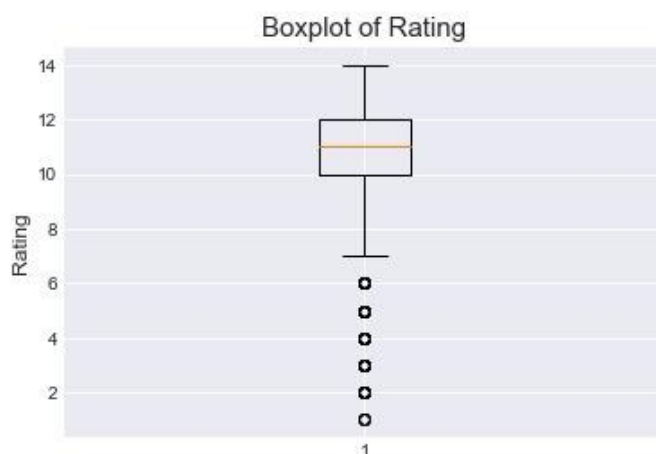


关于 WeRateDogs 推特数据的可视化分析报告

本次分析整理完成了推特用户 @dog_rates 的推文数据，包括每条推特的ID，推文内容，推文中狗子的评级、评分数据，转发数，点赞数，来源以及图像预测结果等。通过对清洁数据的可视化探索分析，初步得到以下分析结果：

1. 狗子评分的分布

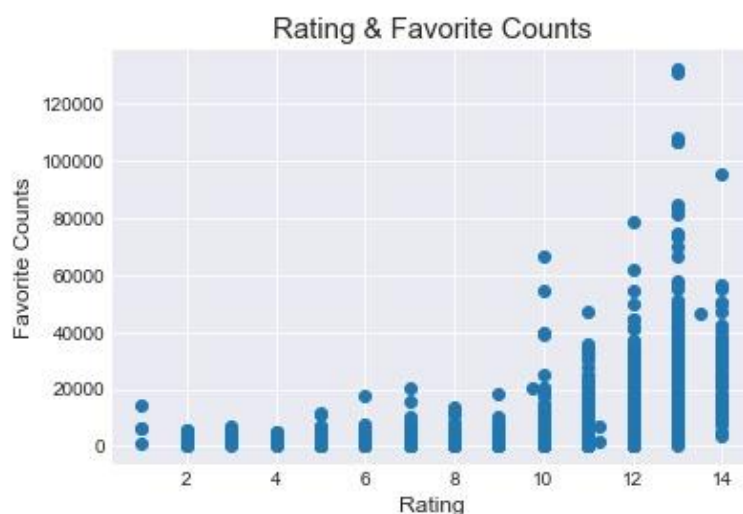
WeRateDogs 以诙谐幽默的方式对人们的宠物狗评级，这些评级通常以 10 作为分母。对清洁数据的分子值做统计概览，得到狗子评分的值域为13.00，最大值为14.00，最小值为1.00，中值为11.00，众数为12.00，平均值为10.57，分布呈现“左偏态”的特点。箱线图如下示：



由箱线图可以看出，狗子评分分子的值分布呈现“左偏态”的特点。

2. 狗子评分值与点赞数量的相关性

狗子评分值与点赞数量的散点分布如下：



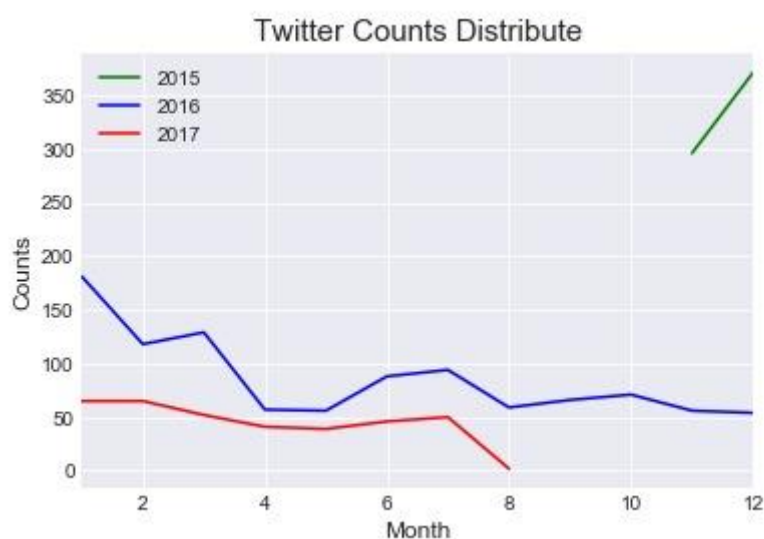
通过在Python中建立狗子评分与点赞数线性回归模型，得到如下结论：

- 推特的点赞数与狗子评分的相关系数为0.417，为中度正相关；
- 点赞数17.4%的可变性可通过狗子的评分进行解释，剩下的82.6%的可变性则决定于其它推特因素；
- 预测推特的点赞数与狗子的评分具有统计相关性，狗子的评分每增加1分，则该推特点赞数会增加2340个。

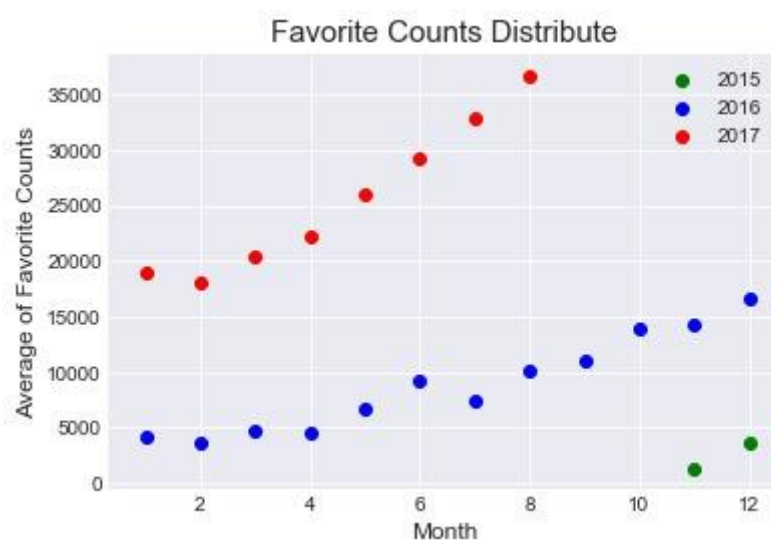
3. 推特发布数量与获得点赞数的时间分布

按年、月为单位，对推特的发布数量，每月的平均或的点赞数进行可视化得到：

1) 推特发布数量随时间散点分布



2) 推特月平均点赞数量随时间的分布



由图可以看出，随着时间（年或者月）的推移，推特的发布数量呈下降的趋势；而推特点赞的月平均数量随时间的推移呈现数量增长的趋势。