

WeRateDogs 推特数据描述性分析和可视化分析

一、项目背景

本次项目提供了三份数据集： WeRateDogs 的推特档案、通过推特 API 获取附加数据、图像预测数据，包含了截止到 2017 年 4 月 1 日的 5000 多条给狗狗评分的推特。

二、数据说明

用来分析的数据集是上述三份数据集经过清洗整理后合并的主数据集，共有 8 个维度变量：

- tweet_id: 推特发文编号
- times: 推特发文时间
- dog_type: 狗狗的种类
- name: 狗狗的名字
- rating: 狗狗获得的评分
- stages: 狗狗获得的“地位”评级
- favorite_count: 推特发文获得的喜爱数
- retweet_coun: 推特发文获得的转发数

三、数据分析

3.1 提出问题

在 WeRateDogs 的推特档案里：

- 1、被评分最多的狗狗种类是哪些？
- 2、哪些狗狗容易获得高评分？

3、哪些狗狗在推特中更受喜爱？

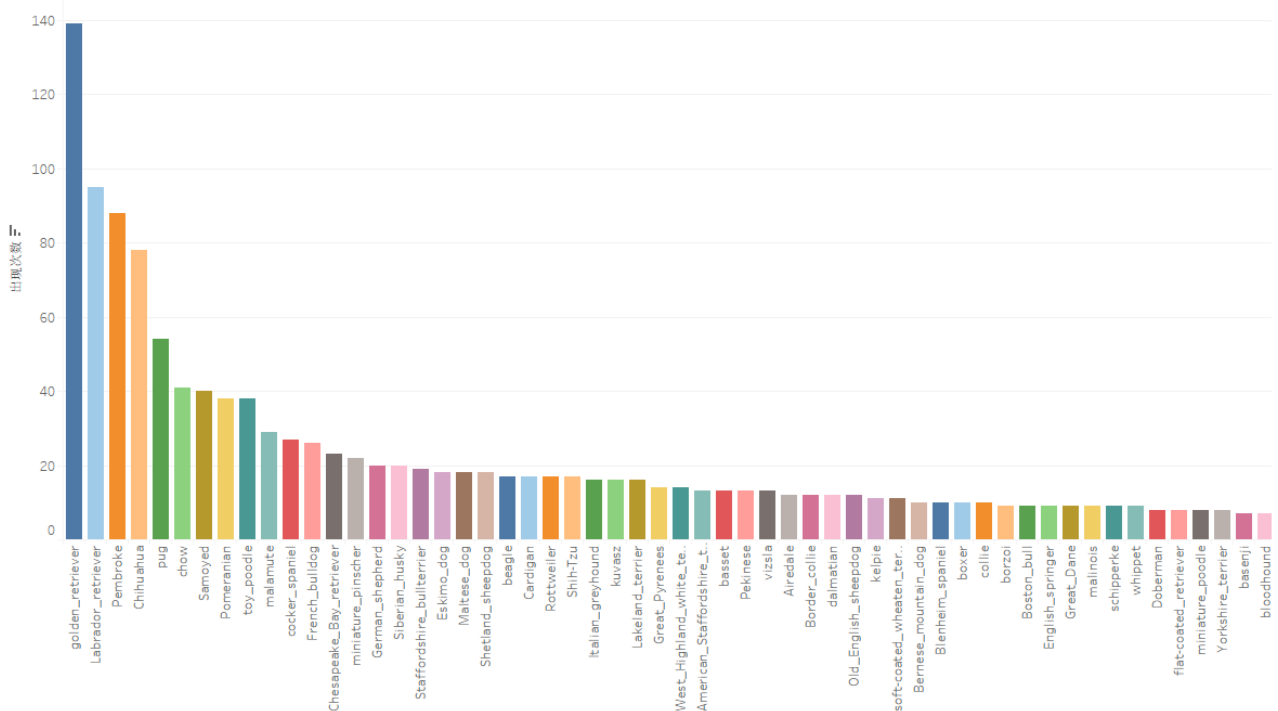
4、哪些狗狗更容易被转发？

3.2 分析问题

问题 1：被评分最多的狗狗种类是哪些？

解答：从 WeRateDogs 的推特档案里找到所有被评分的狗狗，分析它们的种类，发现被评分最多的狗狗是金毛猎犬（golden_retriever），多达 139 次；紧随其后的是拉布拉多寻回犬（Labrador_retriever）和柯基犬（Pembroke），分别是 95 次和 88 次。

WeRateDogs 的推特档案里出现最多的狗狗品种



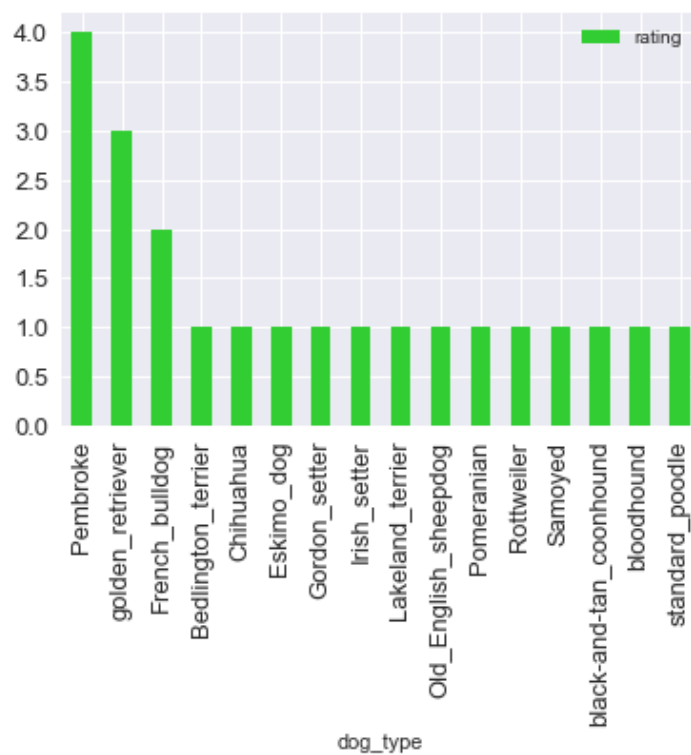
问题 2：哪些狗狗容易获得高分？

解答：在参与评分的狗狗中，因为数量最多的前三名是金毛猎犬（golden_retriever）、拉布拉多寻回犬（Labrador_retriever）和柯基犬（Pembroke），所以总评分最多的也是它们，分别是：1849.50、1321.00、1007.00。

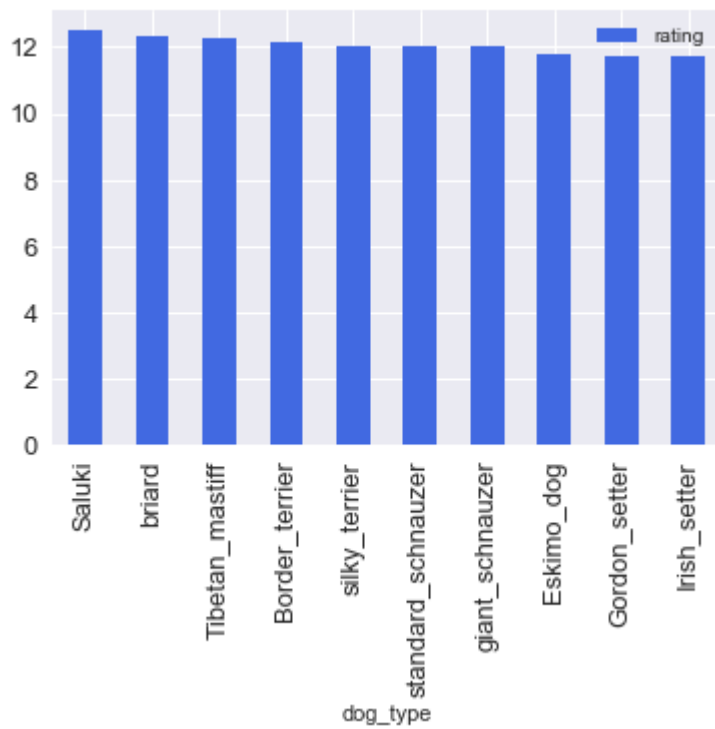
不同种类的狗狗获得的总评分



但是得分最高（14 分）的 22 个狗狗中，柯基犬（Pembroke）最多，有 4 条。其次是金毛猎犬（golden_retriever）3 条和法国斗牛犬（French_bulldog）2 条。



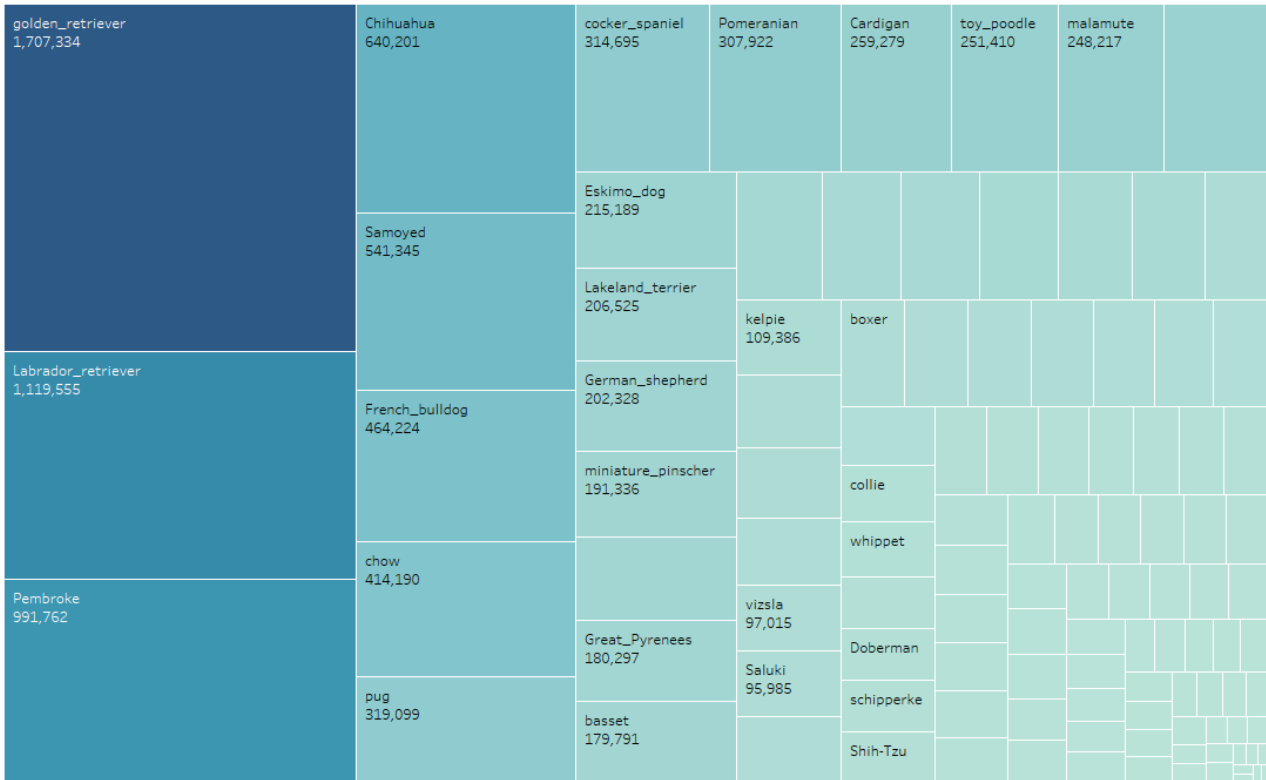
按照狗狗的平均评分来看，最受欢迎的是萨卢基狗(Saluki)、伯瑞犬(briard)和藏獒(Tibetan_mastiff)，评分分别是：12.5、12.33 和 12.25。



问题 3：哪些狗狗在推特中更受喜爱？

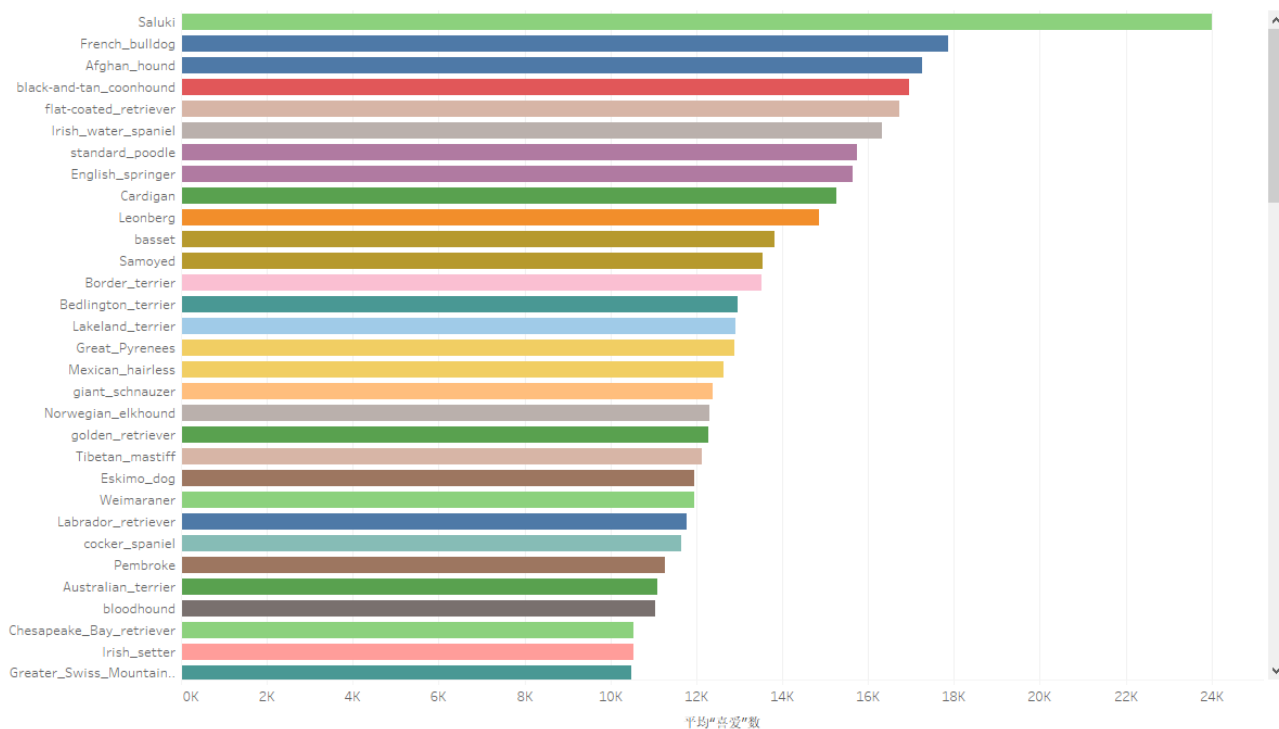
解答：在参与评分的狗狗中，获得喜爱（点赞）数量最多的前三名是金毛猎犬（golden_retriever）、拉布拉多寻回犬（Labrador_retriever）和柯基犬（Pembroke），分别是：1707334.00、1119555.00、991762.00，与参与评分的狗狗数量排名一致。

不同种类的狗狗获得的总“喜爱”数



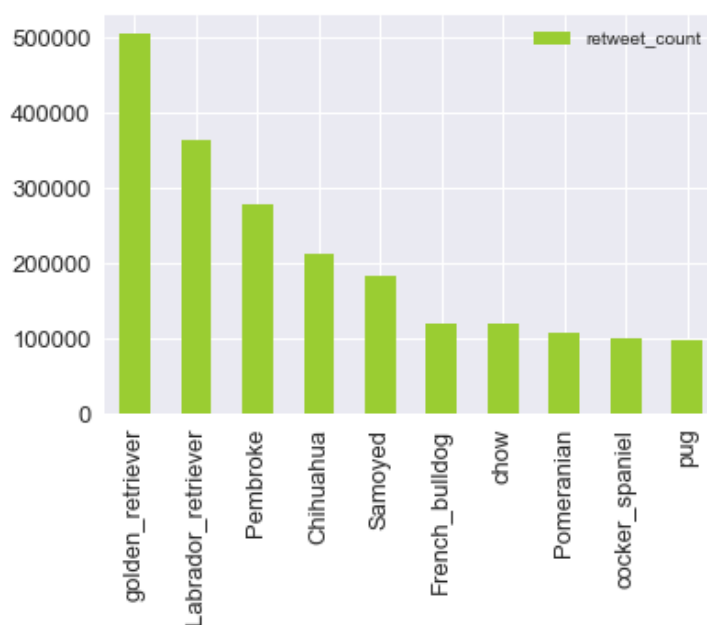
查看不同种类的狗狗的平均喜爱（点赞）数量，前三名分别是萨卢基狗（Saluki）、法国斗牛犬（French_bulldog）和阿富汗猎犬（Afghan_hound）。他们获得的平均喜爱（点赞）数分别是：23996.25、17854.77 和 17250.67。

不同种类的狗狗获得的平均“喜爱”数



问题 4：哪些狗狗更容易被转发？

解答：在参与评分的狗狗中，被转发最多的前三名狗狗分别是：金毛猎犬（golden_retriever）、拉布拉多寻回犬（Labrador_retriever）和柯基犬（Pembroke）。它们获得的转发总数分别是：505153.00、364632.00、278975.00，转发总数排名与参与评分的狗狗数量排名一致。



查看不同种类的狗狗的平均转发数量，前三名分别是标准狮子狗（standard_poodle）、英国史宾格犬（English_springer）和阿富汗猎犬（Afghan_hound）。他们获得的平均转发数分别是：6602.14、5948.67 和 5946.00。

