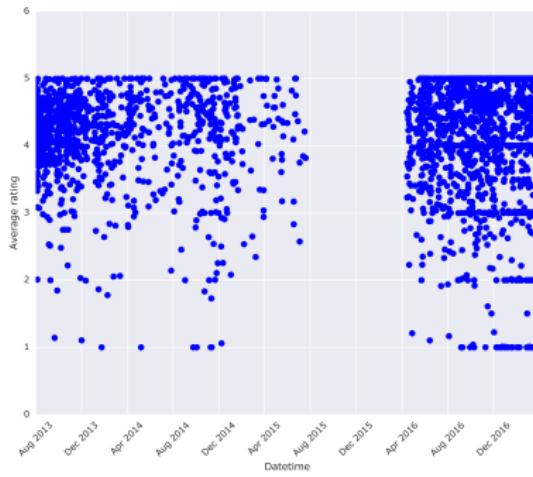


Отчет по заданию «Закономерности в ticketland»

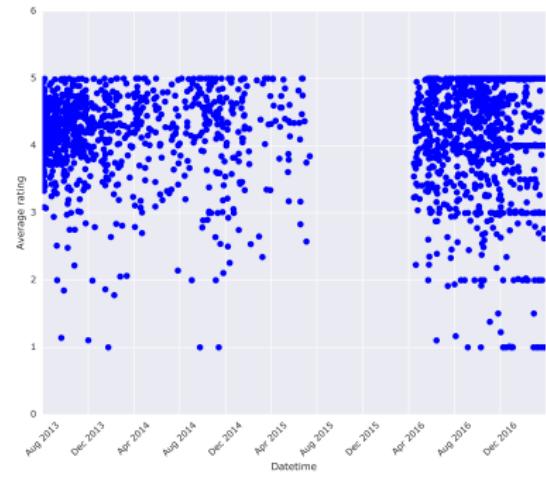
Таскынов Ануар, 517 группа

Курс «Прикладные задачи анализа данных»

15 ноября 2017 г.



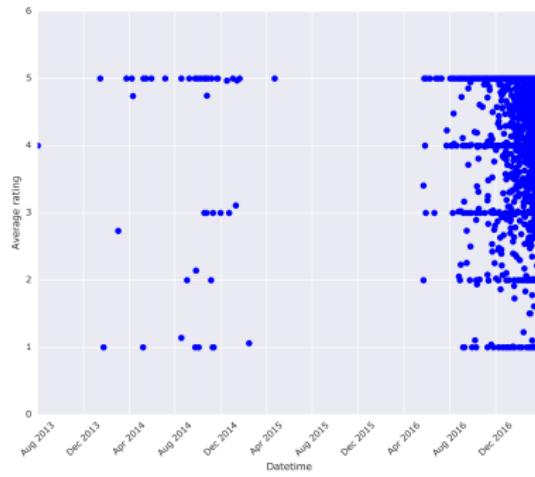
Обучающая выборка.



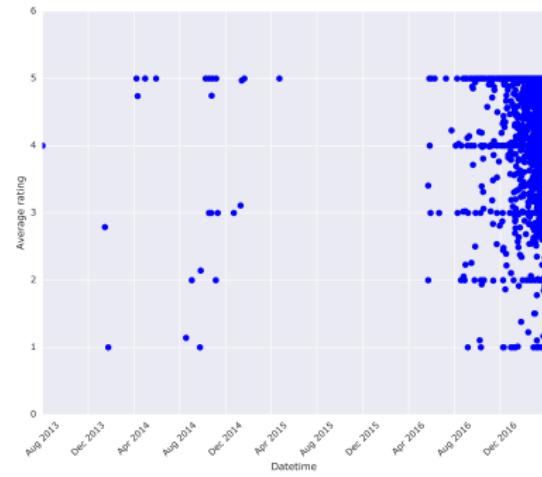
Тестовая выборка.

По вертикали средняя оценка за шоу по горизонтали стоит время, когда впервые начали оценивать шоу.

ПОЧЕМУ не оценены шоу с 2015 по 2016? Средняя оценка обычно завышена.



Обучающая выборка

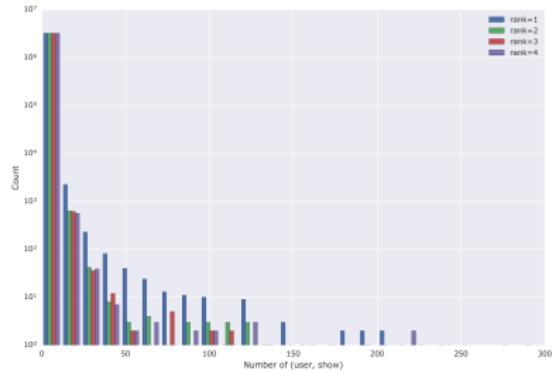


Тестовая выборка

По вертикали средняя оценка за шоу по горизонтали стоит время, когда в последний раз оценивали шоу.

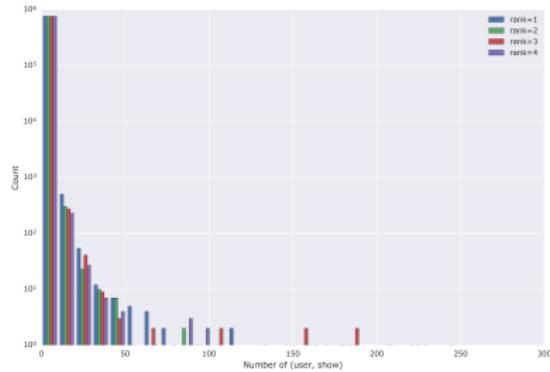
Здесь также пропадает период с 2015 по 2016...

Количество одинаковых пар (user, show)



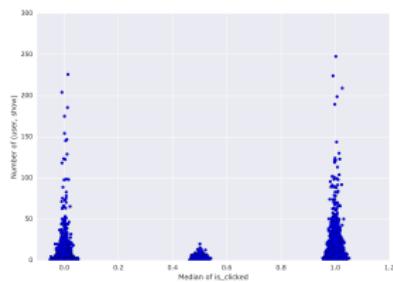
Обучающая выборка

В данных гистограммах изображено количество пар (user, show) и сколько раз в поисковой выдаче данное show выдавалось user с данным rank.
Видно, что количество появлений (user, show) для rank=1 выше, чем для остальных значений.

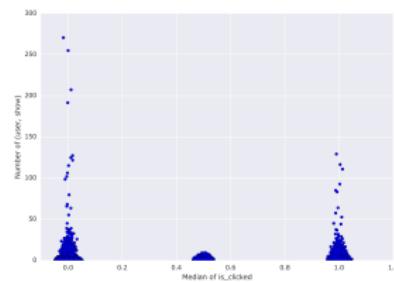


Тестовая выборка

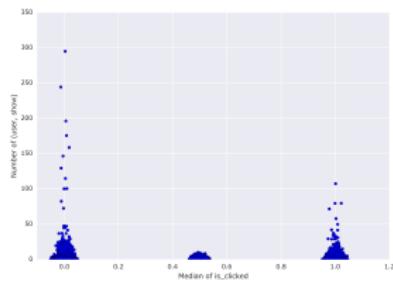
Связь между медианой clicked и количеством появлений (user, show).



rank=1

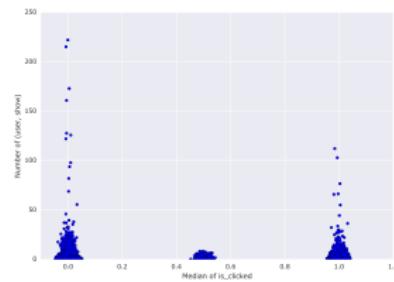


rank=2



rank=3

Видно, что при rank=1, чем больше данное show появляется данному user, тем большее вероятность clicked. Для остальных rank ситуация иная.



rank=4