Requirement:
1. Apache Kafka
2. Druid
3. Imply (Analytics tool)
4. Zookeeper

Setup:

Below is the steps to setup the data collection and visualisation dashboard. Here, I will be explaining two dashboard that can be created.

1. **Imply Setup:**
   a. download impy from **:** https://static.imply.io/release/imply-2.6.11.tar.gz
   b. Extract imply: **tar -xzf imply-2.6.11.tar.gz**
   c. Change directory to imply
   d. After changing directory to imply, you need to start the imply : **bin/supervise -c conf/supervise/quickstart.conf**

      NOTE: This command will start the different services of druid and you can see the detail log of each service in var/sv directory

      NOTE:
      1. Verify imply-UI started : go to localhost:9095 in your browser
      2. check druid coordinator: localhost:8090

2. **Kafka Setup:**
   a. Download kafka : wget http://www-us.apache.org/dist/kafka/0.10.2.0/kafka_2.11-2.0.0.tgz
   b. Extract kafka : **tar -xzf kafka_2.11-2.0.0.tgz**
   c. Change to kafka directory: **cd kafka_2.11-2.0.0**
   d. Start kafka broker using following command : **./bin/kafka-server-start.sh config/server.properties**
   e. Now create a topic using following command: **./bin/kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic affable**

      Note: Once you have created the topic, now you can publish data under this topic and then Druid will listen to the kafka topic.

3. **Kafka Druid Ingestion:**
   a. Now, as IMPLY will also start the druid and zookeeper, so now we can start kafka data ingestion to the druid.
   b. We can start the kafka ingestion with following command: **curl -XPOST -H'Content-Type: application/json' -d @../mockstagram_supervisor_spec.json http://localhost:8090/druid/indexer/v1/supervisor**

NOTE: use mockstagram_supervisor_spec.json, this file contains the supervisor config which have the configuration about data and also have data source config.

Once you are finished with all the above setup, then we are good to go create the dashboard. As I have already mention there can be two different dashboard can be created.

DashBoard 1 - Imply Visualisation/analytics dashboard:
1. You can run this dashboard on port 9095 in your browser:
   http://<druid-server-ip>:9095 (For me it is druid-server-ip = localhost)
2. In this, you can see all the data source ingested to the druid. Data source can be real time streaming from kafka (can be done from setup step-3).
3. Also, Imply provides sql interface to query druid which very fast and efficient.
4. You can create data dashboard for different kind of data sources.

NOTE : Please see the dashboard images added to the github:
https://github.com/gauravshelangia/kafka-Druid-Imply-Dashboard

DashBoard 2 - Mockstagram Dashboard:
We can also create our own custom dashboard. Below is the requirement for this:
1. Write queries to get data from druid database - can use druid query(https://www.npmjs.com/package/druid-query)
2. Then send all the data to the UI to populate on dashboard