**Spark Streaming with Twitter and Kafka Summary**

We used ‘Democrat’ as our query string to be analyzed for the sentiment for a period of over 2 hours. To visualize the data, we used ‘Timelion’ which is a time series data visualizer in Kibana.

As seen in the graphs below there are a maximum number of tweets that are classified as negative followed by neutral tweets and then positive tweets. Over the period of time, the number of negative tweets were fluctuating but started decreasing eventually. The number of neutral and positive tweets remained fairly constant throughout.

We can conclude that a greater number of negative tweets were observed for this query string.

X axis: Timestamp

Y axis: Count of tweets

Interval: Visualized for every 10 seconds

Chart, histogram

Description automatically generated

1. Sentiment analysis graph for the word ‘Democrat’

Chart, histogram

Description automatically generated

1. Using TimeLion, at any given timestamp, we can see the number of negative, positive and neutral tweets.

Chart, histogram

Description automatically generated

1. To visualize only the negative tweets over period of 2 hours.

Chart

Description automatically generated

1. To visualize only the positive tweets over period of 2 hours.

A picture containing chart

Description automatically generated

1. To visualize only the neutral tweets over period of 2 hours.