**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 №2

Выполнила:

Студентка группы ИУ5-63Б

Желтова Александра

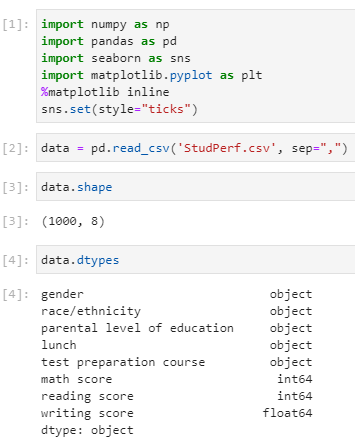
Москва, 2021

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

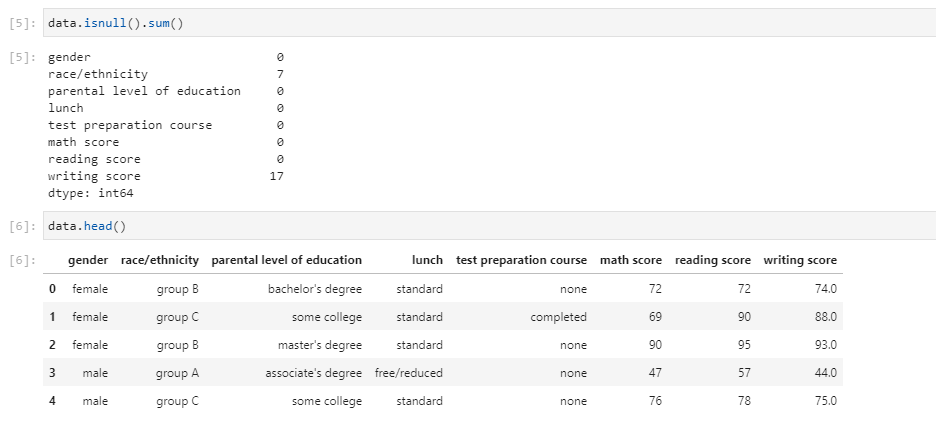
**Задание:**

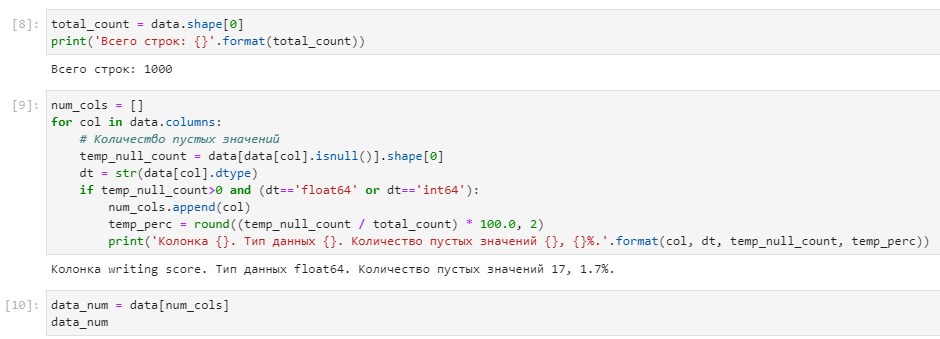
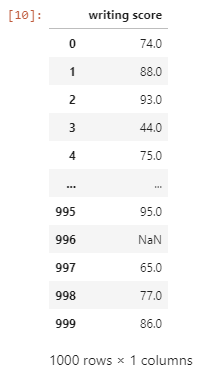
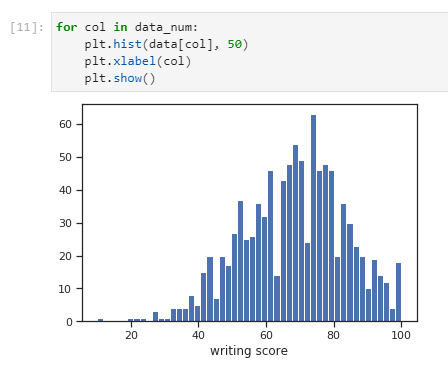
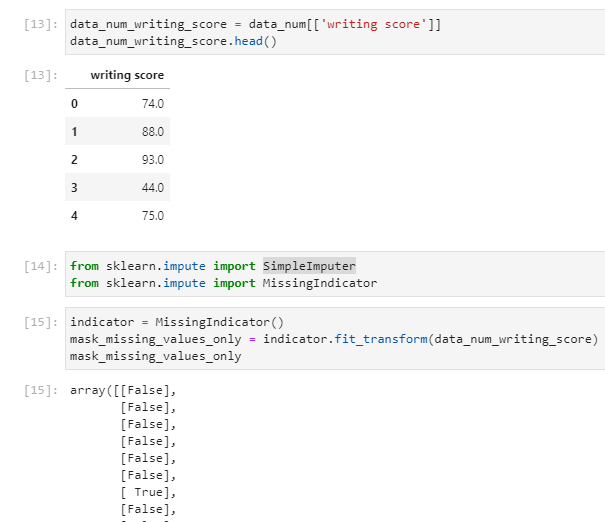
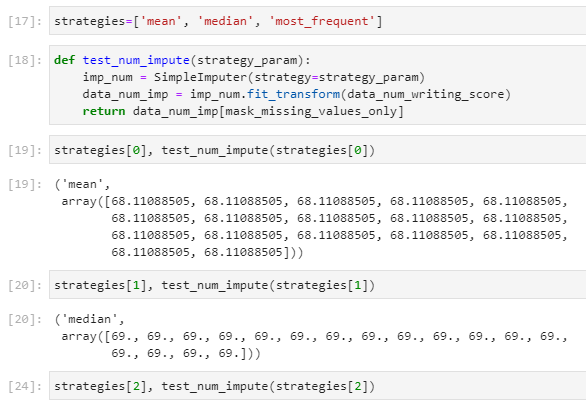
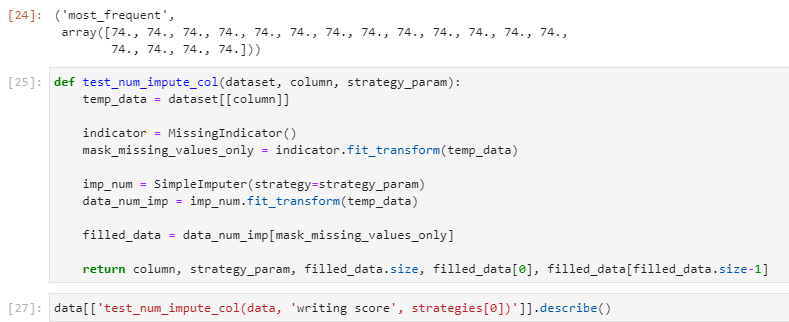
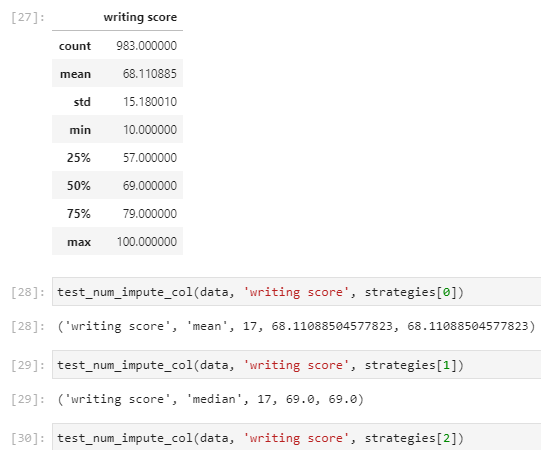
