Московский государственный технический университет им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Автоматизированные системы обработки информации и управления»

**“Методы машинного обучения”**

**Отчет по лабораторной работе №1**

**“Разведочный анализ данных. Исследование и визуализация данных”**

Выполнил:

Буклин С.В.

Группа ИУ5-21м

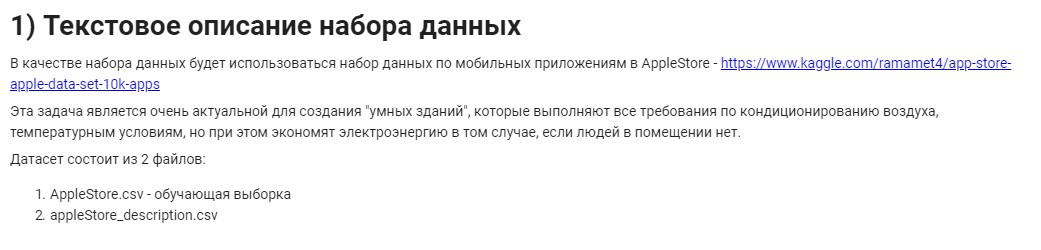
**Москва 2018**

1. **Задание**

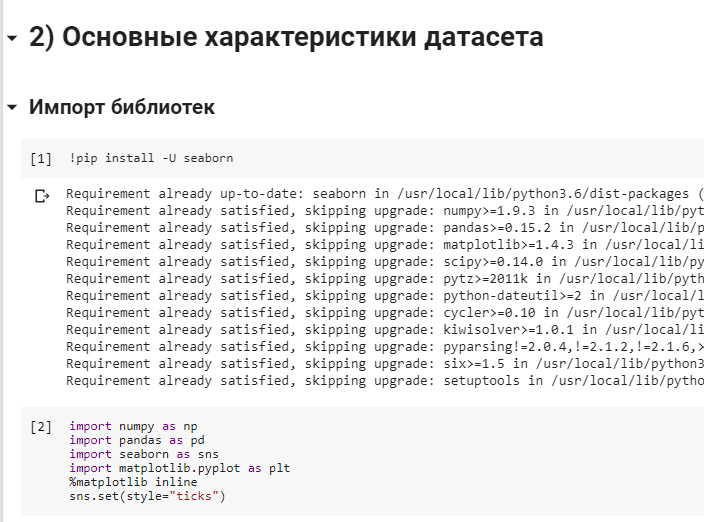
* Выбрать набор данных (датасет).
* Создать ноутбук, который содержит следующие разделы:
  + Текстовое описание выбранного Вами набора данных.
  + Основные характеристики датасета.
  + Визуальное исследование датасета.
  + Информация о корреляции признаков.
* Сформировать отчет и разместить его в своем репозитории на github.

1. **Текст программы и экранные формы**

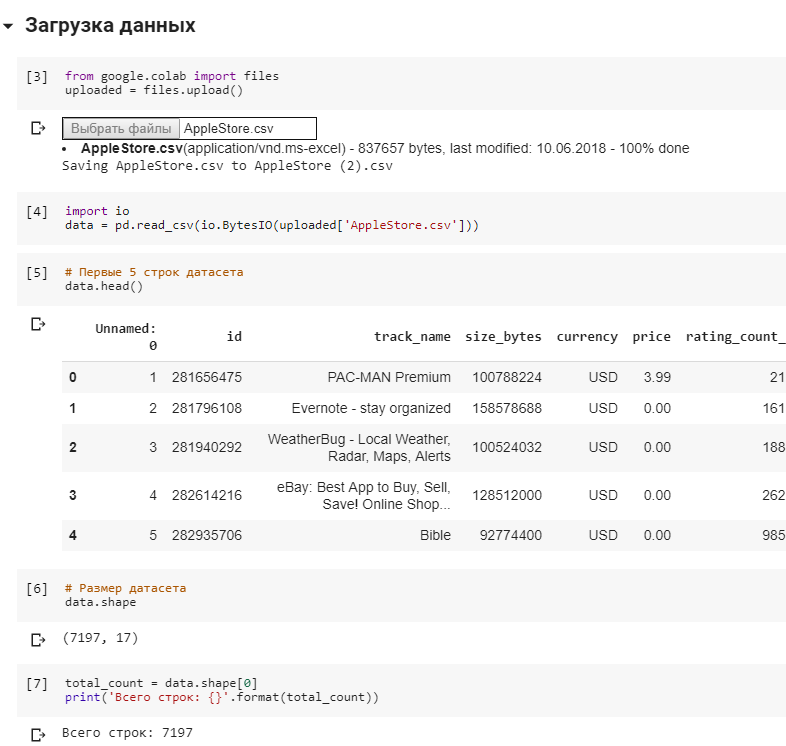
Описание набора данных

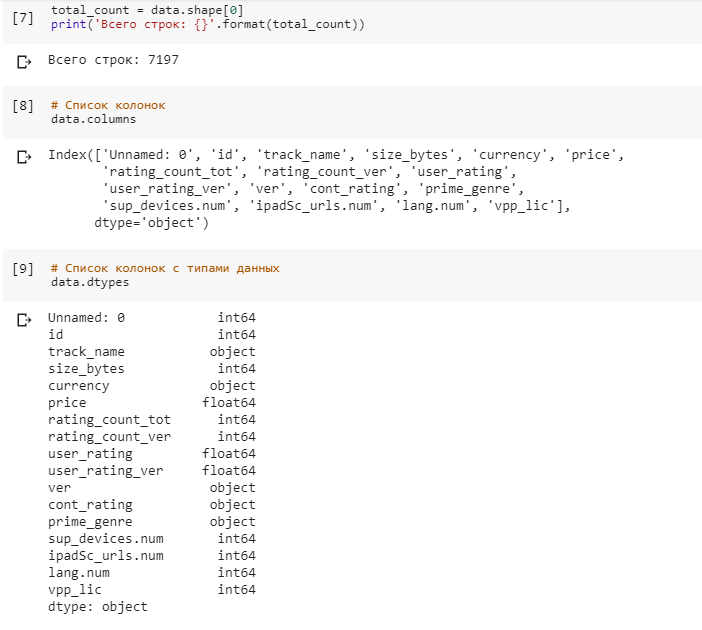
****

Импорт библиотек

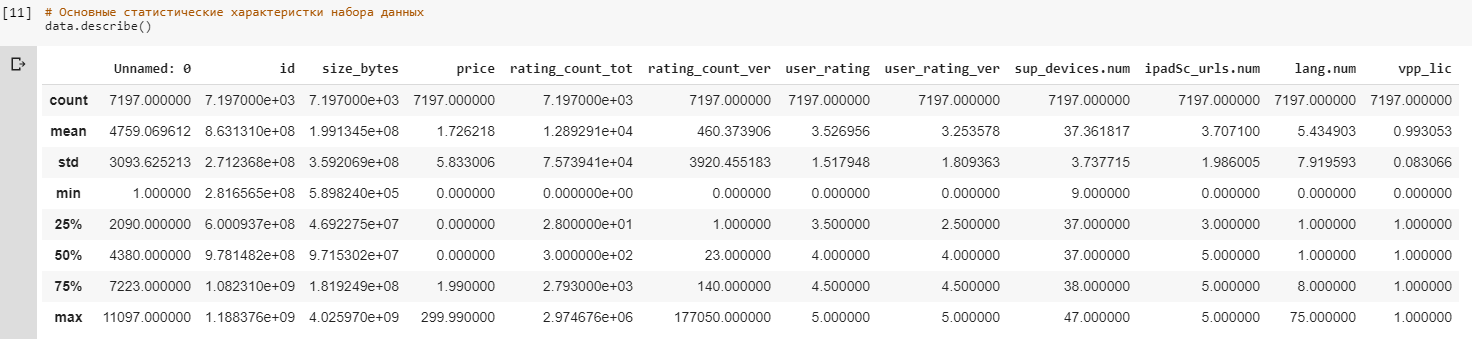


Загрузка данных



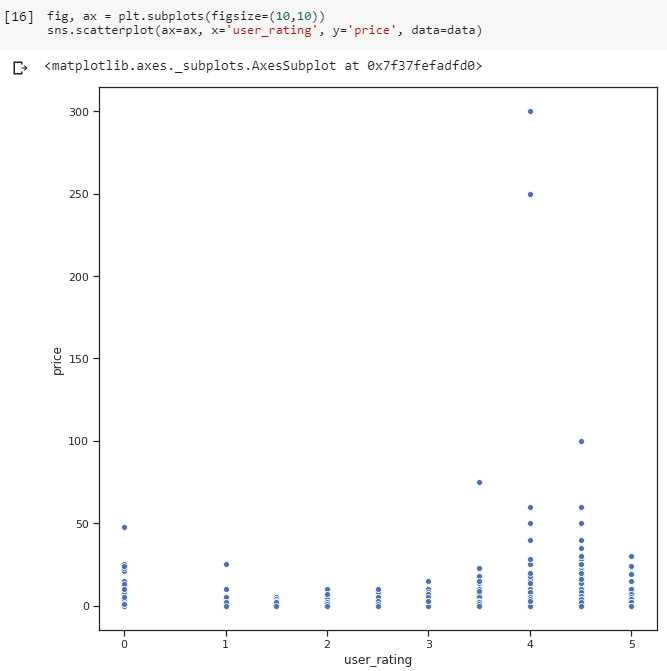
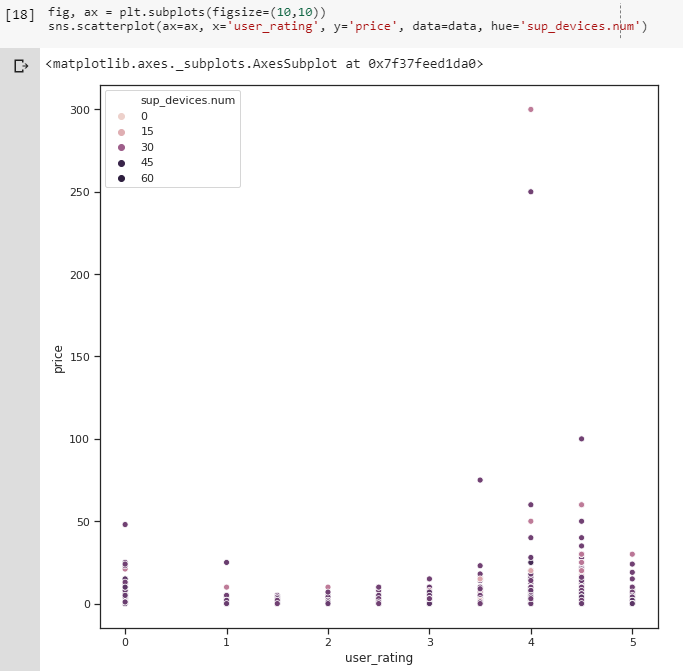


Основные статистические характеристики датасета

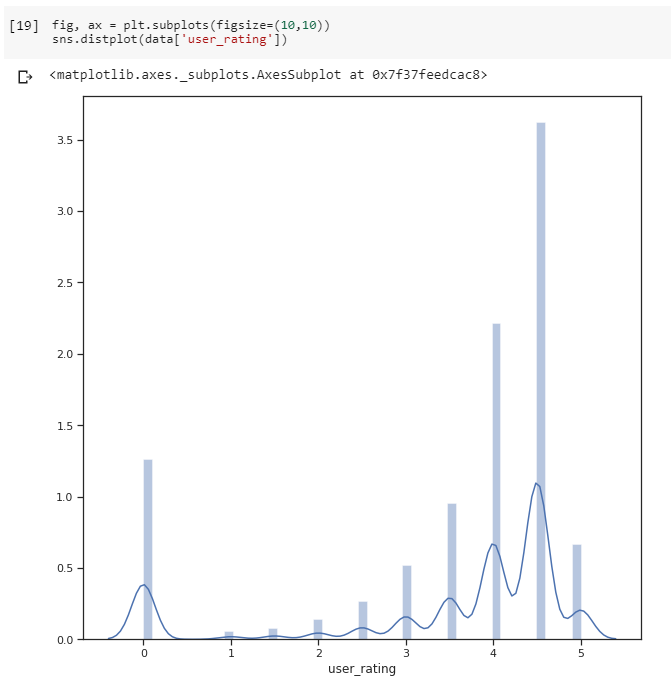


**Визуальное исследование датасета**

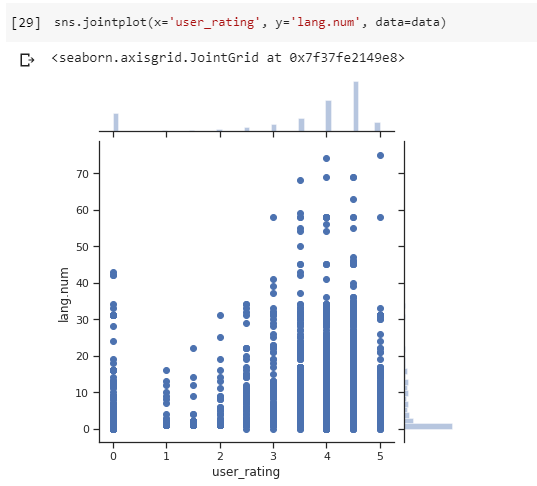
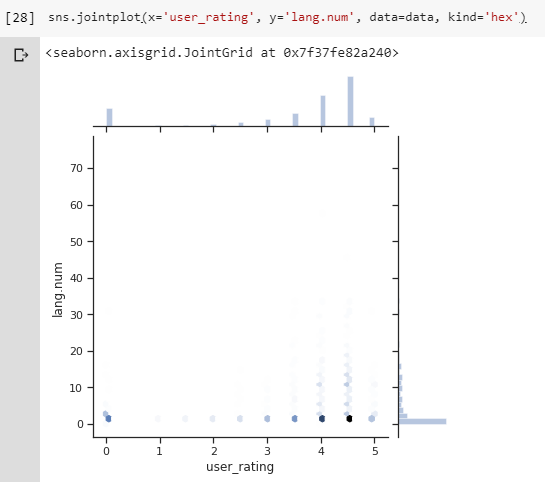
Для визуального исследования могут быть использованы различные виды диаграмм, мы построим только некоторые варианты диаграмм, которые используются достаточно часто.

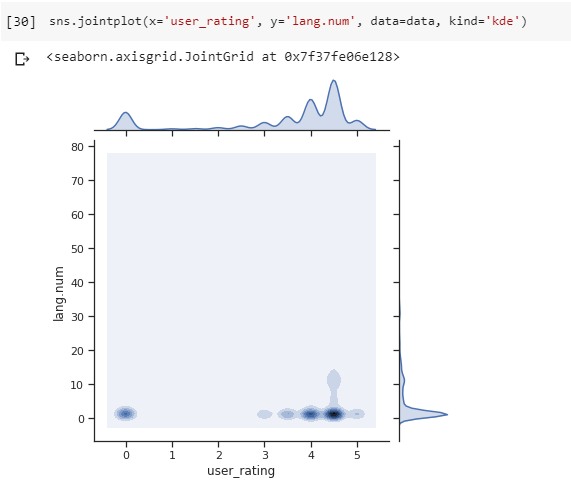
 

**Гистограммы**

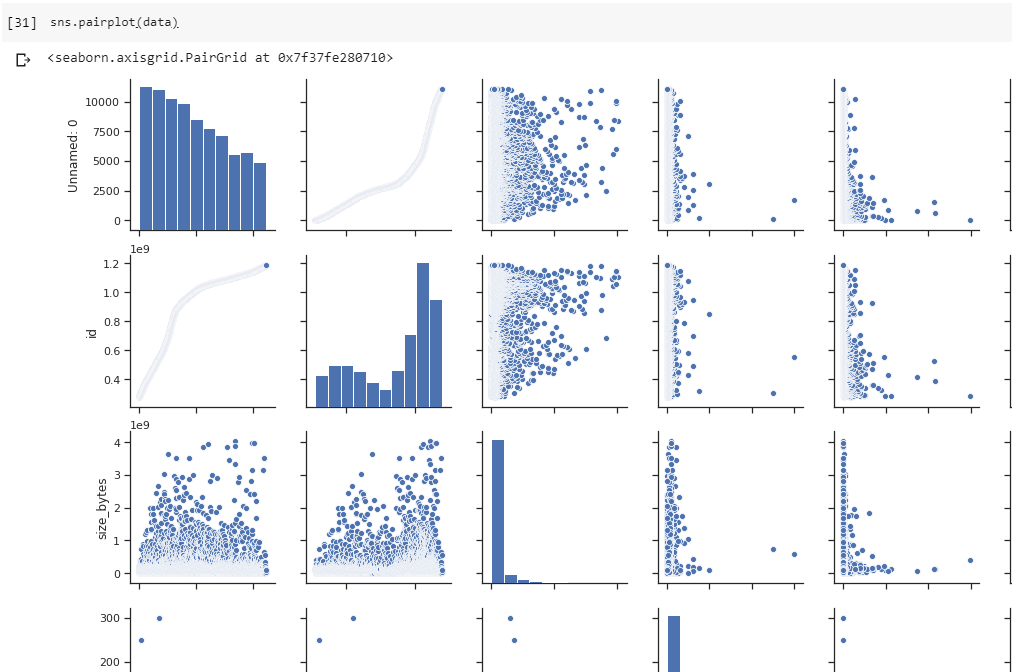
****

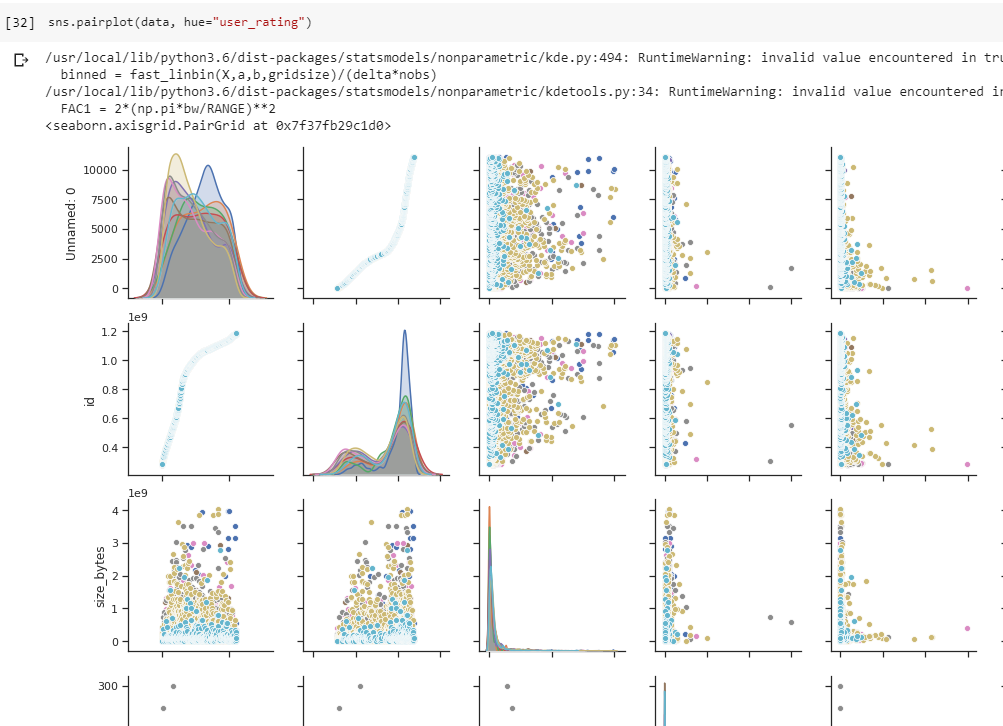
**Joinplot**

** **

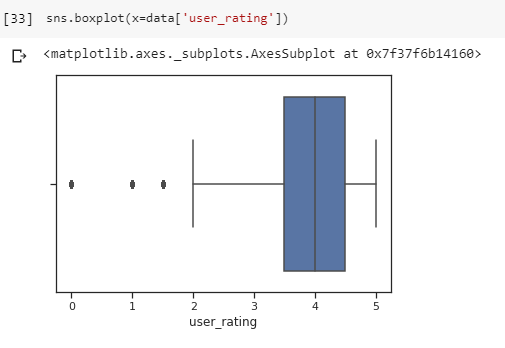
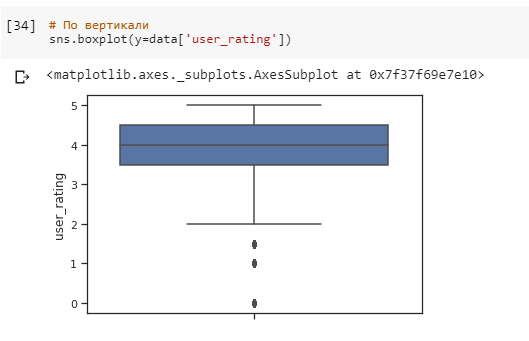
****

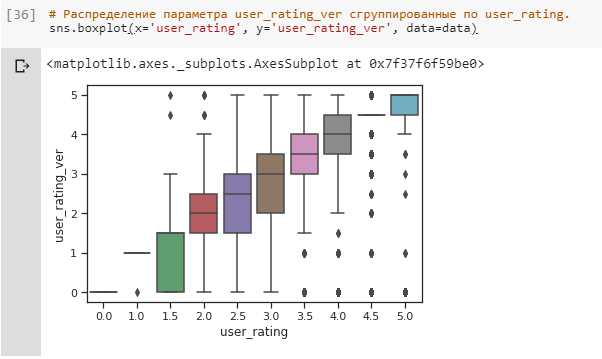
**Парные диаграммы**

****

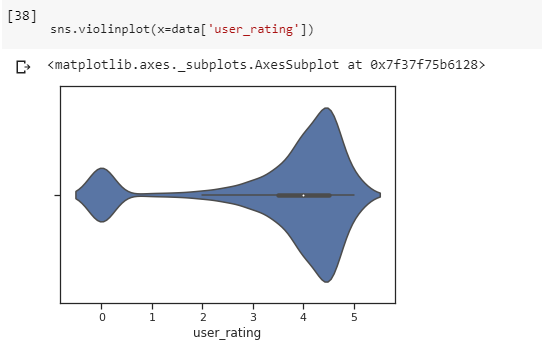
****

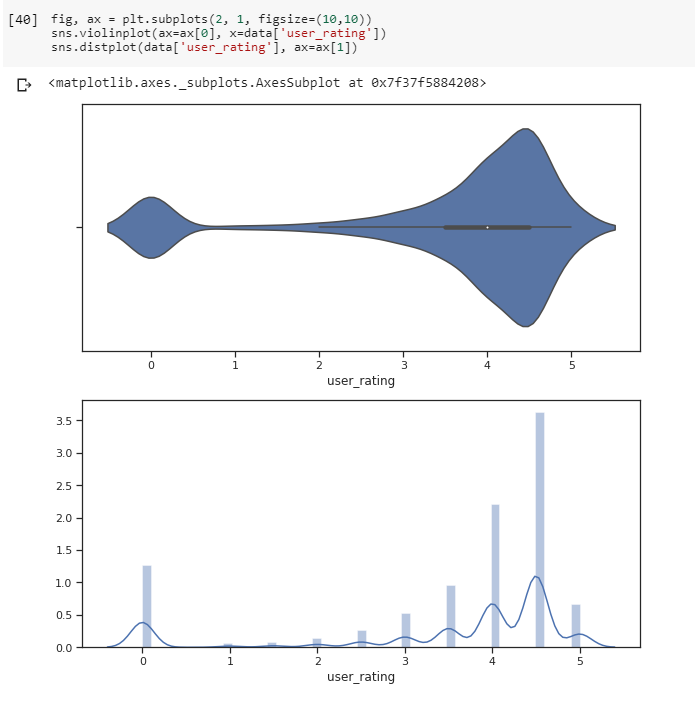
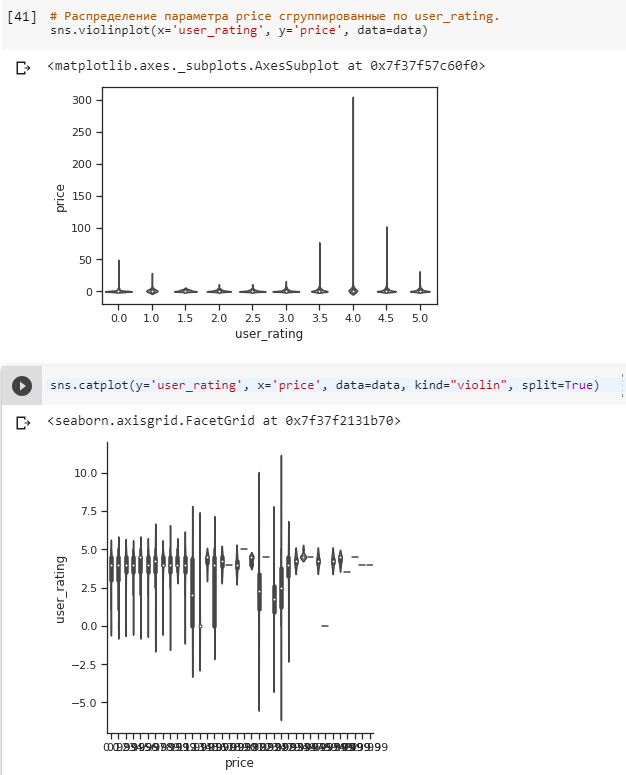
**Ящик с усами**

** **

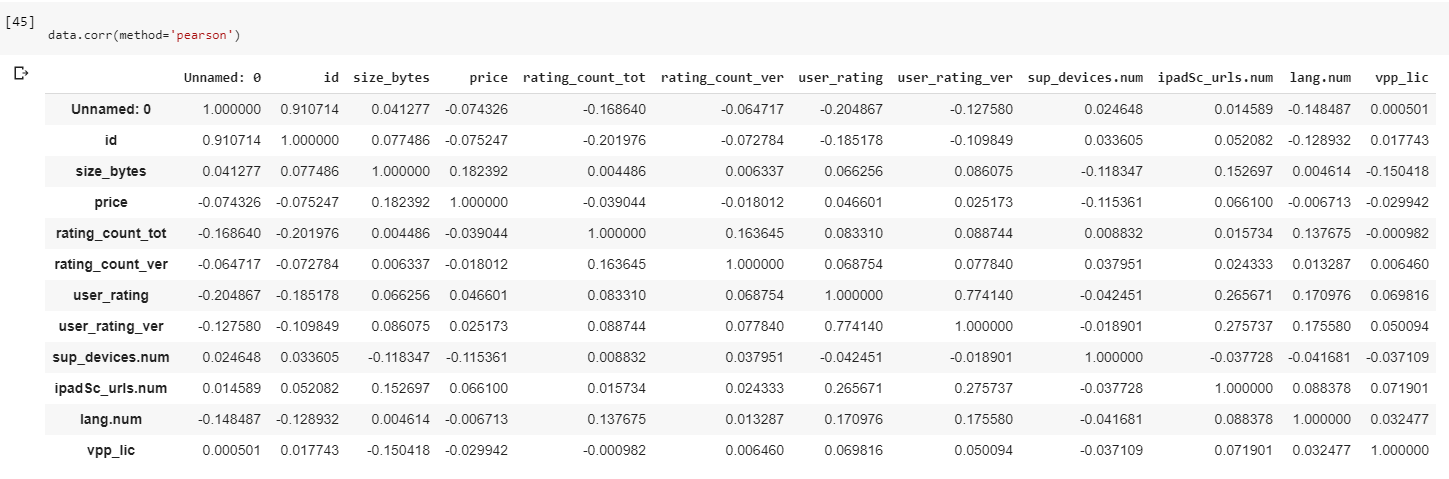
****

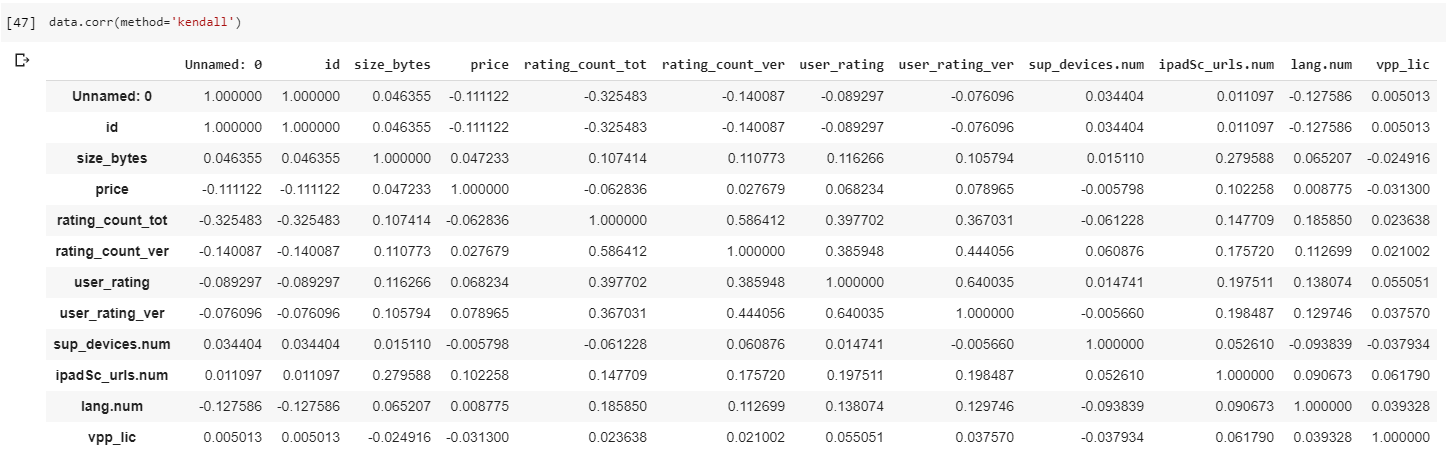
**Violin plot**

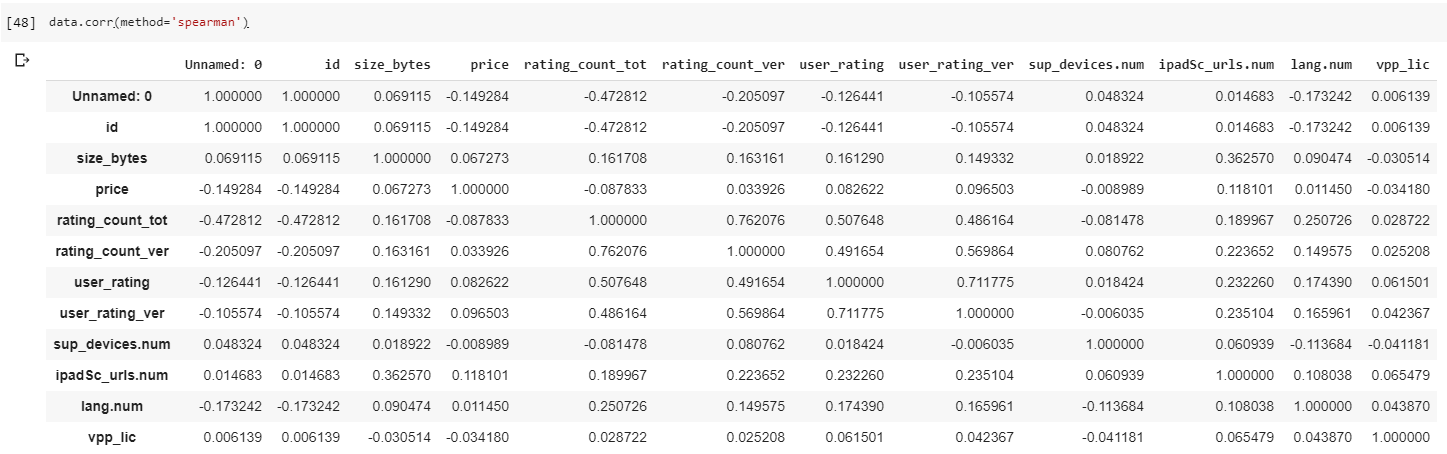
****

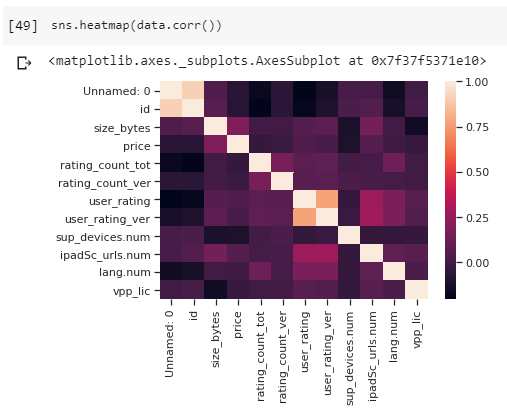
** **

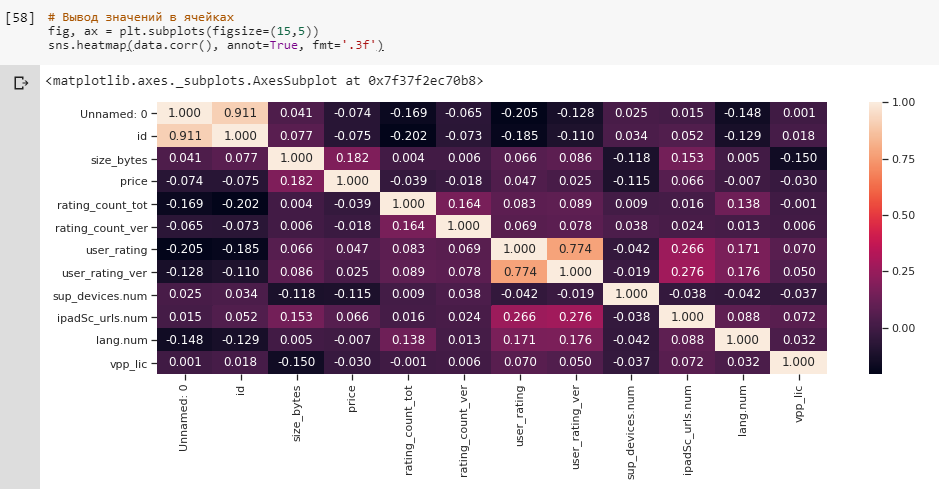
**Корреляционные матрицы**

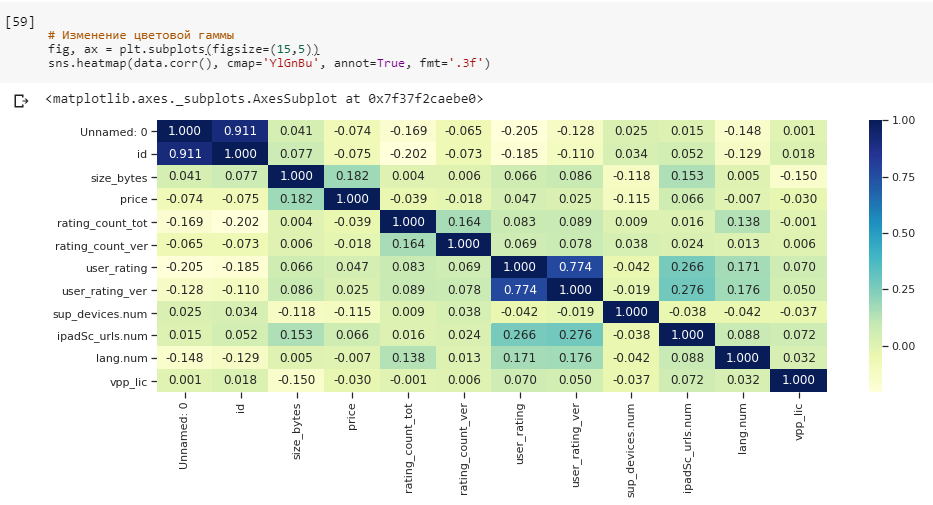
****

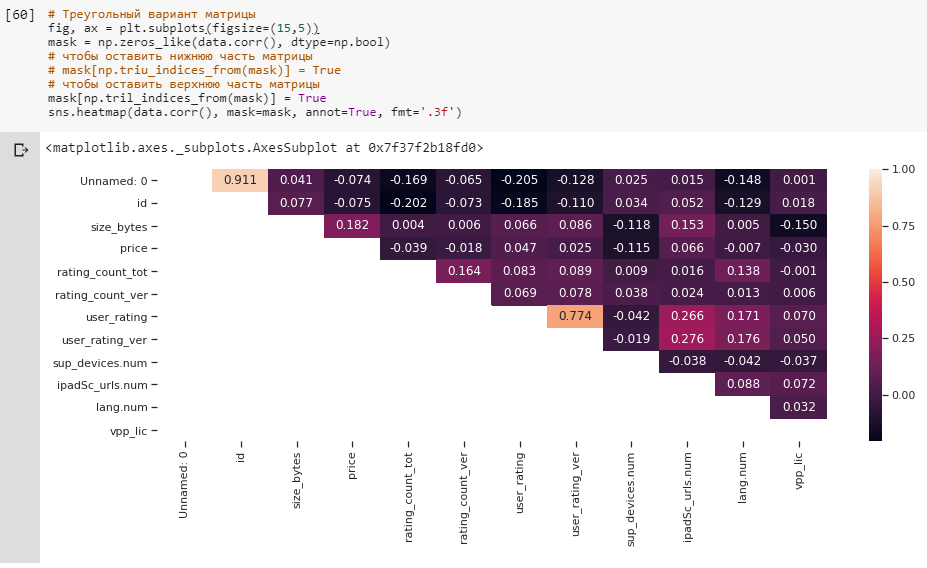
****

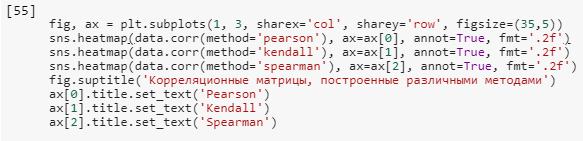
