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

Курс «Технологии машинного обучения»

Отчёт по рубежному контролю №1

«Технологии разведочного анализа и обработки данных.»

Вариант № 23

Выполнил: Проверил:  
Падалко К.Р. Гапанюк Ю.Е.  
группа ИУ5Ц-84Б

Дата: 12.03.25 Дата:

Подпись: Подпись:

Москва, 2025 г.

**Задание:**

Номер варианта: **23**

Номер задачи: **3**

Номер набора данных, указанного в задаче: **7** (https://www.kaggle.com/datasets/lava18/google-play-store-apps)

Для студентов групп ИУ5-64Б, ИУ5Ц-84Б - для произвольной колонки данных построить график "Скрипичная диаграмма (violin plot).

**Задача №3.**

Для заданного набора данных произведите масштабирование данных (для одного признака) и преобразование категориальных признаков в количественные двумя способами (label encoding, one hot encoding) для одного признака. Какие методы Вы использовали для решения задачи и почему?

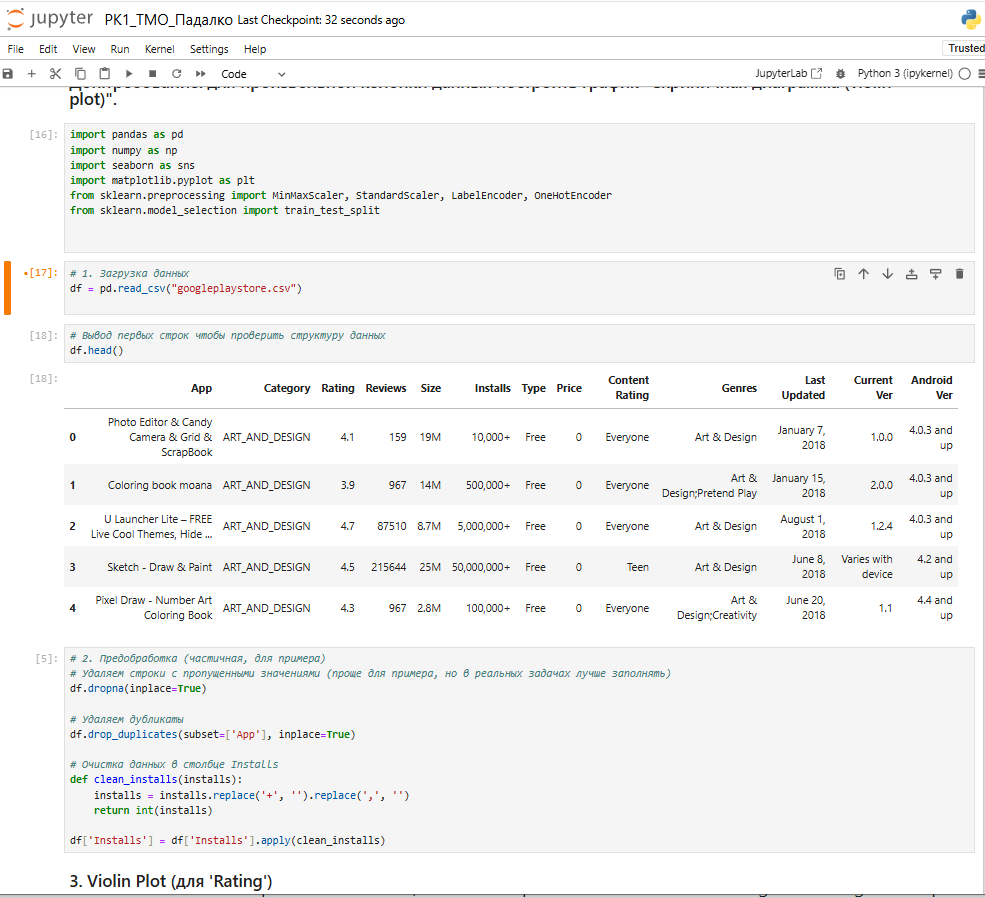
**1. Введение**

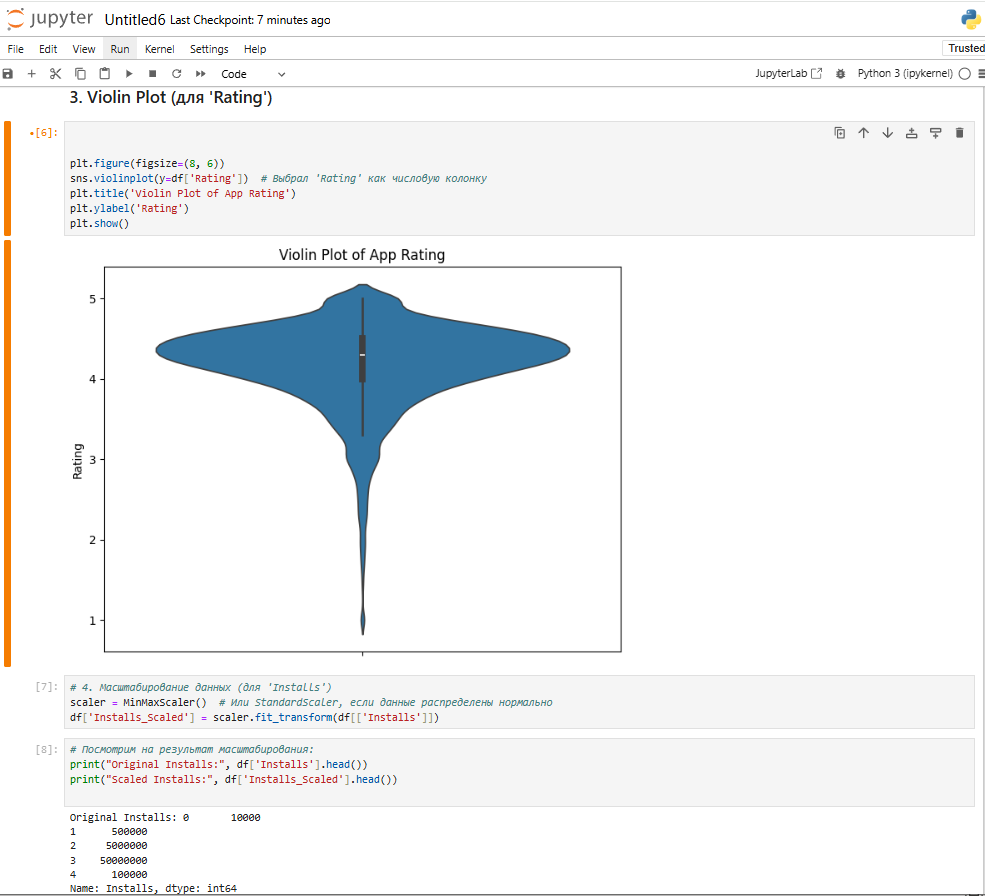
В рамках рубежного контроля была проведена работа с набором данных Google Play Store Apps. Целью работы являлось масштабирование числового признака, преобразование категориального признака в количественный формат двумя методами (Label Encoding и One-Hot Encoding), а также построение скрипичной диаграммы (violin plot) для визуализации распределения данных.

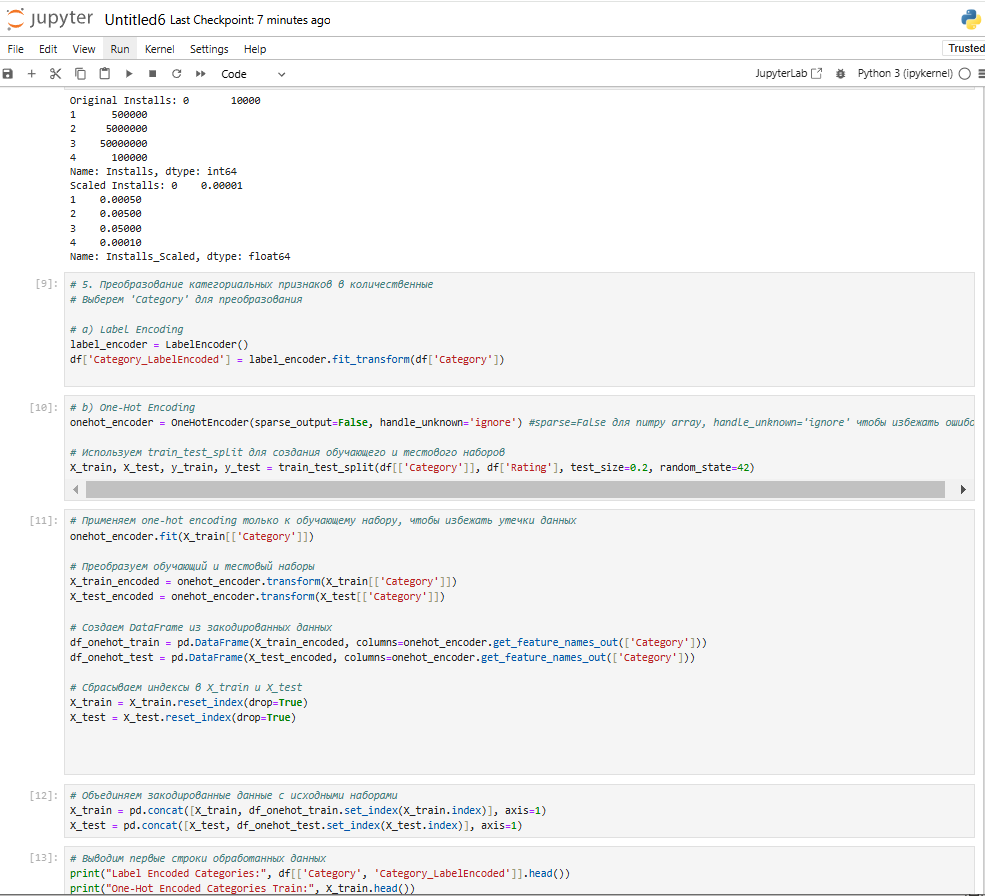
### ****2. Описание исходных данных****

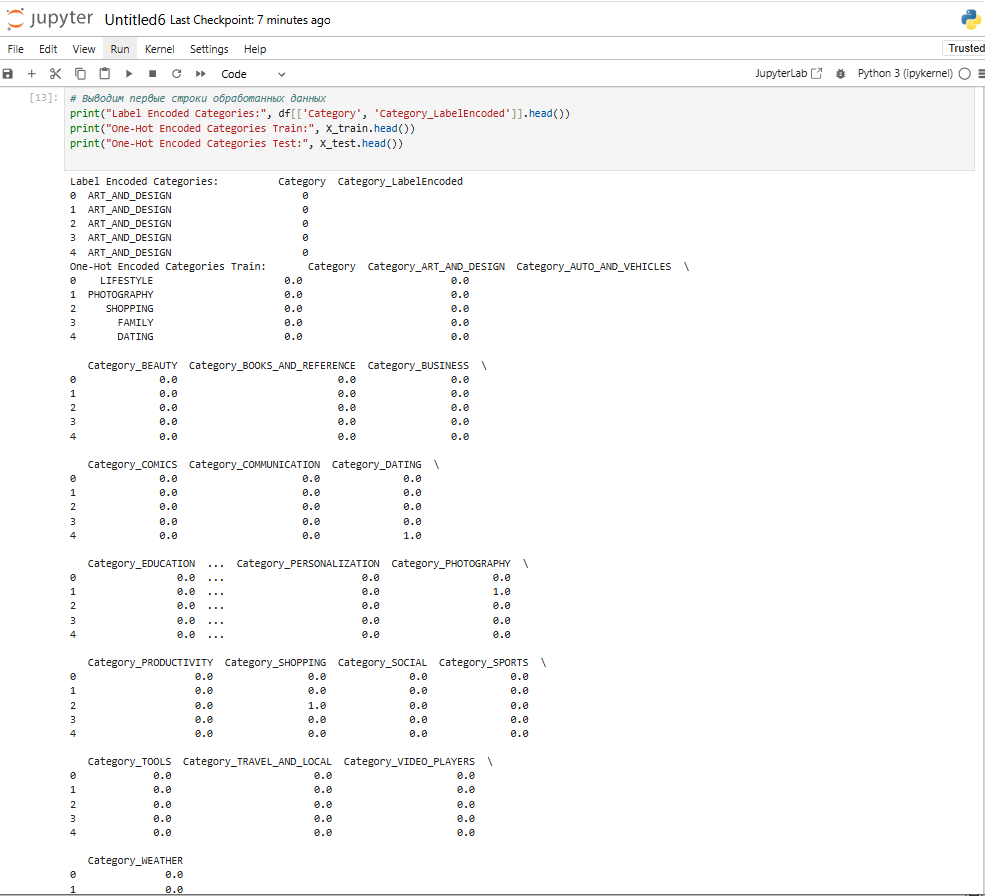
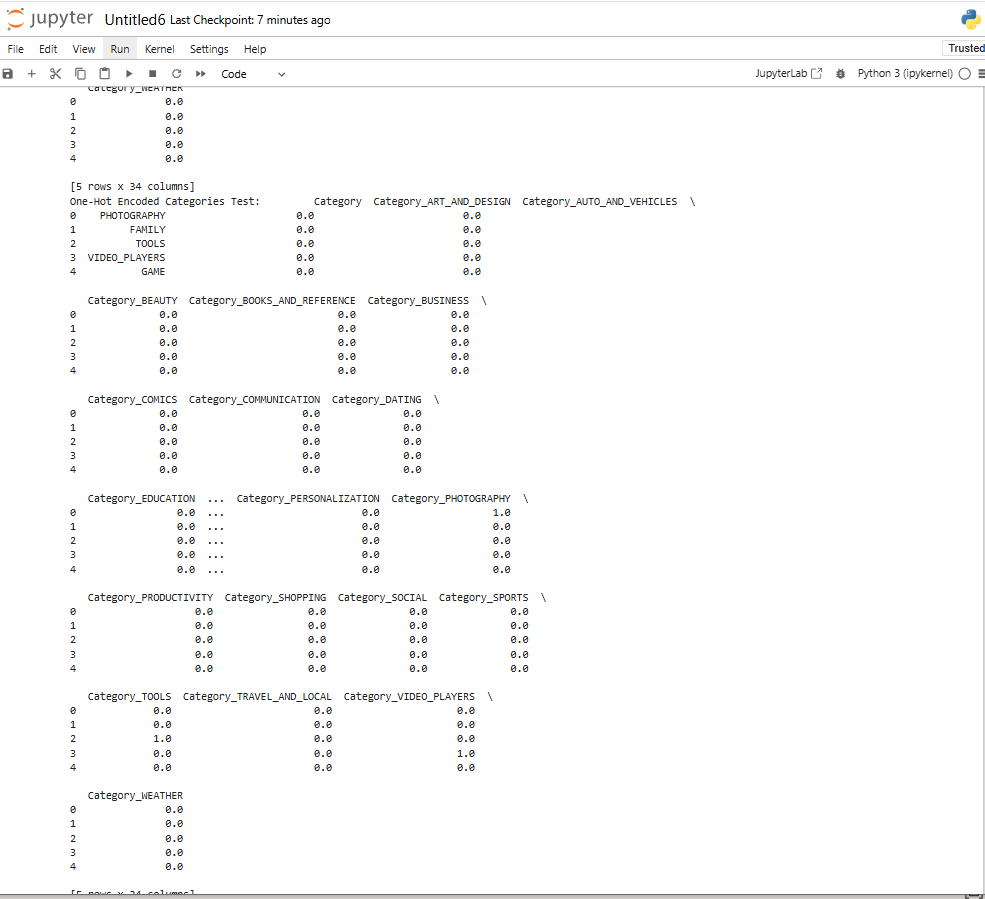
Исходный датасет содержит информацию о приложениях в Google Play Store, включая их категорию, рейтинг, количество отзывов, размер, количество установок и другие параметры. В данных присутствуют пропущенные значения и текстовые форматы чисел, требующие обработки.

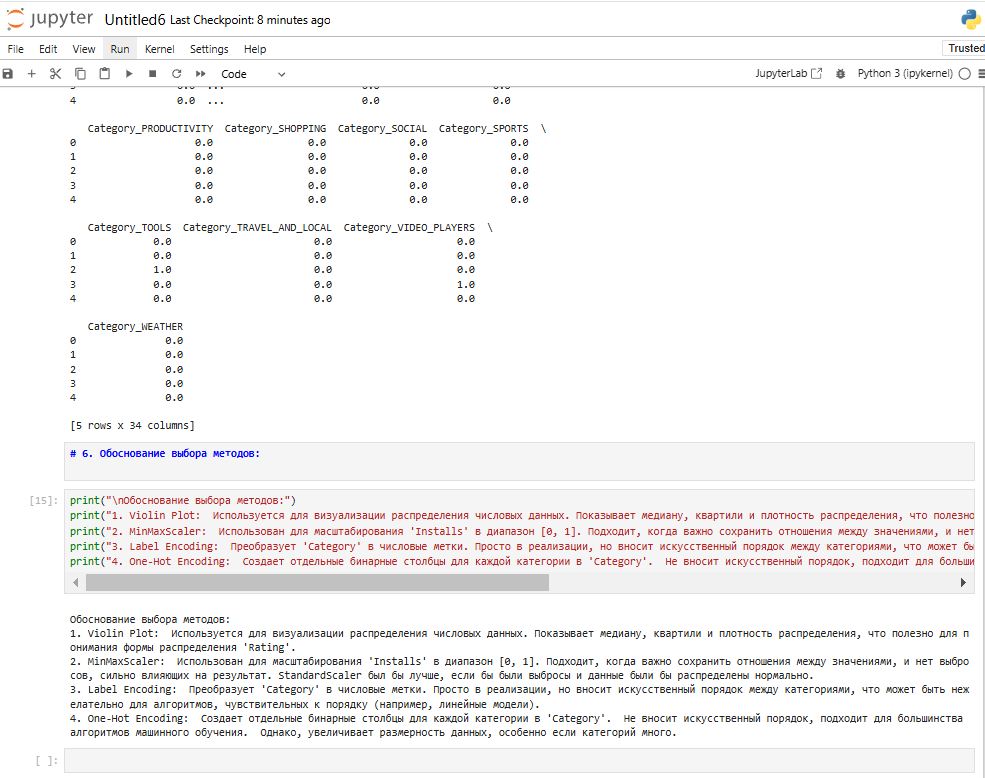
**3. Ход выполнения:**









### ****4. Выводы****

В ходе работы были успешно выполнены задачи по обработке данных, масштабированию числового признака, кодированию категориальных данных и визуализации. Обнаружены зависимости между категориями приложений и их рейтингами. Обработаны пропущенные значения и текстовые числовые данные.