Федеральное государственное бюджетное   
образовательное учреждение высшего   
профессионального образования

Московский государственный технический   
университет имени Н.Э. Баумана   
(МГТУ им. Н.Э. Баумана)

**ЛАБОРАТОРНАЯ РАБОТА 3  
«Обработка пропусков в данных, кодирование категориальных признаков, масштабирование данных»**

**ПО ДИСЦИПЛИНЕ «Технологии машинного обучения»**

Выполнил(а): Орлова С. М.   
 (Фамилия И.О. студента)

РТ5-61   
 (Индекс группы)

Проверил(а): Гапанюк Ю. Е.   
 (Фамилия И.О. преподавателя)

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_ 2020 г.

ЗАЧТЕНО / НЕ ЗАЧТЕНО \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
 (нужное выделить) (подпись)

Москва, 2020

1. Описание задания лабораторной работы

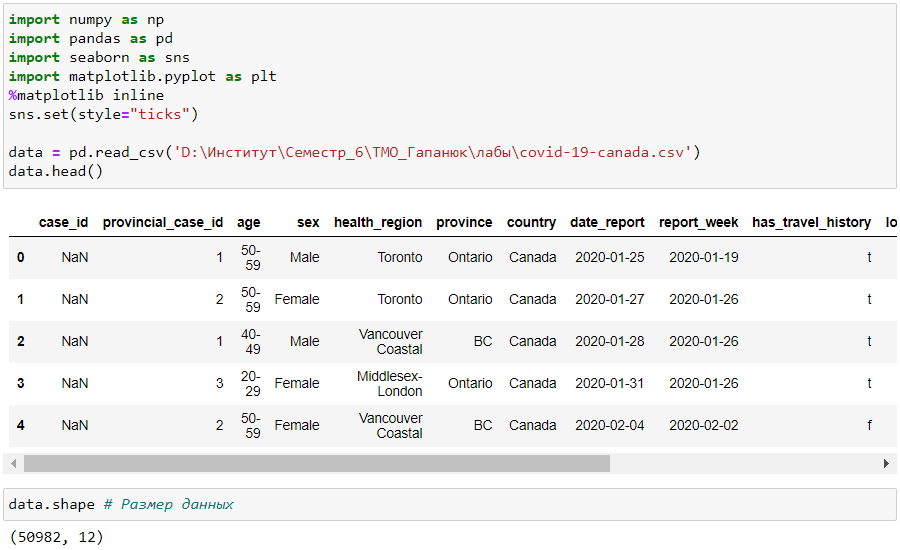
Цель лабораторной работы:  
Изучение способов предварительной обработки данных для дальнейшего формирования моделей.

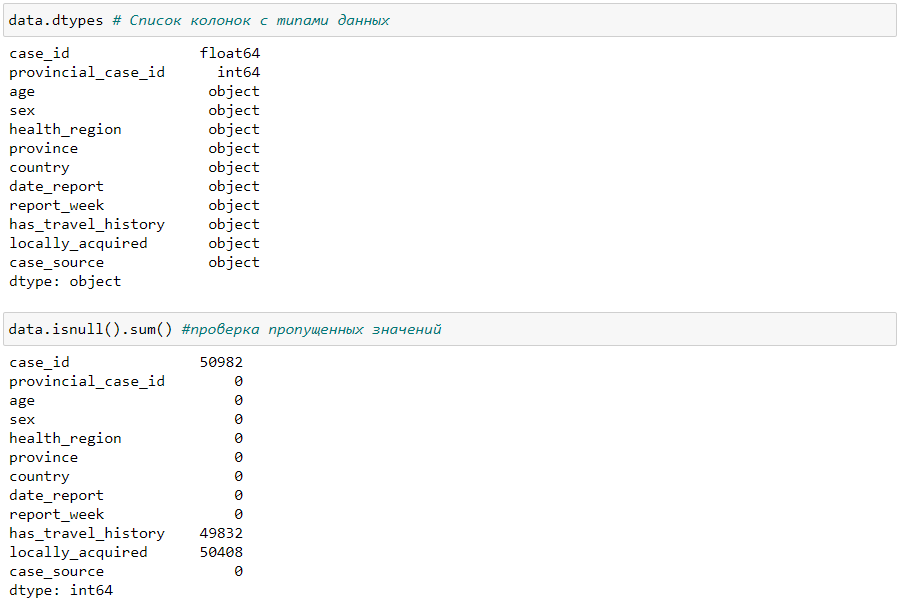
Задание:  
Выбрать набор данных (датасет), содержащий категориальные признаки и пропуски в данных. Для выполнения следующих пунктов можно использовать несколько различных наборов данных (один для обработки пропусков, другой для категориальных признаков и т.д.).  
Для выбранного датасета (датасетов) на основе материалов лекции решить следующие задачи:

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

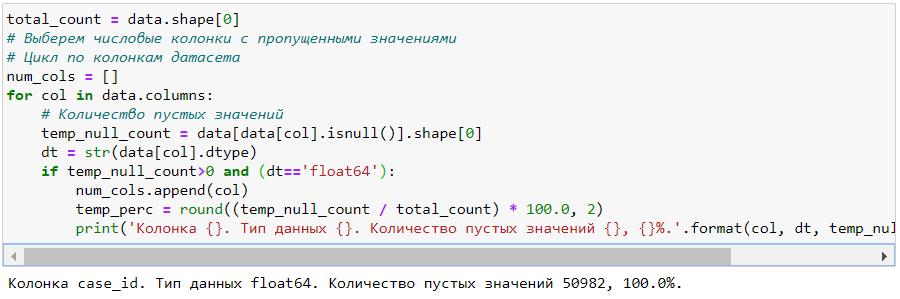
1. Текст программы и экранные формы с примерами выполнения программы

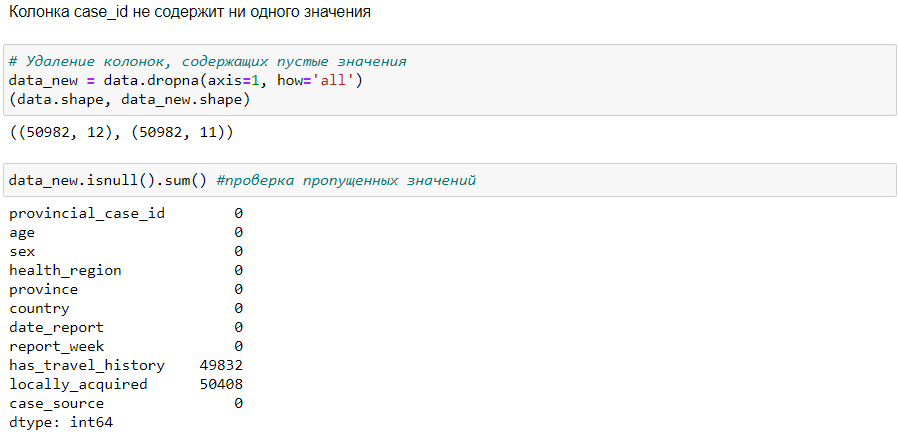
Загрузка и первичный анализ данных:

