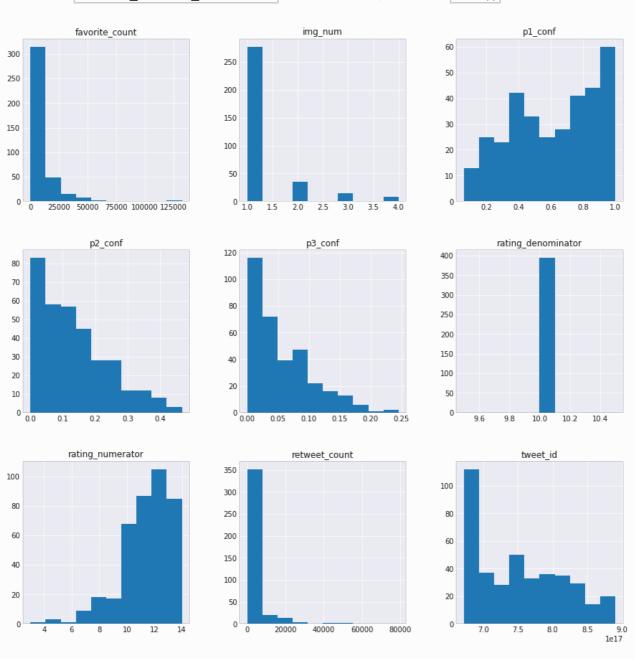
狗狗可视化报告

一、基础情况

对整合后的 twitter_archive_master.csv 进行可视化分析, 首先调用 hist() 展示整体的轮廓:



可以看出,有关狗狗评分的推特都特别活跃——大多数的点赞和转发都在5000-2000之间,最高值分别达到了125000和80000。

同时,在十分制下狗狗的评分也很有趣,大多数在10-14分之间,绝大多数狗狗都突破了10分的限制, 真是可喜可贺。当然,被打低分的是什么情况?比如谁被打了最低分3分?哦,别在意,这是一只猫。



Meet Phil. He's big af. Currently destroying this nice family home. Completely uncalled for. 3/10 not a good pupper

❸ 翻译推文



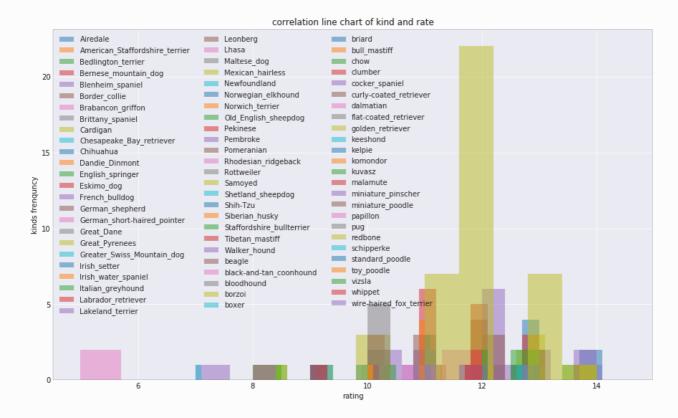
天呐,我们是不是应该把预测中各种奇奇怪怪的东西都先排除出去,保证评分的公平性。

说到狗狗的图像预测,可视化分析告诉我们就首个预测来说一半以上案例可信度都超过0.5;随后的第二、第三次预测大部分可信度都在0.2以下。

二、变量影响

1、预测狗狗品种对打分的影响

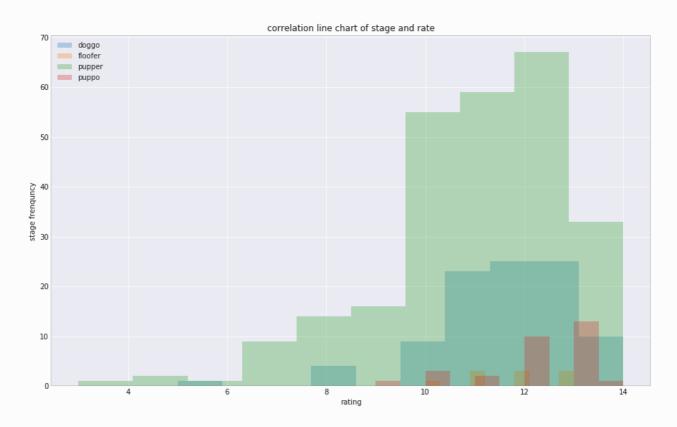
一开始我的猜测是,某种特定类型的狗狗比如金毛犬或者吉娃娃可能会获得更高的分数,于是使用直方 图对两者的关系进行探索:



但在驳杂的品种分类学里我完全迷失了,可见金毛除了数量比较多有优势之外,其他犬种在得分评价上相差无几。

2、狗狗地位对打分对影响

猜测狗狗的地位可能会影响得分,小狗看起来更萌可能得分更高。



事实证明pupper由于数量较多,确实稍占优势,但就总体比率形状来看,大狗doggo的得分也相差无几。

三、推文词云分析

希望了解在每一条推文中所表达情绪和关键信息。

简单浏览可以看到,核心词汇是狗狗的地位pupper等等,此外还有描述狗狗外貌的"look","appears","eye","leg"等等,描述情绪的词汇比如"happy","good","absolute"等等,可以感知到推文的积极热烈。

