# Real-time data pipeline

Downloaded docker desktop

Download docker desktop from official site ,benefit of using docker desktop it now comes with inbuilt linux feature as WSL(windows subsystem for linux) this will get installed along with the installation of docker

Tweek some setting so that WSL is recognized by VM’s

Download Ubuntu to run the docker command

Setting kafka on docker

To setup kafka on docker we will create a docker-compose.yml file

Once yml file is configured depending on what version we want yml will vary; if we are using kafka with zookeeper which is the version below 7.0.1

Configure yml like this

version: '3.8'

services:

zookeeper:

image: confluentinc/cp-zookeeper:7.0.1

container\_name: zookeeper

ports:

- "2181:2181"

environment:

ZOOKEEPER\_CLIENT\_PORT: 2181

ZOOKEEPER\_TICK\_TIME: 2000

kafka:

image: confluentinc/cp-kafka:7.0.1

container\_name: kafka

ports:

- "9092:9092"

environment:

KAFKA\_BROKER\_ID: 1

KAFKA\_ZOOKEEPER\_CONNECT: zookeeper:2181

KAFKA\_ADVERTISED\_LISTENERS: PLAINTEXT://kafka:29092,PLAINTEXT\_HOST://localhost:9092

KAFKA\_LISTENER\_SECURITY\_PROTOCOL\_MAP: PLAINTEXT:PLAINTEXT,PLAINTEXT\_HOST:PLAINTEXT

KAFKA\_INTER\_BROKER\_LISTENER\_NAME: PLAINTEXT

KAFKA\_OFFSETS\_TOPIC\_REPLICATION\_FACTOR: 1

depends\_on:

- zookeeper

If using kraft configure doker-compose.yml like this

environment:

KAFKA\_PROCESS\_ROLES: broker,controller

KAFKA\_NODE\_ID: 1

KAFKA\_CONTROLLER\_QUORUM\_VOTERS: 1@kafka:9093

KAFKA\_LISTENERS: PLAINTEXT://kafka:29092,CONTROLLER://kafka:9093,PLAINTEXT\_HOST://localhost:9092

KAFKA\_ADVERTISED\_LISTENERS: PLAINTEXT://kafka:29092,PLAINTEXT\_HOST://localhost:9092

KAFKA\_LISTENER\_SECURITY\_PROTOCOL\_MAP: PLAINTEXT:PLAINTEXT,CONTROLLER:PLAINTEXT,PLAINTEXT\_HOST:PLAINTEXT

KAFKA\_INTER\_BROKER\_LISTENER\_NAME: PLAINTEXT

docker compose down -v to stop the process running

docker compose up -d to start the process

check running container – docker ps