1. Выбрать набор данных (датасет), содержащий категориальные признаки и пропуски в данных. Для выполнения следующих пунктов можно использовать несколько различных наборов данных (один для обработки пропусков, другой для категориальных признаков и т.д.)
2. Для выбранного датасета (датасетов) на основе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

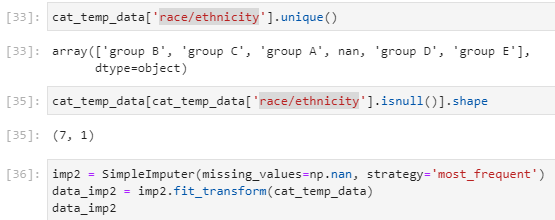
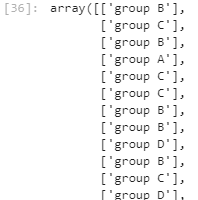
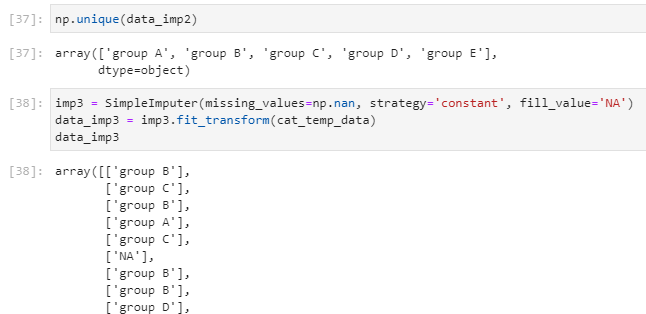
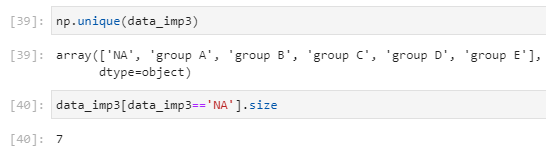
**Выполнение лабораторной работы:**

