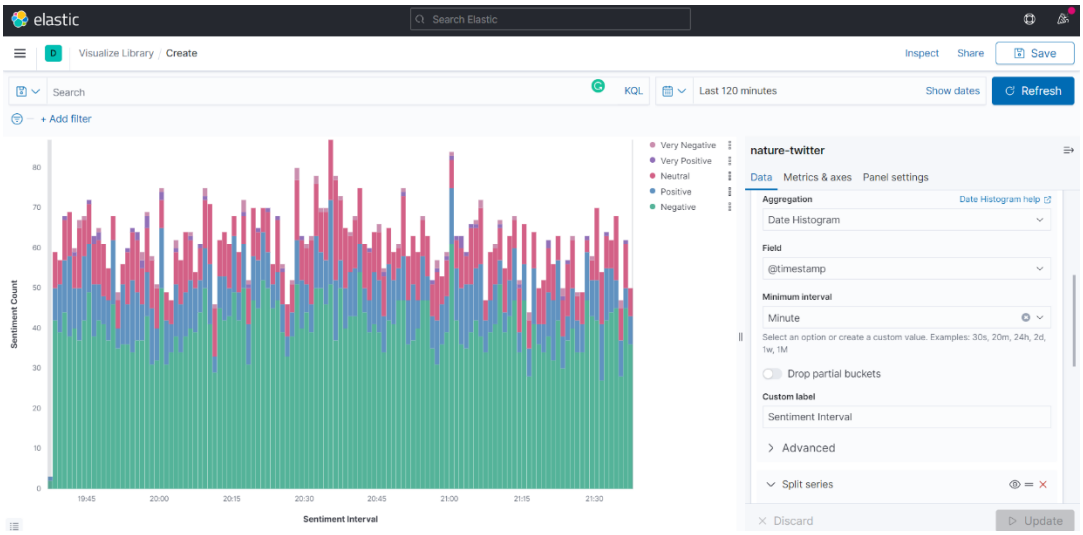
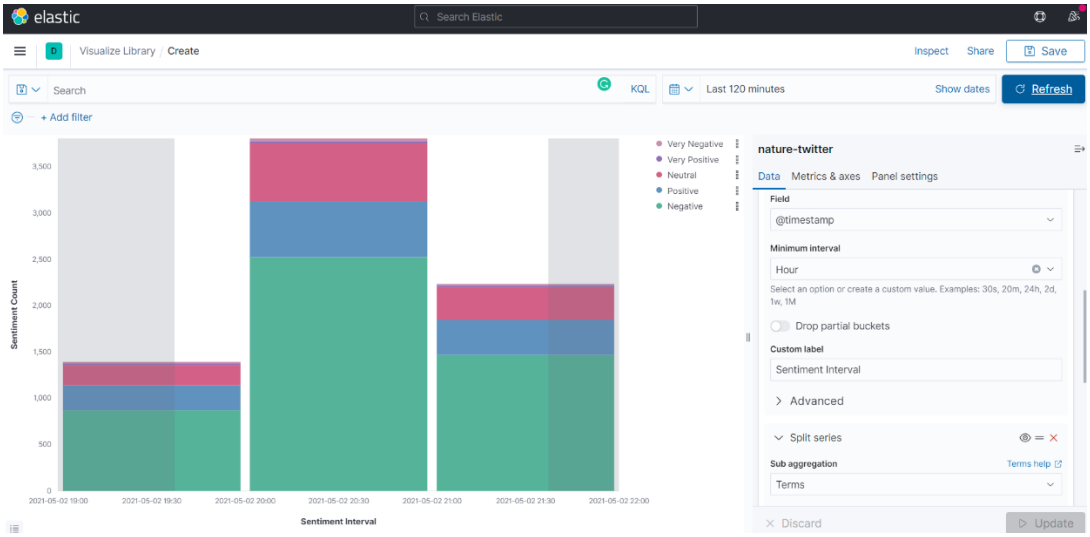
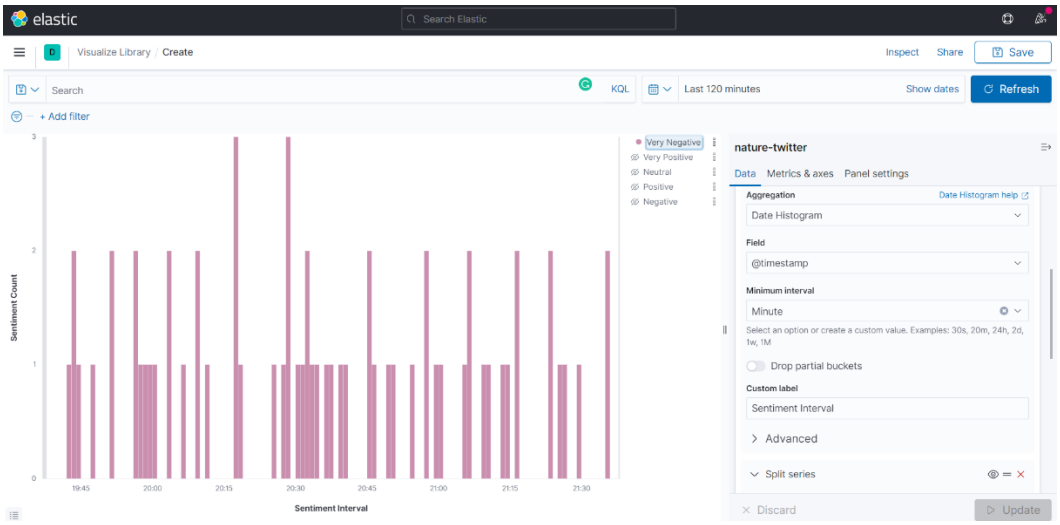
Part-1 : Spark Streaming with Twitter and Kafka

