LAPORAN UJIAN TENGAH SEMESTER
MATA KULIAH BIG DATA
Jobsheet 7 – Spark Docker

Dosen Pengampu:
M. Hasyim Ratsanjani, S.Kom., M.Kom.

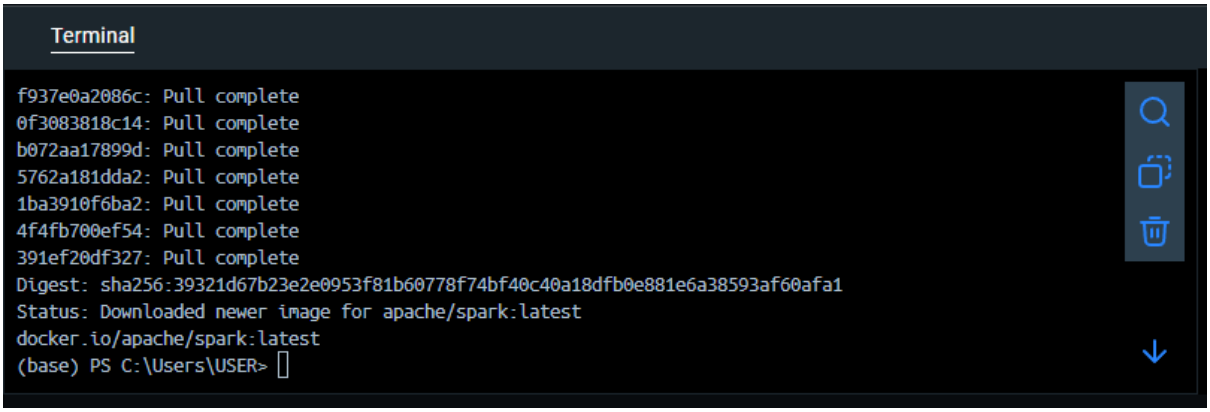
Disusun Oleh:

Yuma Rakha Samodra Sikayo **NIM.2241720194**

PROGRAM STUDI D4 TEKNIK INFORMATIKA
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI MALANG
2025

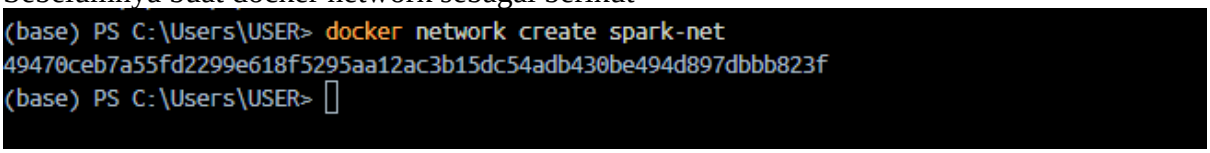
Langkah Langkah

1. Pull Image Spark Resmi

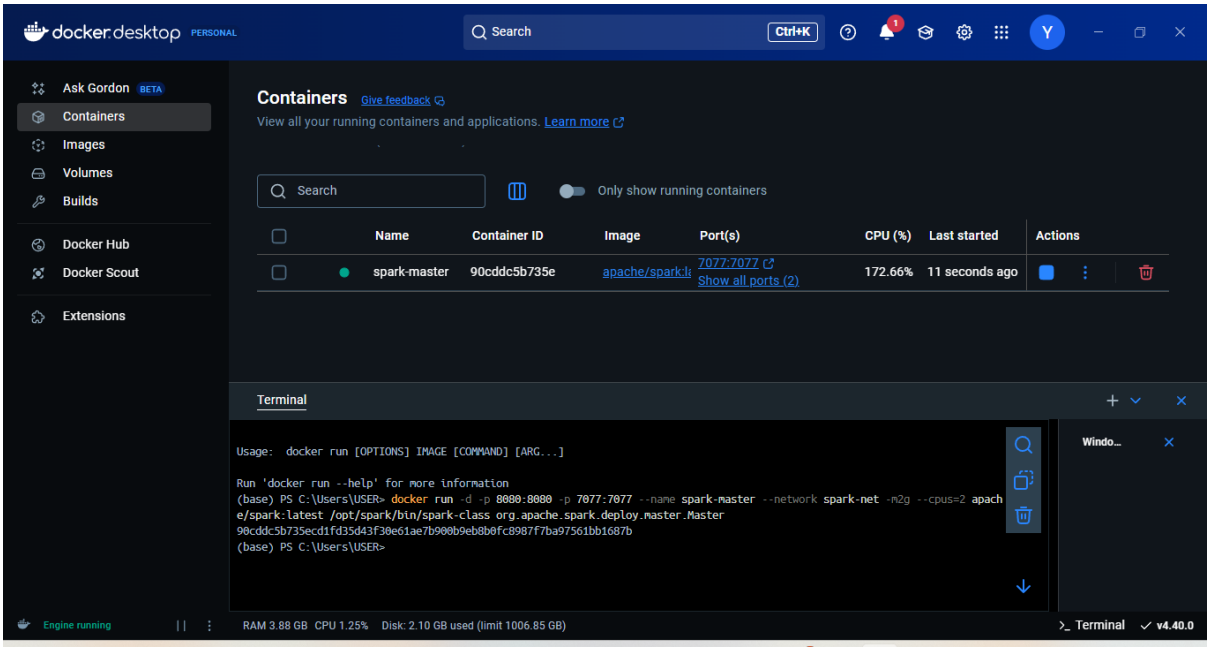


2. Menjalankan Spark Master

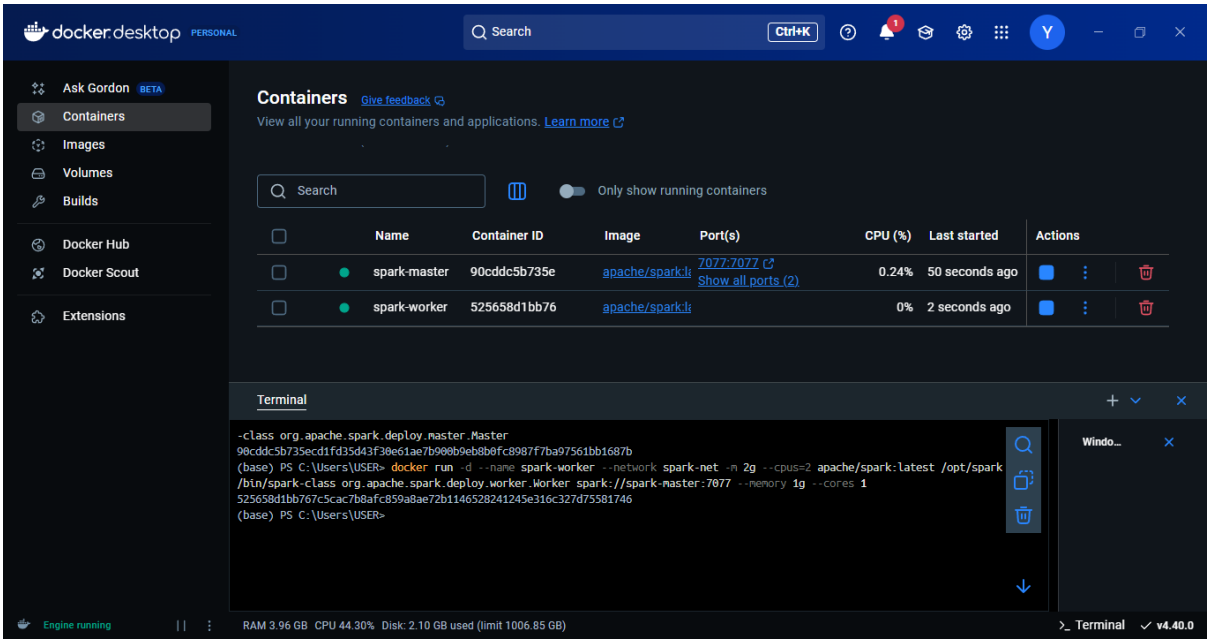
Sebelumnya buat docker network sebagai berikut



