

WHAT ARE RETWEETS?

A retweet allows Twitter users to rebroadcast specific tweets of interest by incorporating all or part of the original tweet into their own.

We hypothesize that retweets play a role in introducing new connections to a network, so that individuals who do not follow an individual may see retweets of their tweets and eventually become followers.



WHAT ARE RETWEETS?



A Retweet is an action taken by a Twitter user to share another user's Tweet without alteration, using Twitter's explicit Retweet functionality.

A Retweet retains information about the user who posted the original Tweet, as well as the user who Retweeted them.

Retweets are an important part of Twitter's platform – they permit content to be shared rapidly and with attribution, and are the most easily measured form of content engagement on the platform. Many social analytics tools use the number of Retweets a particular Tweet receives in calculating its impact or reach (i.e. its importance). However, to do so, your app must be able to accurately identify Retweets.

HOW TO IDENTIFY A RETWEET

For example, here is an excerpt from the root-level of a Retweet:

```
{
  "id": "tag:search.twitter.com,2005:299935329132105728",
  "objectType": "activity",
  "actor": {...},
  "verb": "share",
  ...
}
```

And here is an original Tweet:

```
{
  "id": "tag:search.twitter.com,2005:403224522679009280",
  "objectType": "activity",
  "actor": {...},
  "verb": "post",
  ...
}
```

INTEGRATING RETWEETS

Each Retweet contains two layers: 1) an outer layer, which holds data related to the Retweet action itself, and the user who performed the Retweet, and 2) an inner layer which holds data about the original Tweet, including data about the user who posted it. The outer layer exists at the root-level of the JSON Tweet object – the root-level – while the inner layer is contained within the root-level "object" field. More clearly, the Retweet is a Tweet object, which contains another whole Tweet object within the "object" field.



EXAMINING RETWEET PATTERNS

The approach we'll take to find the most popular retweets is to simply iterate over each status update and store out the retweet count, originator of the retweet, and text of the retweet if the status update is a retweet.

```
retweets = [
    # Store out a tuple of these three values ...
    (status['retweet_count'],
     status['retweeted_status']['user']['screen_name'],
     status['text'])

    # ... for each status ...
    for status in statuses

    # ... so long as the status meets this condition.
    if status.has_key('retweeted_status')
]

# Slice off the first 5 from the sorted results and display each item in the tuple

pt = PrettyTable(field_names=['Count', 'Screen Name', 'Text'])
[ pt.add_row(row) for row in sorted(retweets, reverse=True)[:5] ]
pt.max_width['Text'] = 50
pt.align = 'l'
print pt
```

Count	Screen Name	Text
23	hassanmusician	RT @hassanmusician: #MentionSomeoneImportantForYou God.
21	HSweethearts	RT @HSweethearts: #MentionSomeoneImportantForYou my high school sweetheart 💕
15	LosAlejandro_	RT @LosAlejandro_: ¿Nadie te menciona en "#MentionSomeoneImportantForYou"? JAJAJAJAJAJAJAJAJAJAJAJAJAJAJAJAJAJAJA Ven, ...
9	SCOTTSUMME	RT @SCOTTSUMME: #MentionSomeoneImportantForYou My Mum. Shes loving, caring, strong, all in one. I love her so much 💕💕💕
7	degrassihaha	RT @degrassihaha: #MentionSomeoneImportantForYou I can't put every Degrassi cast member, crew member, and writer in just one tweet....

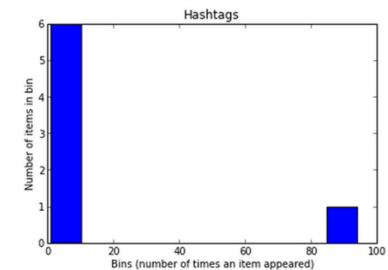
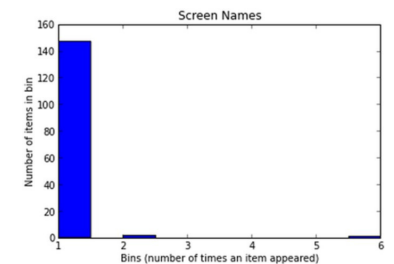
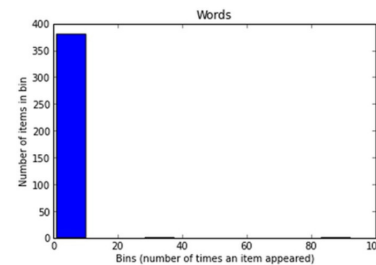
HI STOGRAMS

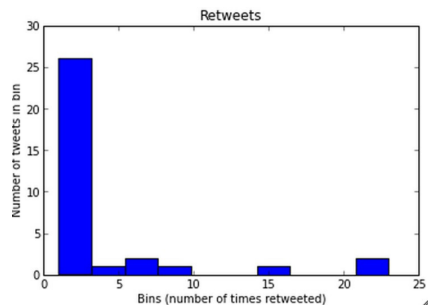


FREQUENCY REPRESENTATION

A histogram is designed for precisely this purpose and provides a convenient visualization for displaying tabulated frequencies as adjacent rectangles, where the area of each rectangle is a measure of the data values that fall within that particular range of values.

A histogram gives us insight into the underlying frequency distribution, with the x-axis corresponding to a range for words that each have a frequency within that range and the y-axis corresponding to the total frequency of all words that appear within that range.





Your Tweet activity

Your Tweets earned **1,825 impressions** over the last **24 hours**



CODE TO GENERATE HISTOGRAMS

We use IPython Notebook

```
for label, data in (('Words', words),
                   ('Screen Names', screen_names),
                   ('Hashtags', hashtags)):

    # Build a frequency map for each set of data
    # and plot the values
    c = Counter(data)
    plt.hist(c.values())

    # Add a title and y-label ...
    plt.title(label)
    plt.ylabel("Number of items in bin")
    plt.xlabel("Bins (number of times an item appeared)")

    # ... and display as a new figure
    plt.figure()
```

CODE TO GENERATE HISTOGRAMS

```
# Using underscores while unpacking values in
# a tuple is idiomatic for discarding them

counts = [count for count, _, _ in retweets]

plt.hist(counts)
plt.title("Retweets")
plt.xlabel('Bins (number of times retweeted)')
plt.ylabel('Number of tweets in bin')

print counts
```



Engagement rate
3.3%



Jan 21
4.2% engagement rate

Link clicks
774



Jan 21
1 link click

On average, you earned **28 link clicks** per day

Retweets
2.2K



Jan 21
6 Retweets

On average, you earned **28 Retweets** per day