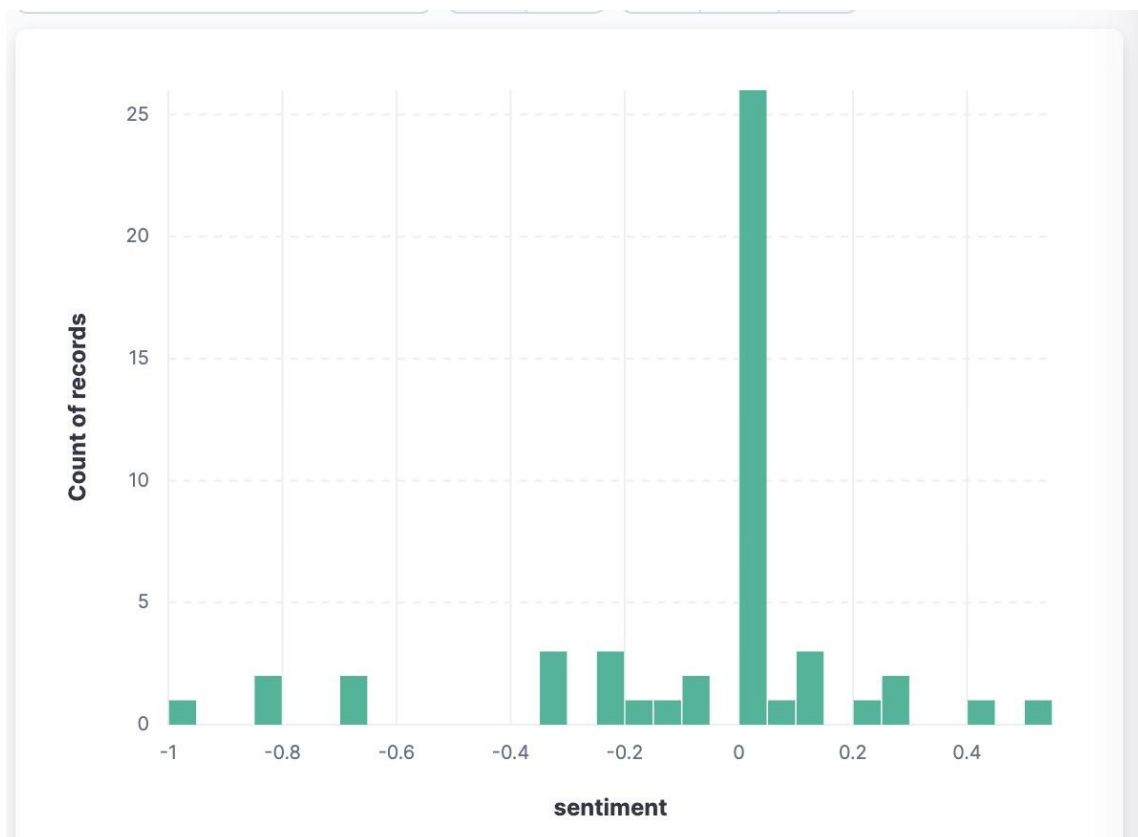


For this project we are trying to analyze the tweets related to the keyword “Republicans”. We have used PySpark’s TextBlob library for this sentiment estimation and analysis. The sentiment analysis code can be viewed in **sentiment\_analyser.py**.

