Московский государственный технический университет им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Автоматизированные системы обработки информации и управления»



**Отчет**

**Лабораторная работа № 2**

**По курсу «Технологии машинного обучения»**

**ИСПОЛНИТЕЛЬ:**

Группа ИУ5-61Б

Овчинников С.С.

"17" февраля 2020 г.

**ПРЕПОДАВАТЕЛЬ:**

Гапанюк Ю.Е.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

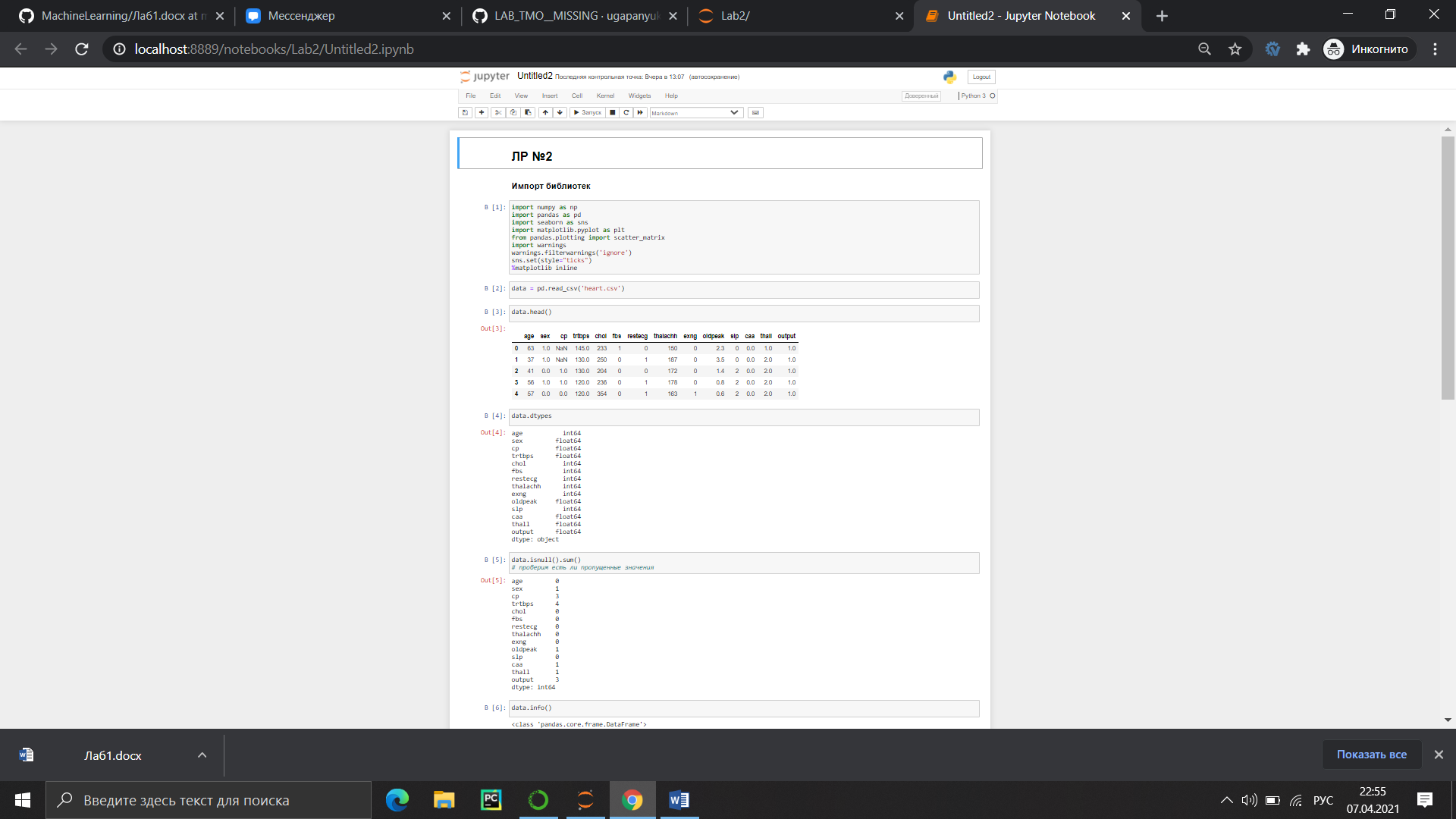
"\_\_"\_\_\_\_\_\_\_\_\_\_\_2021 г.

Москва 2021

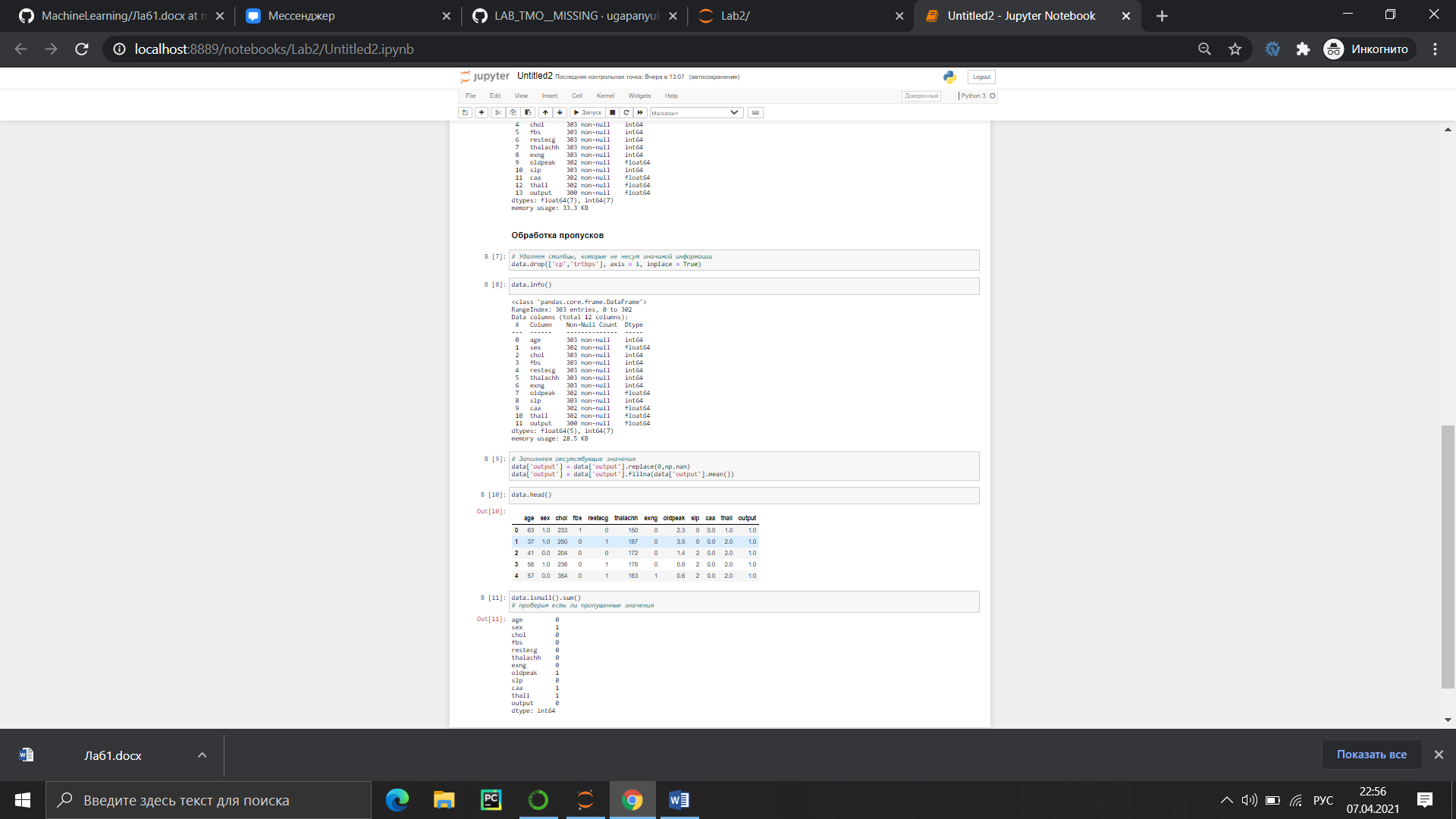
1. **Задание**

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:  
a. обработку пропусков в данных;  
b. кодирование категориальных признаков;  
c. масштабирование данных.