****

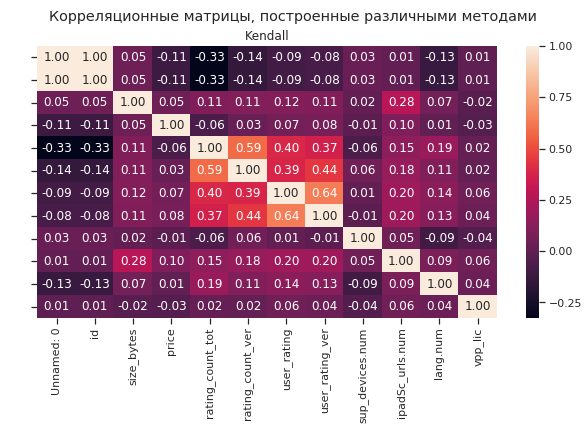
****

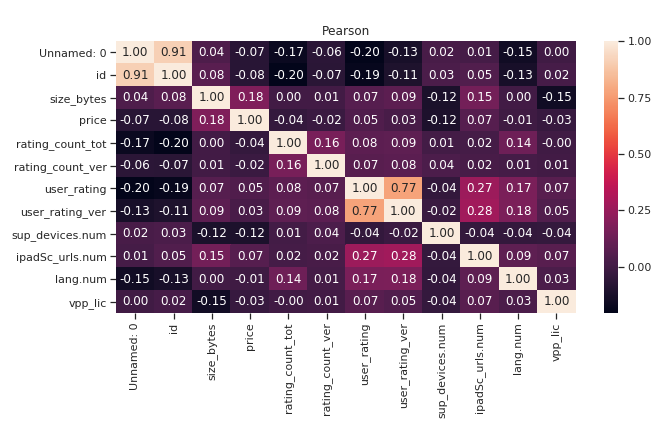
****

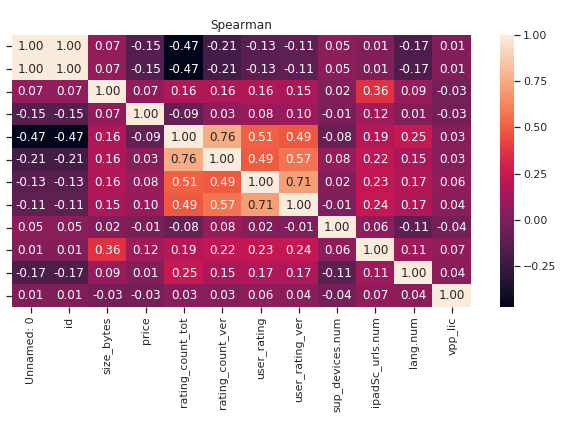
****

****

****

****

****

****