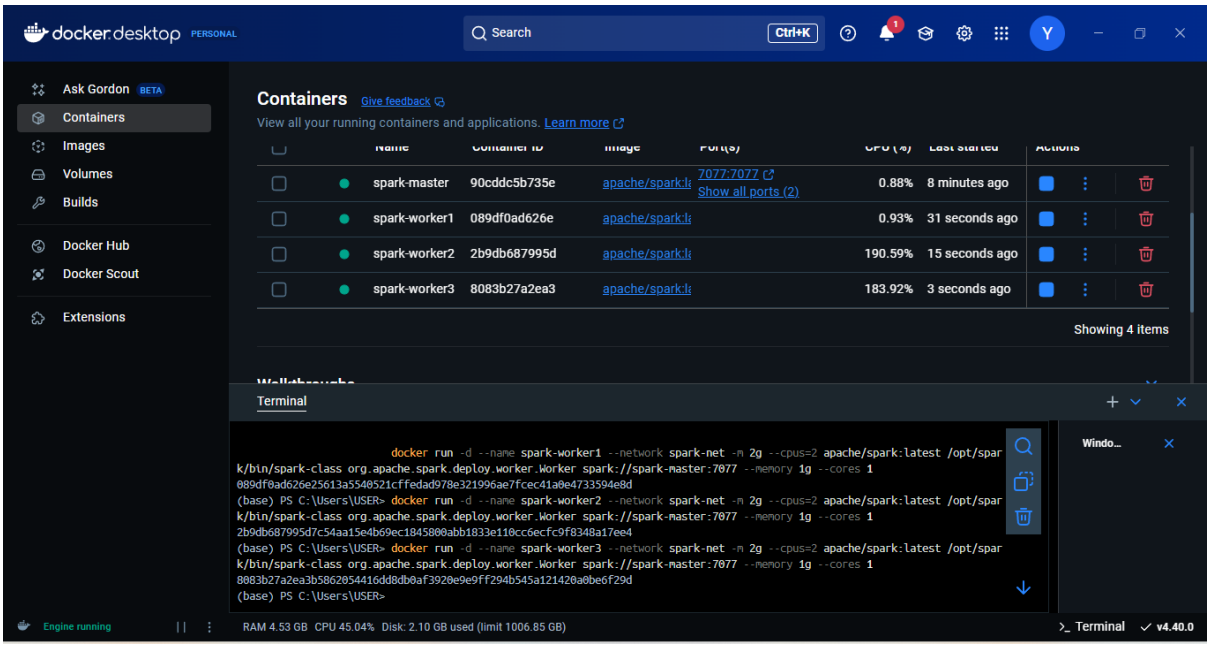
Menjalankan spark-master dalam network tersebut



