

WeRateDogs 推特数据探索分析

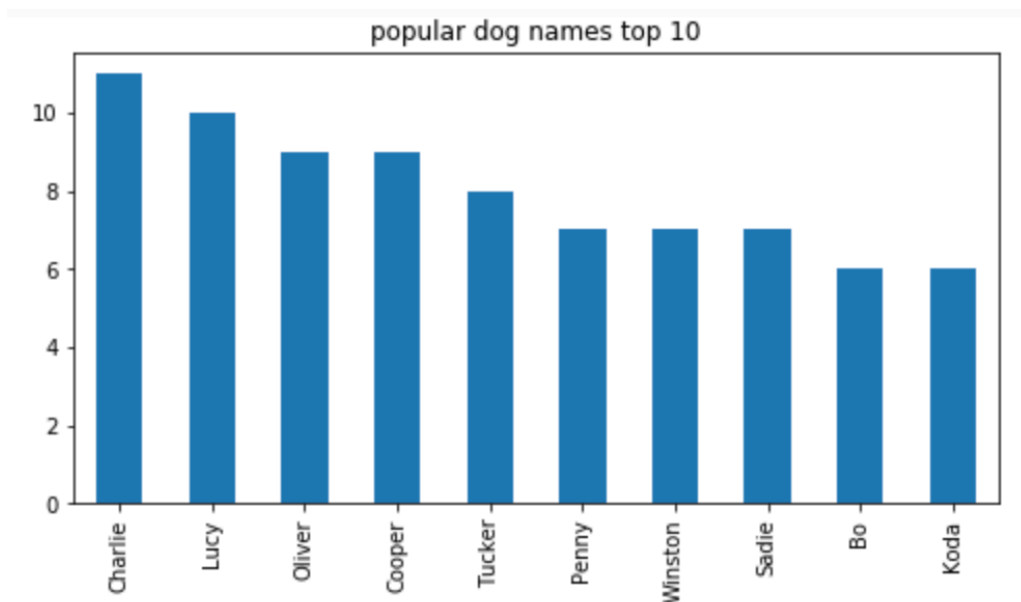
WeRateDogs 是一个推特主，他以诙谐幽默的方式对人们的宠物狗评分。此次我们对 WeRateDogs 的推特信息做了探索分析，并发现了一些有意思的结论。

本次探索我们提出了三个问题：

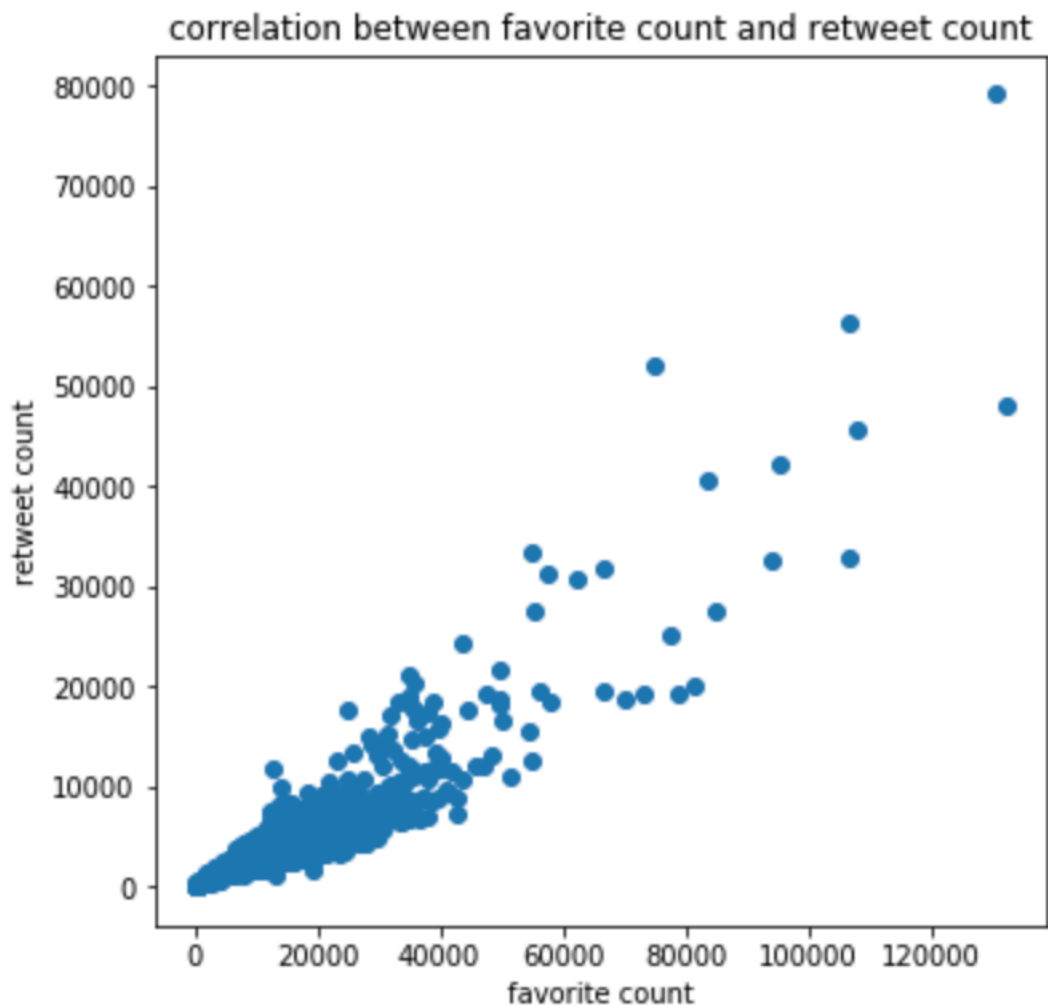
1. 最人气的狗狗名字前十名是什么
2. favorite_count 与 retweet_count 是否相关
3. 在有狗狗地位的数据中，评分前三的狗狗都是什么地位？

结果如下

1. 对于要探索的第一个问题，我们用数据生成了如下的柱状图，我们可以看到最人气的狗狗名字前十名分别是 Charlie, Lucy, Cooper, Oliver, Tucker, Sadie, Winston, Penny, Daisy, Bo. 请仔细看下这几个名字，搞不好你的哪个朋友刚好跟最人气的狗狗共享一个名字；而相关的社会学研究表明，取名也会受当时流行大事件的影响，加入一些社会学方面的知识，没准有些更有意思的发现



2. 我们取 favorite_count 数值与 retweet_count 的数值生成了散点图，如下图所示，我们可以看到，随着 x 轴 favorite_count 数值增大，y 轴 retweet_count 数值也跟着增大，favorite_count 与 retweet_count 正相关。果然越受欢迎越容易产生病毒式传播，这里也有点马太效应的意味了吧。



3. 我们单独提取出狗狗地位和狗狗评分分子两列，然后筛选出包含狗狗地位的数据，按狗狗评分进行降序排列，发现评分最高的前三名评分分子分别为 27，14，14；对应的狗狗地位分别为 pupper, doggo,pupper，如下图所示
- 我们试图据此探索是否某种地位的狗狗具有压倒性优势，但考虑到数据有限，可能探索的结果跟实际情况会有实际偏差。

	stage	rating_numerator
543	pupper	27
9	doggo	14
748	pupper	14