Орищенко Виталий

@Vitosh992

Группа 1

**Архитектор ИИ**

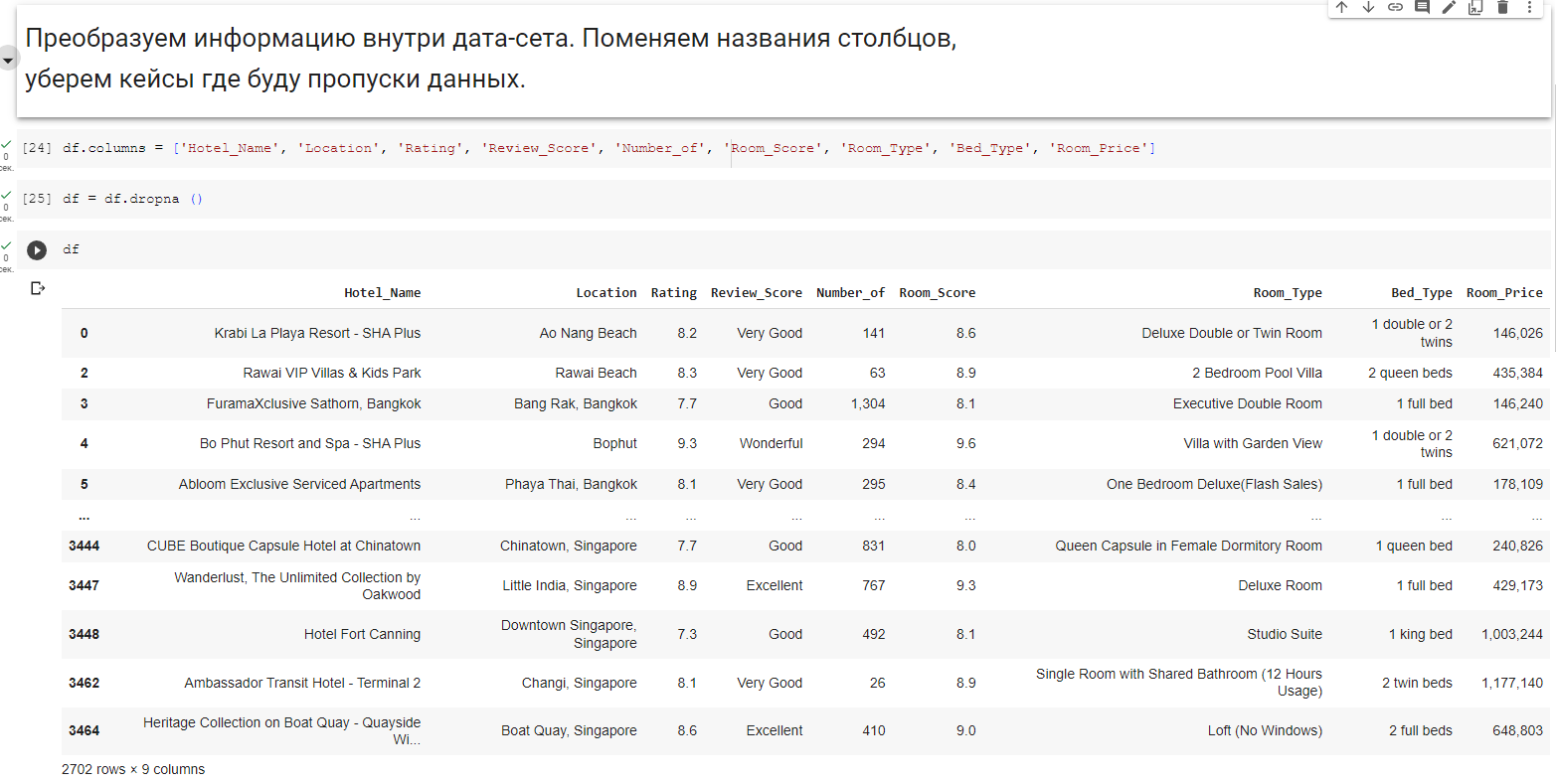
**Задание 2**

В рамках нашего задания нам необходимо выбрать свой дата-сет, сформировать на его основании некоторые гипотезы (на численных данных) и проверить их на истинность.

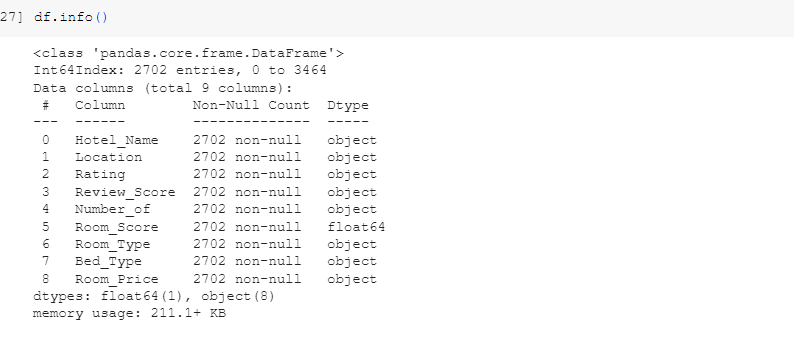
В качестве рассматриваемого дата-сета был выбран список отелей с сайта Booking с такой структурой данных:



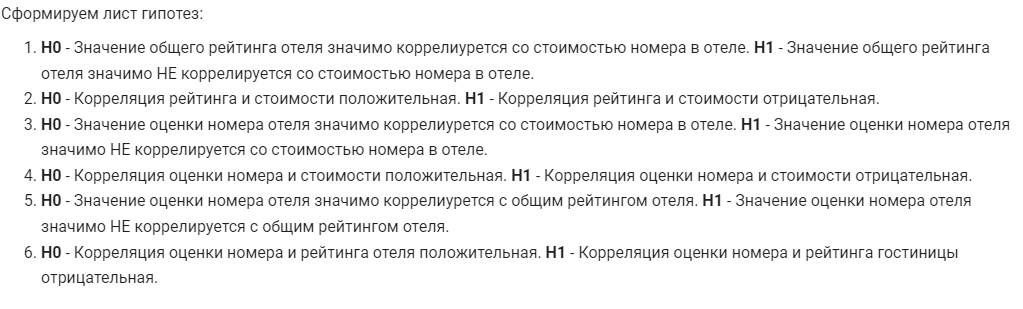
После загрузки данных была проведена работа по их обработке: изменены названия столбцов и очищены кейсы с пропущенными значениями:



В итоге мы получили дата-сет с такими параметрами:



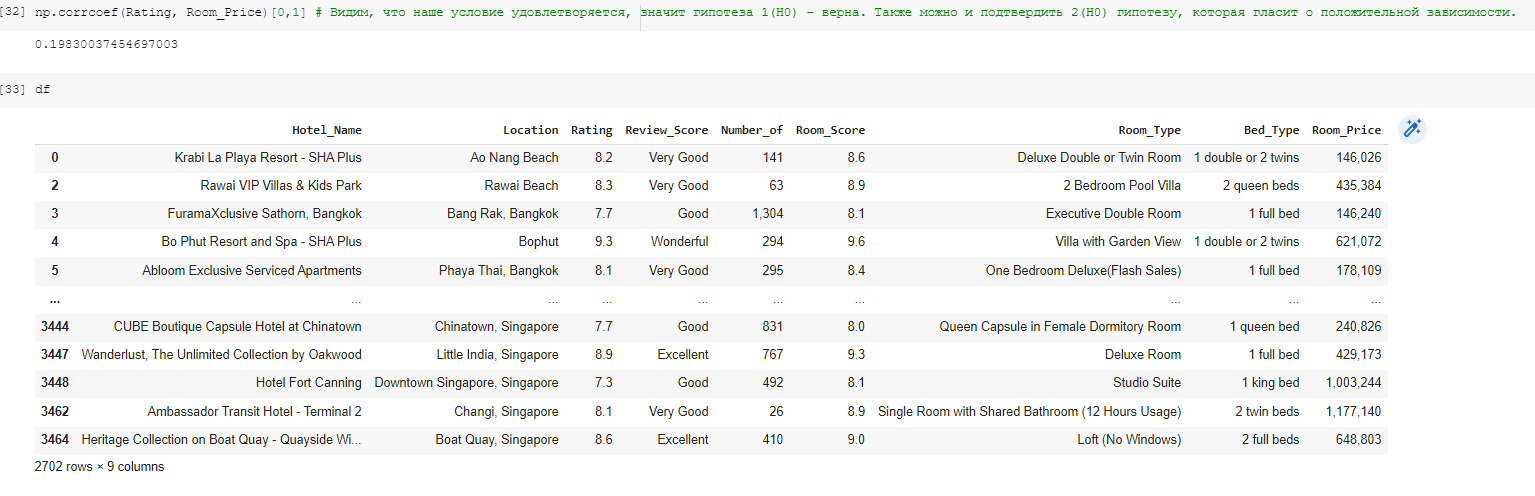
Далее был сформирован лист гипотез. Так как из числовых данных в нашем дата-сете только 3 параметра, следовательно, изучать мы будем именно их.



