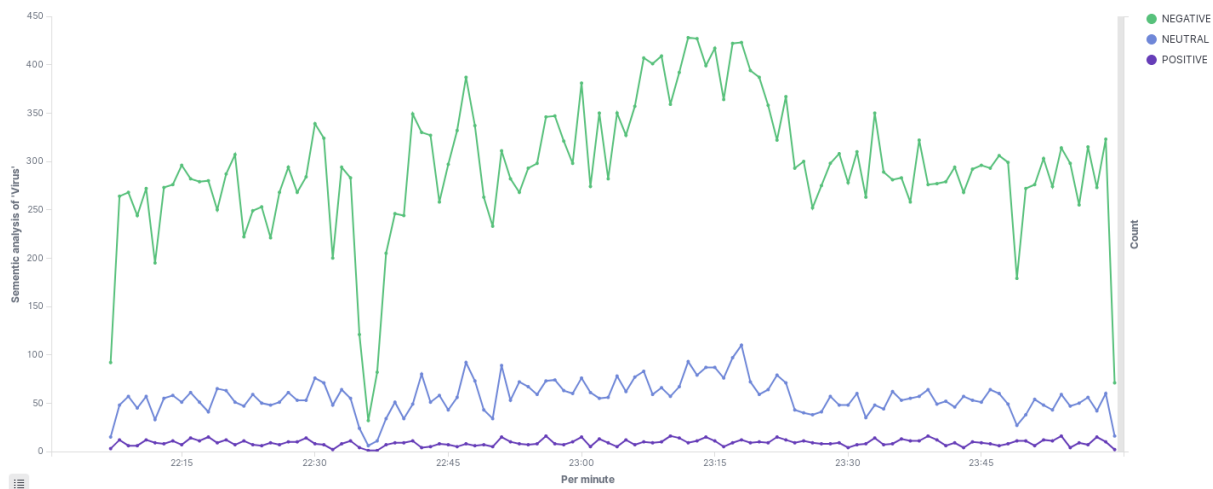


CS 6350

Spark Streaming with Twitter and Kafka

This project is to use Kafka as the middle message queue to transmit message from twitter api to elasticsearch with the help of logstash. The message is the sentiment analysis of a certain topic from real time twitter stream. And Kibana is employed to visualize the trend of the numbers for different types of sentiments including POSITIVE, NEGATIVE and NEUTRA.

Specifically, we chose 'Virus' as our topic interested and analyzed the sentiment for about 2 hours.



The POSITIVE has 1,028 hits (2.5%), NEGATIVE has 33,528 hits (81.8%), and NEUTRAL has 6,413 hits (15.7%). To sum up, there is no explicit trend for the sentiment with the time frame. But we can see that NEGATIVE tweets dominates.

The jar file can be downloaded from <https://bigdata-assign3-p1.s3.amazonaws.com/Sentiment-assembly-0.1.jar>