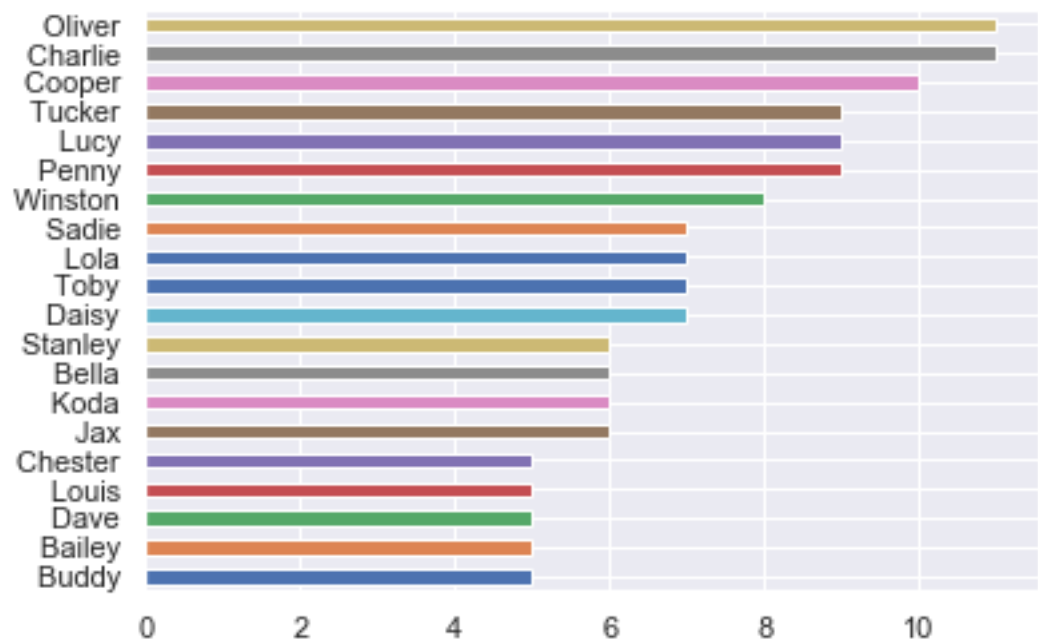


# WeRateDogs 数据可视化

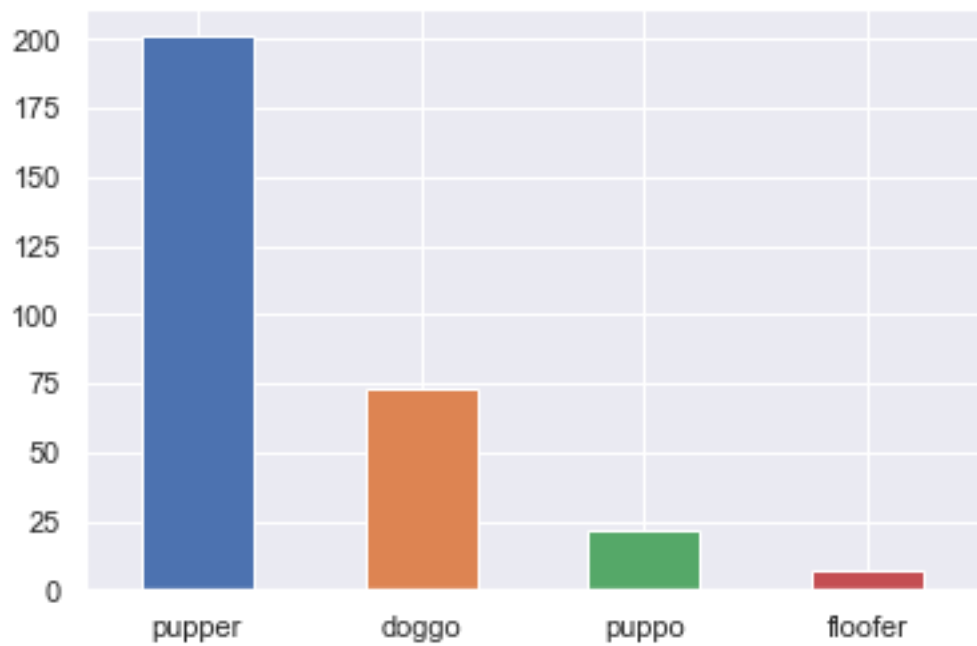
## 1. 单变量统计分析:

