Readme - Part 1

Files:

Producer.py

KafkaConnection.py

logstash.config

Open WSL, follow the following steps:

Step 1:- Start Zookeeper: (Terminal 1)

  bin/zookeeper-server-start.sh config/zookeeper.properties

Step 2:- Start Kafka: (Terminal 2)

  bin/kafka-server-start.sh config/server.properties

Step 3:- Create Topics: (Terminal 3)

 bin/kafka-topics.sh --create --topic tp1 --bootstrap-server localhost:9092

 bin/kafka-topics.sh --create --topic tp2 --bootstrap-server localhost:9092

Step 4:- Configure ELK (by downloading and untar them each one in different Terminals)

Step 5:- Run python files, ELKs (logstash with config file)

Step 6:- Goto localhost:5601 login and create dashboards and do analysis

Packages used: tweepy, json, tensorflow, pipeline, spark, configparser

To Run .py files, use these commands:

1. Producer.py: python3 Producer.py

2. KafkaConnection.py: spark-submit --packages
org.apache.spark:spark-sql-kafka-0-10_2.13-3.2.0 KafkaConnection.py localhost:9092 subscribe
tp1

For running ELK Stack ones :

1. Elasticsearch (In the directory of Elasticsearch) : /bin/elasticsearch

2. Kibana (In the directory of Kibana) :/bin/kibana

3. Logstash (In the directory of Logstash) : /bin/logstash -f ./logstash.config

** Mostly, everything with the password will work. If doesn't work, I can reset or share the specific password again

Mostly run all commands in new different terminals **