

云吸狗也有啊哈时刻

——一颗吃饱没事干强行撸猫老被咬的卤蛋对搞笑萌宠推特WeRateDogs的过度分析

仅以此文献给高冷狂傲拽炸天的蛋蛋：



（一脸怨念地）如果不是因为我每次去撸你你都狠狠咬我的手，这个项目早八百年前就撸完了（cai guai）~

1. WeRateDogs是什么？可以补钙吗？



WeRateDogs是推特上受欢迎的萌宠推主之一。其主要风格为搞笑，具体表现形式为使用推文为一张狗狗（well，理论上都是）的照片进行十分制的打分。尽管其满分为10分，但推主常常打出10分以上的高分（13，14等），并一本正经的配上推文说明理由，十分有趣。推主有时也会一本正经给一些并不是狗的照片打分，喜剧效果十足。某些时候某些狗狗会被刻意打出低分，同样配上令人捧腹的推文。

WeRateDogs创建于11月15日。创始人时坎贝尔大学的（辍）学生Matt Nelson，当时读该校高尔夫管理专业的大二学生。当时，他和他的朋友们一起在一家苹果蜜蜂餐厅(Applebee's) 的餐桌旁，用WeRateDogs账号发出了第一条推特。

如同天底下所有真实的创业故事那样，这个独特的推特账号的发展并非一帆风顺。15年11、12月是最初的激情，很快挑战便接踵而至：

- 大学繁重的课业使得Nelson疲惫不堪，严重睡眠不足，甚至不得不放弃了高尔夫球（well，作为一名高尔夫管理专业的学生，这好像有点，emmm.....too much）；
- 16年2月，甚至有人给他下套，声称WeRateDogs侵犯了图片版权；
- 16年春季，他招聘并最终开除了两名员工；其中一位还想买下他的账号。

这里当然没有主人公强撑这，凭着极度的自信、过人的胆识强撑着运营，并最终成功逆袭的故事。身心俱疲的Nelson当然想过放弃。他联系了John Ricci，一位拥有丰富社交媒体运营经验的市场人，想要卖掉账号一了百了，没想到后者居然拒绝了。Rucci看到了账号的潜力，极力劝说Matt改变方式坚持运营，并在16年某个未知的时间点开始帮助他运营。Rucci每天帮Matt筛选出20-30张各地粉丝发来的请求打分的照片，而Nelson则挑出两张，分别在中午和晚上8：00左右发出即可。慢慢的，Nelson重新找到了节奏。17年某个未知的时间点，升入大四的Nelson退学，全心投入推特账号的运营。

现在，WeRateDogs主要依靠自营**电商**盈利，向通过推特账号积攒的粉丝售卖写有推特账号标志性语录的狗狗围巾、卫衣等。Nelson自己透露，每个月的收入在1万美元左右。

综合自下列报道

- [This 21-Year-Old Quit College to Rate Dogs on the Internet. He's Now Making Six Figures](#),
- [Jerk's One-on-One with the Guy Who Makes a Living Rating Dogs](#),
- [Pet Project](#),
- [WeRateDogs Is Coming to Cambridge](#),
- [I Started the WeRateDogs Twitter Account and Here's My Career Story](#),
- [They're Good Dogs Brent](#)

2. 我们为啥闲着没事儿干会对一个推特账号感兴趣？我们究竟看上它哪点了？以及我们打算做什么？

2.1 为什么会吃饱了没事干研究一个推特账号

Well，首先，这是一份作业。命题作文的命题如此，不研究它还能研究啥。

不过嘛，这个推特账号也不是特别没有意思。显然，WeRateDogs误打误撞探索出的这一定位十分有趣。如果我们拍一拍脑门调取一下我们的（未经检验、直觉构成的）常识，我们不难认为在传统上，1）萌宠推主千千万万，2）绝大多数萌宠的推文都停留在“啊！好萌！”的层次上，3）萌宠推主们之间的竞争主要依靠谁更萌获胜，4）这样的生态体系显然很容易导致严重的同质化，进而演变为洗稿和相互抄袭；5）给狗狗评价和打分的博主又太过严肃显得不近人情；6）传统的搞笑博主又几乎没有与宠物结合起来的。

通过直觉，我们不难得出马后炮式的经典结论：系统性的发掘狗狗的萌点，并通过带有喜剧色彩的文案强化，最后配上打破传统令人捧腹的打分和评价，令人耳目一新。这并不令人意外（说得就好像我们也能搞出来一样）。

但是，正如开心麻花的喜剧，在让观众开怀大笑之余，也有商业和受欢迎程度的考量一样，WeRateDogs的内容质量显然也是有部分更受欢迎的。问题是，**究竟是怎样的内容更受欢迎呢？**

在进行分析之前，我们需要明确，什么是“更受欢迎”。传统意义上，我们需要知道一条推特在发出的时候有多少关注者，这条推特有多少浏览量，多少点赞数和多少转发数，来一步一步计算转化比例。在这一框架中，更多的浏览量可能意味着火爆，但运营者们显然更关注点赞和转发的数据。

在这篇过度分析中，我们将围绕两个维度进行受欢迎程度的探索：

- **点赞量**，在拿不到关注量和浏览量这两个数据的前提之下，点赞量(数据集中叫favorite,Teitler上实际叫like)是最能直接反映一条推文得到关注和赞赏的指标；
- **转赞比** (Retweet/Favorite Ratio)，衡量的是转发用户占点赞用户的比例，用于衡量核心用户占整体点赞用户的比例；

转赞比可以用来衡量核心用户占比的理论依据如下：

- 一个用户在社交网络上发出的内容，本质上都是他/她的社交货币；
- 社交货币是建立人设的一部分，是用户与周遭用户交流的一种前置条件，因此相比起点赞，转发更加困难，因其牵扯到个人形象问题；

- 社交货币依靠转发和原创内容生成；点赞不涉及分享，绝大多数社交网络（嗯我没有说微信朋友圈）里看不到点赞的详细信息，并不构成社交货币；
- 转发，意味着用户愿意将推文内容作为自己人设的一部分；这可能意味着他/她是这一话题的核心粉丝，并且不介意周围人知道这一点，哪怕，在此案例中，这枚社交货币是娱乐用途。

为了研究我们感兴趣的点，我们做了什么

3. 推特账号运营的基本情况一览

在尝试使用数据解答任何问题之前，需要使用数据对全局进行大概的了解。不然任何分析都是无源之水无本之木。

3.1 WeRateDogs的哈哈哈时刻

最棒宝贝：勤奋好学大柯基！14分！（最高评分）



This is Cassie. She is a college pup. Studying international doggo communication and stick theory. 14/10 so elegant much sophisticate <https://t.co/t1bfwz5S2A>

最佳人气：妇女之友！（最多点赞）



Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10 <https://t.co/nIz3FtorBc>

最多铁粉：超萌小柯基！（最高转赞比）



Retweet the h*ck out of this 13/10 pupper #BellLetsTalk <https://t.co/wBmc7OaGvS>

欸喂喂喂，Matt你这样构成诱导分享啊！不怕被平台封吗？！

特别·最具惊喜奖：最新发现的神秘品种！



Cool dog. Enjoys couch. Low monotone bark. Very nice kicks. Pisses milk (must be rare). Can't go down stairs. 4/10 <https://t.co/vXMKrJC81s>

以及，特别·最佳粉丝奖（& 最少铁粉奖）：天生慧眼Kelvin Hill！



Kelvin Hill
@0_kelvin_0



[@dog_rates](#) I love all animals, but I'm starting to question the legitimacy of we rate dogs because I'm seeing more than just dogs here

3/11/17, 5:23 PM



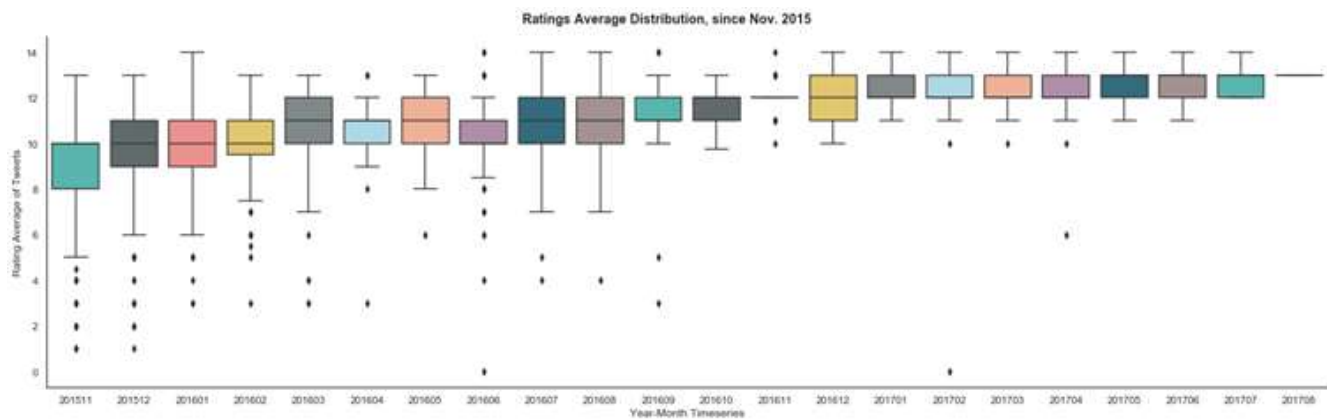
Kelvin Hill @0_kelvin_0 · 23m



[@dog_rates](#) never mind I was mistaken and I will retract my earlier statement they are all good dogs

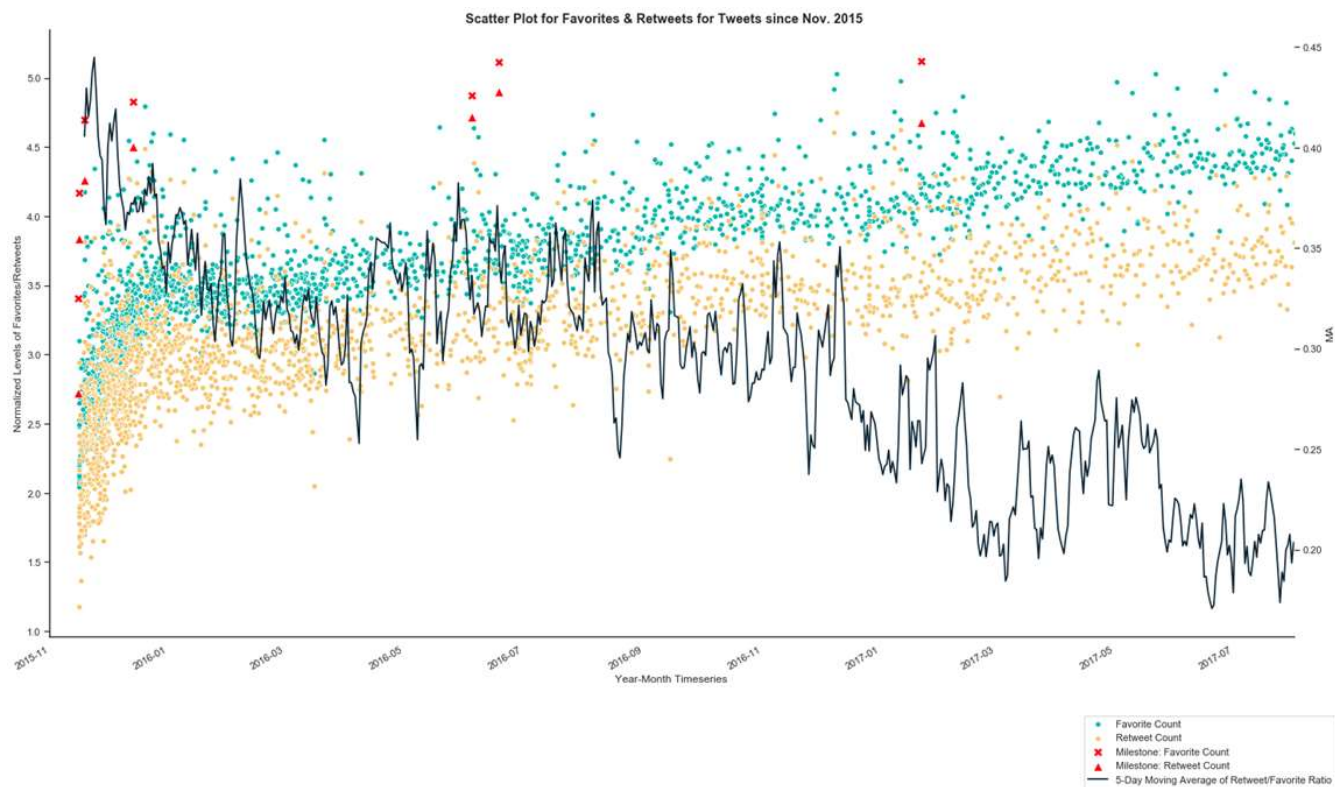


3.2 分数越来越高，当然是因为超棒狗狗越来越多啦！



将账号打出的分数按月汇集，绘制为箱型图。平均分数越来越高；单月的分数区间越来越小

3.3 成长的烦恼：转赞率一路下滑

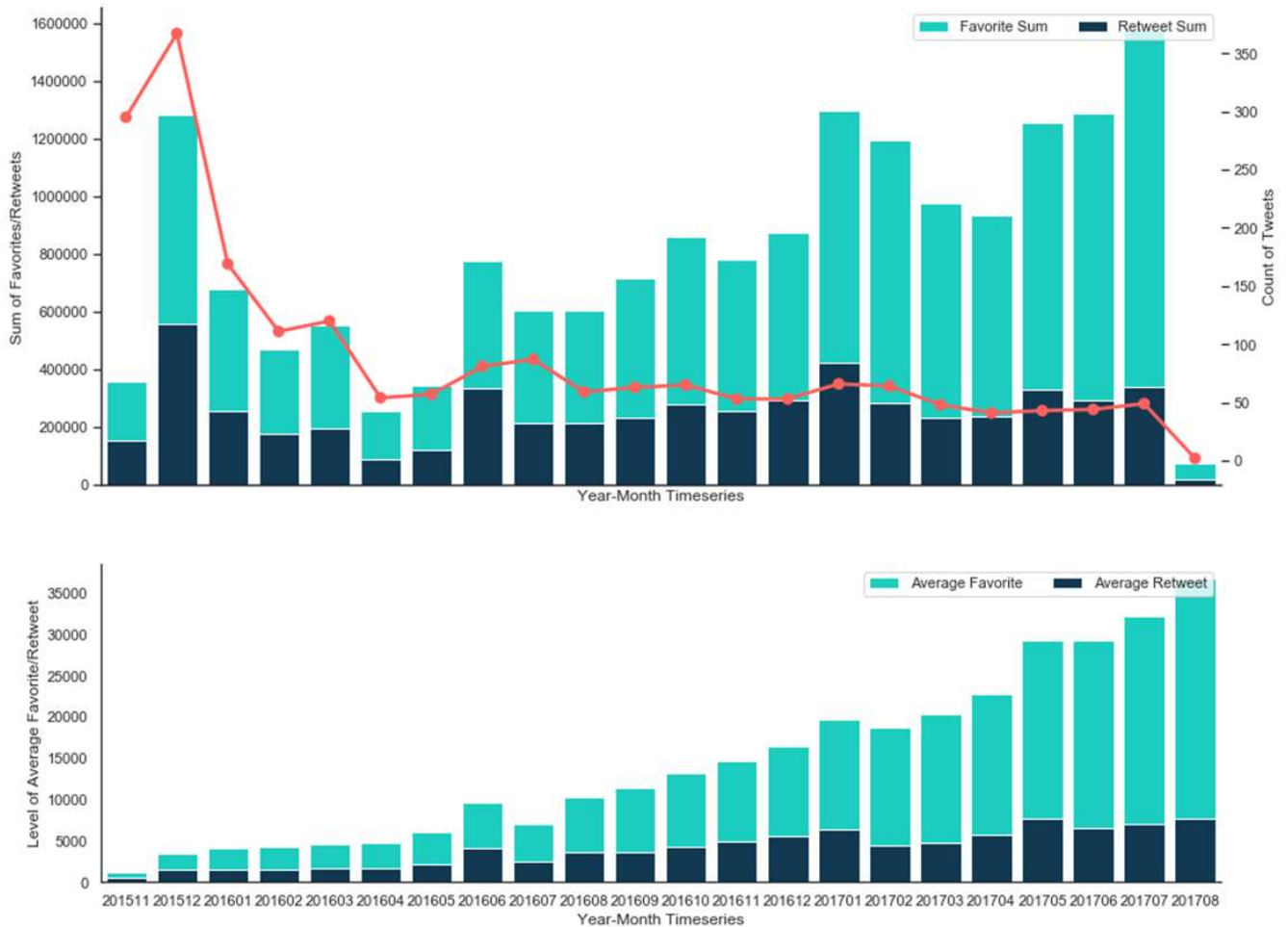


将每一条推特信息的点赞量（绿）和转发量（黄）取10的对数后绘制成散点图；将转赞率取5日移动平均值绘制为折线图；取账号点赞量创下新高的推特，用红色做特别标注。三者沿同样的时间序列排序。

随着时间的推移，虽然账号的点赞量和转发数都在上涨，但新高间隔越来越长，转赞比一路下降的趋势不可逆转。

3.4 爆发—跌倒—站稳脚跟：艰难的16年4月，账号历史上的至暗时刻

Favorites, Retweets & Retweet/Like Ratios per Month: Total vs. Average

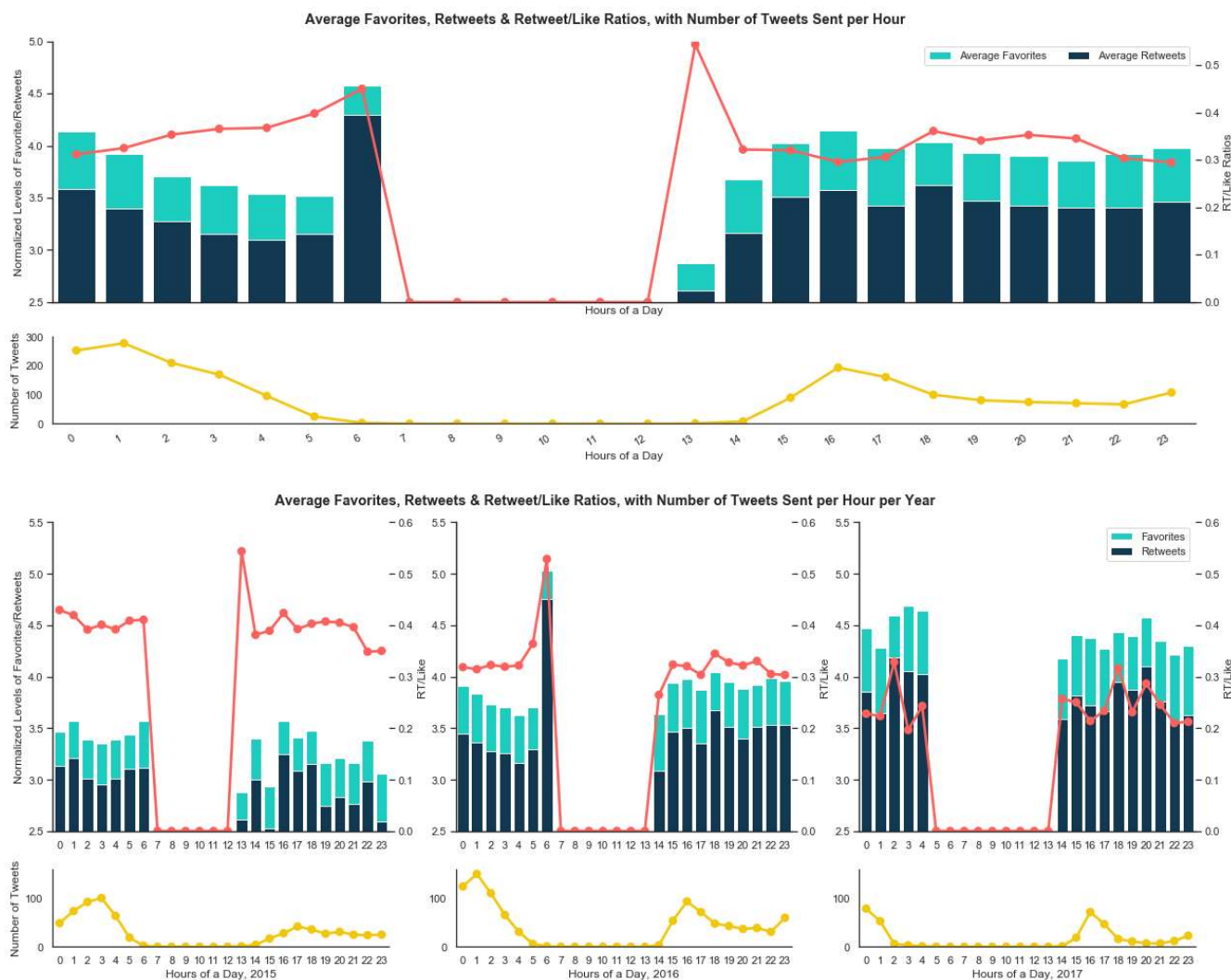


将账号的推特条数（红色折线），点赞量和转发量按月汇总，分别绘制加（上）和和平均值（下）的柱状图。

在初期的爆发之后，16年春季推特账号的点赞量和推特条数不断下滑。结合媒体的报道，当时账号确实正在经受初期的失败和团队动荡。不过在如此的压力之下，Matt的推文质量（平均点赞量和转发量）还能保持，非常厉害！