◇ 狗的名字占比的统计



狗的名字 Top5 是: Oliver, Charlie, Cooper, Tucker, Lucy

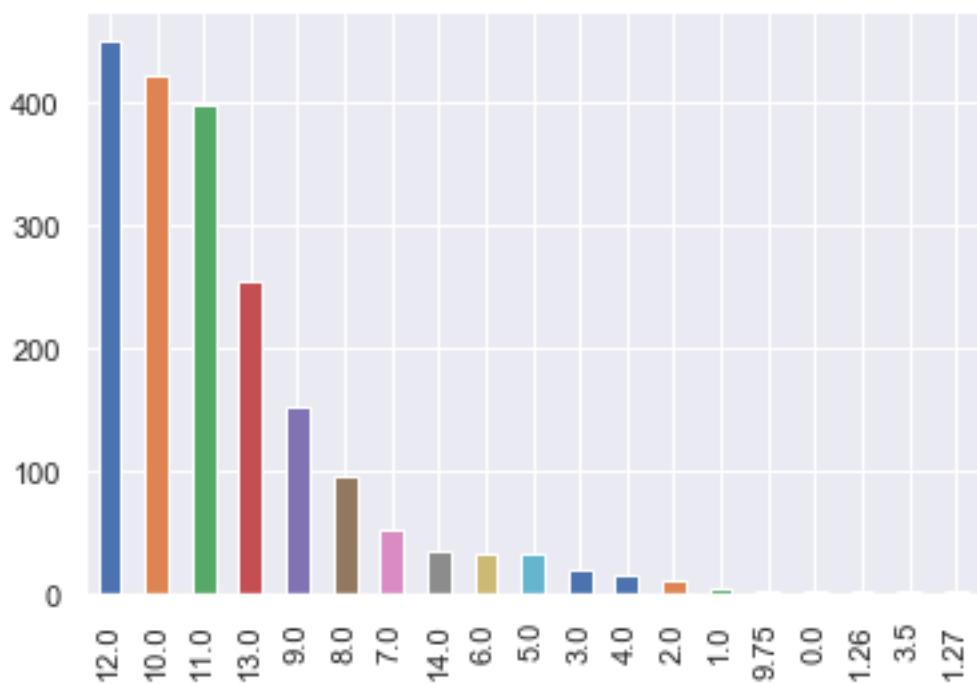
◇ 狗的特征占比的统计



推主提到的 pupper 最多，果然小狗最受欢迎。

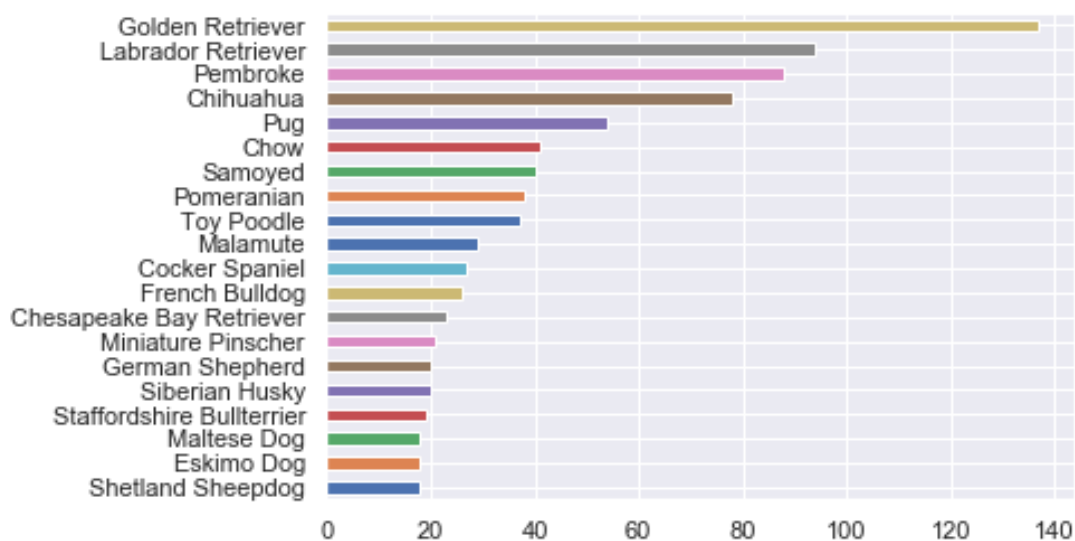
#### ✧ 狗的分数的占比统计

注：在统计之前，先排除掉分数极高的，只取分数小于 20 的进行分析。



推主给大多数狗的评分数 10-13 分，绝大多数都在 2-14 之间。

### ◇ 狗的种类的占比统计

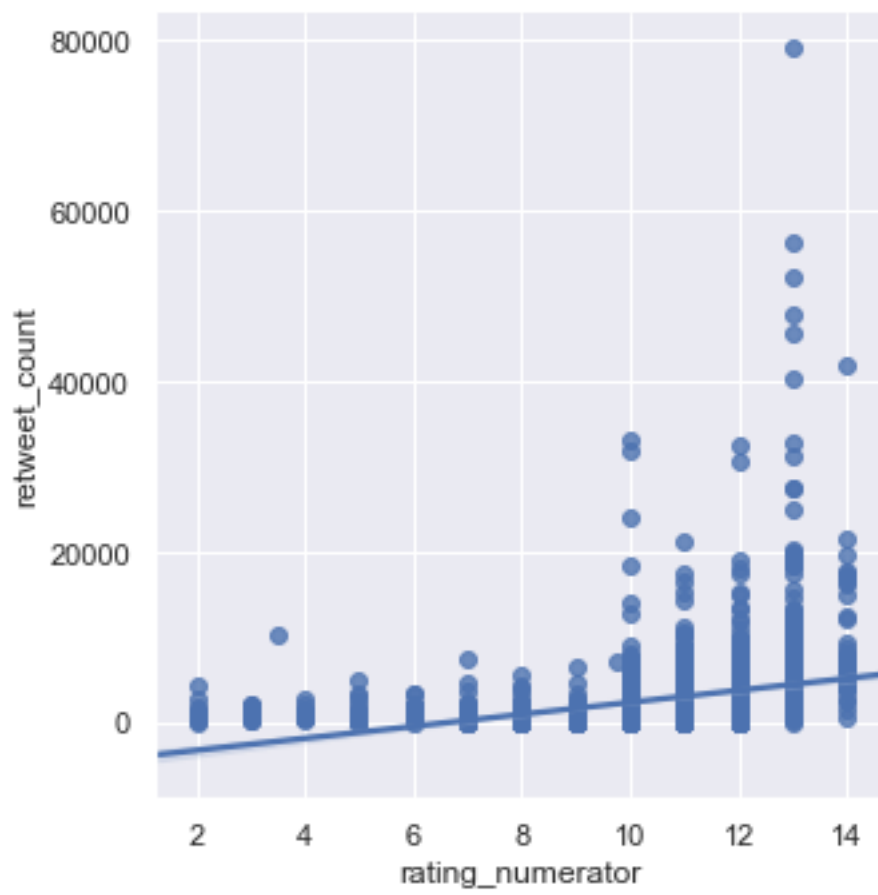


推主评分的狗种类 Top5 是：金毛，拉布拉多，Pembroke 柯基，吉娃娃，八哥。

## 2. 相关性分析：

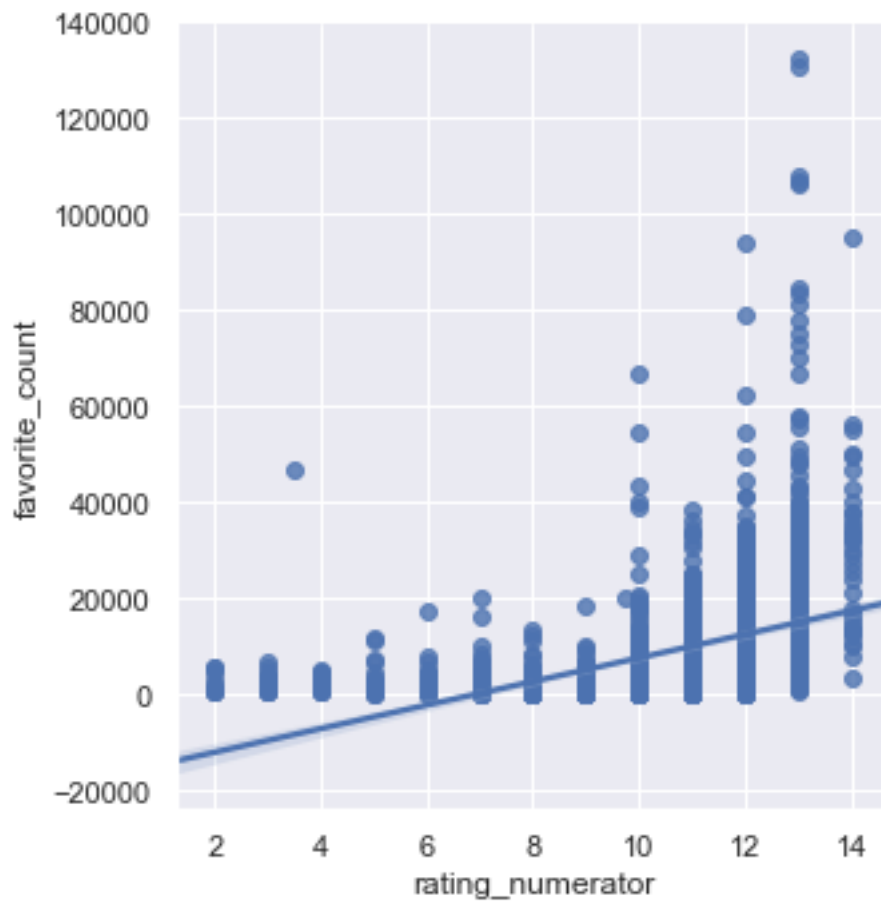
### ◇ 转推数与狗评分的关系

狗的分数大多集中在 2-14，评分过低或过高意义不大。所以只分析这些数据。



从结果上看，转推数与评分有一定的正相关关系。狗的评分越高，相应的转推数就会越多。

#### ✧ **favorite** 与狗评分的关系

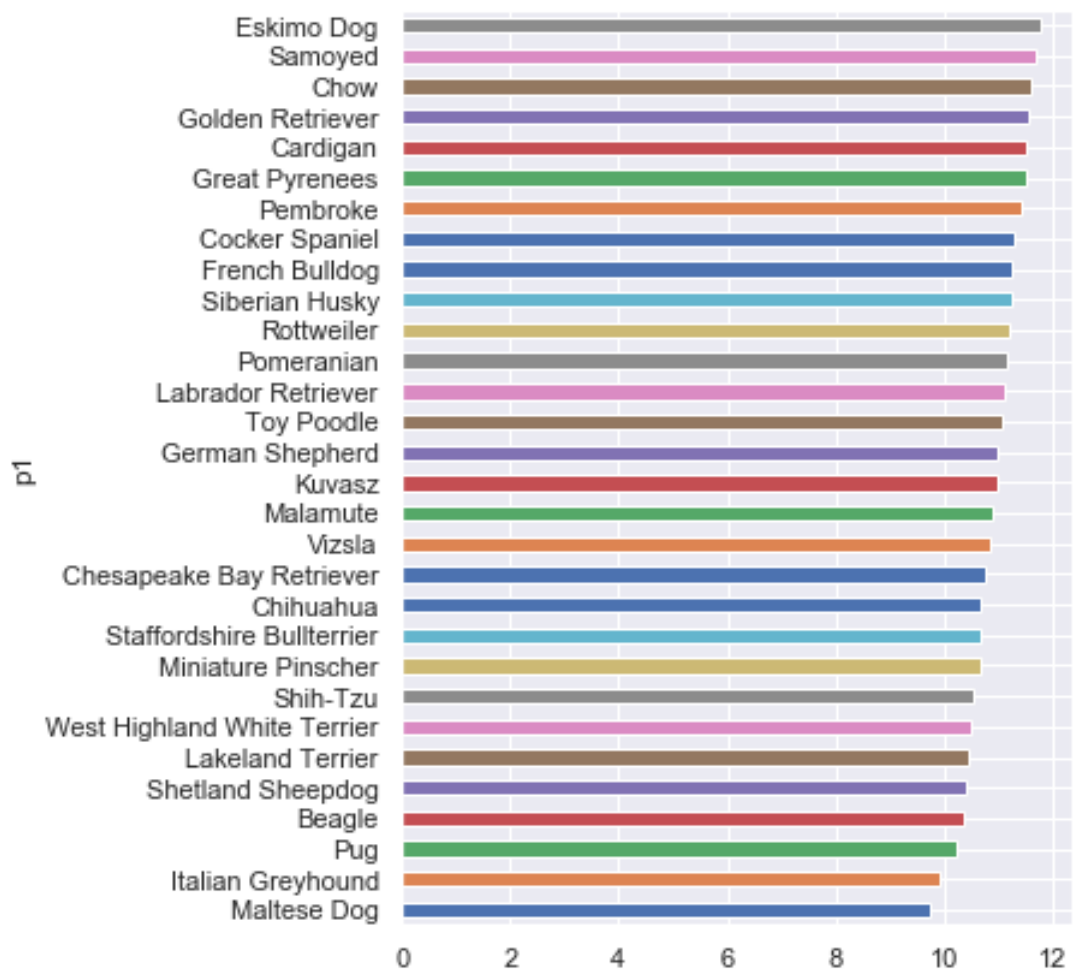


从结果上看，**favorite** 与评分有一定的正相关关系。狗的评分越高，相应的 **favorite** 数就越多。

#### ✧ 品种与狗评分的关系

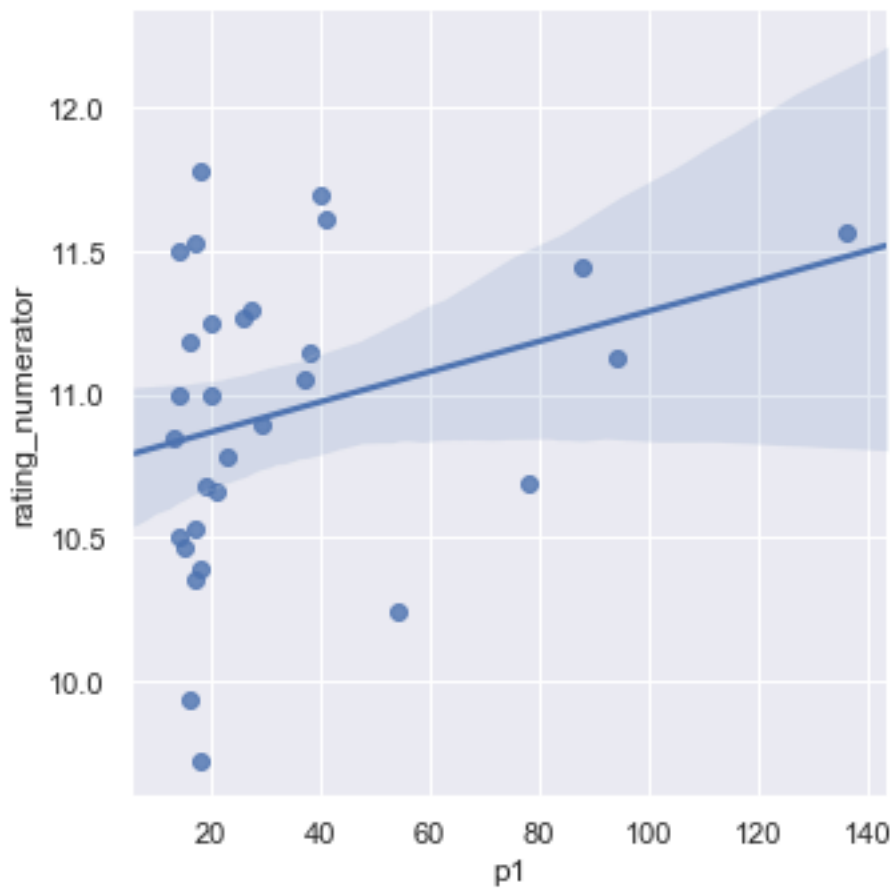
由于狗的品种太多，只选择数量排名前 **30** 的品种进行分析。

首先对不同种类狗的评分进行可视化分析：



结果显示，评分高的狗品种 Top5 是：爱斯基摩犬，萨摩耶，松狮犬，金毛猎犬和 Cardigan 柯基。看来金毛作为出现频次最高的狗评分也是相当高的。

然后分析某个品种狗的数量与评分之间的关系：

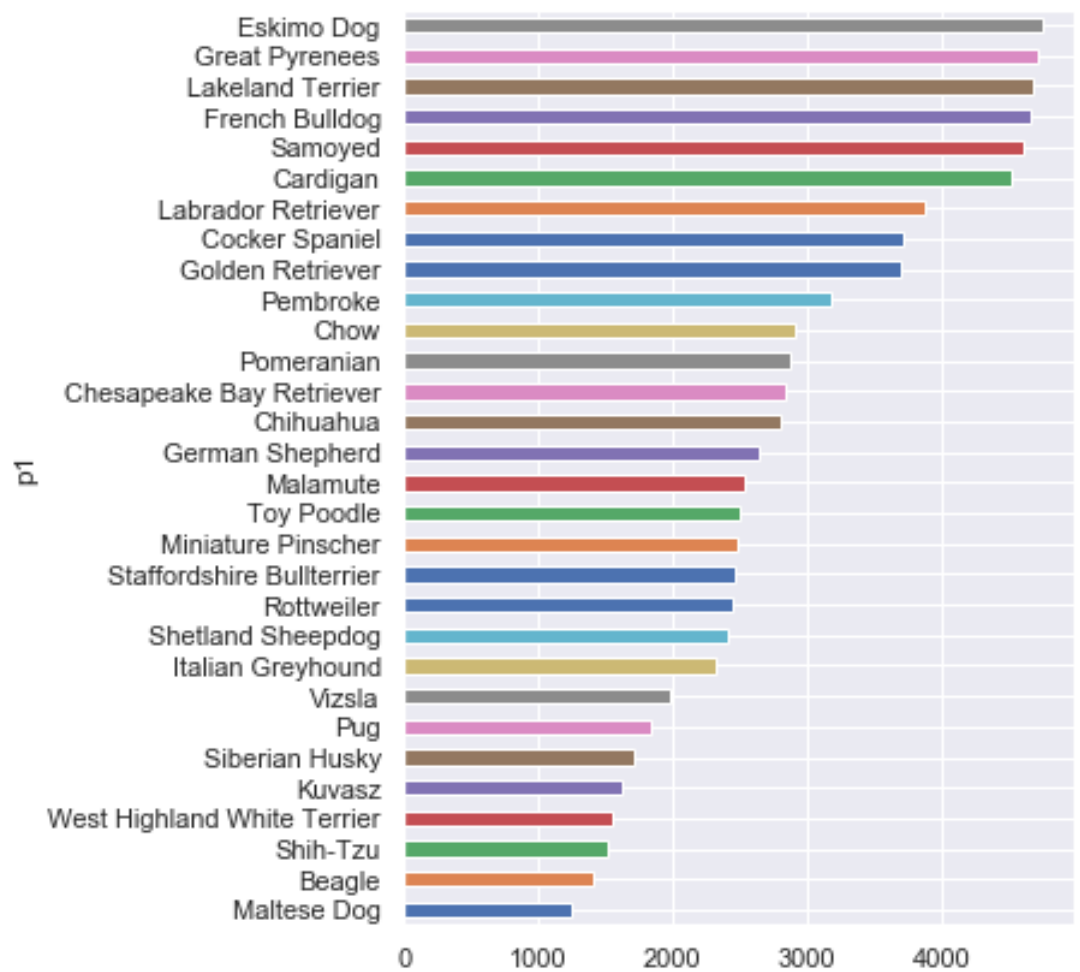


从上图看出，某个品种的狗的数量与其平均得分之间呈正相关。  
即推主评分较多的狗品种其分数相对较高。

✧ 品种与转推和 **favorite** 之间的关系

◆ 品种与转推之间的关系

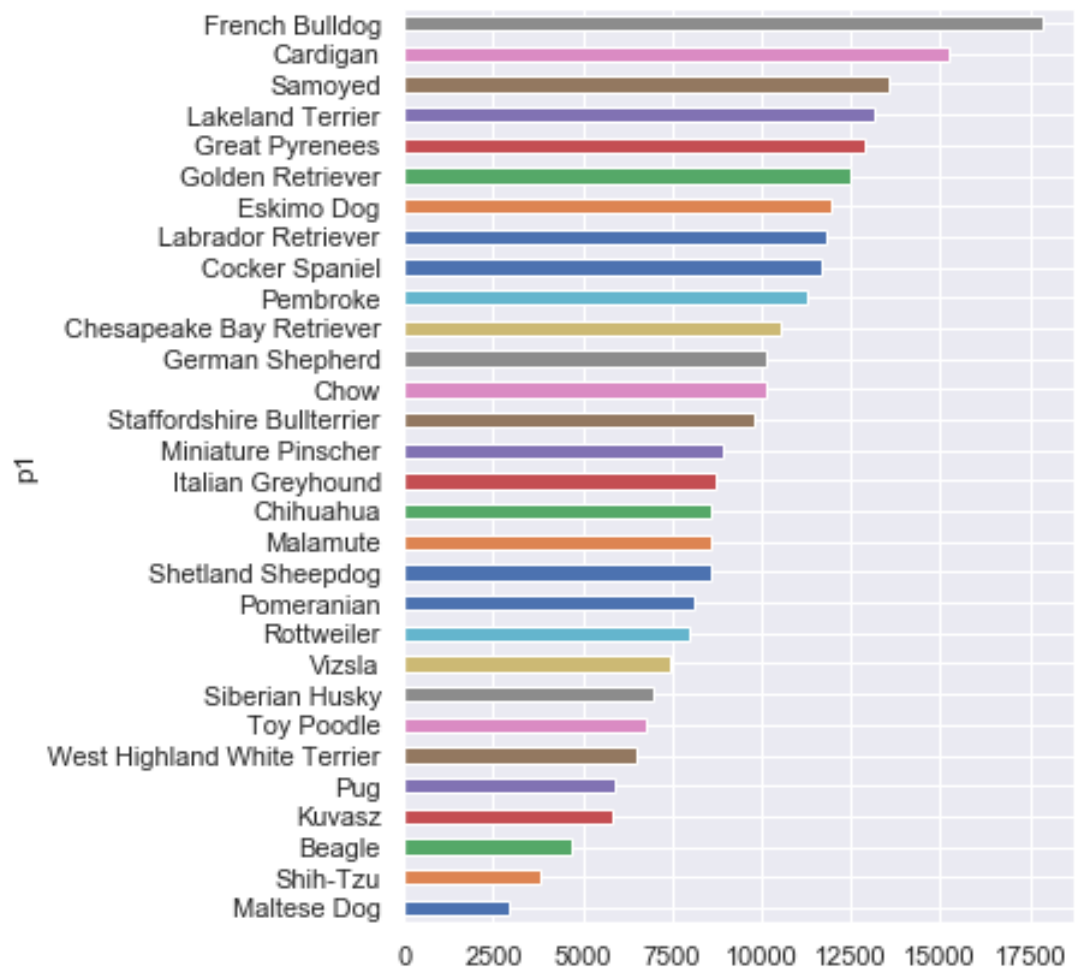
首先是不同品种狗的转推数对比：



结果：转推数前 5 的狗品种是：爱斯基摩犬，大白熊犬，湖畔梗，法斗，萨摩耶。

品种与 favorite 之间的关系





