МГТУ им. Н. Э. Баумана, кафедра ИУ5

курс “Технология машинного обучения”

Лабораторная работа №3

# «Обработка пропусков в данных, кодирование категориальных признаков, масштабирование данных»

ВЫПОЛНИЛ:

Матюнин да Вейга Р.А.

Группа: ИУ5-61Б

ПРОВЕРИЛ:

Гапанюк Ю.Е.

Москва 2020

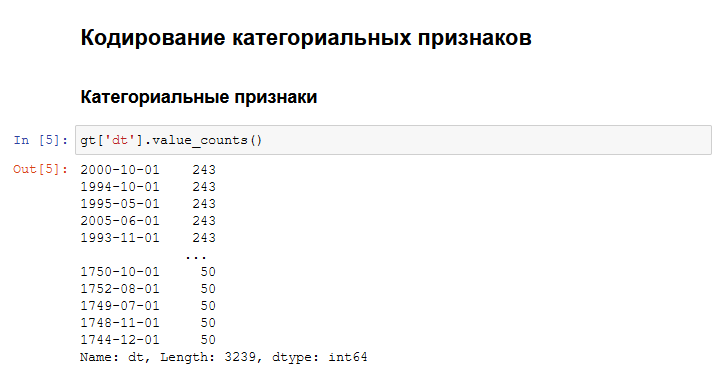
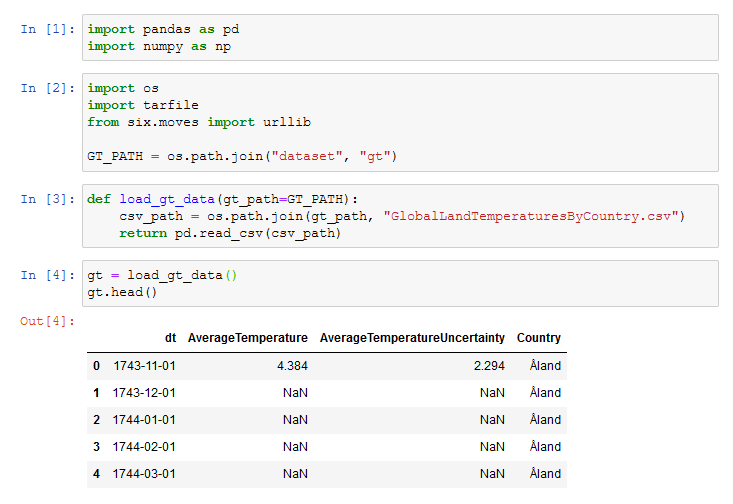
**Цель лабораторной работы:** изучение способов предварительной обработки данных для дальнейшего формирования моделей.

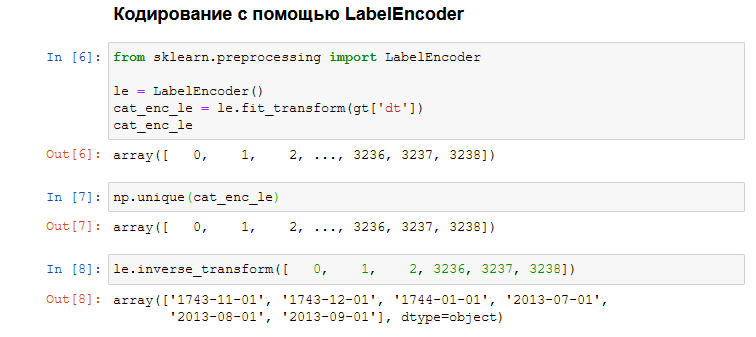
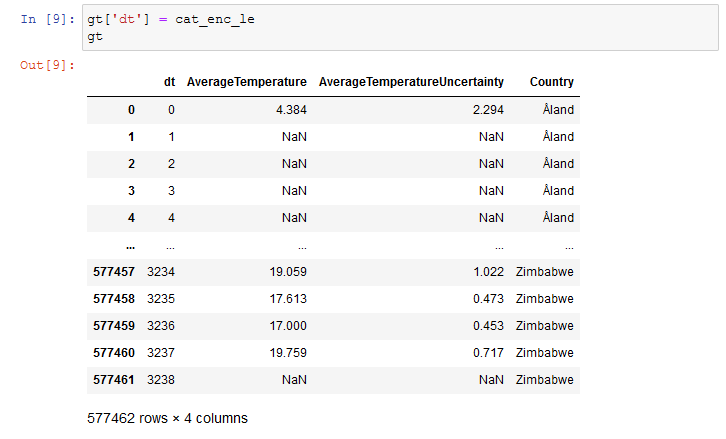
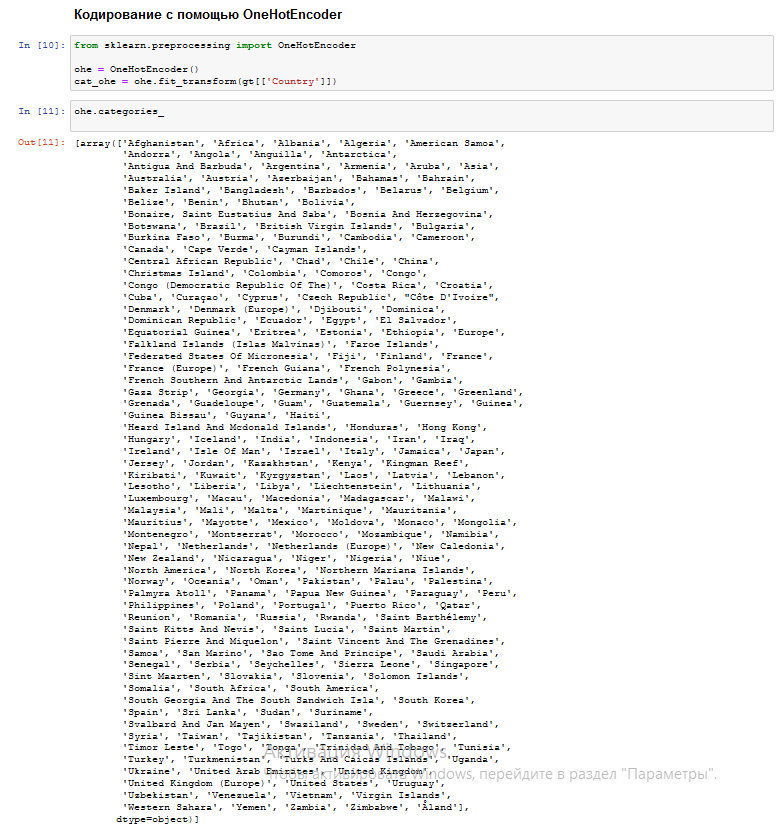
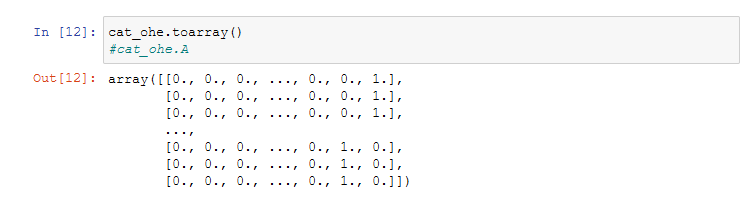
**Задание:**

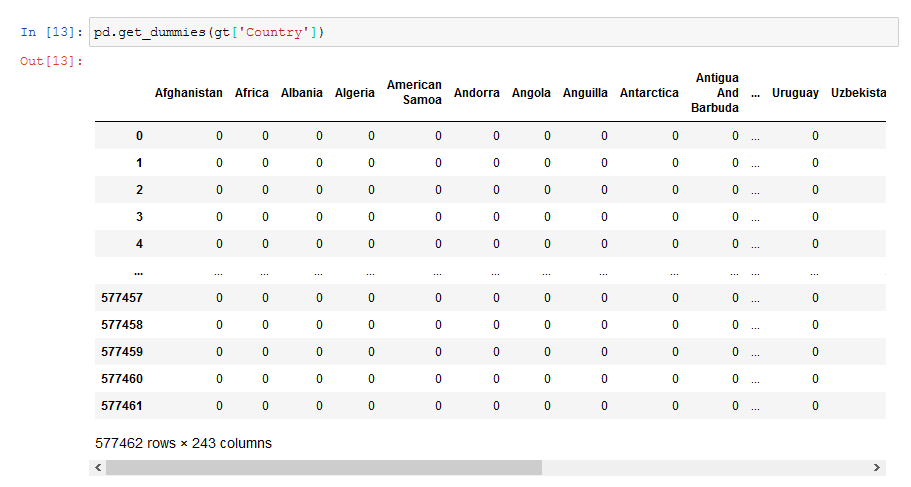
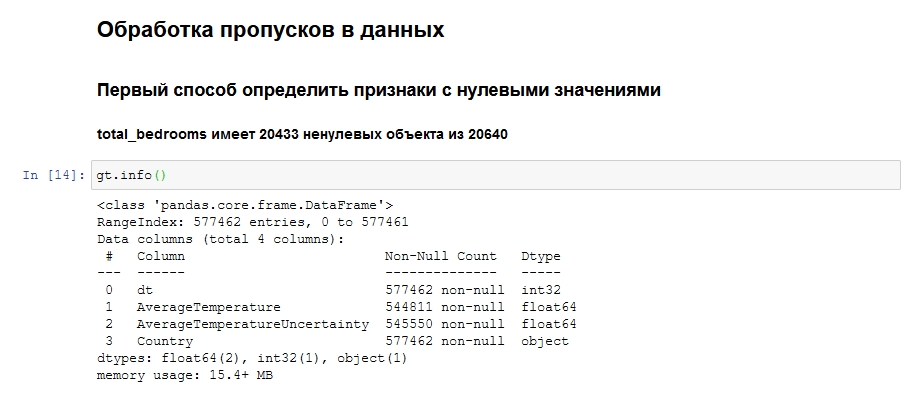
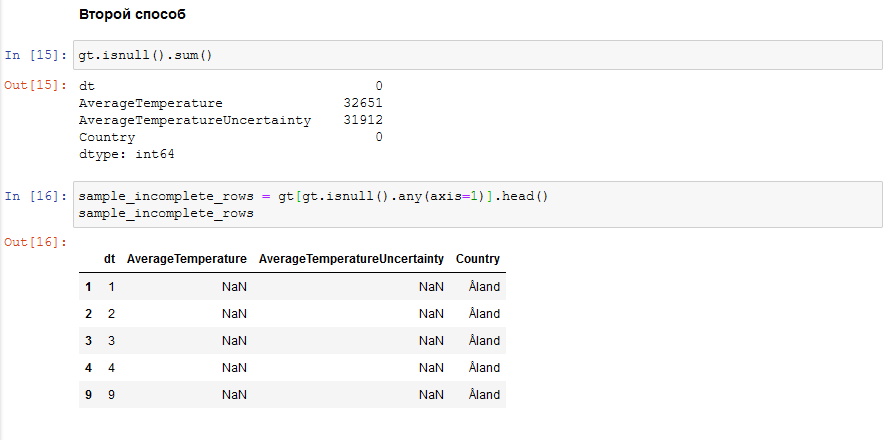
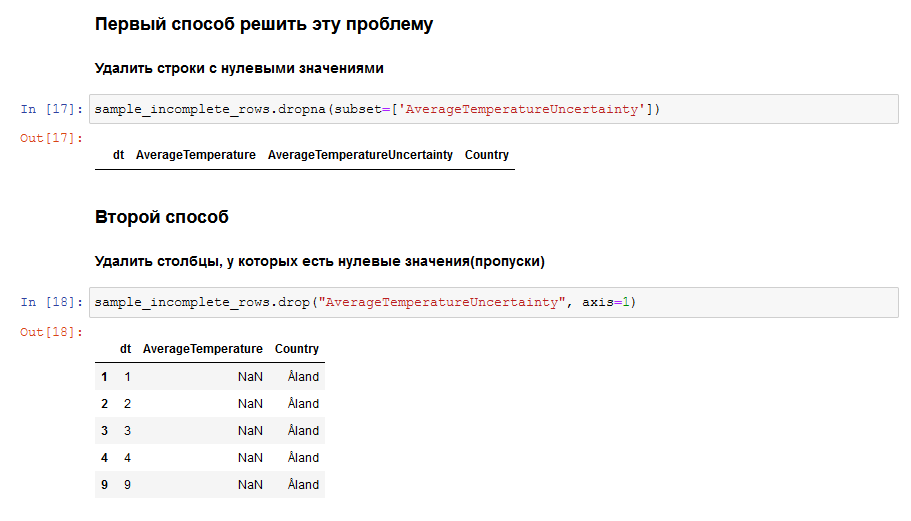
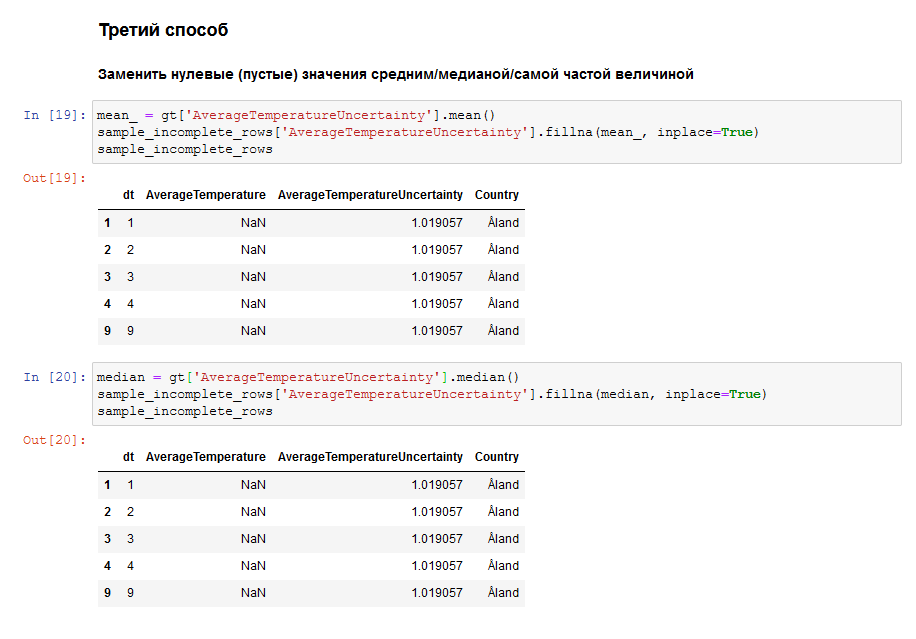
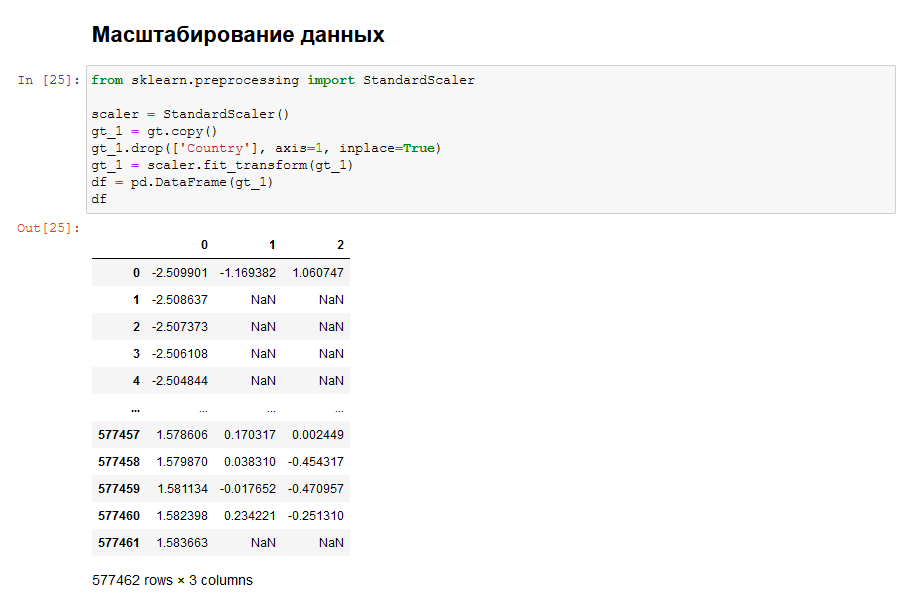
1. Выбрать набор данных (датасет), содержащий категориальные признаки и пропуски в данных. Для выполнения следующих пунктов можно использовать несколько различных наборов данных (один для обработки пропусков, другой для категориальных признаков и т.д.)
2. Для выбранного датасета (датасетов) на основе материалов [лекции](https://github.com/ugapanyuk/ml_course_2020/blob/master/common/notebooks/missing/handling_missing_norm.ipynb) решить следующие задачи:

* обработку пропусков в данных;
* кодирование категориальных признаков;
* масштабирование данных.

### **Выполненная работа**



* Ноутбук с выполненной работой и отчет размещены в репозитории на github: <https://github.com/Yorati/TMO>