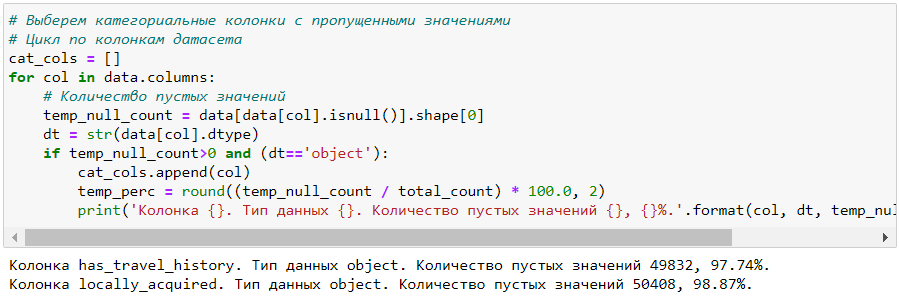


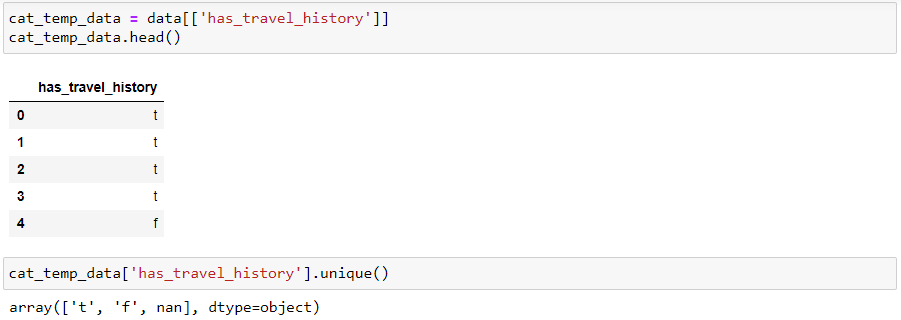


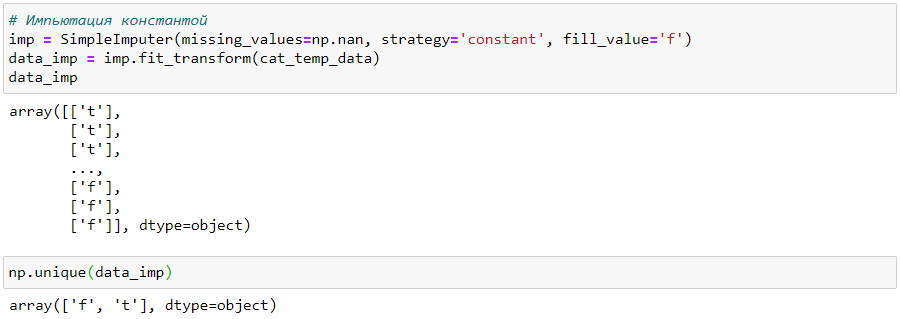
Обработка пропусков в данных:



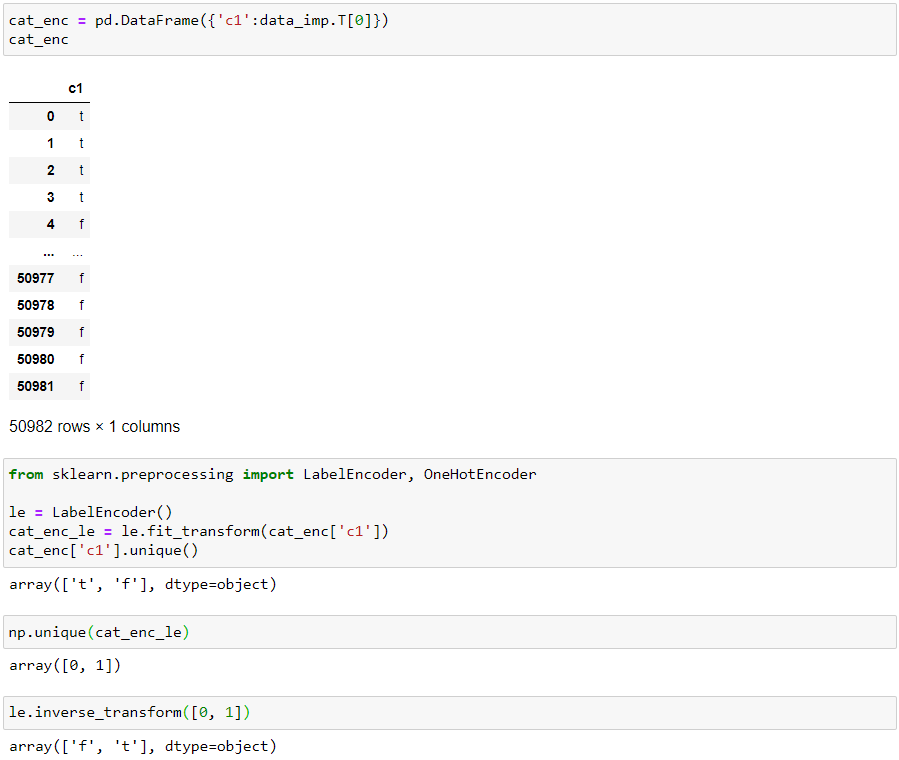








Преобразование категориальных признаков в числовые:



Масштабирование данных:

