Московский государственный технический университет им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Автоматизированные системы обработки информации и управления»



**Отчет**

**Лабораторная работа № 1**

**По курсу «Технологии машинного обучения»**

**ИСПОЛНИТЕЛЬ:**

Группа ИУ5-65Б

Филатова А. Е.

"18" апреля 2021 г.

**ПРЕПОДАВАТЕЛЬ:**

Гапанюк Ю.Е.

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

"\_\_"\_\_\_\_\_\_\_\_\_\_\_2021 г.

Москва 2021

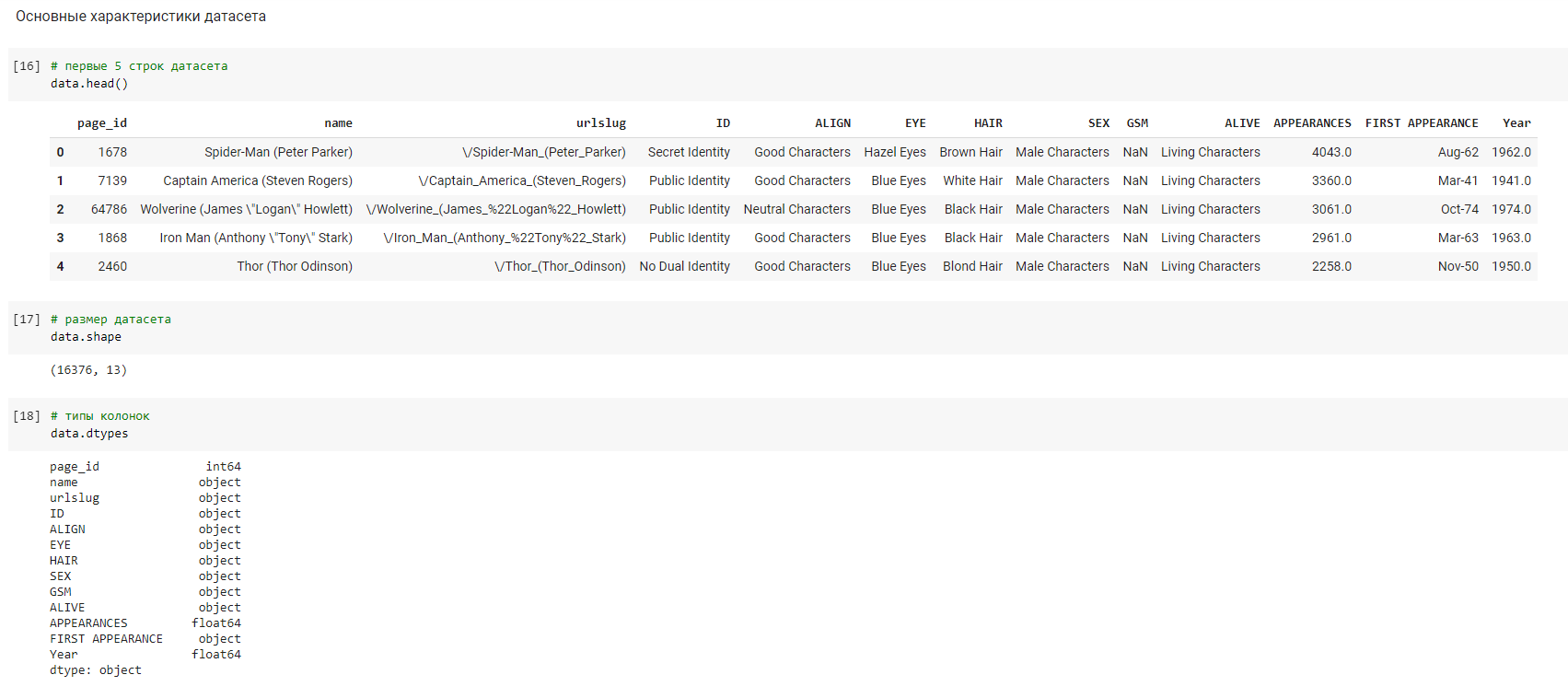
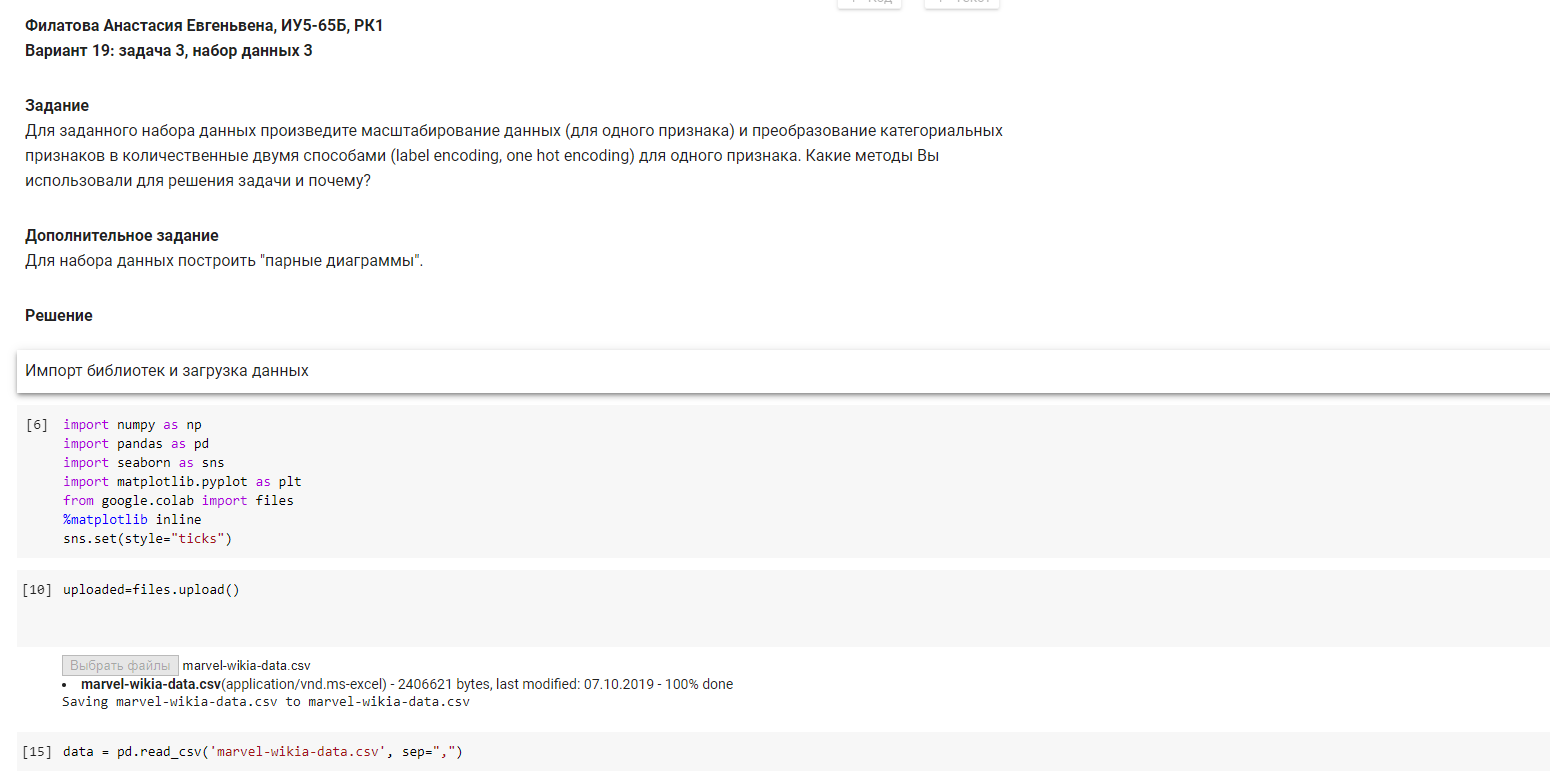
1. **Задание**