4. 推文几点，星期几发都行：客观因素对推文受欢迎程度影响不大

4.1 Matt Nelson是个夜猫子，越夜越疯狂那种



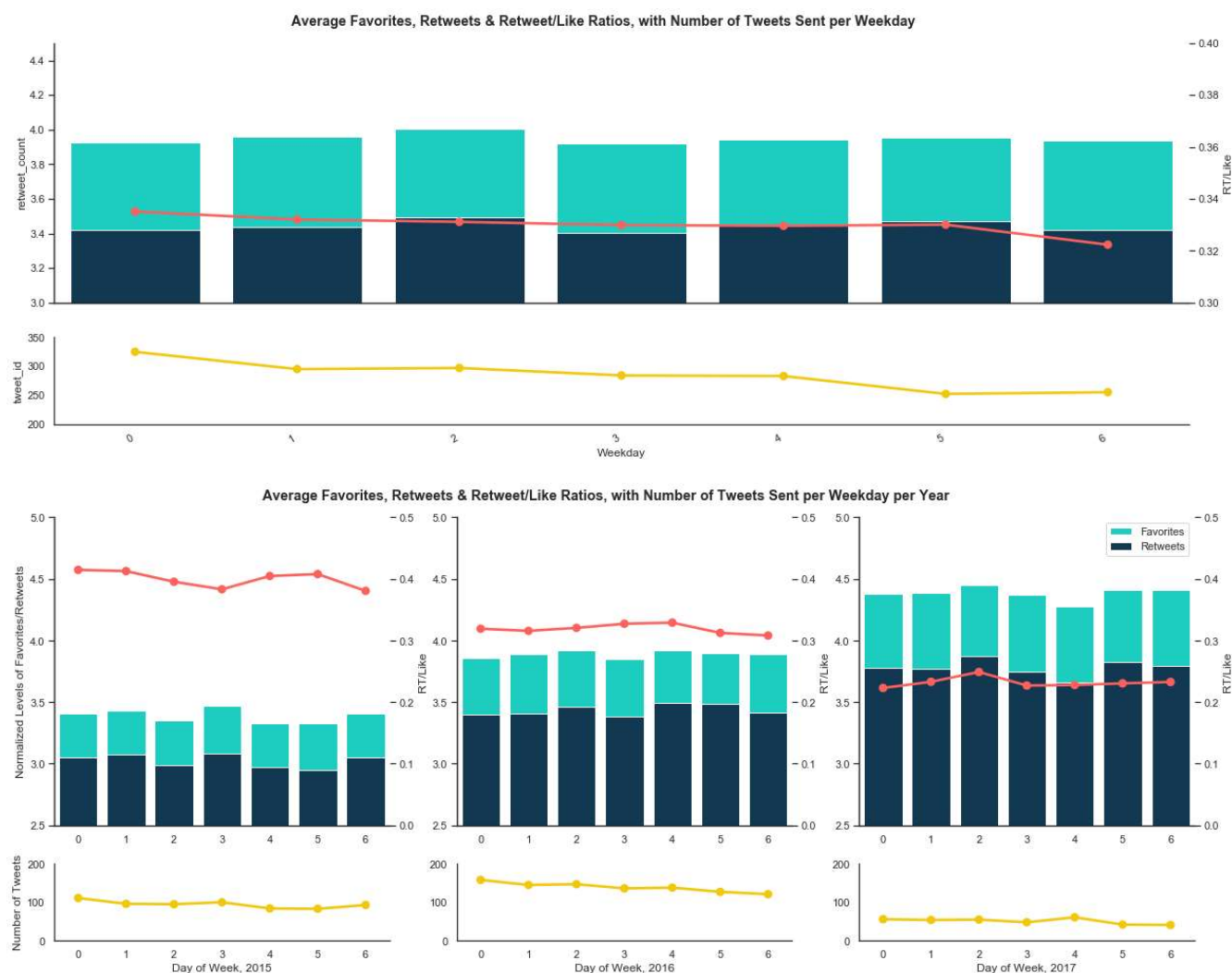
将推文数据按小时聚合，绘制点赞量、转发量，转赞比和推文条数。

我们发现：MattNelson喜欢在下午和晚上发推。

不过，为他创下了最高点赞量的式凌晨6点的一条推文。就是这条推文，使我们不得不给这幅图的y轴主坐标取对数。不得不说，他老人家真的是越夜越疯狂啊~

不过，17年他再也没有在5点到13点之间发过推特了。被妈妈骂了？

4.2 周一生产力爆棚，但对推文的受欢迎程度影响不大



将推文的发帖量，平均点赞量和转发量按星期汇集。

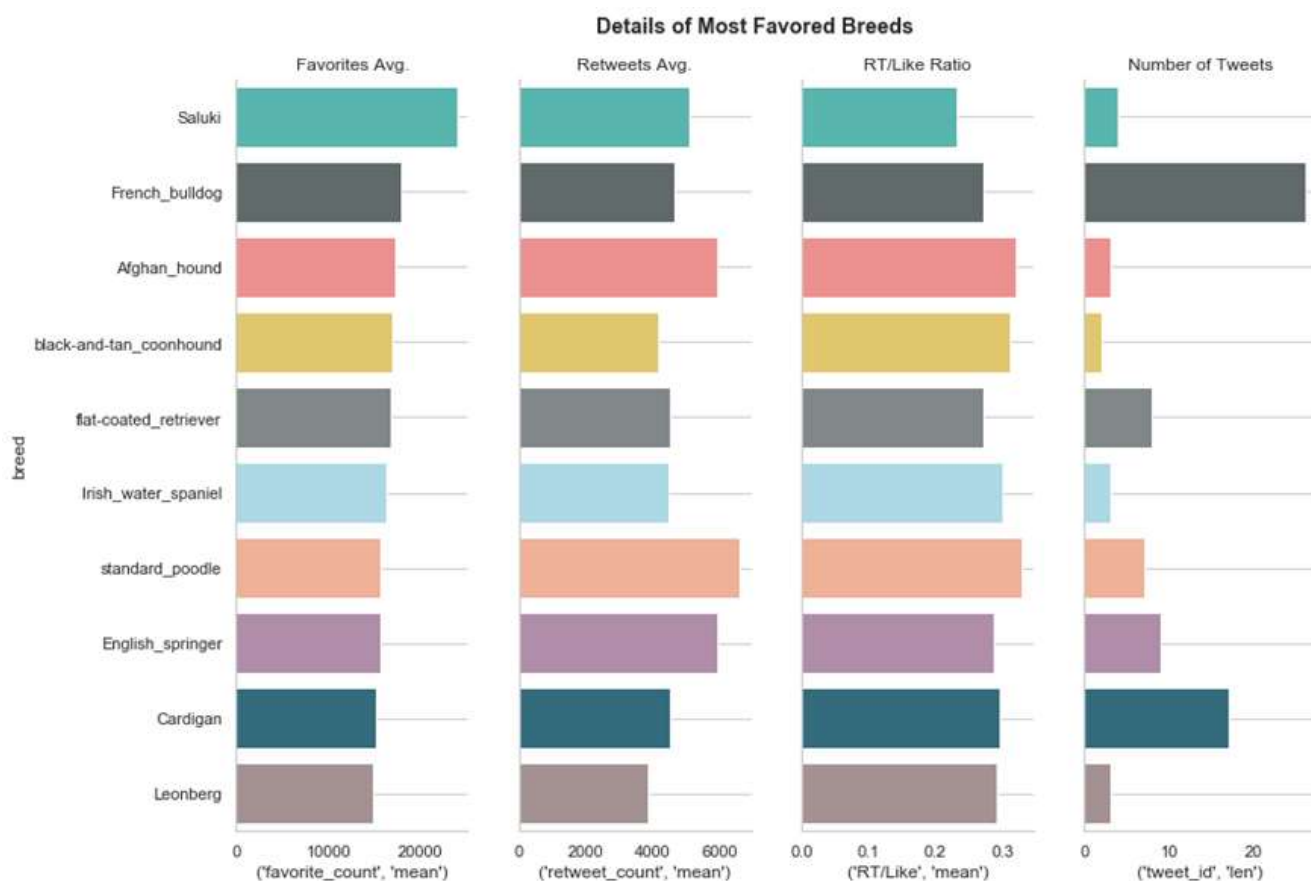
长期来看：周一Matt的发帖量略高，周三平均点赞量更高，但总体趋势平缓。我们不认为这其中蕴藏着什么秘密。

当然，我们有可能是错的。如果将数据去趋势化再聚集，说不定就是另外一个故事了。

5. 众汪平等，小狗疲劳，与高低有别——对推文特征的观察

在我们提取了一系列推文特征之后，我们最终得出了这三个有意思的结论。

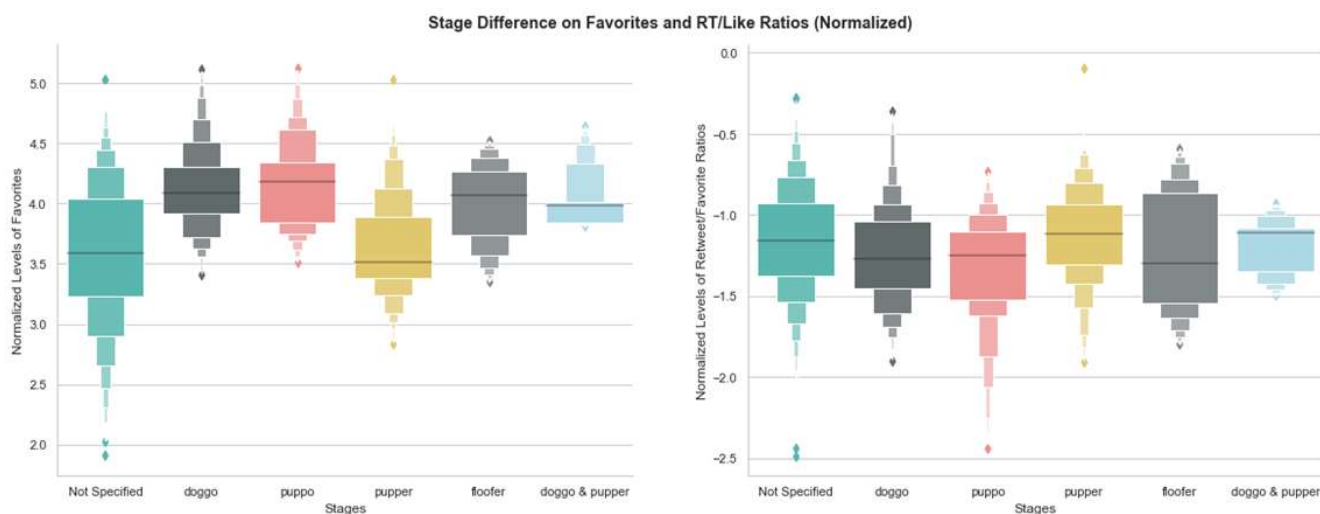
5.1 众汪平等



筛选出数据集中被机器学习判断“一定是狗”的条目。按照品种字段聚合，计算平均点赞量，平均转赞比和推特条数。选出前10名，按照平均点赞量由高到低排列。可以看到：萨卢基猎犬，法国斗牛犬等品种有相对更高的平均点赞量。

但是，我们进一步的检验表明，这些品种中存在的平均点赞量和转赞比的区别并没有统计学意义上的区别：根据已有的条件判断，随着推特条数的增多，这些品种的平均点赞量和平均转赞比很可能趋于一致。也就意味着，我们没有足够的证据认为他们的受欢迎程度是不同的。