1. **Скрины jupyter notebook**

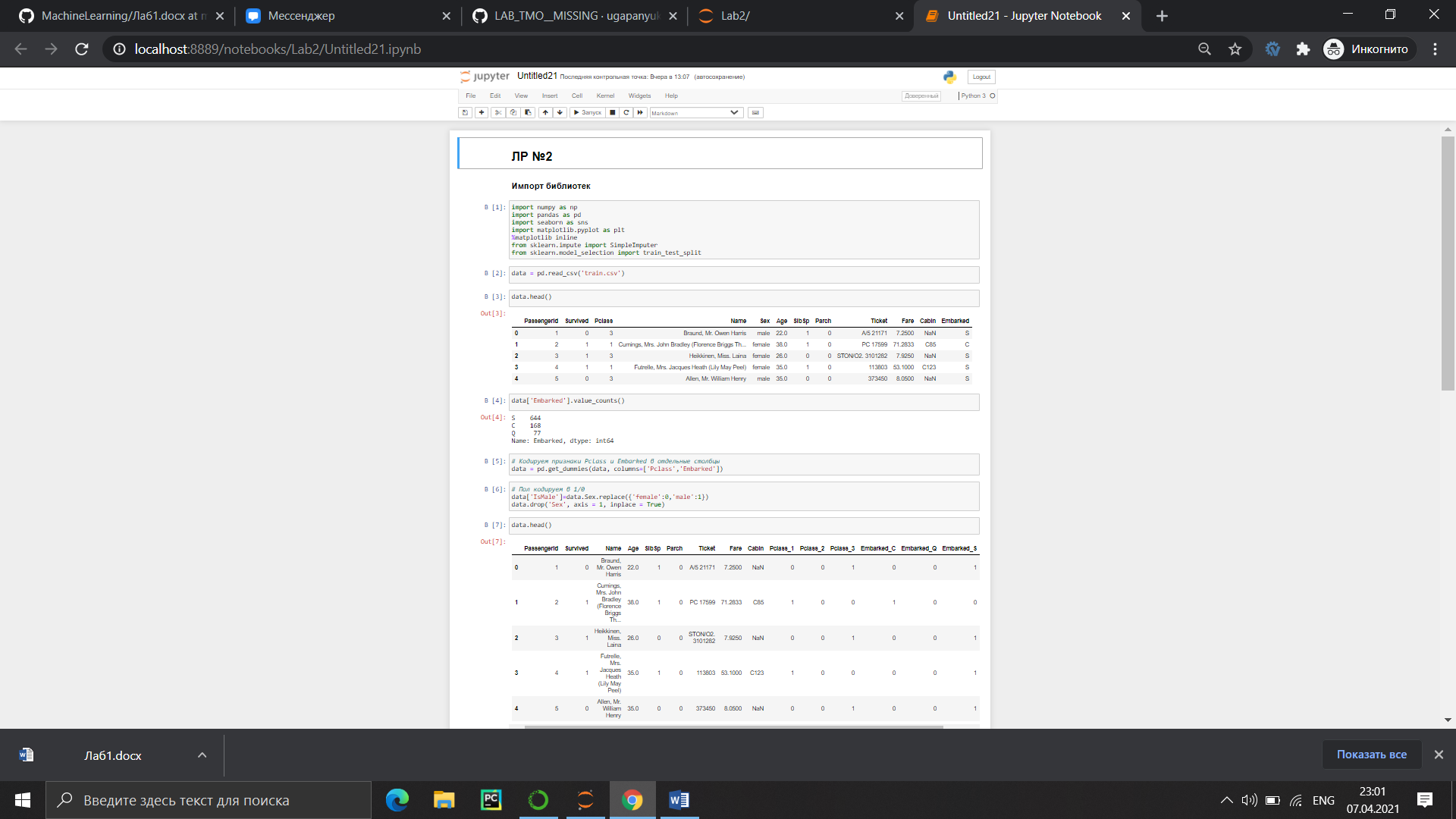


**Обработка пропусков**  
Некоторые пропуски оставлены для наглядности, удалял столбцы cp и trtbps, заменял значения в столбце output, столбцы sex, caa, oldpeak оставил без изменений с пропусками.



Для второго и третьего пункта взял более удобный датасет

**Кодирование признаков**



**Масштабирование значений**