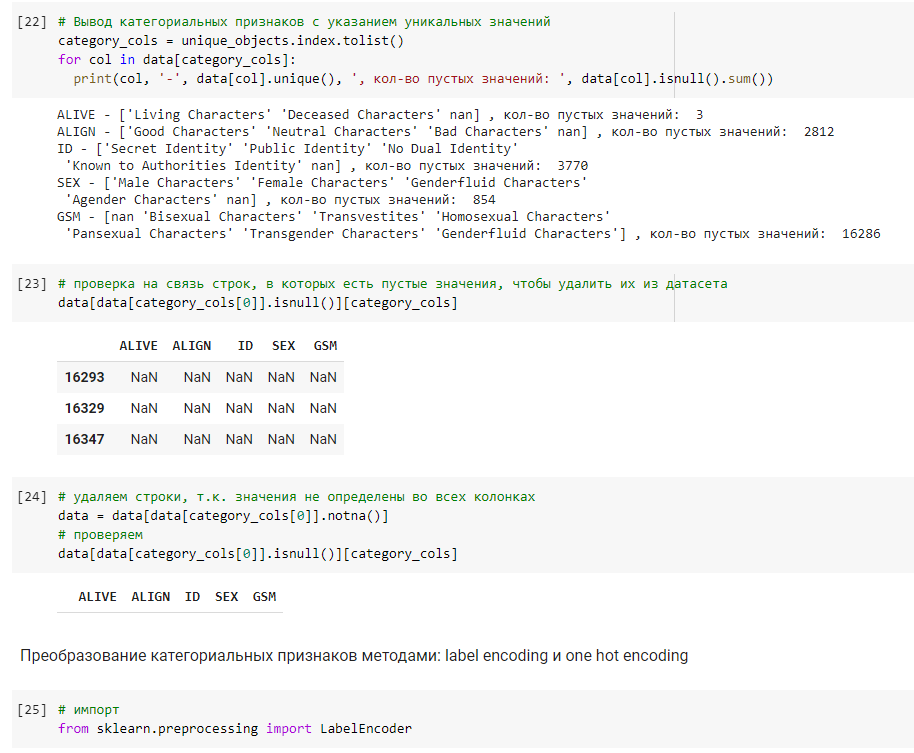
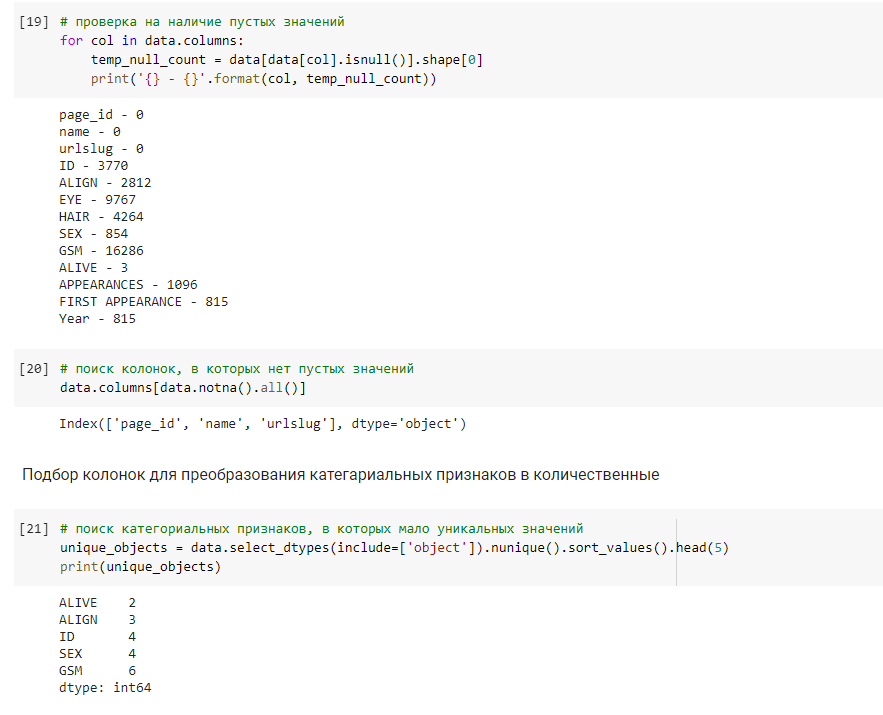
* Выбрать набор данных (датасет). Вы можете найти список свободно распространяемых датасетов [здесь.](https://github.com/ugapanyuk/ml_course_2021/wiki/DSLIST)
* Для первой лабораторной работы рекомендуется использовать датасет без пропусков в данных, например из [Scikit-learn.](https://scikit-learn.org/stable/datasets/toy_dataset.html)
* Пример преобразования датасетов Scikit-learn в Pandas Dataframe можно посмотреть [здесь.](https://github.com/ugapanyuk/ml_course_2021/blob/main/common/notebooks/ds/sklearn_datasets.ipynb)
* Создать ноутбук, который содержит следующие разделы:

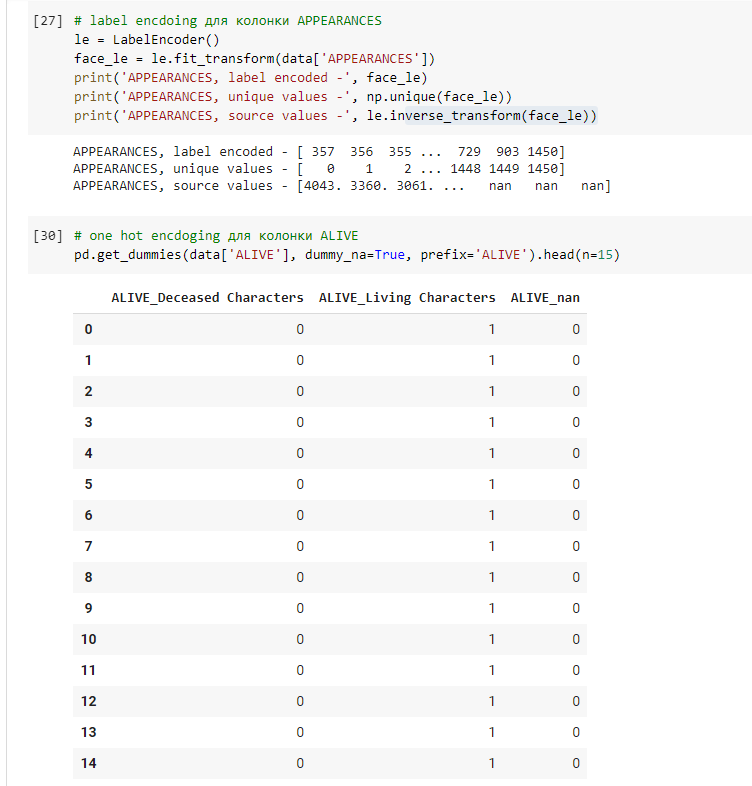
1. Текстовое описание выбранного Вами набора данных.
2. Основные характеристики датасета.
3. Визуальное исследование датасета.
4. Информация о корреляции признаков.