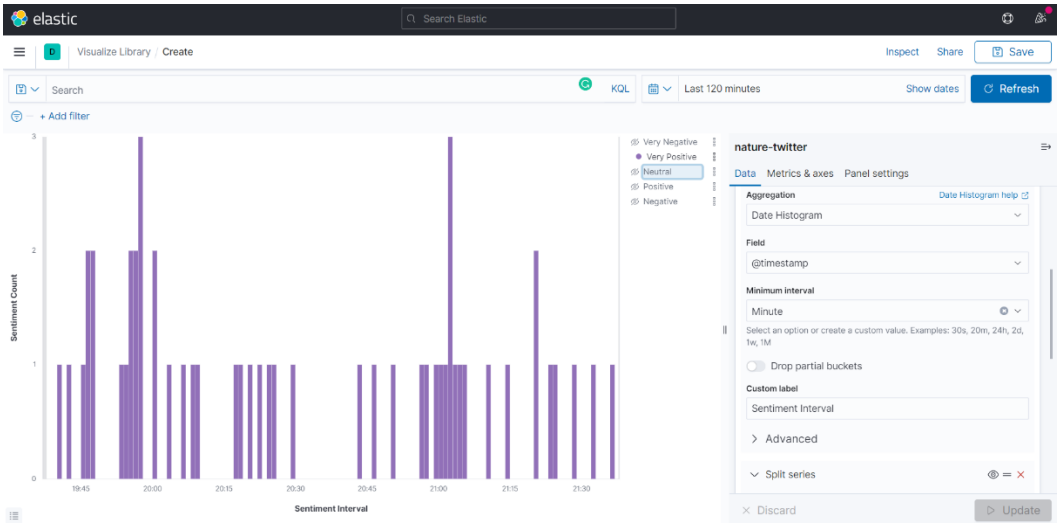
Data taken between time : May 02, 2021 : 7:39 pm to 9:39 pm
Topic : "**nature**"

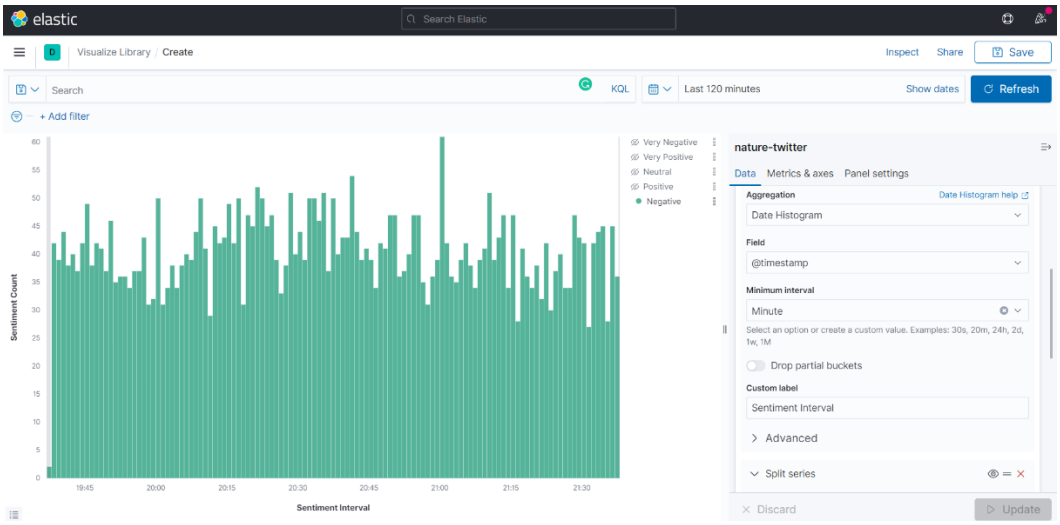
Below are the graph plots for the data:

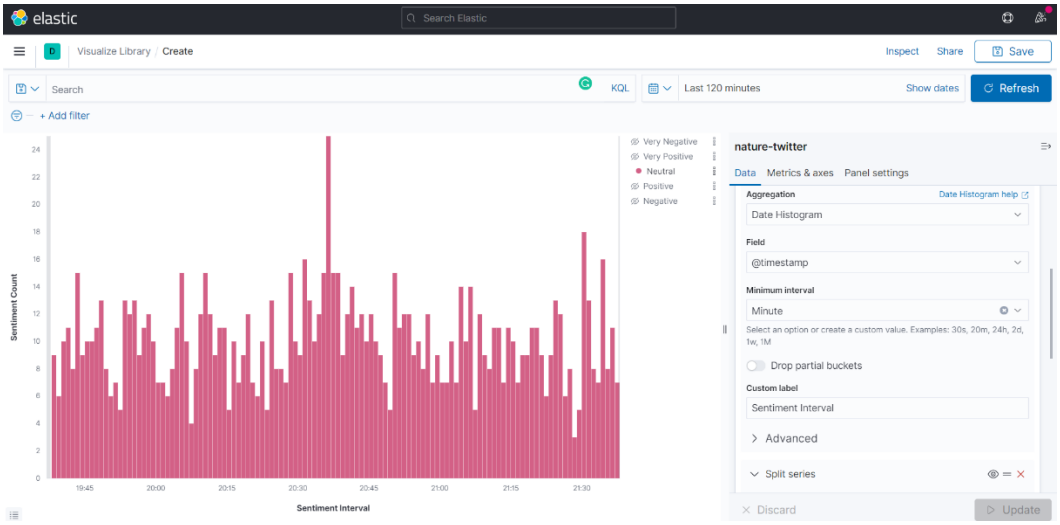


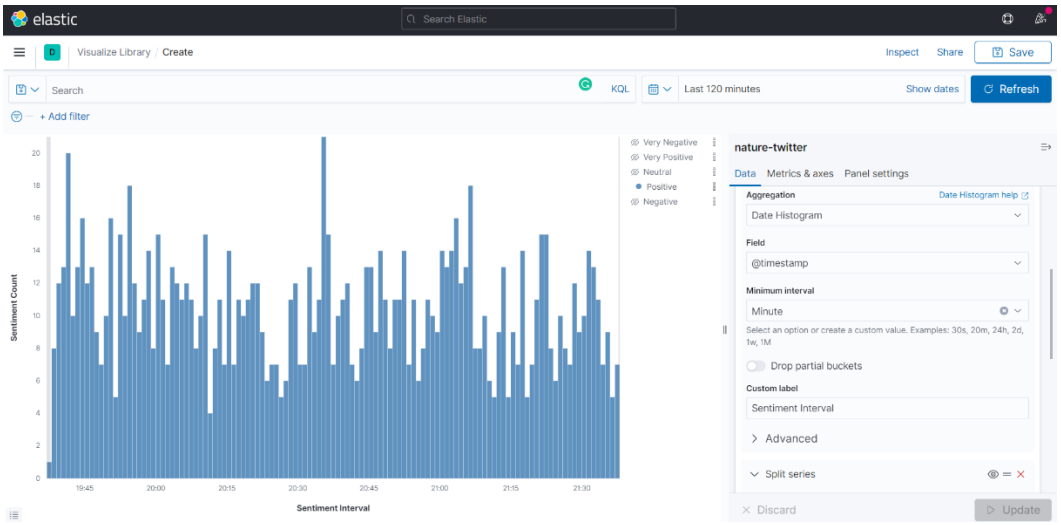



elastic
Search Elastic
Visualize Library / Create
Inspect
Share
Save
Search
KQL
Last 120 minutes
Show dates
Refresh
+ Add filter
Very Negative
Very Positive
Neutral
Positive
Negative
Sentiment Count
Sentiment Interval
nature-twitter
Data
Metrics & axes
Panel settings
Aggregation
Date Histogram help
Date Histogram
Field
@timestamp
Minimum interval
Minute
Select an option or create a custom value. Examples: 30s, 20m, 24h, 2d, 1w, 1M
Drop partial buckets
Custom label
Sentiment Interval
Advanced
Split series
Discard
Update


elastic
Search Elastic
Visualize Library / Create
Inspect
Share
Save
Search
KQL
Last 120 minutes
Show dates
Refresh
+ Add filter
Very Negative
Very Positive
Neutral
Positive
Negative
Sentiment Count
Sentiment Interval
nature-twitter
Data
Metrics & axes
Panel settings
Aggregation
Date Histogram help
Date Histogram
Field
@timestamp
Minimum interval
Minute
Select an option or create a custom value. Examples: 30s, 20m, 24h, 2d, 1w, 1M
Drop partial buckets
Custom label
Sentiment Interval
Advanced
Split series
Discard
Update


elastic
Search Elastic
Visualize Library / Create
Inspect
Share
Save
Search
KQL
Last 120 minutes
Show dates
Refresh
+ Add filter
Very Negative
Very Positive
Neutral
Positive
Negative
Sentiment Count
Sentiment Interval
nature-twitter
Data
Metrics & axes
Panel settings
Aggregation
Date Histogram help
Date Histogram
Field
@timestamp
Minimum interval
Minute
Select an option or create a custom value. Examples: 30s, 20m, 24h, 2d, 1w, 1M
Drop partial buckets
Custom label
Sentiment Interval
Advanced
Split series
Discard
Update


elastic
Search Elastic
Visualize Library / Create
Inspect
Share
Save
Search
KQL
Last 120 minutes
Show dates
Refresh
+ Add filter
Very Negative
Very Positive
Neutral
Positive
Negative
Sentiment Count
Sentiment Interval
nature-twitter
Data
Metrics & axes
Panel settings
Aggregation
Date Histogram help
Date Histogram
Field
@timestamp
Minimum interval
Minute
Select an option or create a custom value. Examples: 30s, 20m, 24h, 2d, 1w, 1M
Drop partial buckets
Custom label
Sentiment Interval
Advanced
Split series
Discard
Update



Summary:

- From the data and graphs, we can see that there is high amount of data on "**nature"** that has negative sentiments.
- There is less count of sentiment data in extremes (very-negative and very positive)

The same can be seen from the table below.

Sentiment --->	Very Negative	Negative	Neutral	Positive	Very Positive
Total Counts --->	64	4876	1198	1261	52

NOTE: The detailed data is in the csv file in output folder

References:

https://www.elastic.co/guide/index.html
https://spark.apache.org/docs/latest/graphx-programming-guide.html
https://www.confluent.io/resources/kafka-the-definitive-guide-v2/