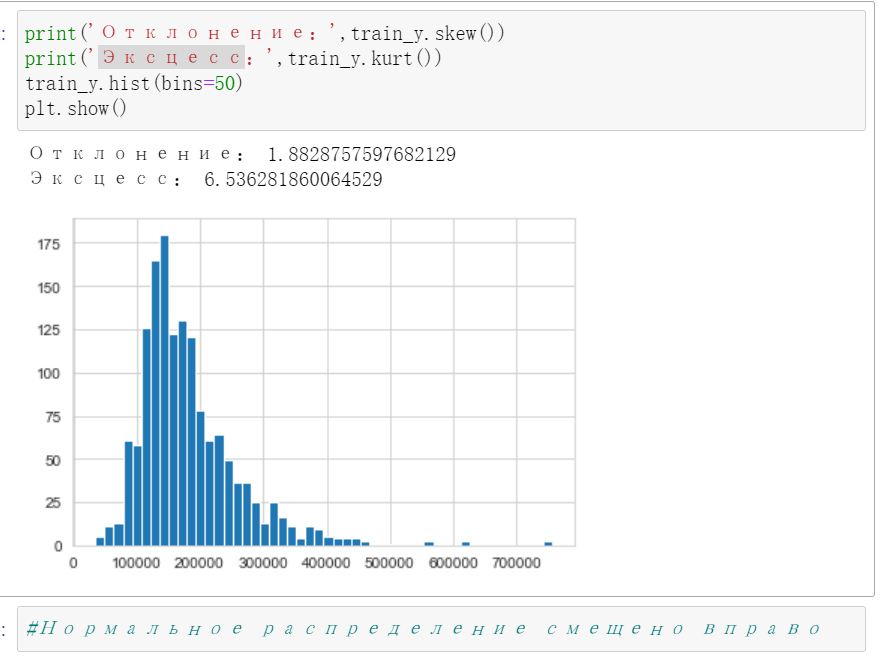


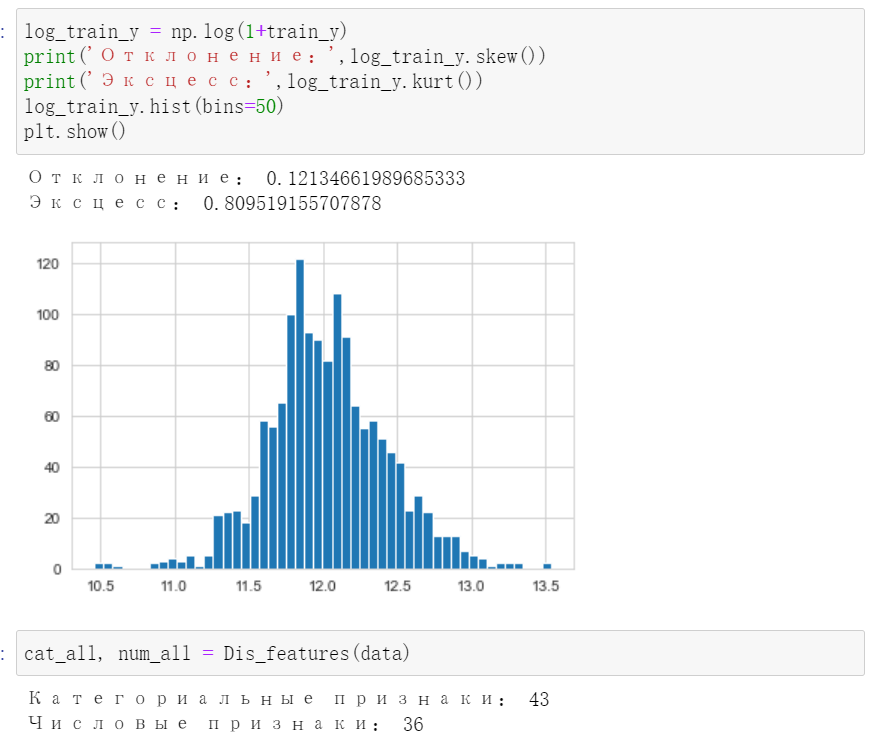


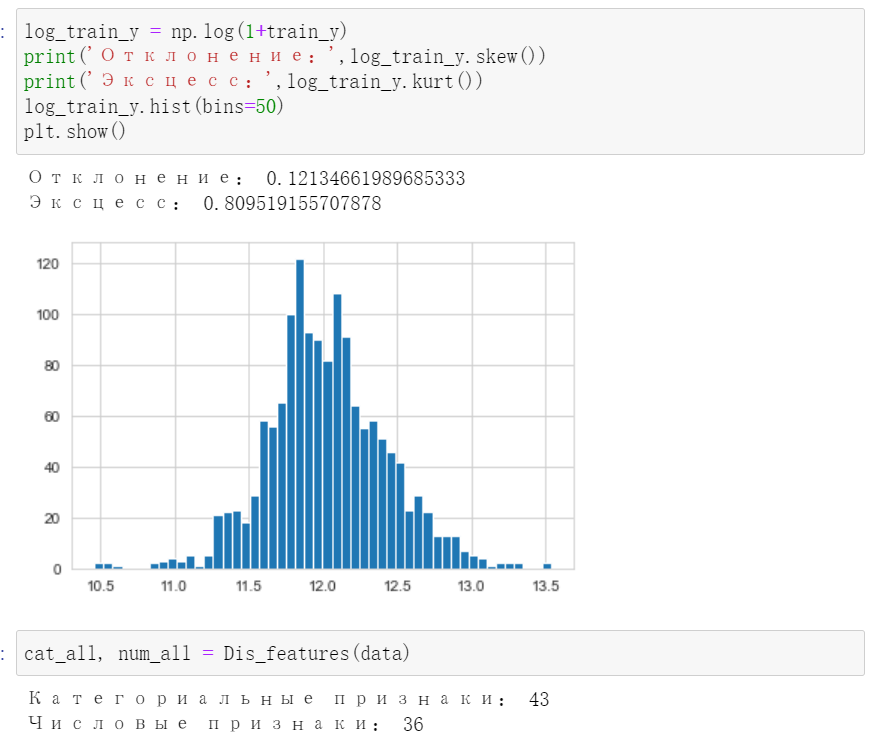


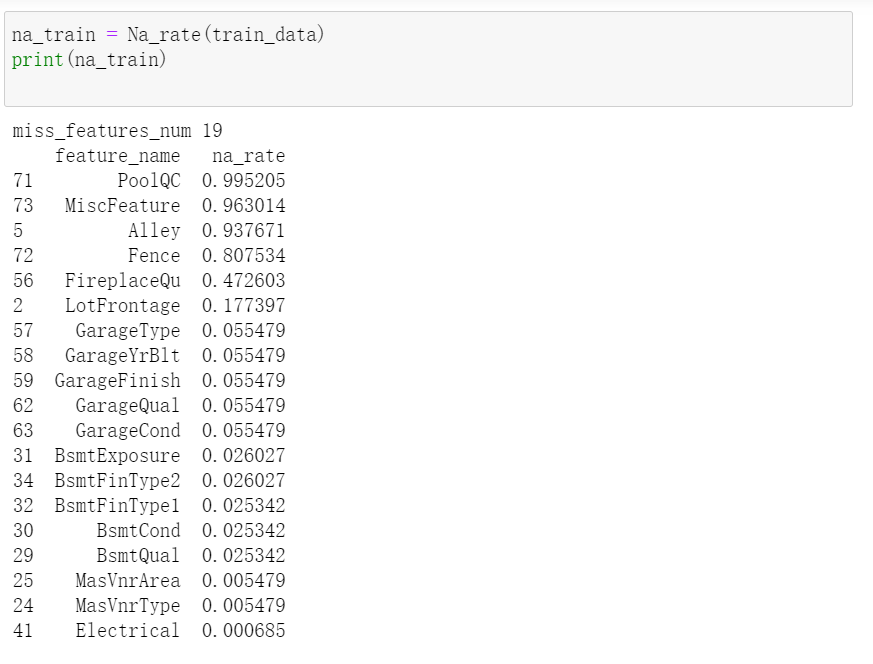
**3.1. Обработка пропусков в данных**

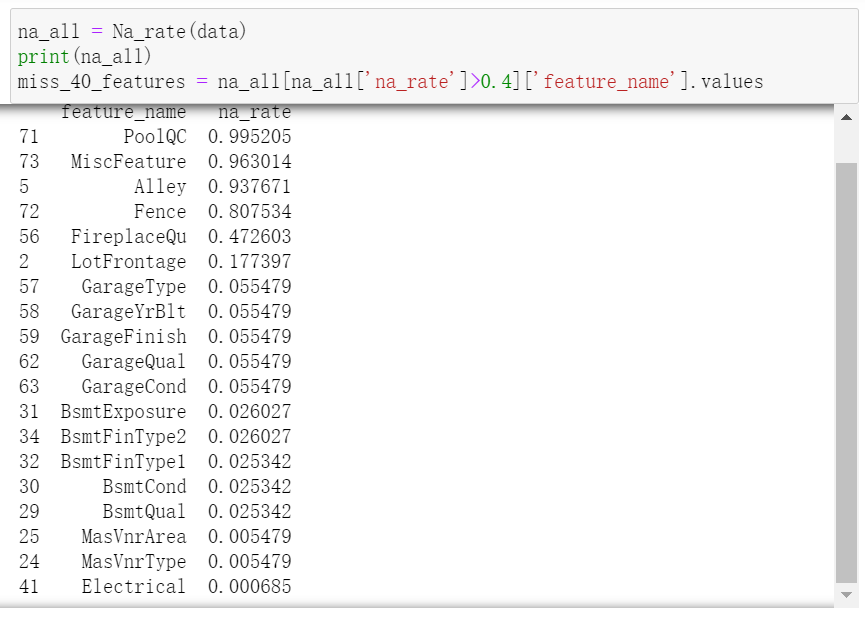
Проверить целевую переменную:

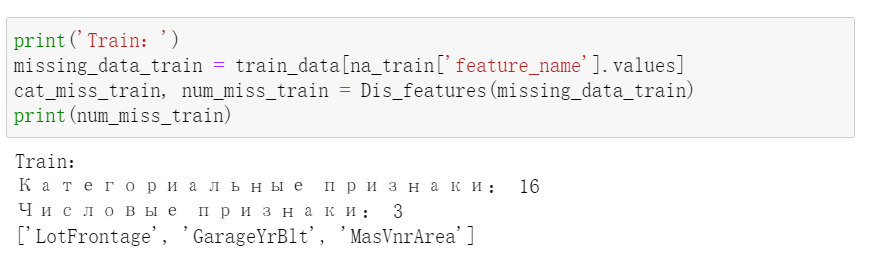






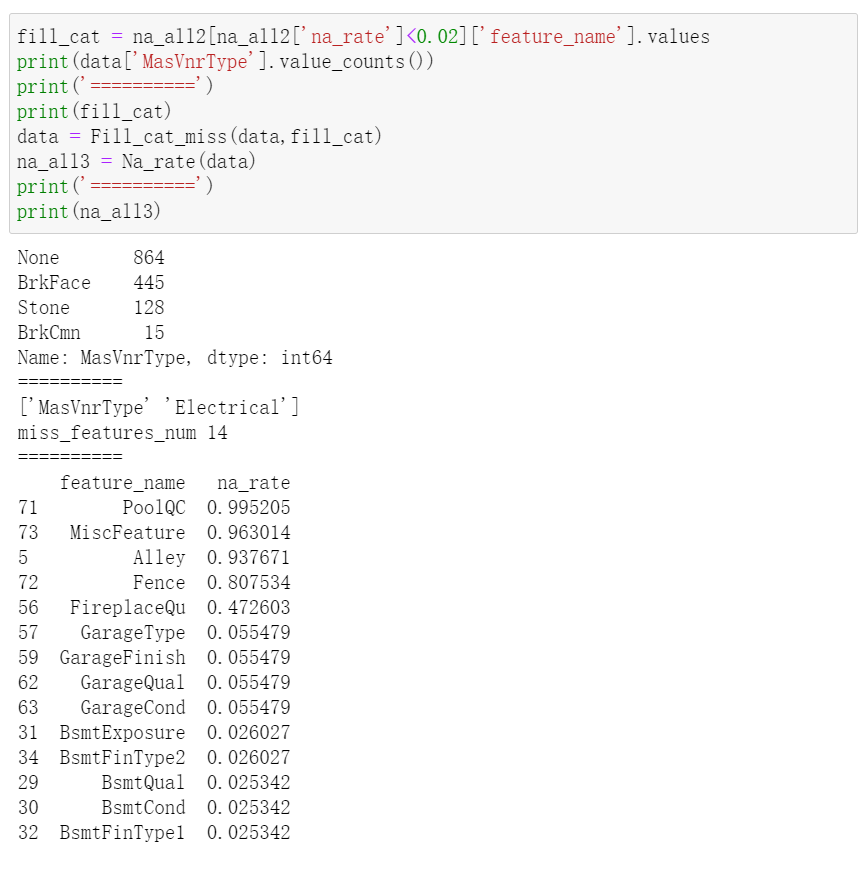












**3.2. Кодирование категориальных признаков**



**3.3. Нормализацию числовых признаков.**





**Список литературы**

[1] Гапанюк Ю. Е. Лабораторная работа «Разведочный анализ данных. Исследование  
и визуализация данных» [Электронный ресурс]

[2] Wes McKinney «Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython»