**Московский государственный технический   
университет им. Н. Э. Баумана**

Курс «Технологии машинного обучения»

Отчёт по лабораторной работе №1

Выполнил: Проверил:  
Здобняков Ф. А. Гапанюк Ю.Е.  
группа ИУ5-64Б

Дата: 07.04.25 Дата:

Подпись: Подпись:

Москва, 2025 г.

**Цель лабораторной работы:** изучение различных методов визуализация данных.

**Краткое описание.** Построение основных графиков, входящих в этап разведочного анализа данных.

Рекомендуемые инструментальные средства можно посмотреть [здесь](https://github.com/ugapanyuk/courses_current/wiki/IDE).

**Задание:**

* Выбрать набор данных (датасет). Вы можете найти список свободно распространяемых датасетов [здесь.](https://github.com/ugapanyuk/courses_current/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/courses_current/blob/main/notebooks/ds/sklearn_datasets.ipynb)

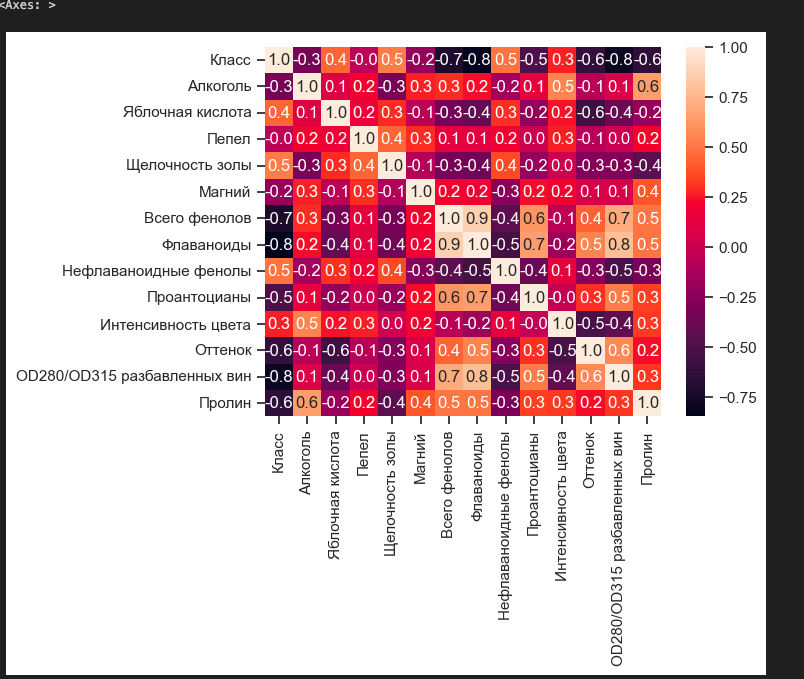
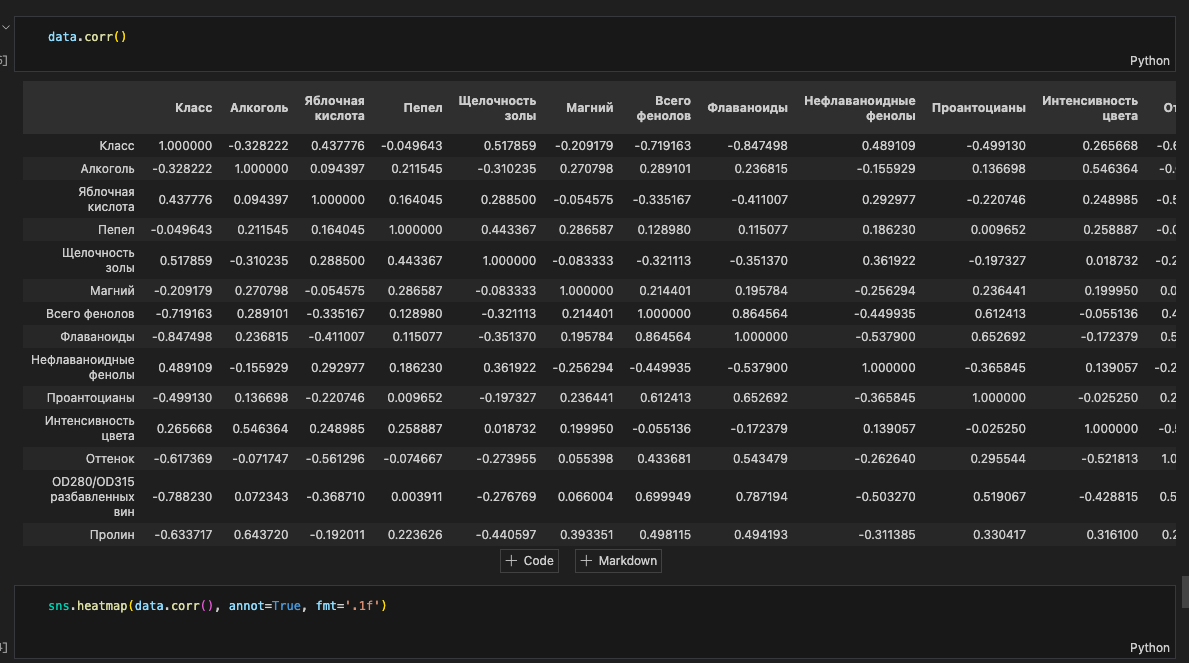
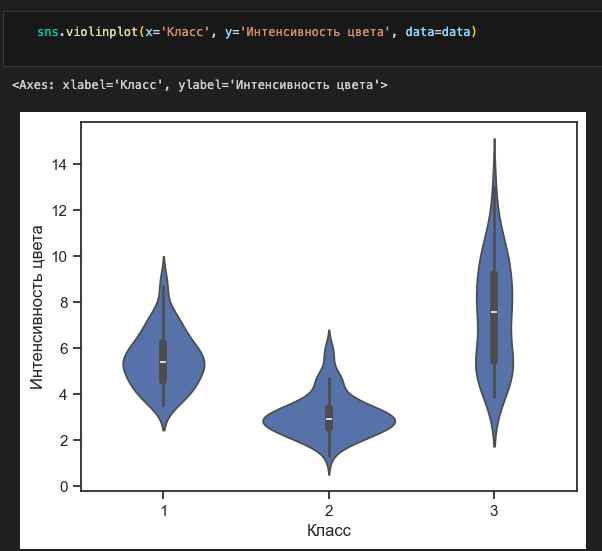
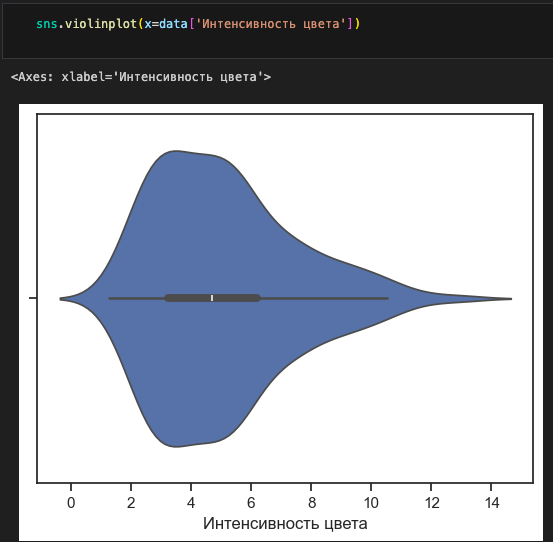
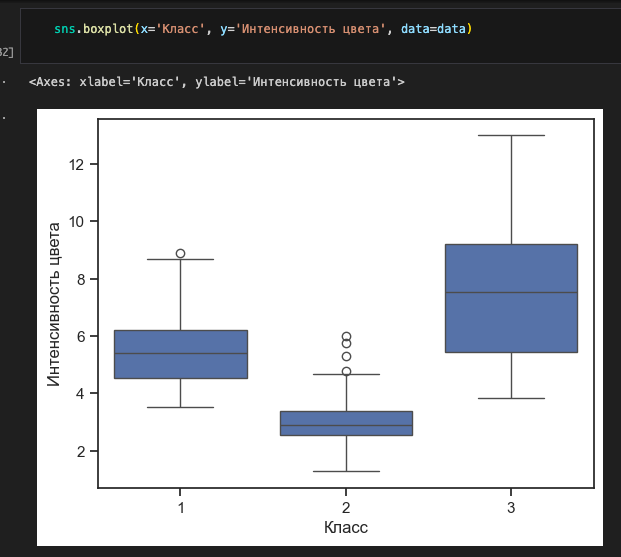
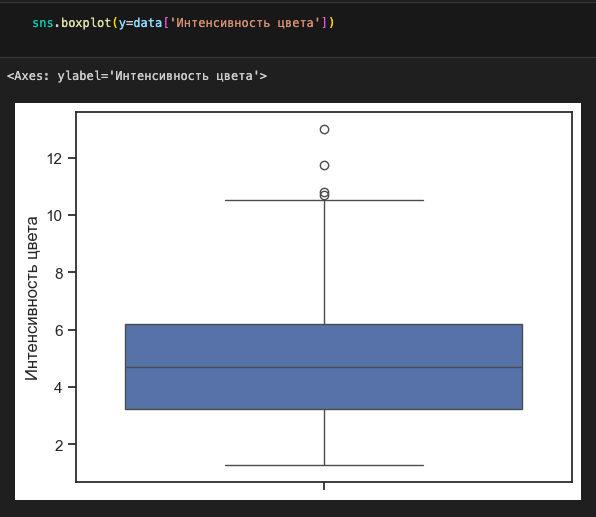
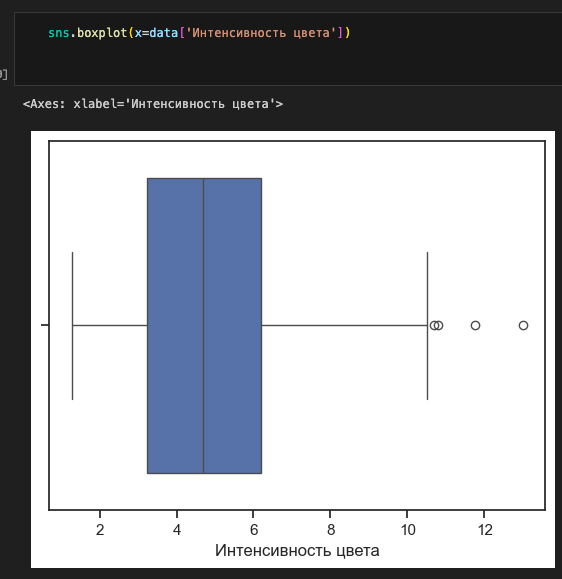
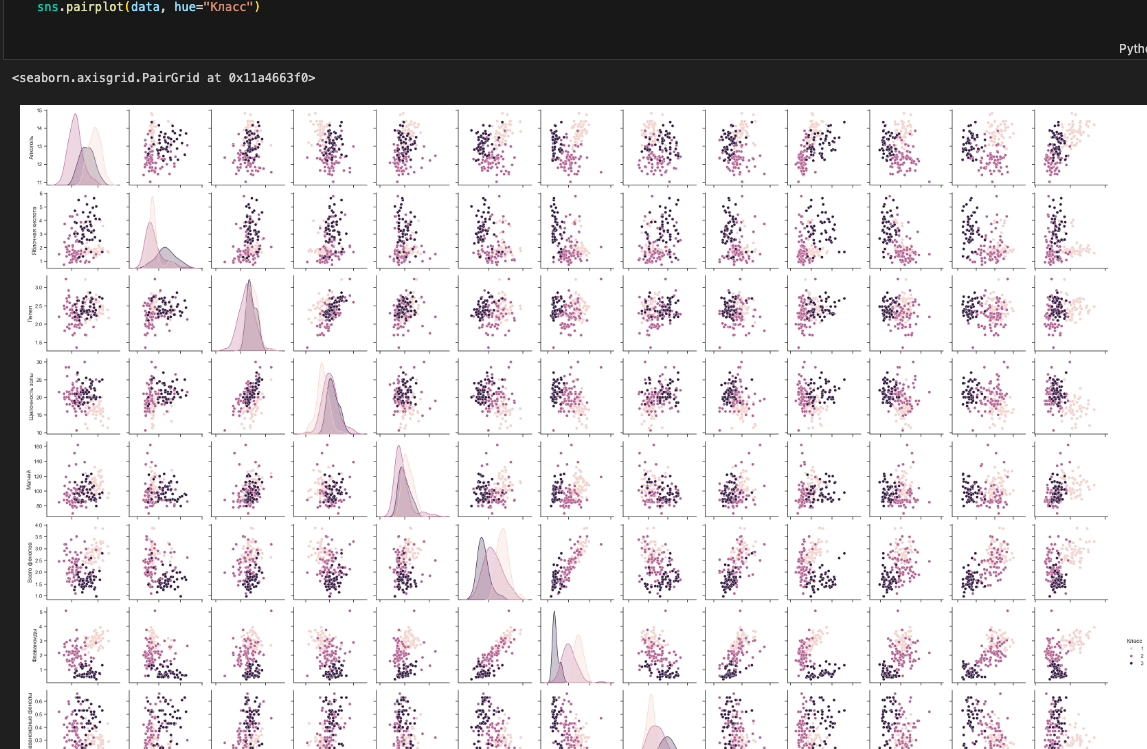
