

# Twitter Data Mining Using ELK

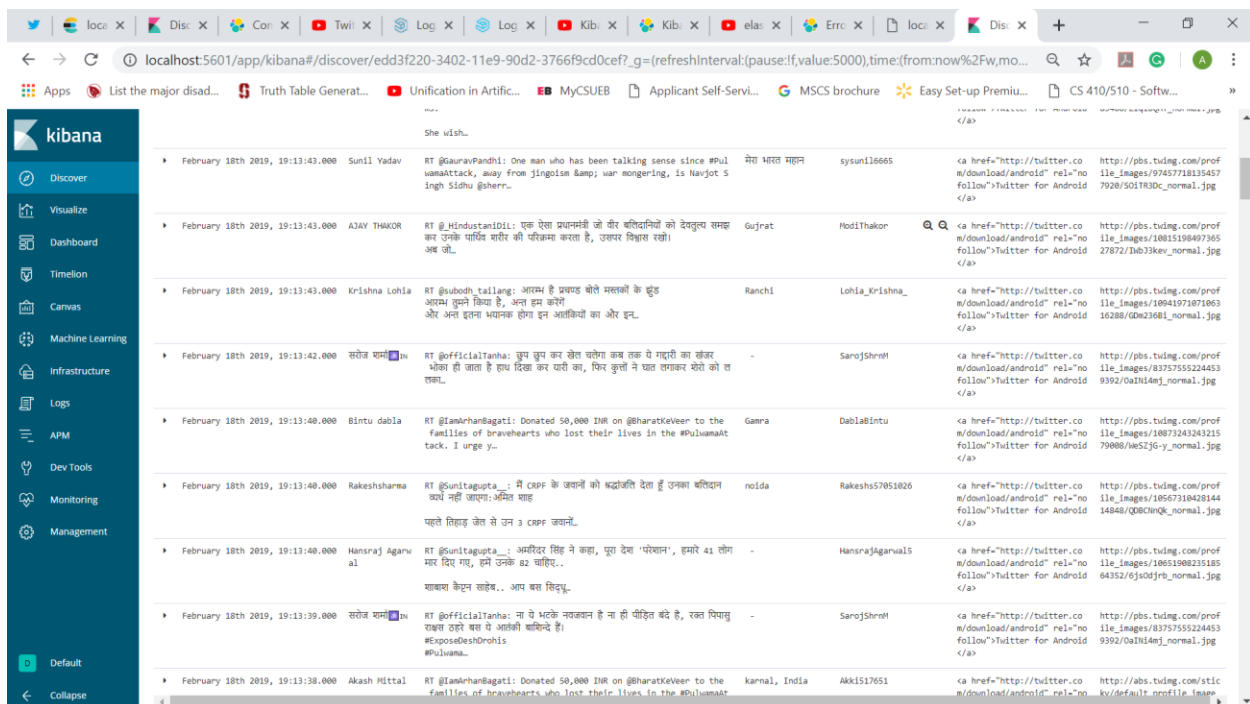
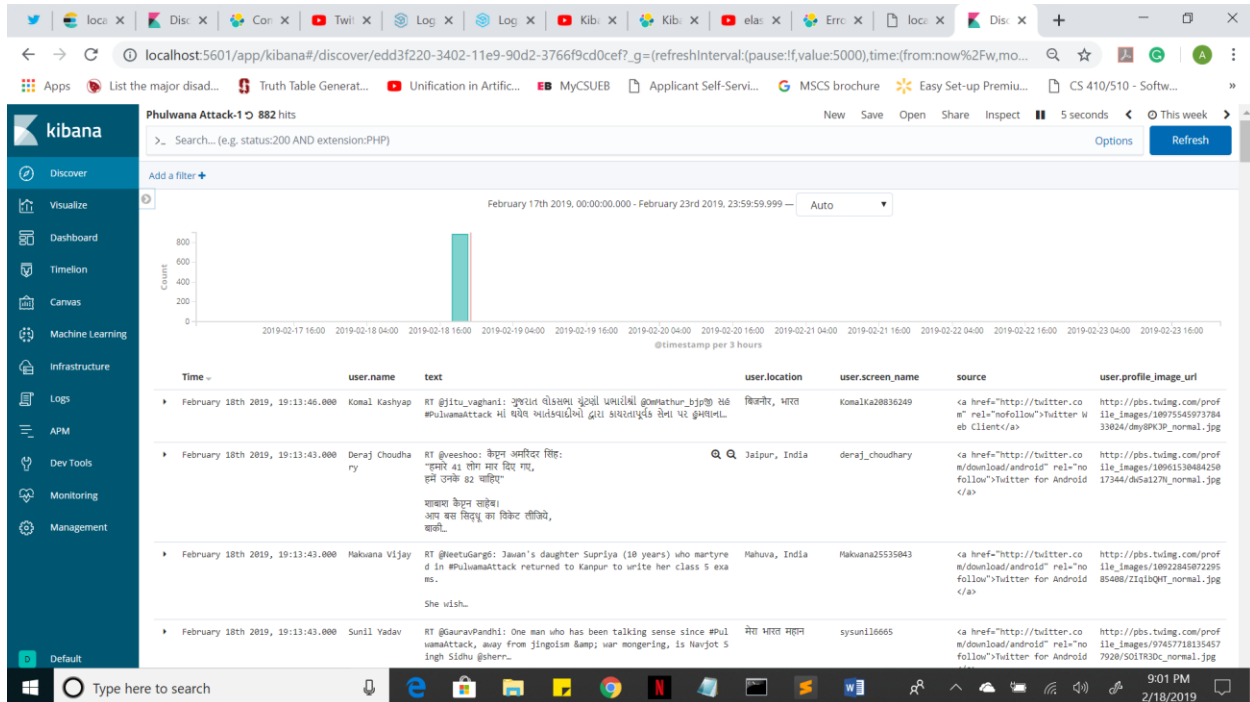
Group: Aishwarya Murkute (210567047)

