**Do Hashtags Really Matter?**

DS710 – Shawn Chapler

It has been quite an exciting week for subscription styling service, Stitch Fix, as it has confidentially filed for an IPO. Stitch Fix is a leader in a newer form of e-commerce which focuses on clothing, shoes, and accessories selected by a personal stylist conveniently delivered to the consumer’s door. With Stitch Fix having adopted data science as a core business function, I was interested in seeing how Twitter hashtag usage drove consumer interest as measured by retweet counts and favorite counts in this specific market.

I used Twitter’s REST API to capture data over a period of eight days. The tweeting activity on these particular hashtags was not active enough to capture tweets with the Streaming API. Here are the profiles of the companies targeted for this study via hashtag usage.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Joined | Tweets | Following | Followers | Likes | Hashtags Queried |
| Stitch Fix | July 2011 | 26.9k | 3,055 | 61.K | 23K | #StitchFix |
| Wantable | May 2012 | 14.5k | 4,950 | 43.2K | 11.6K | #Wantable |
| DailyLook | July 2011 | 4,736 | 91 | 17.5K | 1,711 | #elevateyourstyle  #NewInDL |
| Trunk Club | January 2008 | 12.5K | 1,348 | 69.6K | 7,040 | #TrunkClub |
| Stella & Dot | March 2009 | 12.4K | 1,273 | 55K | 11.1K | #stelladot  #stelladotstyle |

*\*August 4, 2017 Twitter Profiles*

***Is there a correlation between the number of hashtags used in a tweet and retweet counts?***

Analysis of the data reflects there is a correlation between number of hashtags and retweet counts.

In analyzing the data, there was a right skew in retweet counts specifically at two significantly higher retweet points as you can see in the comparisons of Figures 1 and 3 below. Most retweets counts trended between zero and ten. However, there were two data points (149 and 229) which accounted for over 36,000 retweets.

It became apparent bot activity was a likely driver. Drilling down further, individual tweets were generated in rapid date timestamp succession through multiple user accounts utilizing the same hashtag indices. Interestingly, 90% of the activity in the four hashtag category, where these two retweet outliers are found, was driven by Stitch Fix hashtags.

Comparing the blue and red data points in the graphs, StitchFix-focused tweets tend to contain more hashtags than its competitors even after the outlier retweets were removed.

***Is there a correlation between number of hashtags and favorite counts?***

The data reflects a relatively flat trend in the correlation between hashtags and favorite counts. Increased usage of hashtags does not appear to drive favorite counts. Also of note, bot-like activity did not influence favorite counts. Figure 2, which reflects the data trends with the outliers removed, is virtually identical to the graph containing bot-like activity above it in Figure 4.

***Is the mean number of hashtags containing #StitchFix the same as its competitors?***

A Welch Two Sample T-Test was performed on the data. Findings show that at a 95% confidence level, the mean number of hashtags for Stitch Fix (3.18) was not the same as its competitors (1.37) which are in-line with the blue dots in the graphs showing Stitch Fix tends to use more hashtags as compared to its competitors. The findings also held true when the same t-test was run on the data set excluding the data outliers.

***Are hashtags effective as a marketing tool?***

It depends. If an organization wants to get a sense of consumer sentiment without bot-like activity influences then monitoring favorite counts is much more important and less reliant on hashtags for generating interest.

Stitch Fix’s increased use of hashtags appears to reflect a belief that more hashtags result in more retweets which I would argue is comparable to direct mail or email campaigns where there are expectations of relatively low consumer response rates per piece.

Perhaps not surprisingly, Stitch Fix’s tweet count is more than double that of its Twitter-active competitor, Stella & Dot. The bot-style activity apparent in the StitchFix data seem to indicate, Stitch Fix may not be targeting and/or expecting a significant number of individual consumers. A review of these specific Twitter accounts reflect pages devoted to product retweets of various sorts.

In marketing, a primary goal is to measure market interest. Whether an actual consumer reads bot-activity posted on these Twitter pages is debatable. Interestingly, Stella & Dot and Trunk Club follow some of these “digital influencers” pages. And that’s something I’d take note of.