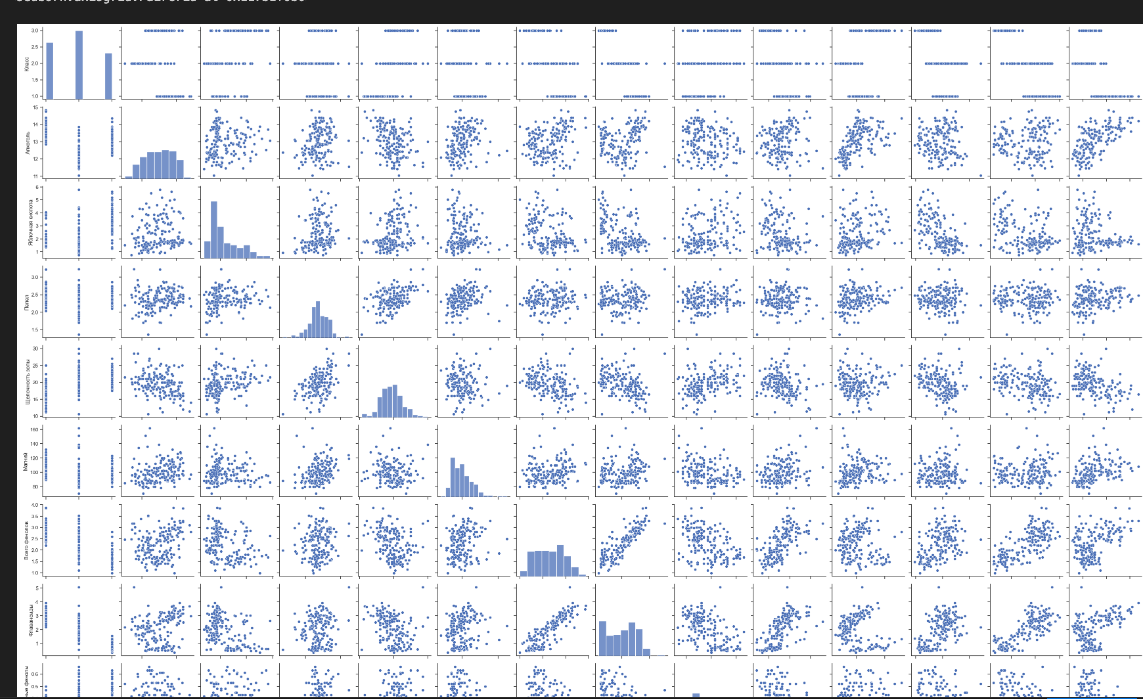
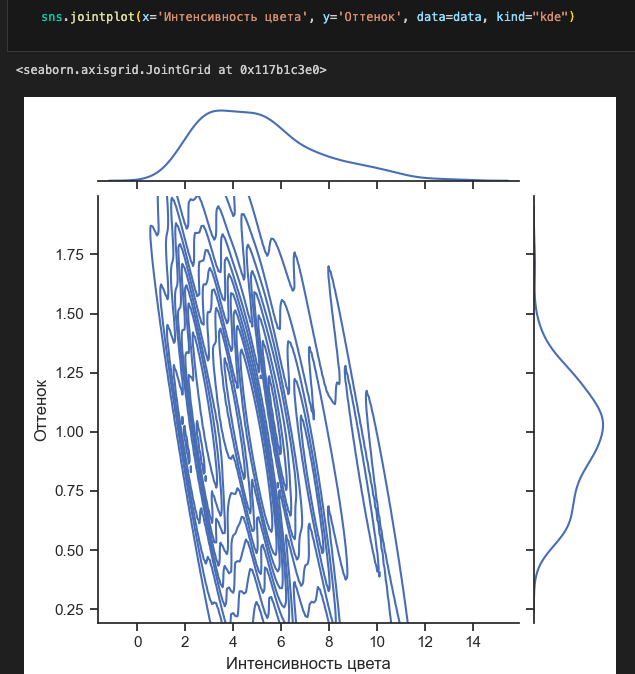
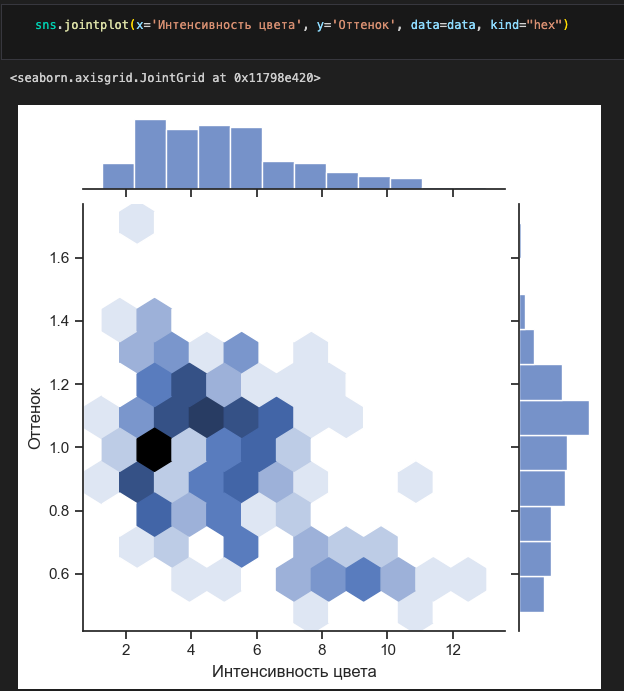
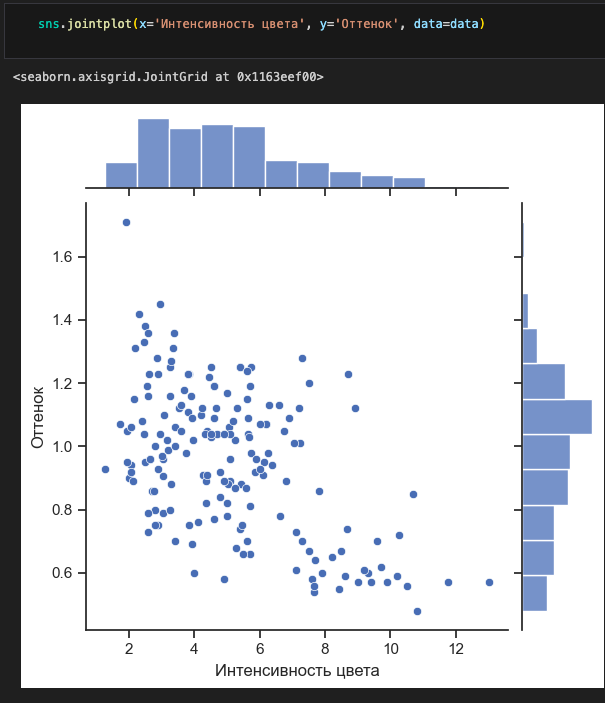
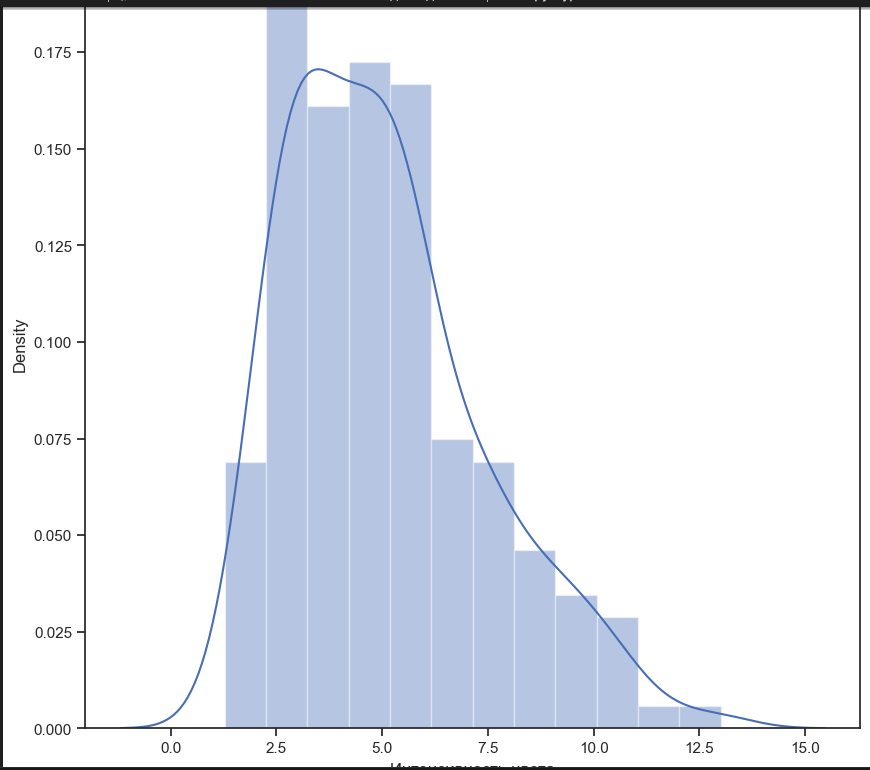
