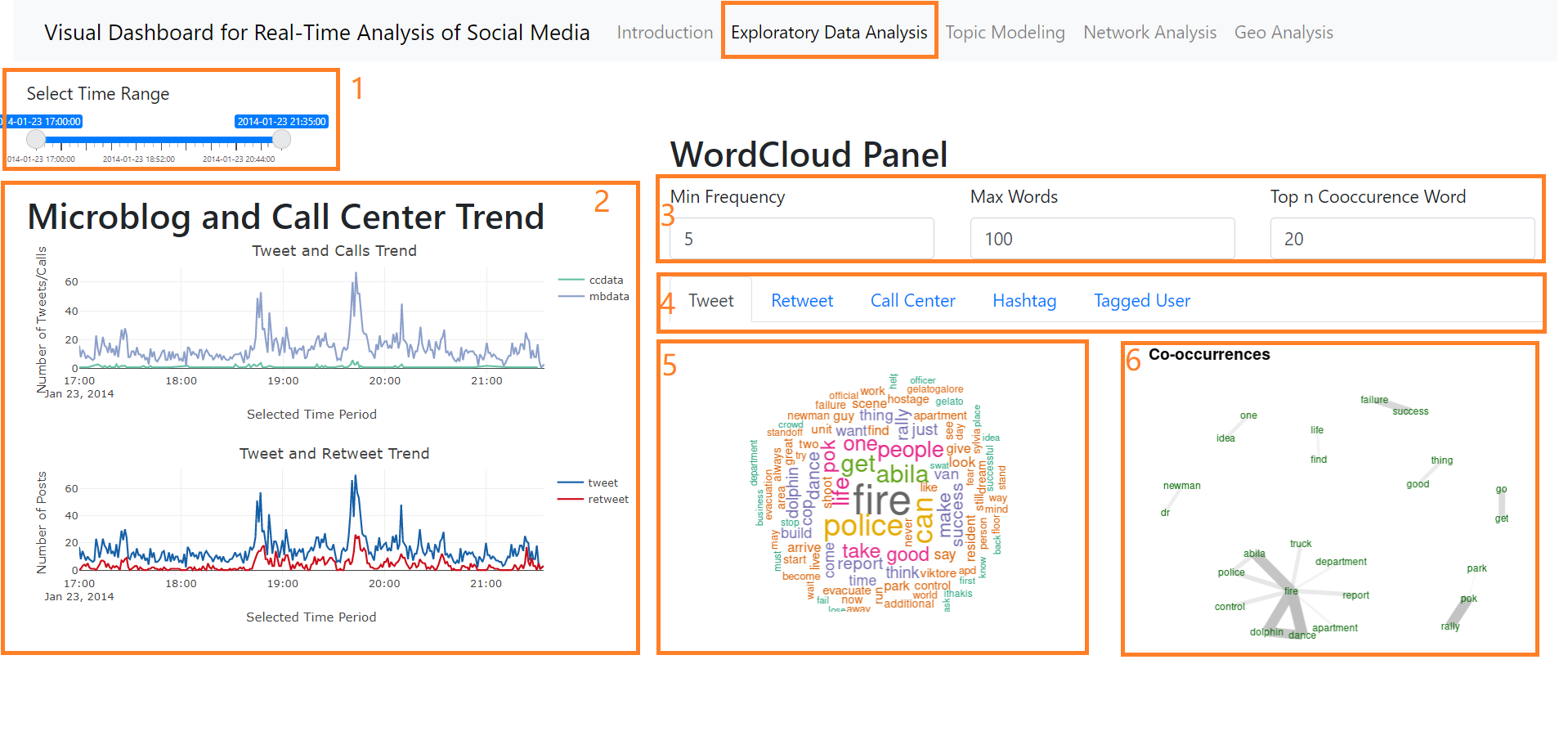
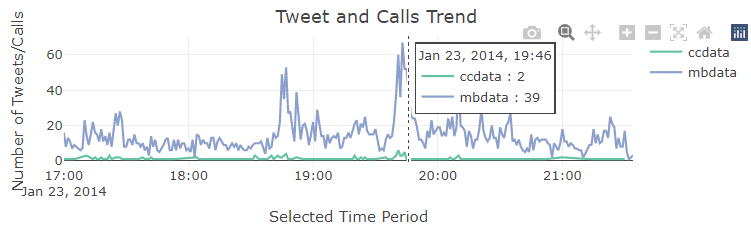
User Guide for Shiny - Visual Dashboard for Real-Time Analysis of Social Media

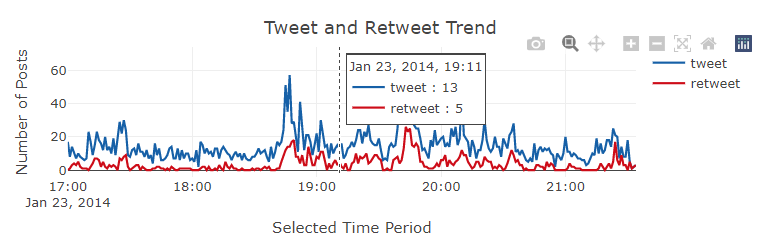
# 1. Introduction Page

The introduction page of the application provides a brief background of the dashboard , design concept and related project deliverable links.