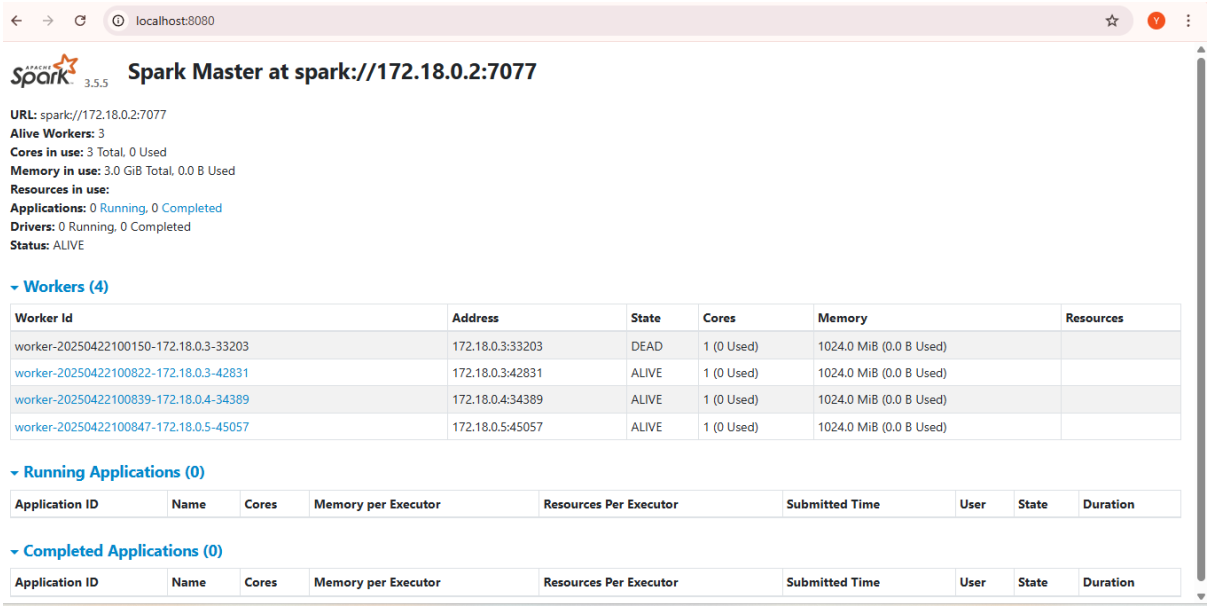
3. Menjalankan Spark Worker



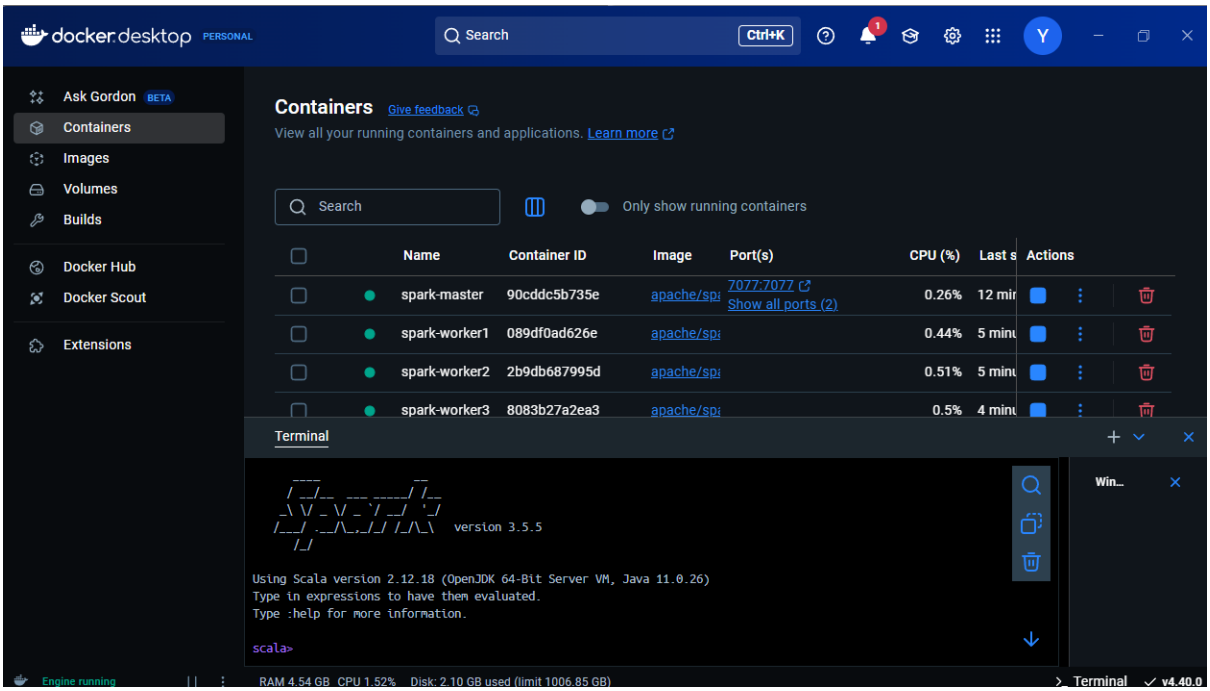
Membuat 3 worker



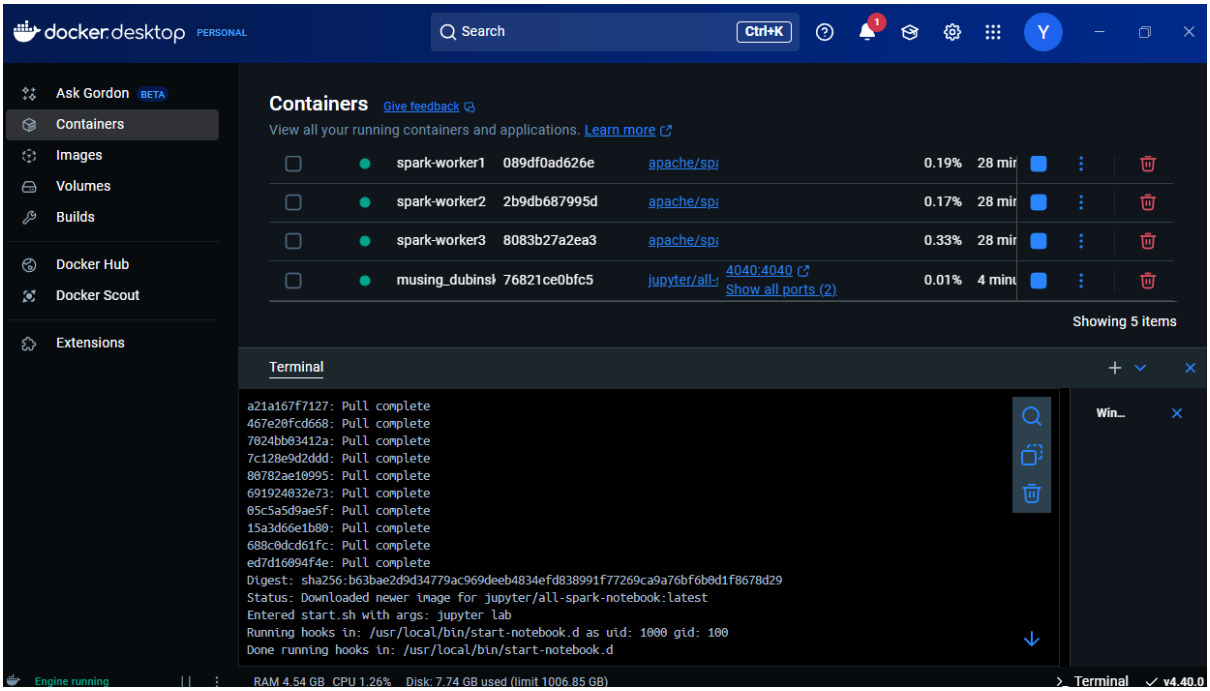
4. Mengakses Spark Web UI

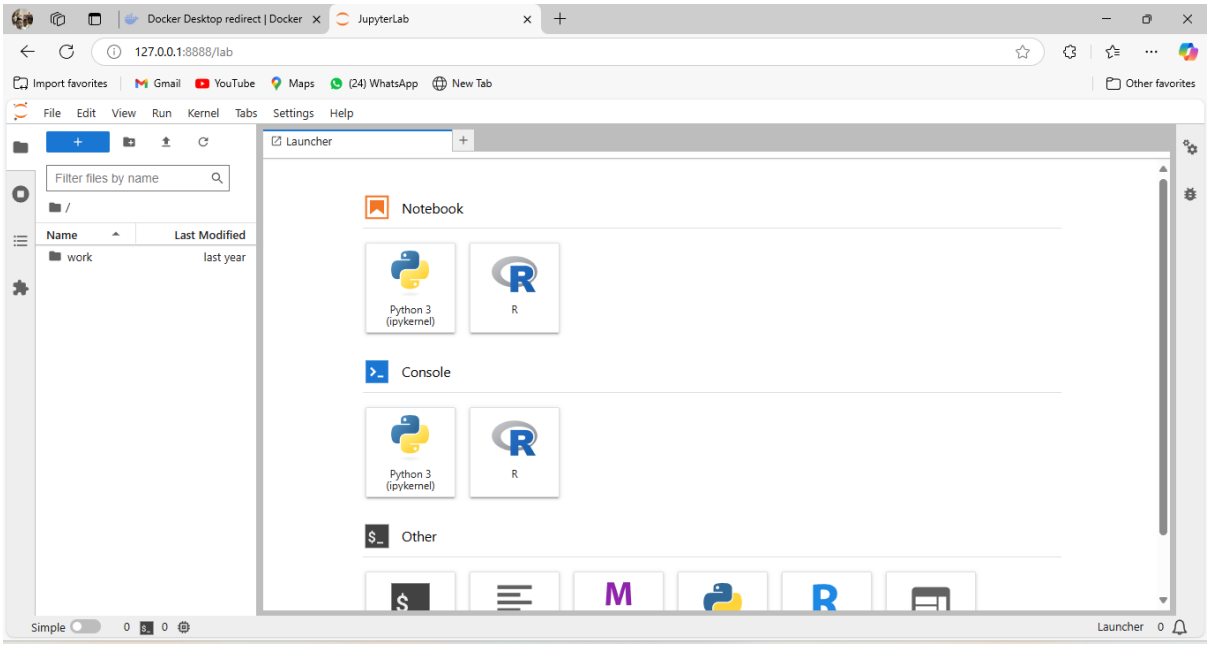