5.2 小狗疲劳：小奶狗居然不吃香



推主Marr Nelson在表述狗狗的时候发明了如下词汇：

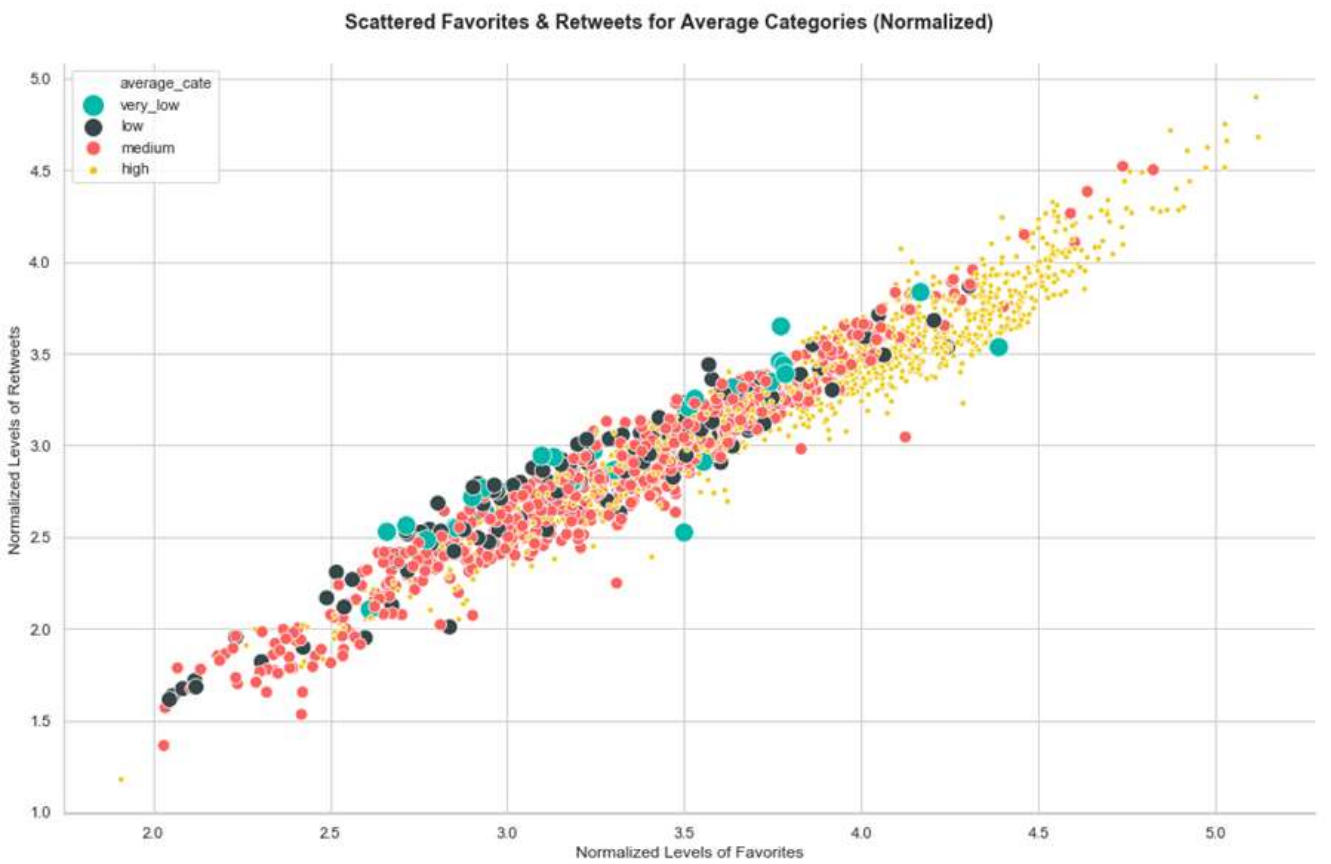
- Doggo，可以简单的理解为大狗（尽管作者表示这并不代表大狗就不是宝宝了）

- Pupper, 可以简单的理解为小狗，主要是体型小，一般年龄也比较小（但这不代表大狗就不是宝宝了）
- Puppo, 可以简单的理解为青春期的狗狗，介于大狗和小狗之间（也可能是表现得像pupper的doggo）
- Floofer, 可以简单的理解为毛好看的狗狗。

这些词汇主要用于形容不同狗狗的生长状态。从词汇构成的角度来看，其主要由一个主词汇和后缀变体组成（例如 doggo = dog + go, pupper = pup + per），营造出一种喜庆但略显低龄化的语言现象（待补充）。上述所谓的“简单理解”，是指在尽可能不曲解原意的情况下，抛弃作者营造的一切搞笑成分，并尽量压缩理解难度。总而言之言而总之一言以蔽之，这个分类比较随意，亲爱的读者你大概知道这么个意思就行。

我们按照上述分类条件将推文汇总。我们发现**小狗和未说明的推特条目似乎没有其他分类平均点赞量高**。进一步的统计分析则支持了上述观察。看来在Matt这里萌萌哒的小奶狗不太行呀~

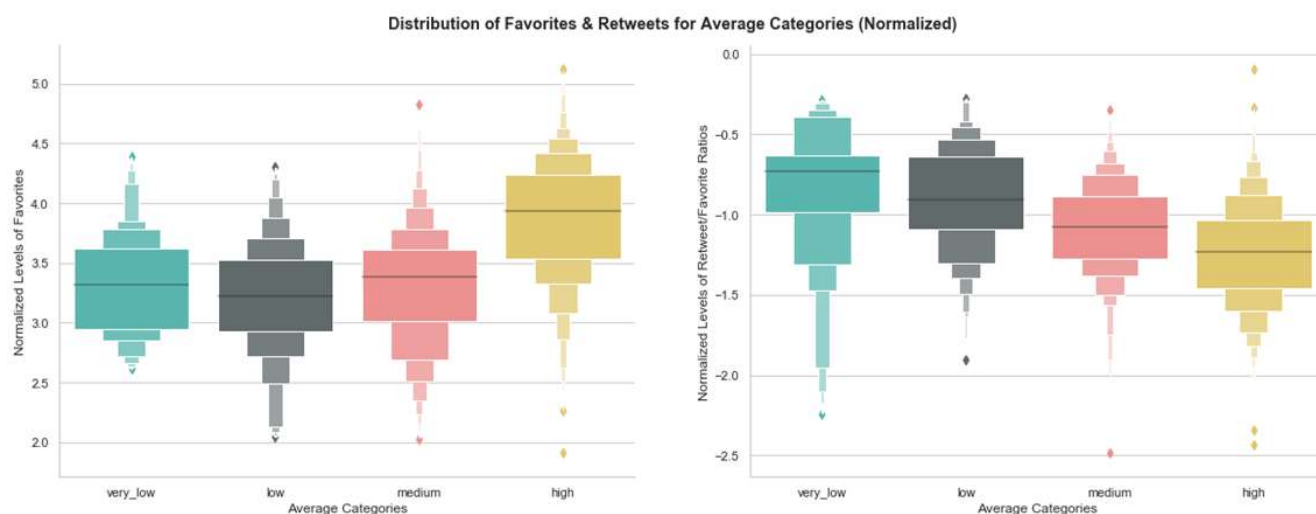
5.3 高低有别：低分有铁粉，高分更吸睛



我们将分数分为4个等级：

分数段	类别（中）	类别（英）
0-3.5	极低	Very Low
3.5-7	低	Low
7-10	中	Medium
10-14	高	High

以推文的点赞量和转发量分别为横、纵坐标绘制散点图，并表明其分数类别，可以看到不同的分数组在分布上有明显的差别。



以上述分数段分组，绘制点赞量和转赞比的箱型图。似乎高分组的平均点赞量更高，低分与超低分组拥有更高的转赞比。后续的统计学检验支持了我们这一结论：**低分有铁粉，高分更吸睛。**

这可能是由于低分组主要集中在账号创立早期，转赞比本身就较高的时代，拉高了平均值。下图显示了分数分组占比的逐年变化情况。

Stacked Percentage of Average Categories, per Year