# 2. Exploratory Data Analysis



1. **Slider** – drag the time range input slider for interested period, all charts in the dashboard will be refreshed to target time period.
2. **Line Graph** – showing microblog and call center trend, rendered by R plotly package. Hover to the chart, a trace line and tooltip with count number will show up. 



1. **Word Cloud Panel Input** – control the parameters of word cloud results.

* Min Frequency - words with frequency below min.freq will not be plotted, range from 1 to 20
* Max Words - maximum number of words to be plotted, range from 10 to 200
* Top n Cooccurrence Words – top n words with highest cooccurrence rate, range from 1 to 200

1. **Tab Panel** – world cloud from various data sources, including

Tweet- data type as mbdata

Retweet – data message with ‘RT @username’ pattern

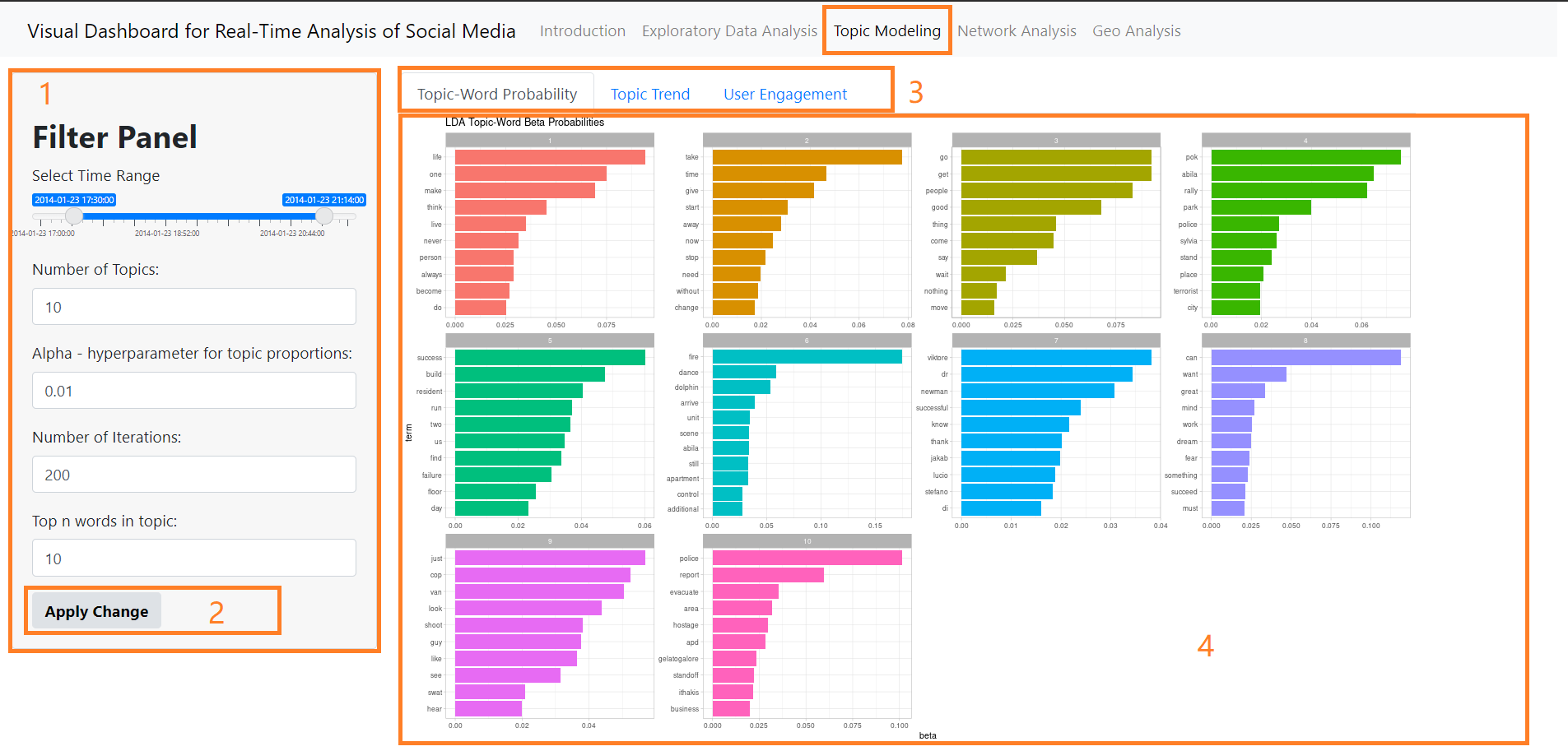
Call Center – data type as ccdata

Hashtag- retrieved ‘#hashtag’ words from tweet message

Tagged User – retrieved ‘@username’ pattern from tweet message

1. **Word** **Cloud** – rendered by R wordcloud package for all tabs
2. **Cooccurrence** **Text Plot** - rendered by R textplot package for message tabs(Tweet, Retweet and Call Center)

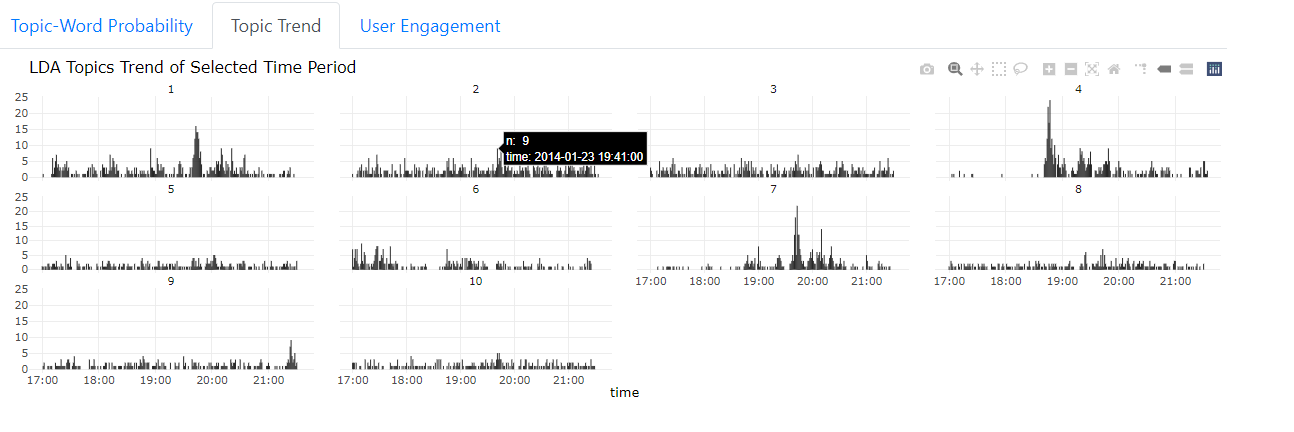
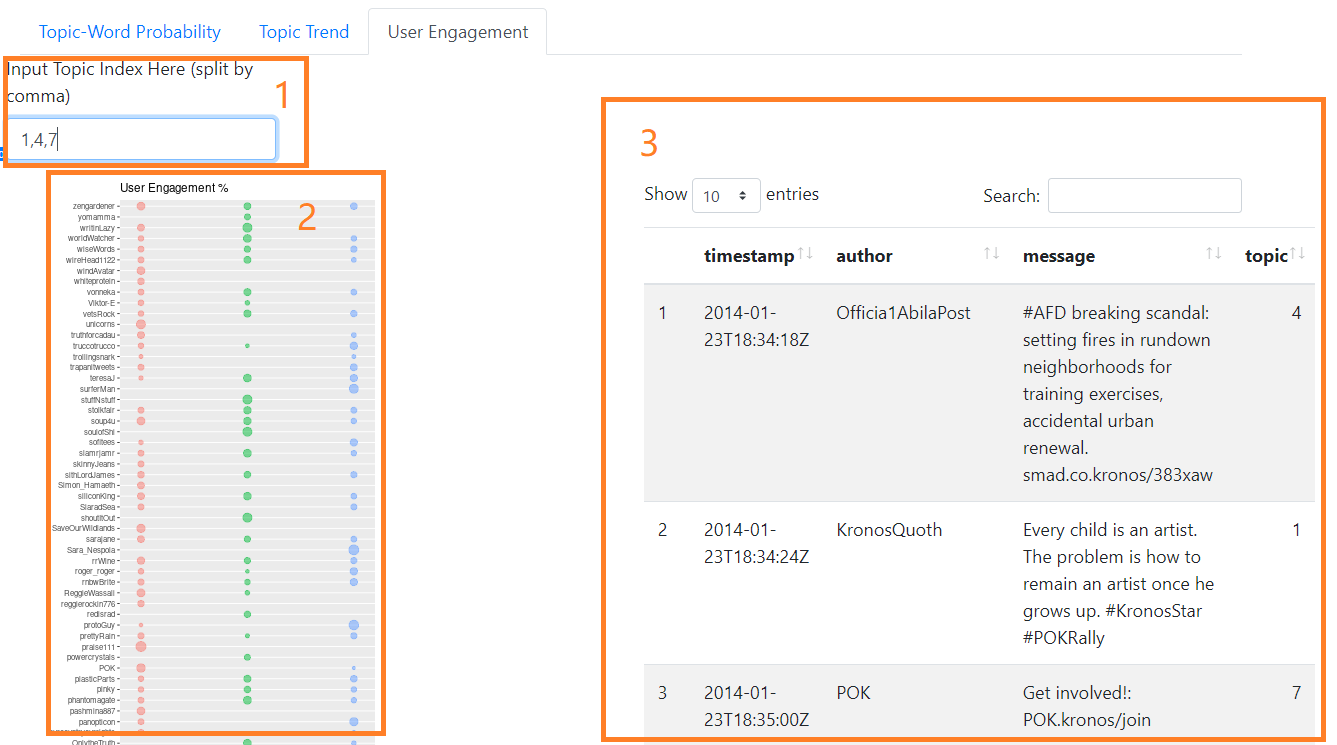
# 3. Topic Modelling