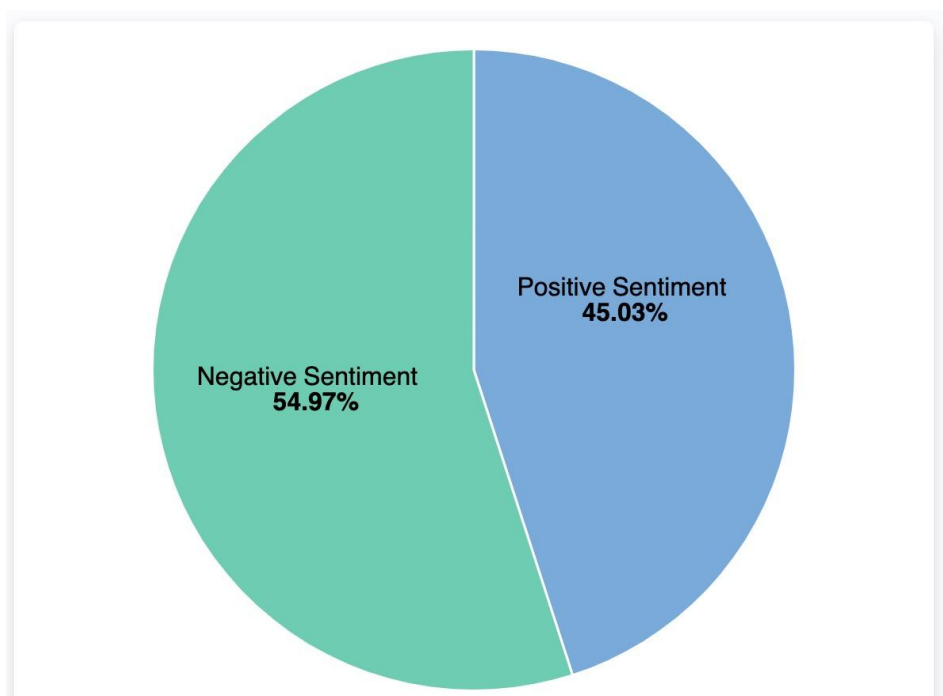
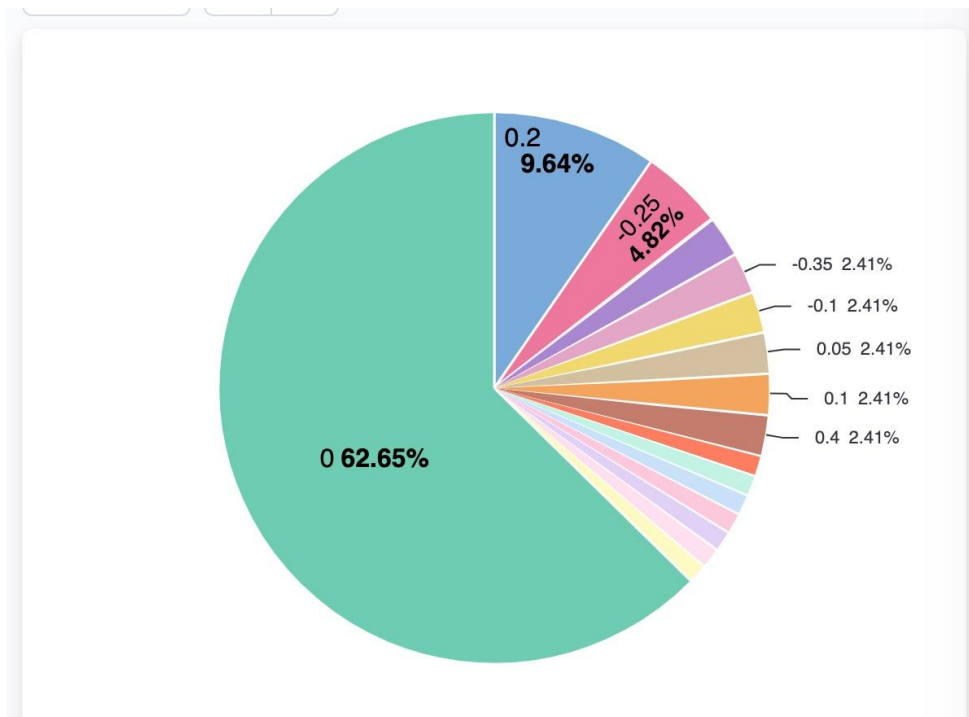
To execute the stream sentiment analysis we need to run **stream\_object.py** which imports the function from sentiment\_analyser.py and does continuous stream analysis and sends the analysis to kafka broker services which in turn is sent to Elastic search via logstash for visualization using Kibana.

Upon starting the kafka zookeeper and server and running the script **python3 stream\_object.py** and visualizing the analysis on about 200 tweets with keyword “Republicans”, we get following insights:

1. The count of tweets which are negative in sentiment are in majority. Most of the tweets have a negative or almost neutral sentiment as shown by the bar chart below. This means that people on twitter have a negative opinion about republicans.



The following pie chart shows the distribution of sentiment polarity scores of tweets.



As stated above, about 54.97% tweets are having a negative sentiment polarity score and 45.03% are having a positive sentiment polarity score. This again reinstates the fact the republicans are pretty unpopular at this point of time on twitter.