

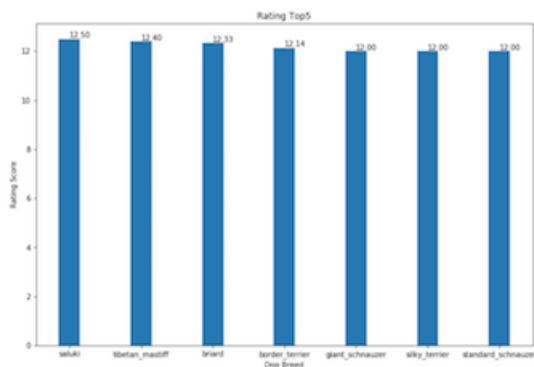
WeRateDogs 的推特档案分析报告

该项目分析了四个问题

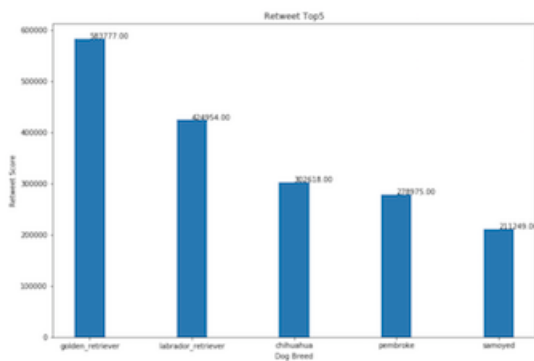
- 哪个品种的狗评分最高，转发数和点赞数更多(取前五名)
- 哪个品种的狗出现的频率最高(取前五名)
- 有名字的狗和没有被命名的狗，哪一类的转发数和点赞数更多
- 哪一个时间段twitter的活跃度更高（取前五名）

问题一：哪个品种的狗评分最高，转发数和点赞数更多(取前五名)

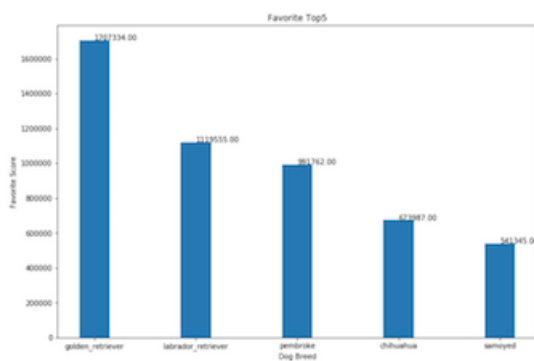
通过对数据进行分析，我们发现在 *WeRateDogs* 中获得评分前5名的品种分别是：第一名 *saluki*，12.5分；第二名 *tibetan_mastiff*，12.4分；第三名 *briard*，12.33分；第四名 *border_terrier*，12.14分；并列第五名 *giant_schnauzer*、*silky_terrier*、*standard_schnauzer*，12分。



获得转发数前5名的品种分别是：第一名 *golden_retriever*，583777；第二名 *labrador_retriever*，424954；第三名 *chihuahua*，302618；第四名 *pembroke*，278975；第五名 *samoyed*，211249。

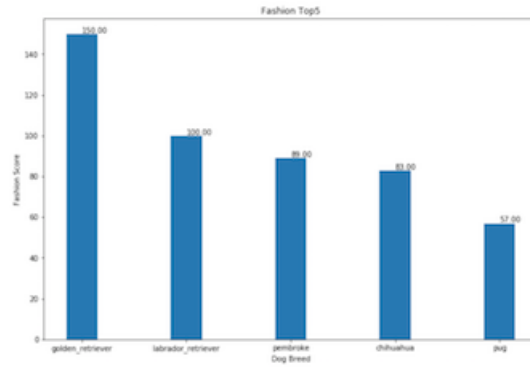


获得点赞数前5名的品种分别是：第一名 *golden_retriever*，1707334；第二名 *labrador_retriever*，1119555；第三名 *pembroke*，991762；第四名 *chihuahua*，673987；第五名 *samoyed*，541345.0



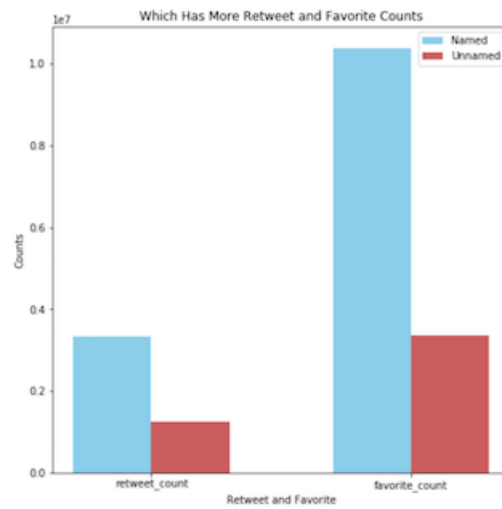
问题二：哪个品种的狗出现的频率最高

通过分析发现，在 *WeRateGogs* 中出现频率排在前5名的品种分别是：第一名 *golden_retriever*，150；第二名 *labrador_retriever*，100；第三名 *pembroke*，89；第四名 *chihuahua*，83；第五名 *pug*，57。



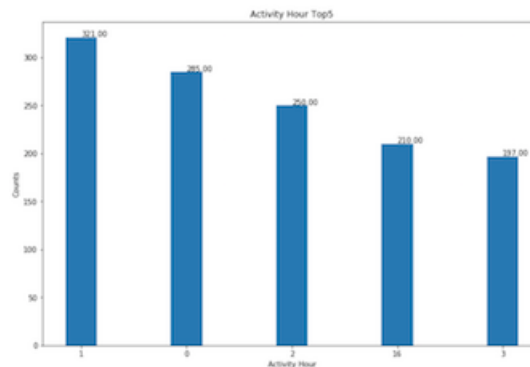
问题三：有名字的狗和没有被命名的狗，哪一类的转发数和点赞数更多

通过分析发现，有名字的狗的转发数和点赞数都要高于没有名字的狗。



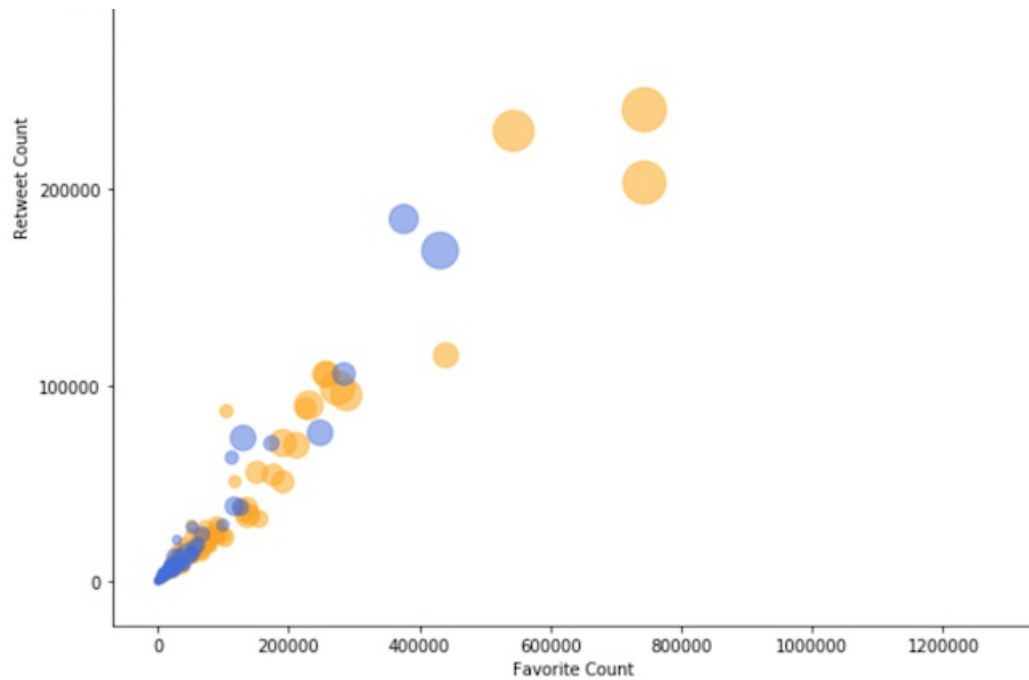
问题四：一天中哪一个时间段twitter的活跃度更高（取前五名）

一天中用户活跃度排在前五名的时间分别是：第一名1:00, 321；0:00, 285；2:00, 250；16:00, 210；3:00, 197。看起来用户们喜爱在后半夜转发他们喜爱的宠物狗，也可能用户们在后半夜使用手机刷 *Twitter* 的时间更多。



最后我们来看看点赞数、转发数，以及狗是否被命名，这三者之间存在的关系





通过上图能够看出

- 点赞数和转发数大体上呈正相关趋势。
- 圆圈的大小代表出现的次数。出现次数较多的狗，favorite和retweet值也相应增加。
- 从分布上看，未命名的狗点赞数和转发数要少于有名字的狗。其中一个原因可能是有名字的狗受到的照料要好于未命名的狗，外观上可能更容易引人喜欢。