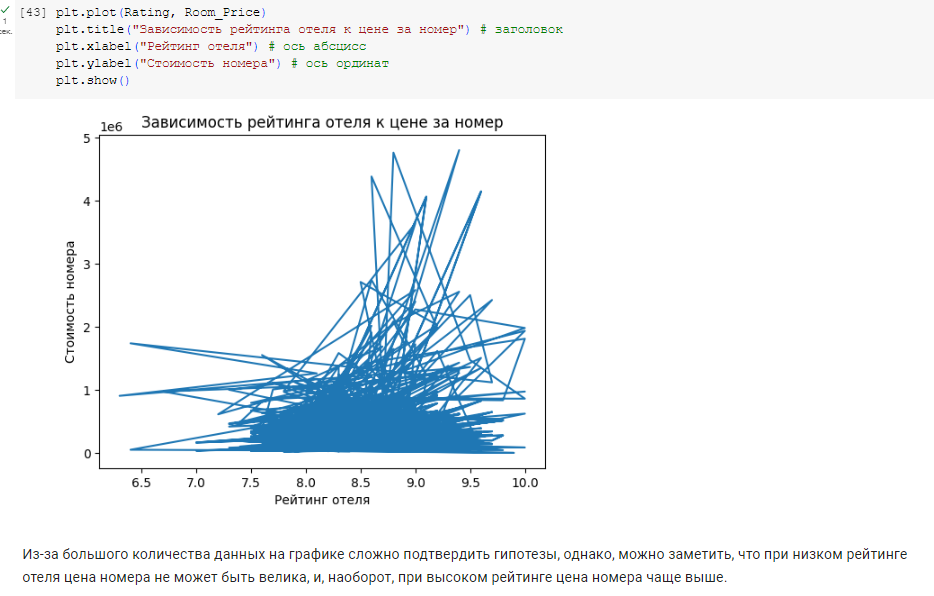
После этого мы приступили к проверке первой и второй гипотезы. В первую очередь в отдельные переменные были скопированы каждые рассматриваемые переменные, переведенные в тип данных float и int(для цены):



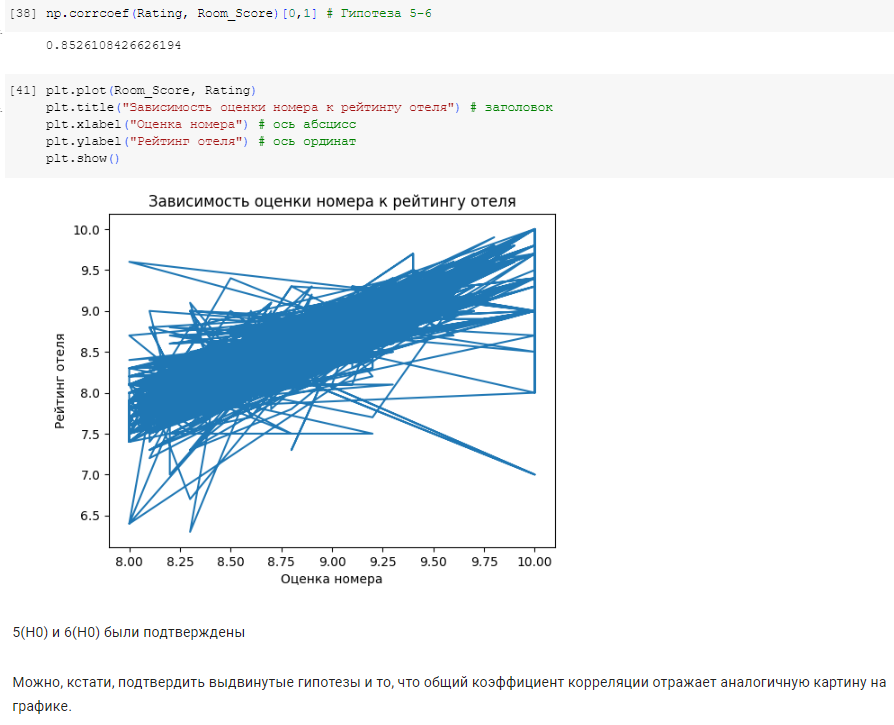
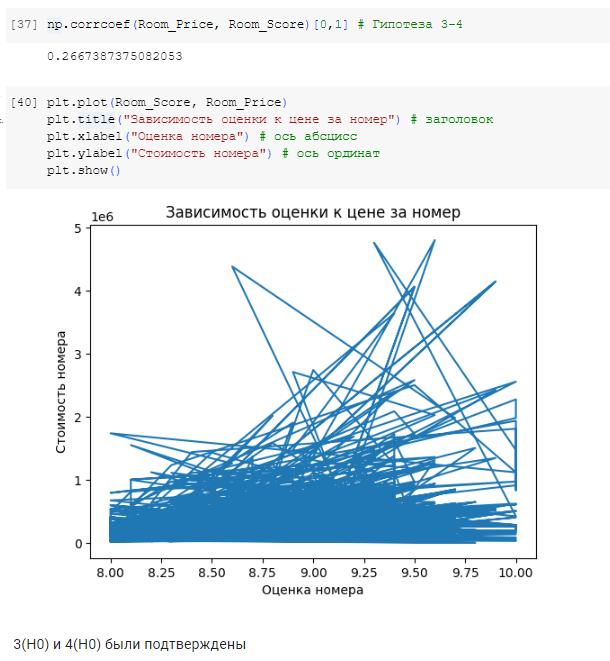
Для проверки гипотез будем находить коэффициент корреляции для каждой пары параметров, сравнивать их с критическим значением Пирсона (0,05).



И для лучшего восприятия будем смотреть на график зависимости этих переменных попарно. Для первой и второй гипотезы:



Аналогично для следующих пар переменных и гипотез 3-6:



Во всех 6 гипотезах были подтверждены основные гипотезы **H0**. На использованном дата-сеты была проведена работа по формированию и проверка возможных гипотез в отношении числовых данных.

ColabURL: <https://colab.research.google.com/drive/1Gx1gP_Urex4n7fp0FPJn4YGPHkQnCiGc?usp=sharing>