5. Menjalankan Spark Shell

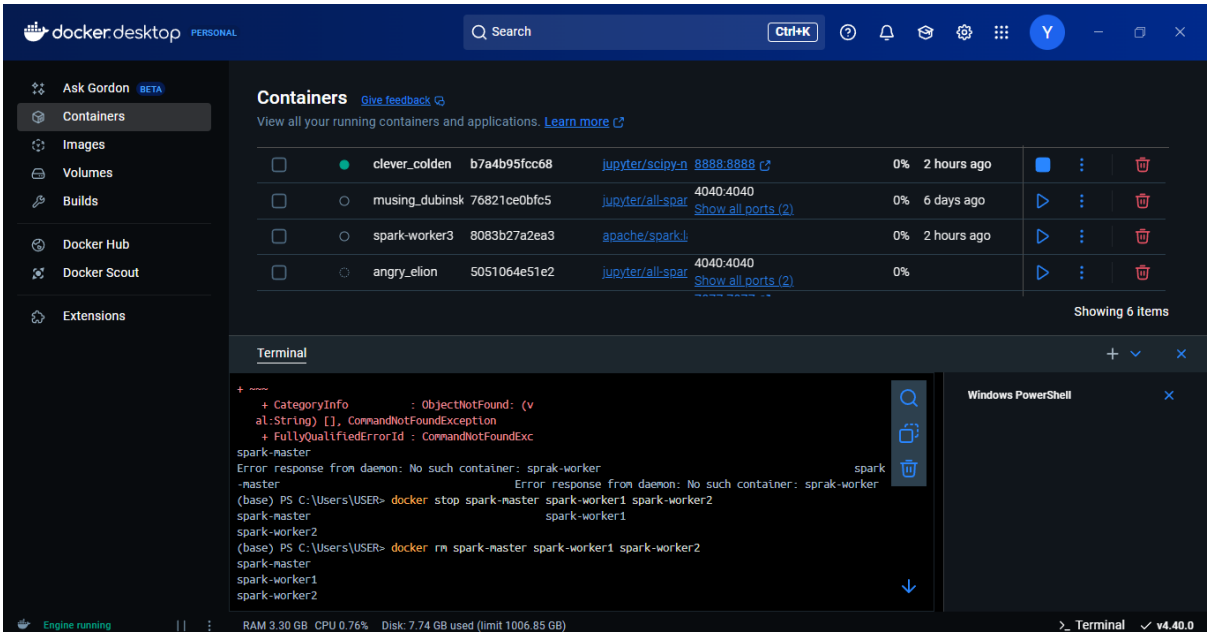


6. Menggunakan Jupyter Notebook dengan Spark