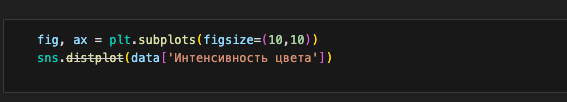
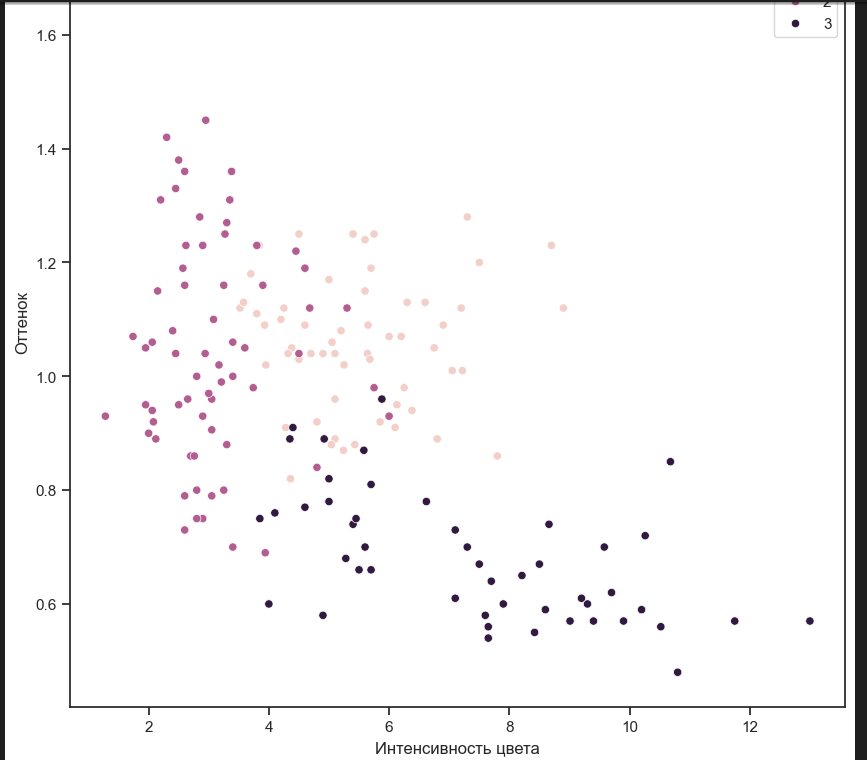
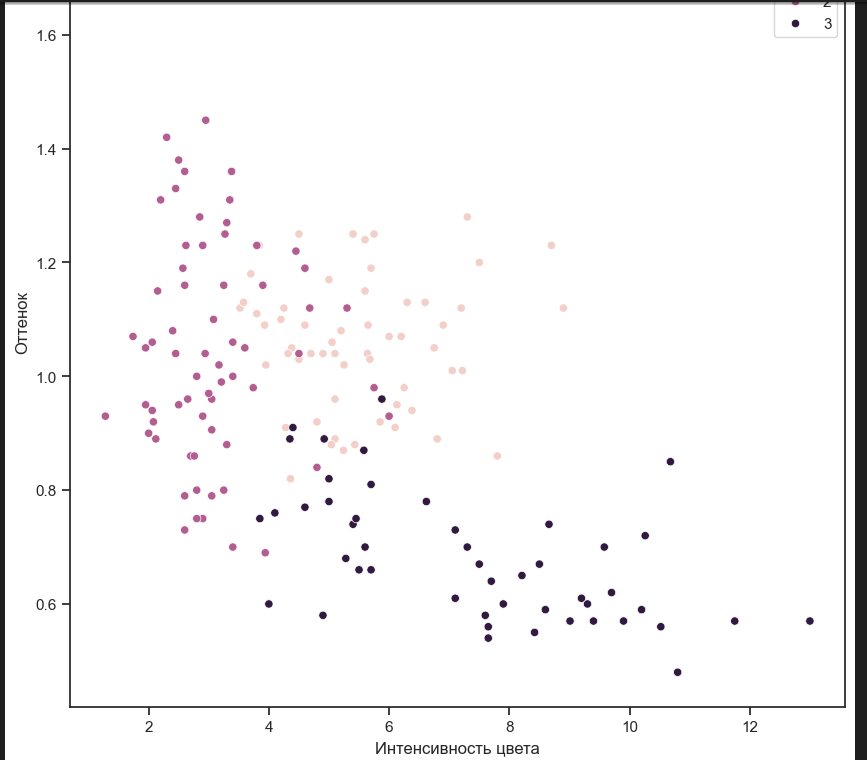
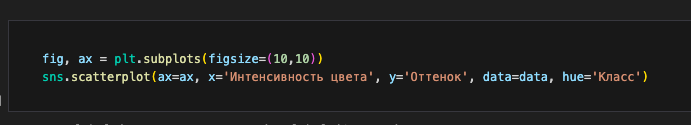
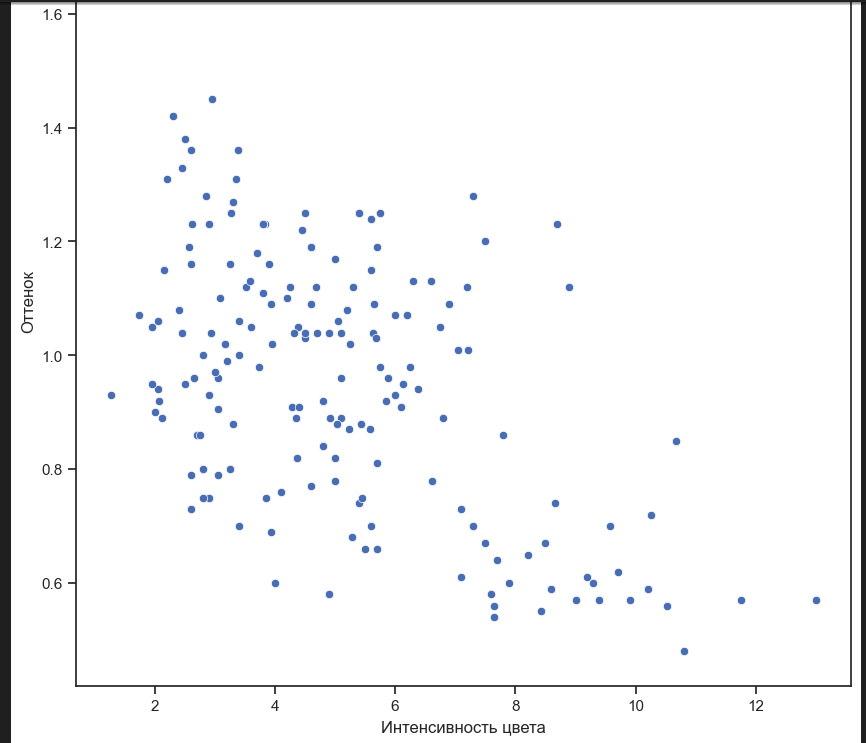
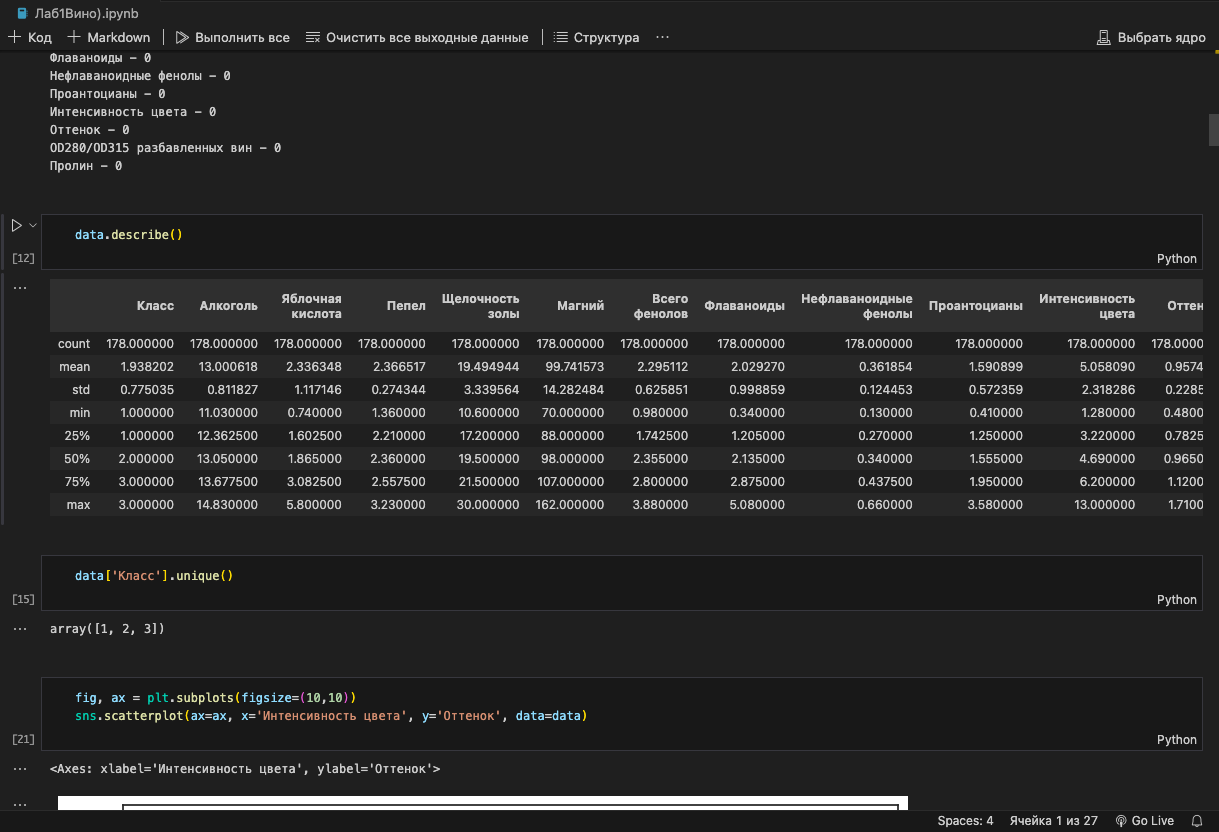
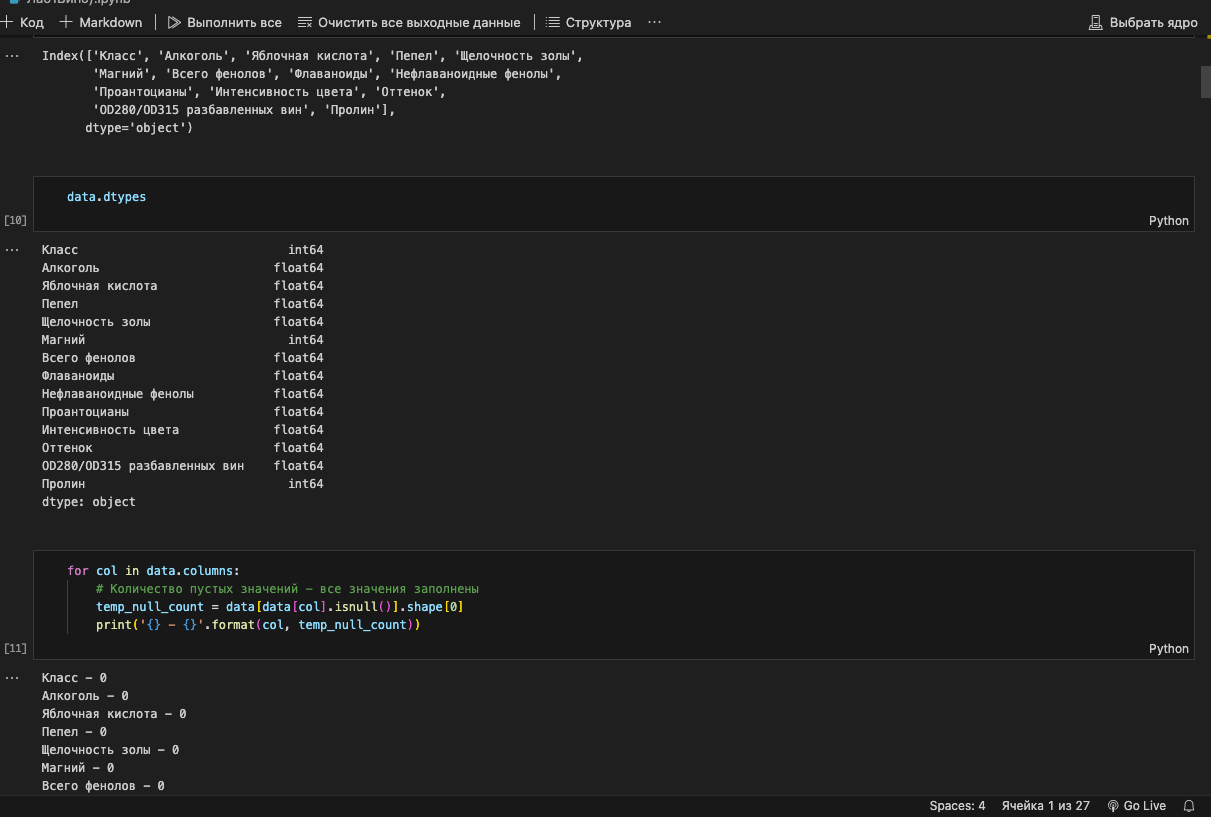
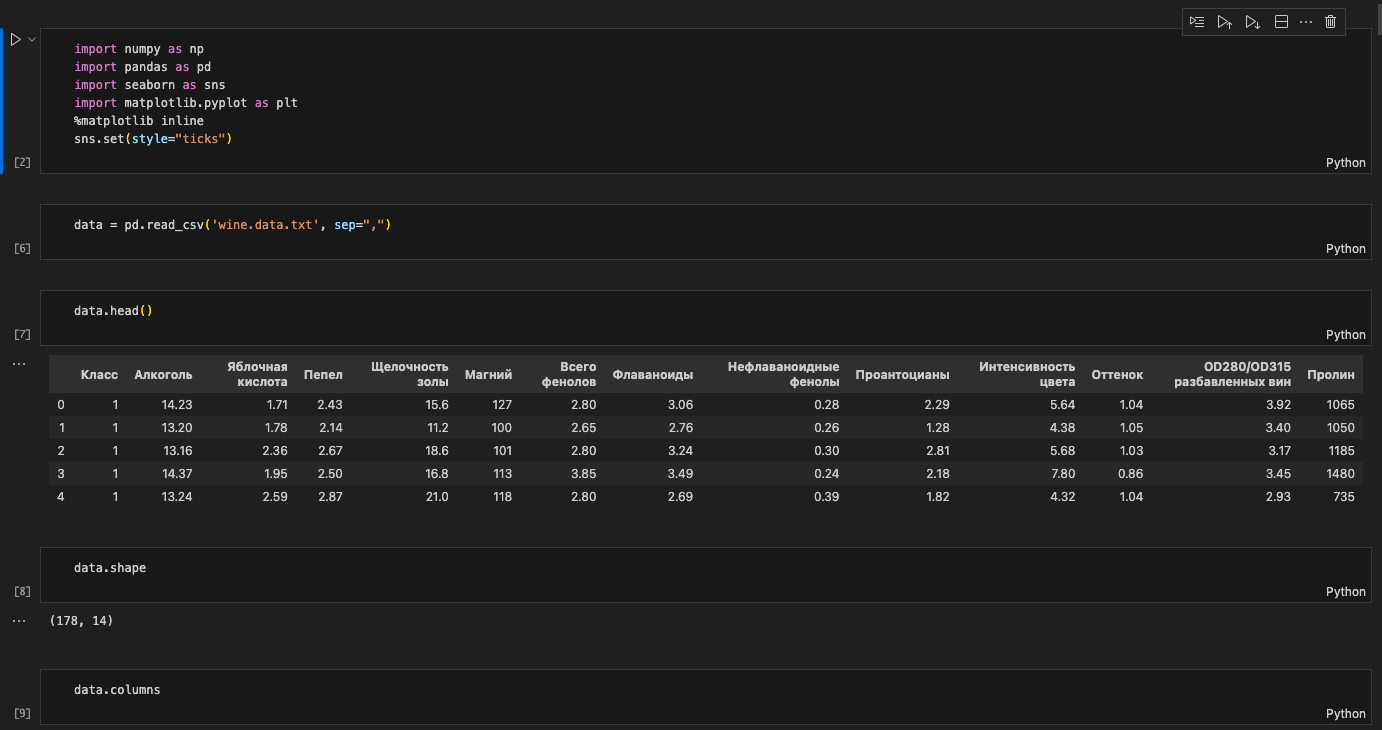
Для лабораторных работ не рекомендуется выбирать датасеты большого размера.

* 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.

**Ход выполнения:**

****