1. **Filter Panel**–

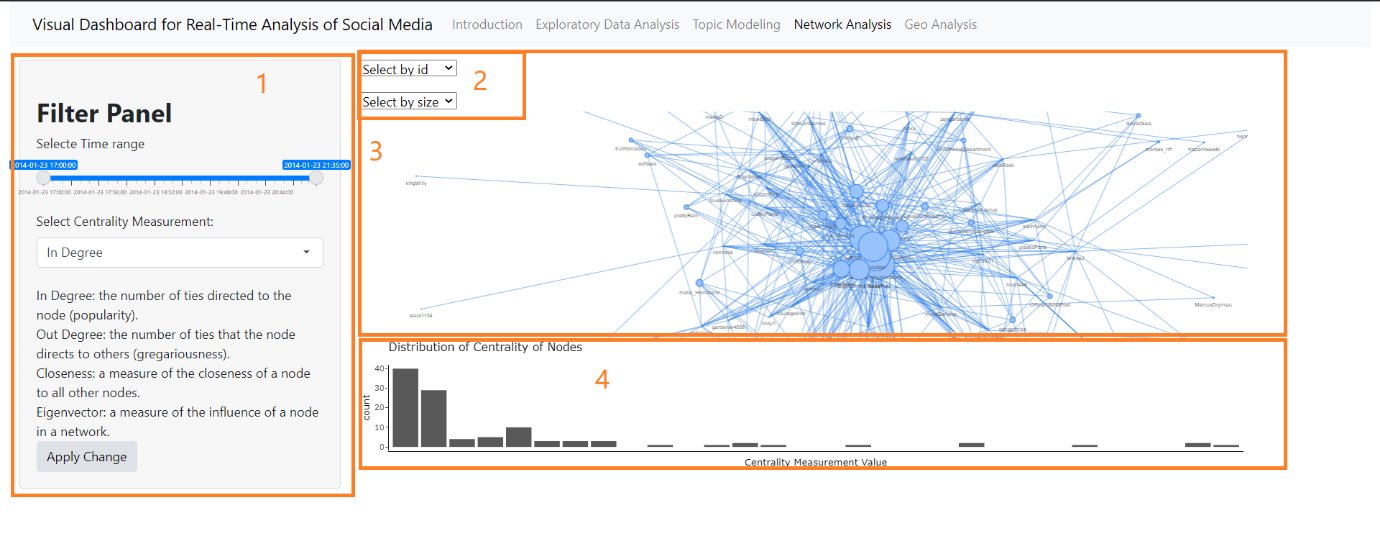
* Select Time Range - drag the time range input slider for interested period
* Number of Topics – input numbers of topics, range from 1 to 20, default as 10
* Alpha - hyperparameter for topic proportions, range from 0.00001 to 1, default as 0.01
* Number of Iterations – input model iteration times, range from 50 to 500, default as 200
* Top n words in topic – input numbers for top n words observation in Topic-Word Probability tab, range from 1 to 20, default at 10

1. **Apply Change button** – after selection in filter panel, click the button and LDA topic model will run backend and refresh results in the tab panel
2. **Tab Panel** – LDA modelling results analysis

* Topic-Word Probability Tab (4)– showing the top n words beta value bar chart, allow users to understand the words contributes most to each topic
* Topic Trend Tab – after assigning topic to each message based on gamma value, the tab shows LDA topics trend of selected time period. Hover mouse to the bar chart, the counts (aggregated for 1 mins) and time will show in the tooltip. Users can observe peaks of abnormal temporal topic pattern through this tab.
* User Engagement Tab – Subsequently, after initial understanding of topic contents and topic trend. User engagement tab allow users to take a closer look of interested topics.

1. **Topic Index Input** – input interested topic index, split by comma value, eg. 1,4,7
2. **User Engagement % Bubble Chart ­**– with topic assigned to each message as well as author, we calculated the percentage of posts for each topic and authors. The size of the bubble represents the participation % of selected topic.
3. **DT Table** – the table on the right is filtered from original dataset with target topic indexes.

# 4. Network Analysis



1. **Filter Panel** –

* Select Time Range - drag the time range input slider for interested period
* Select Centrality Measurement – retweet network nodes measurement through 4 different approaches,

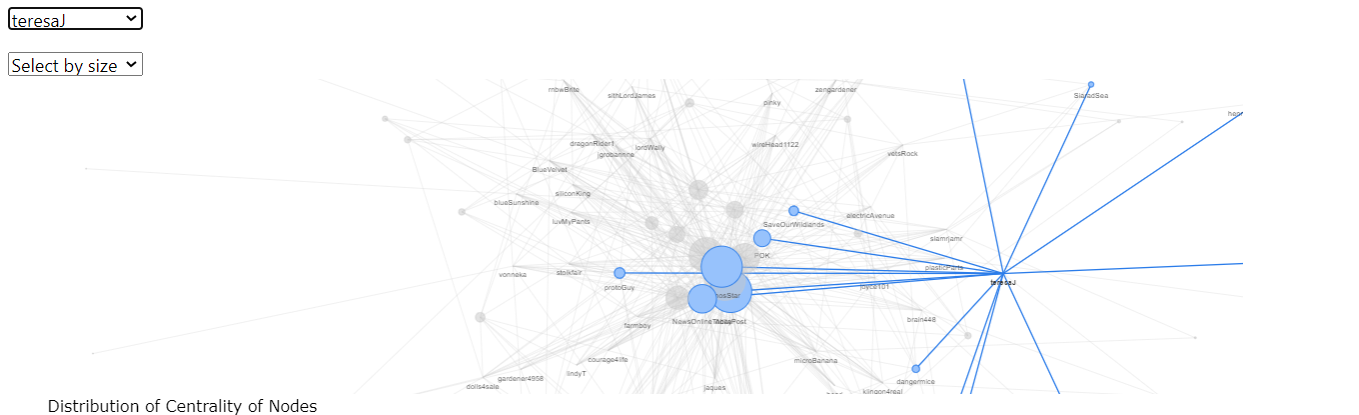
In Degree: the number of ties directed to the node (popularity).

Out Degree: the number of ties that the node directs to others (gregariousness).

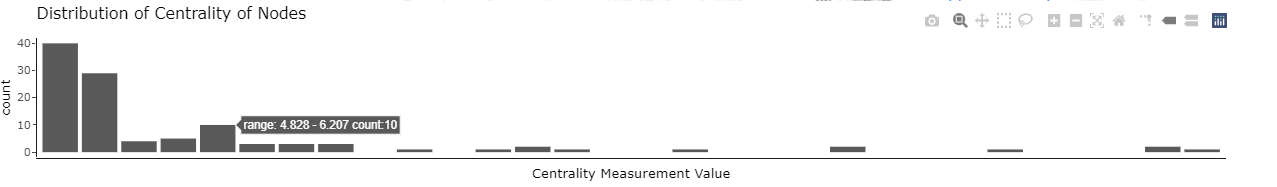
Closeness: a measure of the closeness of a node to all other nodes.

Eigenvector: a measure of the influence of a node in a network.

1. **Vis Network Drop Down List­**– Embedded drop down list in visNetwork, select by id (author) and size (nodes centrality value) to highlight the nodes and edges in network graph.



1. **Vis Network Plot** – Retweets network visualised by R visNetwork package.
2. **Distribution of Nodes Centrality** - Histograms of centrality distribution, right skewed distribution indicating the relationship of users in the community. Hover mouse to the bar to find out details of range and counted value are in the tooltip.



# 5. Geo Analysis