**GT Master Modeler Competition: Glossary**

**Terminology:**

* **Following/Follower:** Following someone on Twitter means: You are subscribing to their Tweets as a follower. Their updates will appear in your Home tab. That person is able to send you Direct Messages.
* **Influencer:** A Twitter user who works with brands to promote products or events throughout the Twittersphere by communicating with his/her followers. An influencer is usually compensated for his/her promotional tweets and is chosen due to their quantity and quality of followers.
* **Interaction:** An event where a Twitter user retweets another user’s tweet or mentions another user in a tweet. For the scope of this competition, we consider only retweets and mentions as interactions. Retweeted tweets have ‘RT’ in front of them, and mentions have “@ username” in them.
* **Active Follower:** Someone who has retweeted or mentioned a particular tweet from an influencer.
* **Passive Follower:** Someone who has *not* retweeted or mentioned a particular tweet from an influencer.
* **Hashtag:** A word or phrase preceded by a hash or pound sign (#) and used to identify messages on a specific topic
* **Verification:** Verification is currently used to establish the authenticity of a Twitter user (person or brand)*.* If a Twitter user meets the criteria put forward by the verification team, his/her profile is said to be a “verified profile”.

**Terminology Visualized:**

Following

Passive Following (No Interaction)

Active Following (Interaction)

Followers of

Passive Follower

Passive

Follower

Followers of

Active Follower

**Influencer**

Followers

Of

Passive Follower3

Followers

Of

Passive Follower1

Followers

Of

Passive Follower2

Followers

Of

Active Follower3

Followers

Of

Active Follower2

Followers

Of

Active Follower1

Passive

Follower

3

2

Passive

Follower

Passive

Follower

1

3

Active

Follower

2

Active

Follower

1

Active

Follower

**Influencer1**

Active

Follower

**Influencer Marketing:**

Influencer marketing is a strategy driven approach, connecting brands with todays:

* Social Celebrities
* Innovative Creators
* Influential brand advocates

**Identifying an Influencer:**

360i has an Influential Marketing Team. They identify the influencers through the following four-step approach:

* Use Internal Research Tools: From reviewing the results of internal

searches on tools like Cision, Crimson Hexagon and 360i’s own proprietary database of 10,000+ influencers, the team comes up with an initial list of emerging influencers.

* Do Detective Work: By knowing the trending/hot topics, research is

directed towards identifying the relevant influencers and reviewing their ‘blog rolls’ to understand who is influencing them.

* Measure their influence: Using various tools such as Alexa, Compete,

Klout, etc. the team analyses how actively the shortlisted influencers post messages, photos, videos on various social platforms. They also take into account the number of fans, followers, posts, and interactions.

* Apply the Lens of Brand Relevancy: The Influencer Marketing team

tries to answer the following questions about the influencer:

* + - Are the bloggers/ shortlisted influencers brand-relevant?
    - Are they brand-detractors?
    - Do they mention brands?

After answering the questions, and finalizing the influencers, the next step is to build relationships (with the influencer), and vetting the influencer. The ultimate aim is to make user that the influencer fits into the lens of brand and campaign relevance.

**List of Influencers:**

For the scope of this competition, we consider **20 ‘Influencers’**. They are:

|  |  |  |
| --- | --- | --- |
| User ID | **User Name** | **Name** |
| 96697085 | AlexisKornblum | Lexi's Clean Kitchen |
| 23241541 | camillestyles | Camille Styles |
| 818155 | caro | Caroline McCarthy |
| 7121092 | Carolineoncrack | Caroline on Crack |
| 39940842 | concoctologist | Tricia Carr |
| 15452859 | DawnSandomeno | Dawn Sandomeno |
| 241844284 | freutcake | Leah Bergman |
| 873828446 | HBHarvest | HalfBakedHarvest |
| 213405395 | inspiredbycharm | Michael Wurm, Jr. |
| 18809624 | jennifer | Jen Rubio |
| 86369703 | jetsetfarryn | Jetset Farryn |
| 21586778 | kellyryanobrien | KELLY RYAN OBRIEN |
| 20253942 | lauratrevey | Laura Trevey |
| 172758851 | LizaUtter | Liza Utter |
| 483195215 | MacCocktail | MacCocktail |
| 15655027 | mattduckor | Matt Duckor |
| 22981009 | mollyyeh | molly yeh |
| 18478105 | stevedolinsky | Steve Dolinsky |
| 260773660 | thefauxmartha | Melissa Coleman |
| 26561883 | theyearinfood | KimberleyHasselbrink |

All the active followers, passive followers, followers of active/passive followers, connecting tweets, etc. revolve around these 20 Influencers.

**Table descriptions:**

1. **Tweet\_Data:**

This table contains the tweets of influencers, active followers, passive followers, and followers of active and passive followers. Some tweets may be missing, so do not assume that this data is complete. The challenge is to identify and attribute the tweets to the right creator.

Fields:

1. Creator\_ID: The unique numeric ID of the tweet creator
2. Tweet\_ID: The unique numeric ID of the tweet
3. Text: The content of the tweet
4. Language: The language of the tweet content
5. Time\_Stamp: The date and time of tweet creation
6. **User\_Profile\_Data:**

This table contains the user profile data of influencers, active followers, passive followers and followers of active and passive followers.

Fields:

1. User\_ID: The unique numeric ID of the twitter user
2. User\_Name: The unique username of the twitter user
3. Name: The twitter user’s name
4. Description: The profile description of the twitter user
5. Location: The twitter user’s location, if he/she has provided one
6. URL: Any URL mentioned in the user’s profile description. This could be the user’s website, personal blog, etc.
7. Followers\_Count: The total number of followers the user has (during the time of data pull)
8. Verified: Whether the user profile is has been verified
9. **Tweet\_Hashtag:**

This table contains the hashtag(s) in a tweet, if a hashtag was used. Not all tweets contain hashtags.

Fields:

1. Tweet\_ID: The unique numeric ID of the tweet
2. HashText: The hashtag mentioned in the tweet
3. **Tweet\_Media\_Data:**

This table contains the media data in the tweet. Twitter media can include photos, video, or animated gifs.

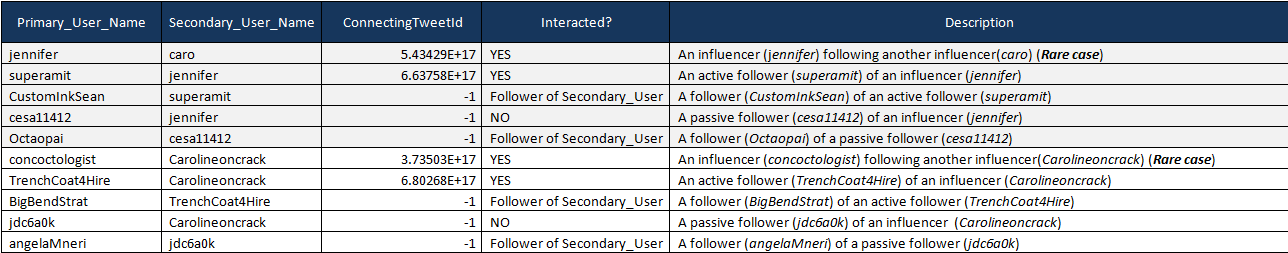
Fields:

1. Tweet\_ID: The unique numeric ID of the tweet
2. ExpandedURL: The full URL of the tweet with the media embedded
3. MediaURL: The URL for accessing the media content
4. MediaType: The type of media file in the tweet
5. **User\_Tweet\_Link:**

This table can be used for:

* Linking the interactions between the users
* Identifying active/ passive followers and their respective followers.
* Isolating influencers, active followers, passive followers, connecting tweets, etc.

A snapshot of various combinations of primary and secondary users is below. For demonstrative purposes, user names are provided in the snapshot; however, the table has User\_IDs rather than user names for the primary and secondary users.



Fields:

1. Primary\_User & Secondary\_User: The primary users & secondary users are those users who make up the following situations:

|  |  |
| --- | --- |
| **Primary User- Anyone who:** | **Secondary User-Anyone who/whose:** |
| * Mentions the influencer in his/her tweets (active follower) | * Is mentioned in the primary user’s tweet(s) |
| * Retweets the influencer’s tweet (active follower) | * Tweets are retweeted by the primary user |
| * Is a passive follower of the influencer | * The influencer himself/herself |
| * Is a follower of the active or passive follower | * The active or passive follower who is followed by the primary user |

1. ConnectingTweetID: The connecting tweet ID can be:
   * 1. The tweet ID that connects the primary user to the secondary user, thus counting as an interaction (active following)
     2. -1 if the primary user is a passive follower of the secondary user
     3. -1 if the primary follower is a follower of the active/passive follower
2. Interacted?: This field can take in the following three values:
   * 1. YES – If there is an interaction between the primary and secondary user (active following)
        + Note: There can be rare cases of influencers interacting with other influencer(s)
     2. NO – If there is no interaction between the primary and secondary user (passive following)
     3. Follower of Secondary\_User – In this case, the primary user is a follower of the secondary user. We do not consider whether they have interacted or not.
3. **Graph\_Attributes:**

Each user in the network is assumed to be a node. This table contains the various graph attributes associated with each node. These features have been calculated based on the whole structure of the network. Let the total number of nodes be ‘n’.

Fields:

1. User\_ID: The unique numeric ID of the twitter user
2. Degree\_Centrality: The degree centrality for a node/vertex -’u’, is the fraction of nodes it is connected to.
3. Closeness\_Centrality: Closeness centrality of a node ‘u’ is the reciprocal of the sum of the shortest path distances from ‘u’ to all *n-1* other nodes. Since the sum of distances depends on the number of nodes in the graph, closeness is normalized by the sum of minimum possible distances *n-1*.

C(u) = \frac{n - 1}{\sum_{v=1}^{n-1} d(v, u)},

Where *d*(v, u) is the shortest-path distance between v and u, and n is the number of nodes in the graph.

1. Degree: The number of graph edges which touch the node/vertex
2. Katz\_centrailty: Katz centrality computes the centrality for a node based on the centrality of its neighbors. It is a generalization of the eigenvector centrality. The Katz centrality for node ‘i’ is:

x_i = \alpha \sum_{j} A_{ij} x_j + \beta,

Where A is the adjacency matrix of the graph G with eigenvalues λ. The parameter α is the Attenuation factor and parameter β controls the initial centrality.

**Table Links & Connections:**

The diagram below is a schematic representation of the possible key connections. Do not limit yourselves to these connections. You are allowed to try other possible connections (even customized connections).