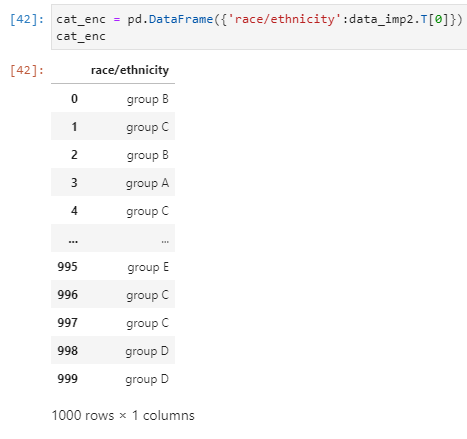


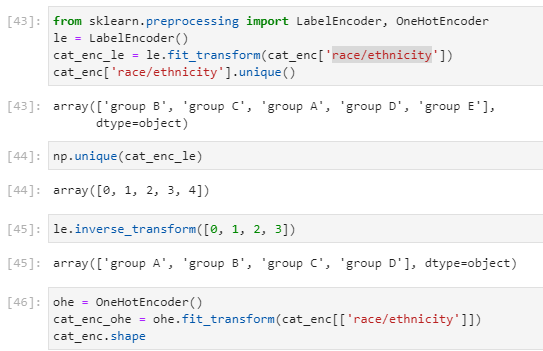
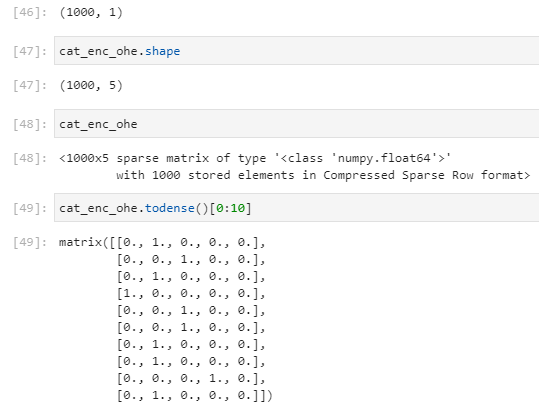
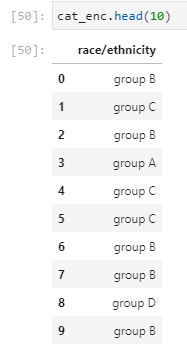
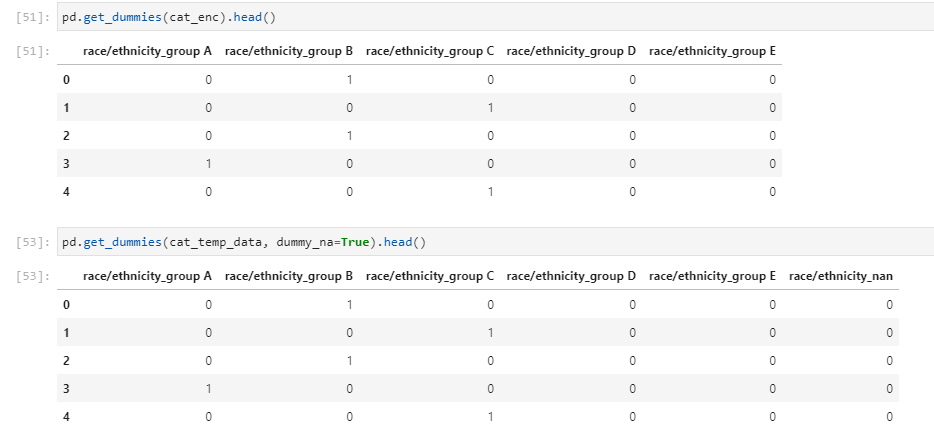
Пропуски в данных:

 Обработка пропусков в числовых данных:

Обработка пропусков в категориальных данных:    

Преобразование категориальных признаков в числовые:

Кодирование категорий числовыми значениями и наборами бинарных значений:   

Масштабирование данных: