**Министерство науки и высшего образования**

**Российской Федерации** **Федеральное государственное бюджетное образовательное учреждение высшего**

**образования** **«Московский государственный**

**технический университет** **имени Н.Э. Баумана**

**(национальный исследовательский университет)»**

**(МГТУ им. Н.Э. Баумана)**

**Факультет «Информатика и системы управления»**

**Кафедра ИУ5 «Системы обработки информации и управления»**

Отчёт по РК1

«Технологии машинного обучения»

Вариант 7

Выполнил:

студент группы ИУ5-63Б

Климов Н.С.

Преподаватель:

Гапанюк Ю. Е.

2024 г.

Задание:

Задача №1.

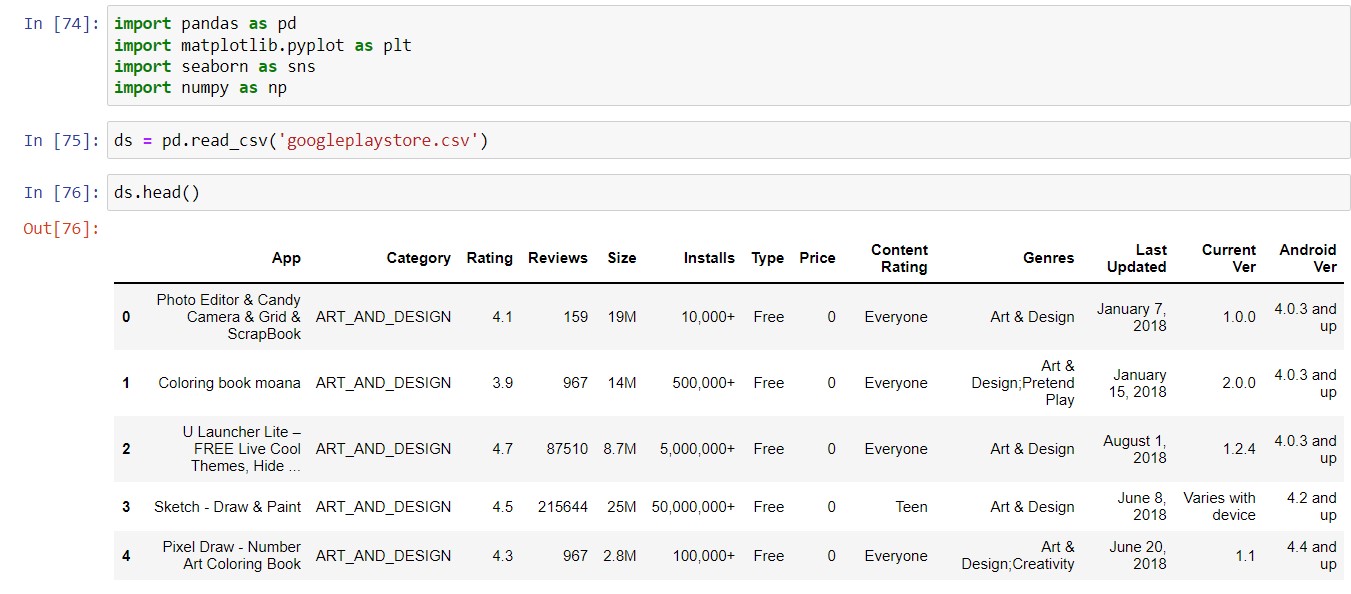
Для заданного набора данных проведите корреляционный анализ. В случае наличия пропусков в данных удалите строки или колонки, содержащие пропуски. Сделайте выводы о возможности построения моделей машинного обучения и о возможном вкладе признаков в модель.

Для студентов групп ИУ5-63Б, ИУ5Ц-83Б - для произвольной колонки данных построить график "Ящик с усами (boxplot)".

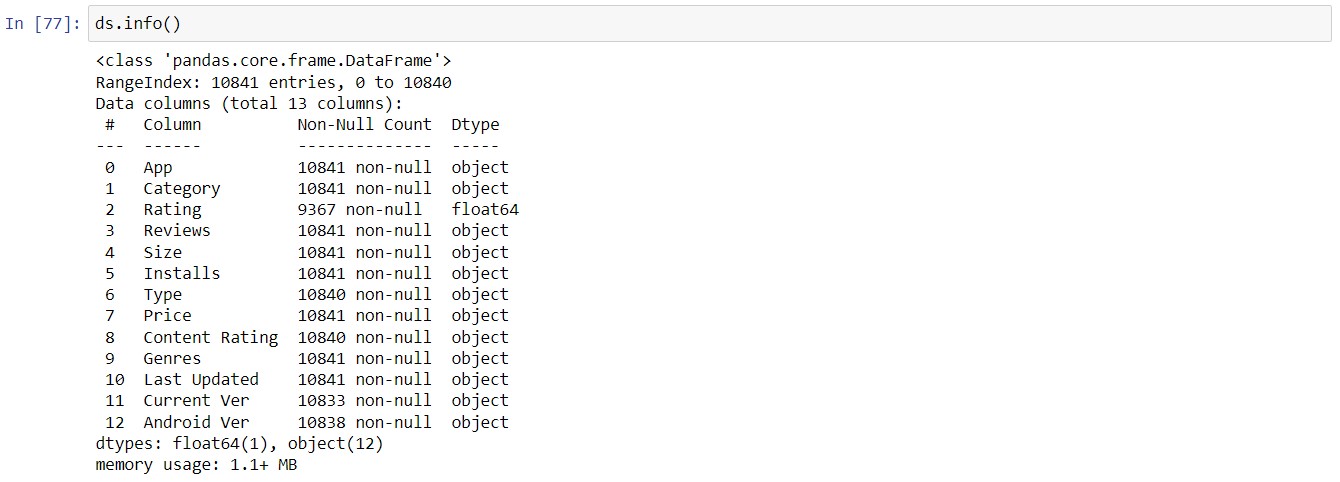
<https://www.kaggle.com/lava18/google-play-store-apps>

Решение:

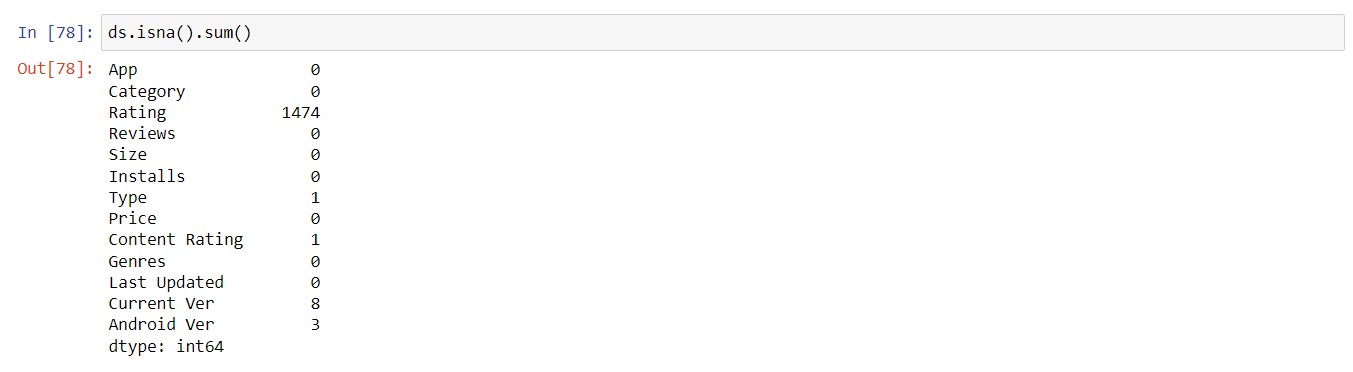
Загружаем датасет и подключаем необходимые библиотеки:



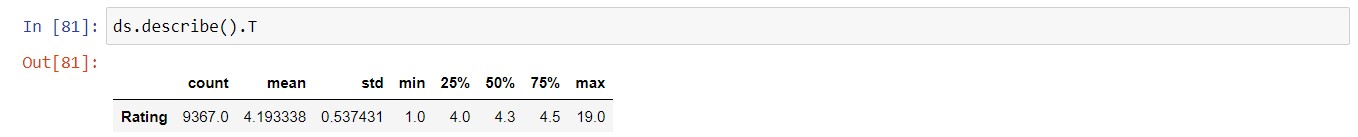
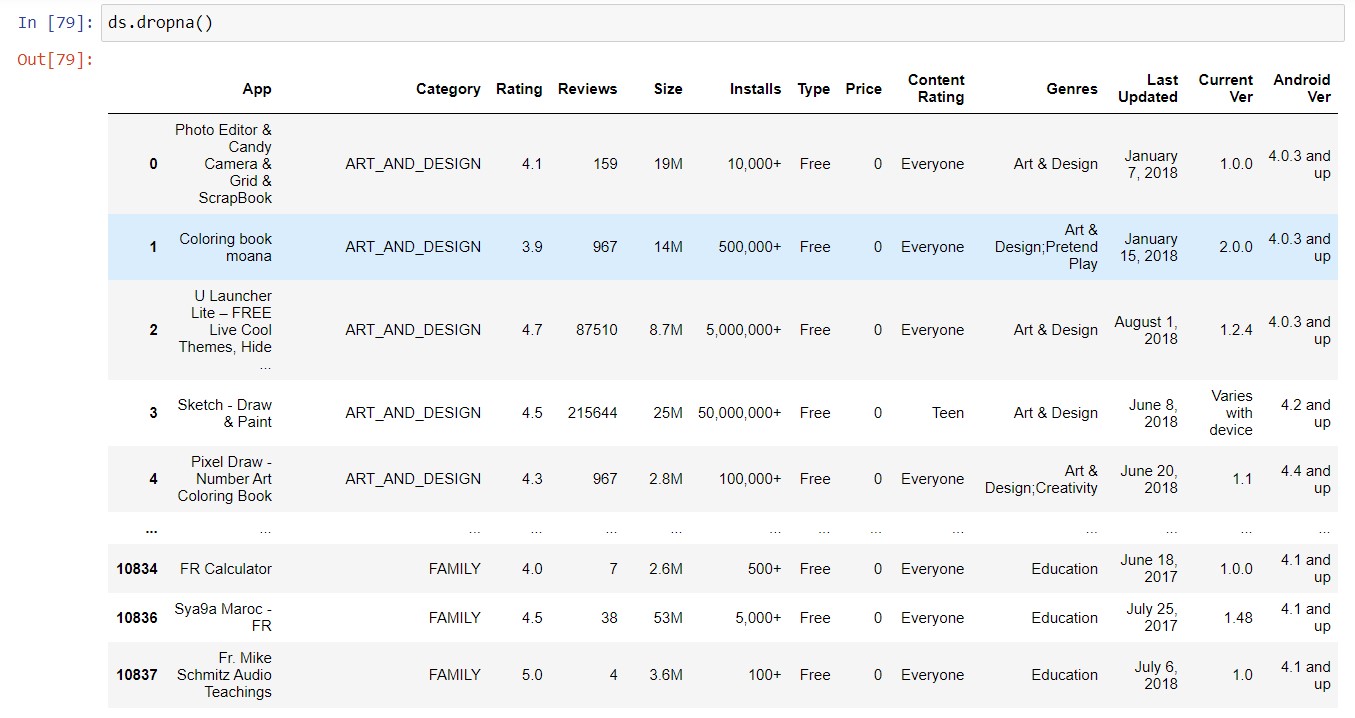
Получим информацию о датасете:



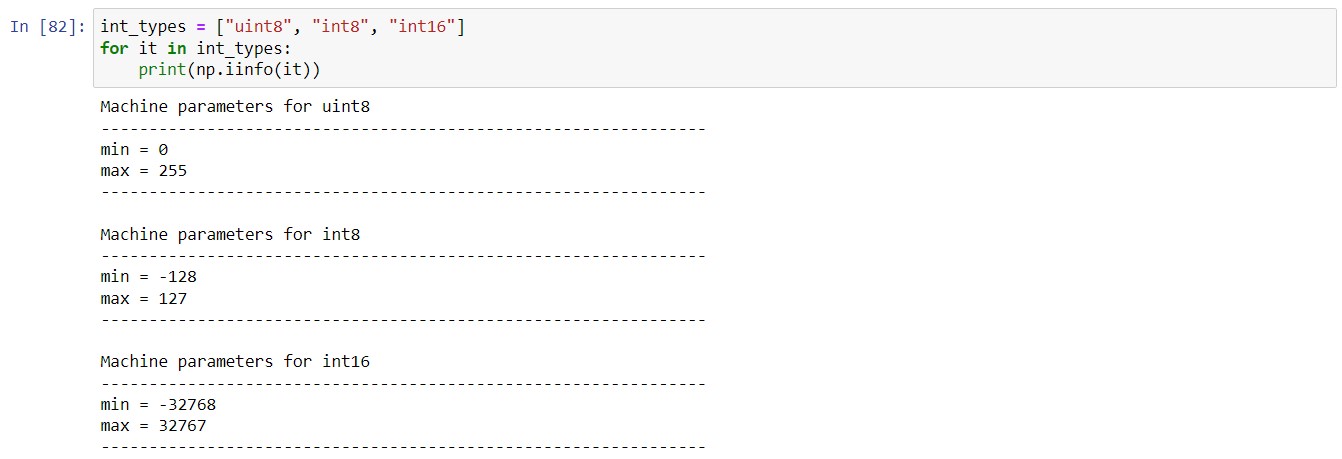
Посчитаем количество пустых значений:

