

Twitter Data set

We work with a twitter dataset related to the 2015 Super Bowl. The tweets have been collected starting from 2 weeks before the game to a week after the game. The dataset contains tweets with the following hashtags defined in the table:

Hashtag Statistics

In this part, we compute some of the statistics of the hashtags listed in the introduction. The statistics computed are:

1. Average number of tweets per hour (ATH)
2. Average number of followers of users posting the tweets (AFU)
3. Average number of retweets (AR)

There are a couple of ways to compute the average number of retweets. The total number of citations as the total number of retweets and then averaged the total number of citations over the total number of tweets with the hashtag. The statistics for the hashtags are given below:

HashTag	ATH	AFU	AR
#gohawks	193.5	1709.4	2.01
#gopatriots	38.4	1558	1.4
#patriots	499.4	1859.7	1.783
#nfl	279.55	4376.6	1.54
#sb49	1419.89	2243.8	2.5
#superbowl	1401.25	4221.4	2.39

The plot depicts most activity happens during the super bowl event. Hence it motivates for task in asset 2.