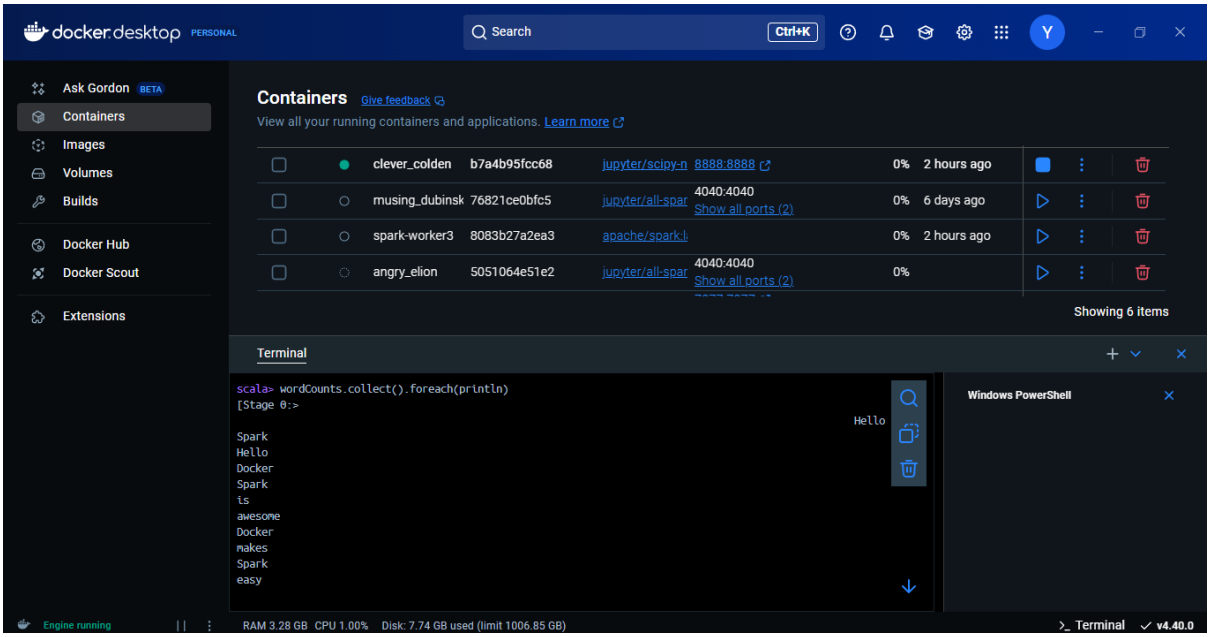


7. Meghentikan container

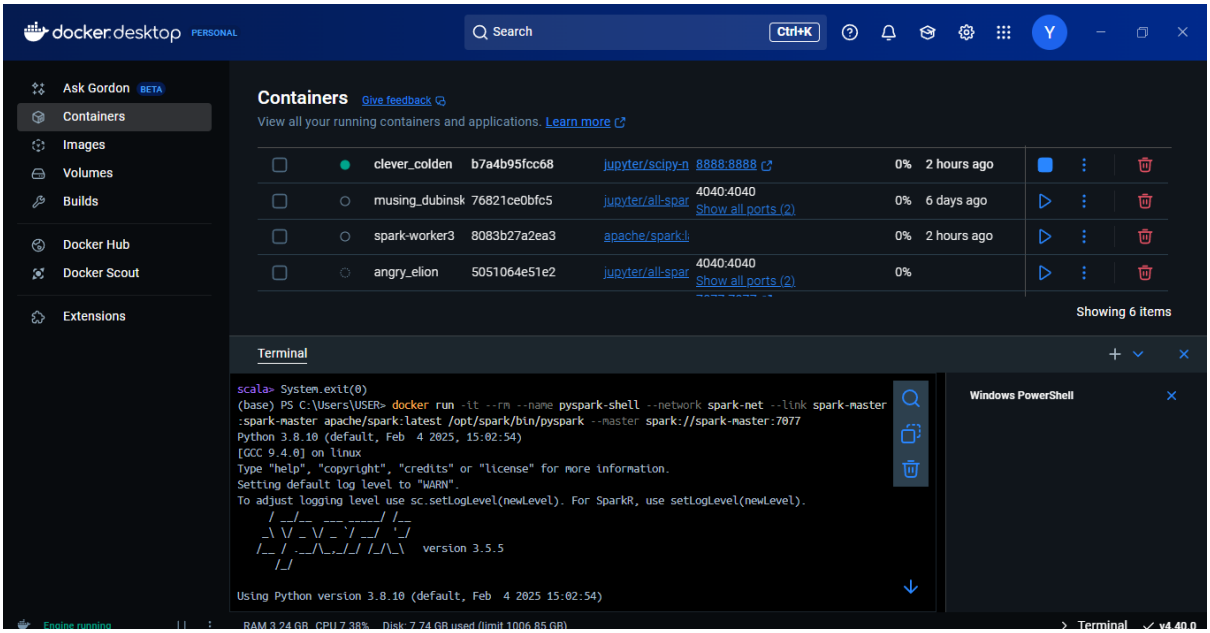


