

# Topics

## Day 1- Introduction to Big Data and Hadoop

### Theory

- Introduction to Big Data and prime techniques to handle/process it
- MapReduce algorithm
- Hadoop Framework, its component and related technologies
- Run through of one or two real use-cases of Hadoop in the industry
- Step-by-step simple word count problem explained in detail

### Lab Sessions

- Accessing HDFS file system
- Implementation of word count problem
- **(and if time permits)** one more simple example

## Day 2 : NoSQL(Elasticsearch) and Real Time Analytics

### Theory

- Big Data Analytics and NoSQL Storage
- Few associated technologies
- Introduction to - Elasticsearch, Logstash and Kibana
- End-2-end data flow with a real use case explained
- Data ingestion and data access mechanism/techniques
- Analytics using Kibana

### Lab Sessions

- Few hands-on example of data insertion and data access from Elasticsearch
- Prepare Analytics Dashboard for Apple Stock Trends (E2E)
- **(and if time permits)** New York City Traffic Incidents navigation

## Day 3 : Fitment of Cloud, Message Broker into Analytics

### Theory

- Message brokers and their fitment into the data flow
- Walk through of Popular Message Brokers
  - ActiveMQ, RabbitMQ, Kafka etc.
- Convergence of Hadoop, Cloud, NoSQL, Message Broker towards Analytics

### Lab Sessions

- Carried over assignments from Day 2
- ActiveMQ
  - GUI Interface
  - Hands-on with sending and receiving messages using producers and consumers.