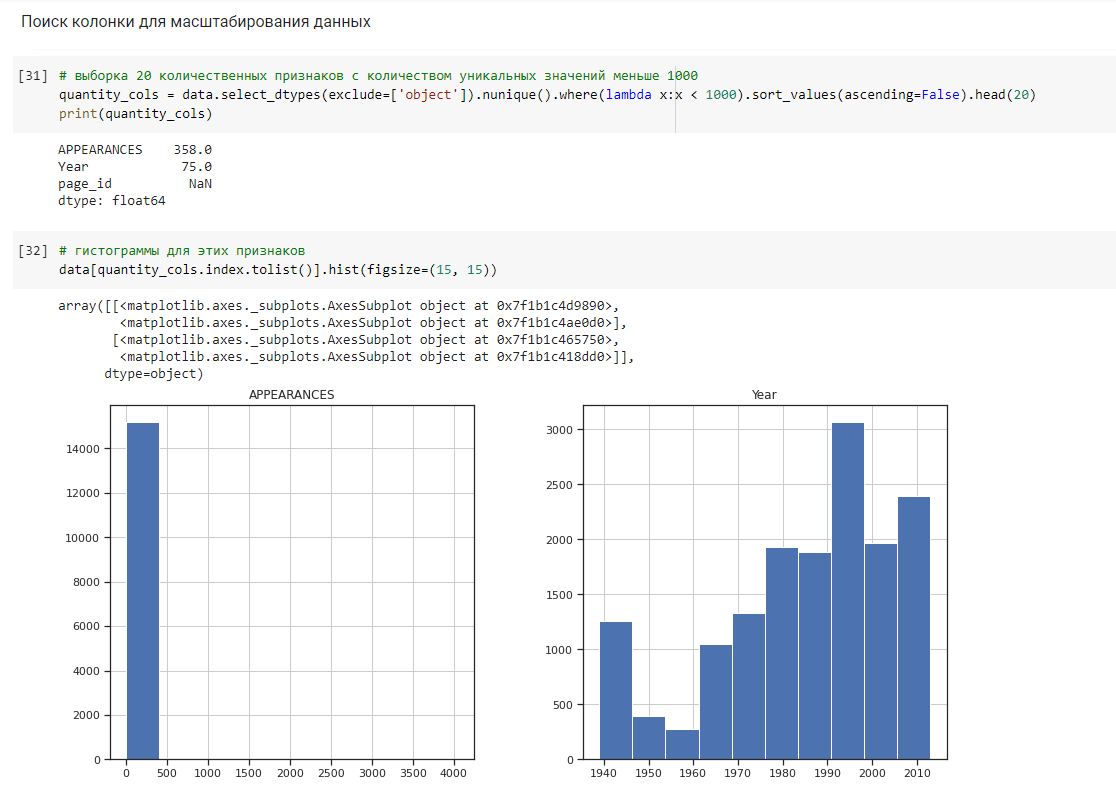
* Сформировать отчет и разместить его в своем репозитории на github.

