Tutorial 2:

* Run zookeeper  
  .\bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties

A screen shot of a computer

Description automatically generated

* Run kafka server:

.\bin\windows\kafka-server-start.bat .\config\server.properties

A screen shot of a computer

Description automatically generated

* Create topic:

kafka-topics.bat --create --bootstrap-server localhost:9092 --topic test

A black and white screen with white text

Description automatically generated

* Create producer:

kafka-console-producer.bat --broker-list localhost:9092 --topic test

A computer screen with white text

Description automatically generated

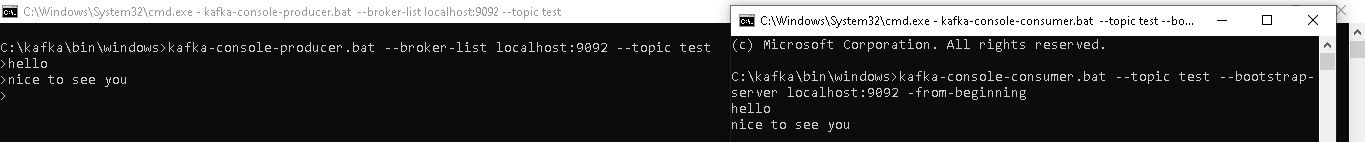
* Create consumer to read from test of producer

kafka-console-consumer.bat --topic test --bootstrap-server localhost:9092 -from-beginning

A computer screen with white text

Description automatically generated

* Result:



Tutorial 6:

* Step1: Start Kafka cluster using Terminal

.\bin\windows\kafka-server-start.bat .\config\server.properties

* A screen shot of a computer

  Description automatically generated
* Step 2: Run KafkaProducer in Jupyter Notebook

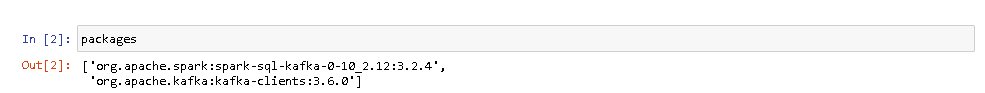
A screenshot of a computer

Description automatically generated

* Open another Jupyter Notebook

A computer screen shot of a computer code

Description automatically generated



* Create dataframe from Kafka data

If facing this issue



downgrade the version of pandas to pip install pandas==1.5.3

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Creating a Kafka Source for Streaming Queries

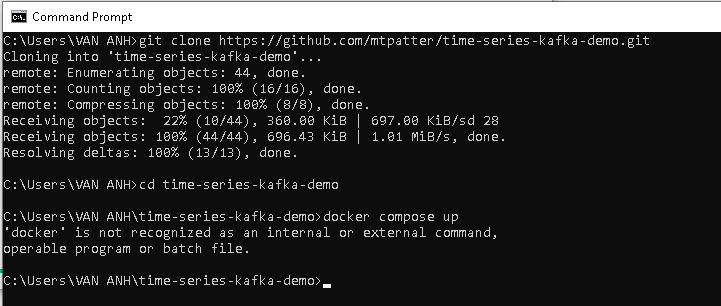
Phần 2 của ví dụ 6 chưa được

A computer screen shot of a computer code

Description automatically generated

Tutorial 8

https://towardsdatascience.com/make-a-mock-real-time-stream-of-data-with-python-and-kafka-7e5e23123582



Cài docker

Cai ubuntu

Bimat182

Cai xong docker

A screenshot of a computer error

Description automatically generated

Dùng machine learning, deep learning để xử lý dl lớn

Xử lý streaming

Giống Vd15

Có tập dl file csv

Stream dữ liệu lên kafka server

Lấy dl từ kafka sever về -> cho vào data frame streaming trong spark-> mô hình machine learning để xl dl -> hiển thị kq real time -> hiển thị kq lên web app (Nodejs,…) thay vì terminal