Aditya Lalchandani (210052091)

Date: 25<sup>th</sup> February 2019

## Mining Results:

We could extract the name of the person tweeting, the user ID Location and device from where the tweet was sent.



# Twitter Data Mining Using ELK

Group: Aishwarya Murkute (210567047)

Aditya Lalchandani (210052091)

Date: 25<sup>th</sup> February 2019

**Logstash was used to pull the data:**

```
C:\Windows\System32\cmd.exe - logstash - f twitter.com
Microsoft Windows [Version 10.0.17134.598]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\Users\adity\Desktop\ELK-Stack\logstash-6.6.0\logstash-6.6.0\bin>logstash -f twitter.conf
Sending Logstash logs to C:/Users/adity/Desktop/ELK-Stack/logstash-6.6.0/logstash-6.6.0/logs which is now configured via log4j2.properties
[2019-02-18T19:36:23.707][WARN ][logstash.config.source.multilocal] Ignoring the 'pipelines.yml' file because modules or command line options are specified
[2019-02-18T19:36:23.738][INFO ][logstash.runner           ] Starting Logstash {"logstash.version"=>"6.6.0"}
[2019-02-18T19:36:31.934][INFO ][logstash.pipeline        ] Starting pipeline {"pipeline_id"=>"main", "pipeline.workers"=>8, "pipeline.batch.size"=>125, "pipeline.batch.delay"=>50}
[2019-02-18T19:36:32.387][INFO ][logstash.outputs.elasticsearch] Elasticsearch pool URLs updated {:changes=>{:removed=>[], :added=>[http://localhost:9200/]}}
[2019-02-18T19:36:32.606][WARN ][logstash.outputs.elasticsearch] Restored connection to ES instance {"url"=>"http://localhost:9200/" }
[2019-02-18T19:36:32.669][INFO ][logstash.outputs.elasticsearch] ES Output version determined {:es_version=>6}
[2019-02-18T19:36:32.669][WARN ][logstash.outputs.elasticsearch] Detected a 6.x and above cluster: the 'type' event field won't be used to determine the document_type {:es_version=>6}
[2019-02-18T19:36:32.700][INFO ][logstash.outputs.elasticsearch] New Elasticsearch output {:class=>"LogStash::Outputs::ElasticSearch", :hosts=>["//localhost:9200"]}
[2019-02-18T19:36:32.716][INFO ][logstash.outputs.elasticsearch] Using mapping template from {:path=>nil}
[2019-02-18T19:36:32.747][INFO ][logstash.outputs.elasticsearch] Attempting to install template {"name"=>"logstash", "version"=>60001, "settings"=>{"index.refresh_interval"=>"5s"}, "mappings"=>{"_default_"=>{"dynamic_templates"=>[{"message_field"=>{"path_match"=>"message", "match_mapping_type"=>"string", "mapping"=>{"type"=>"text", "norms"=>false}}, {"string_fields"=>{"match"=>"*", "match_mapping_type"=>"string", "mapping"=>{"type"=>"text", "norms"=>false, "fields"=>{"keyword"=>{"type"=>"keyword", "ignore_above"=>256}}}], "properties"=>{"@timestamp"=>{"type"=>"date"}, "@version"=>{"type"=>"keyword"}, "geoip"=>{"dynamic"=>true, "properties"=>{"ip"=>{"type"=>"ip"}, "location"=>{"type"=>"geo_point"}, "latitude"=>{"type"=>"half_float"}, "longitude"=>{"type"=>"half_float"}}}}]}
[2019-02-18T19:36:33.309][INFO ][logstash.pipeline        ] Pipeline started successfully {"pipeline_id"=>"main", :thread=>#<Thread:0x67449673 run>}
[2019-02-18T19:36:33.325][INFO ][logstash.inputs.twitter  ] Starting twitter tracking {:track=>"phulwananattak"}
[2019-02-18T19:36:33.356][INFO ][logstash.agent           ] Pipelines running {:count=1, :running_pipelines=>[:main], :non_running_pipelines=>[]}
[2019-02-18T19:36:33.841][INFO ][logstash.agent           ] Successfully started Logstash API endpoint {:port=>9600}
```

This is the configuration file that we used in Logstash:

[illegible]