结果: favorite 数前 5 的狗品种是: 法斗, Cardigan 柯基, 萨摩耶, 湖畔梗, 大白熊犬

## 总结:

- ✧ 狗的名字出现最多的的是 Oliver, Charile, Cooper, Tucker, Lucy
- ✧ 小狗 (pupper) 最多。
- ✧ 推主给大多数狗的评分数 10-13 分, 绝大多数都在 2-14 之

间。

- ✧ 推主评分的狗种类 Top5 是：金毛，拉布拉多，Pembroke 柯基，吉娃娃，八哥
- ✧ 转推数以及 favorite 与评分有一定的正相关关系。狗的评分越高，相应的转推数和 favorite 就会越多。
- ✧ 评分高的狗品种 Top5 是：爱斯基摩犬，萨摩耶，松狮犬，金毛猎犬和 Cardigan 柯基。看来金毛作为出现频次最高的狗评分也是相当高的。
- ✧ 转推数前 5 的狗品种是：爱斯基摩犬，大白熊犬，湖畔梗，法斗，萨摩耶。
- ✧ favorite 数前 5 的狗品种是：法斗，Cardigan 柯基，萨摩耶，湖畔梗，大白熊犬。
- ✧ 转推数和 favorite 与这个品种狗的数量存在一定的正相关关系，即数量较多的狗品种对应的转推数也较多。
- ✧ **综合以上结果，我认为推主最喜欢的狗是金毛，大众最喜欢的狗是法斗。**