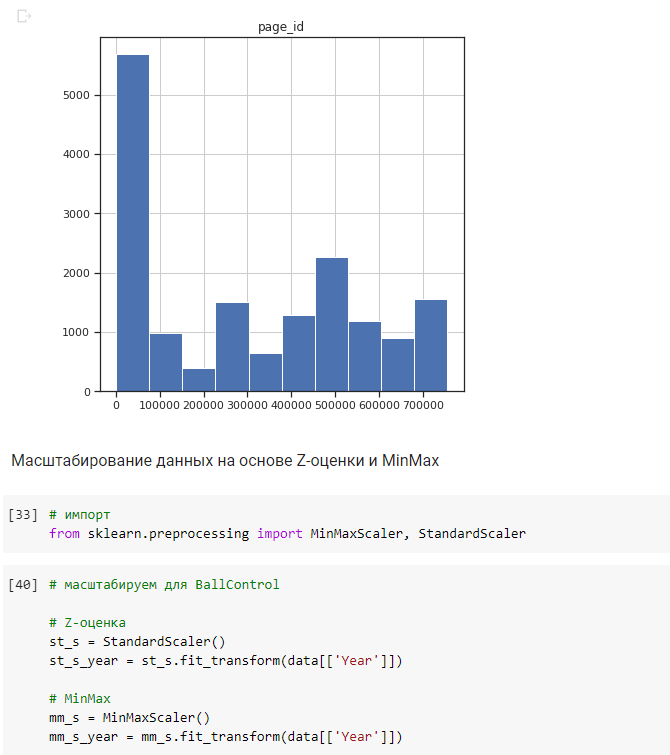
1. **Скрины jupyter notebook**

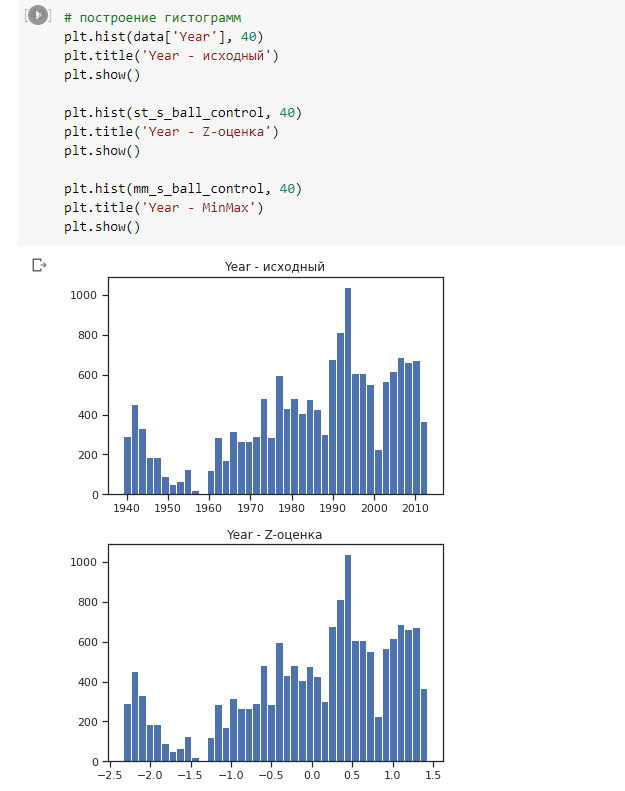


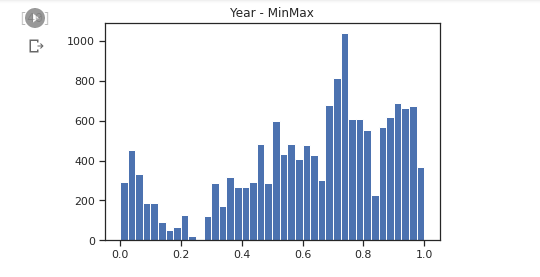


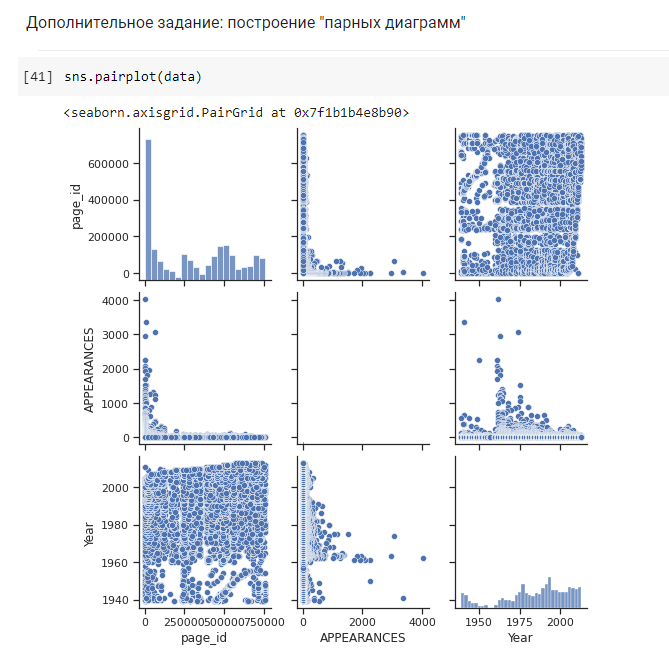












**Вывод**  
Масштабирование данных и преобразование категориальных признаков в количественные были проведены с учетом характеристик датасета. Для масштабирования были использованы методы MinMax и на основе Z-оценки. Для преобразования категориальных признаков в количественные были использованы методы label encoding и one hot encoding. Были подобраны признаки, которые наглядно продемонстрировали работу этих методов. Для реализации и визуализации были использованы функции из библиотек: pandas, sklearn, seaborn, matplotlib.