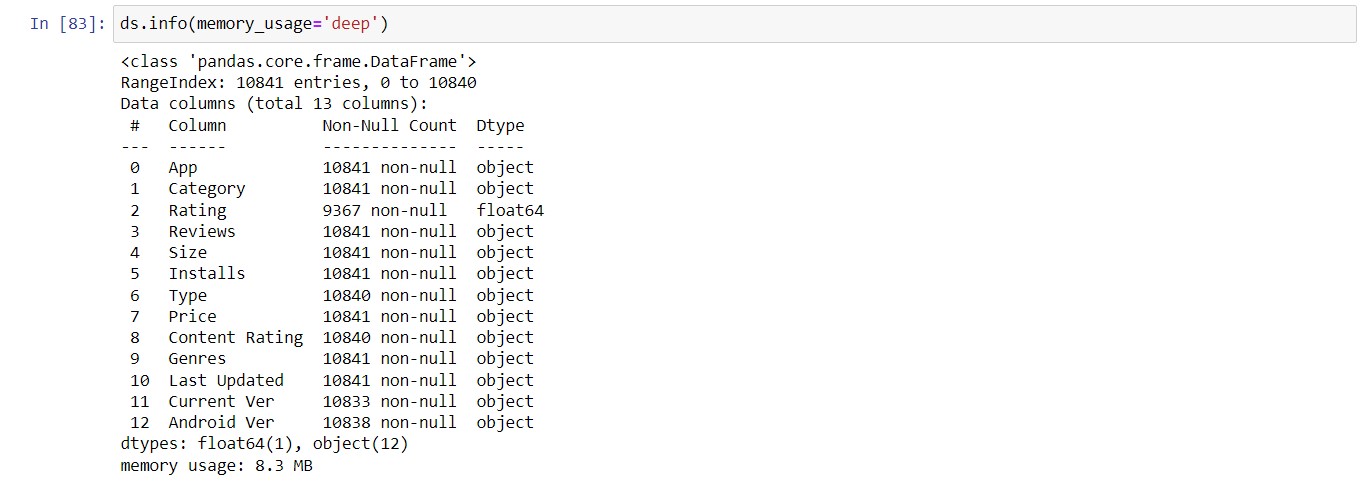


Удалим строки с пустыми значениями:

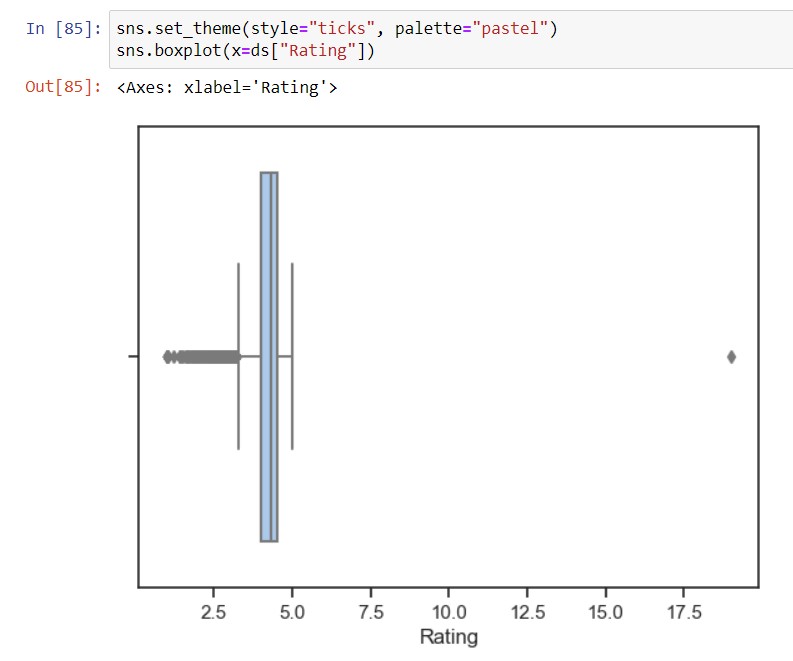


Оптимизируем память и изменим типы данных на более корректные:





По колонке “Rating” сделаем boxplot (Ящик с усами):



Корреляционный анализ:

Данный датасет нельзя использовать для корреляционного анализа, так как он почти полностью состоит из неподходящих типов данных. Значит, построение такой модели машинного обучения невозможно.