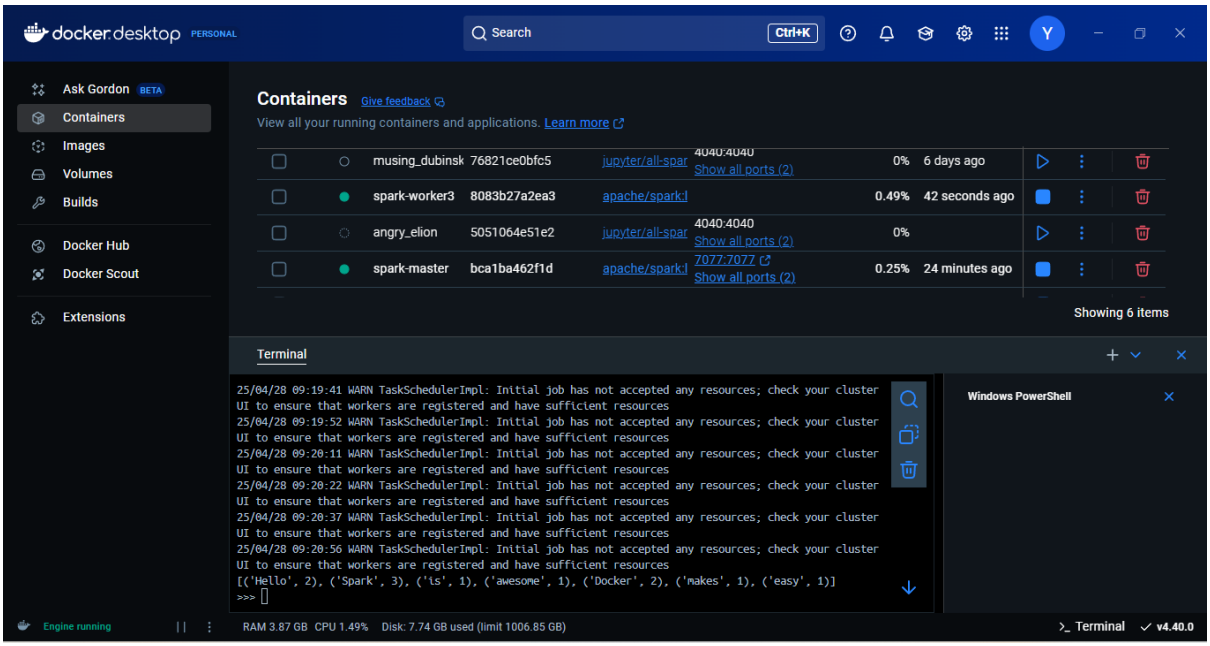
Contoh Program Word Count dengan Spark di Docker

1. Menggunakan PySpark Shell

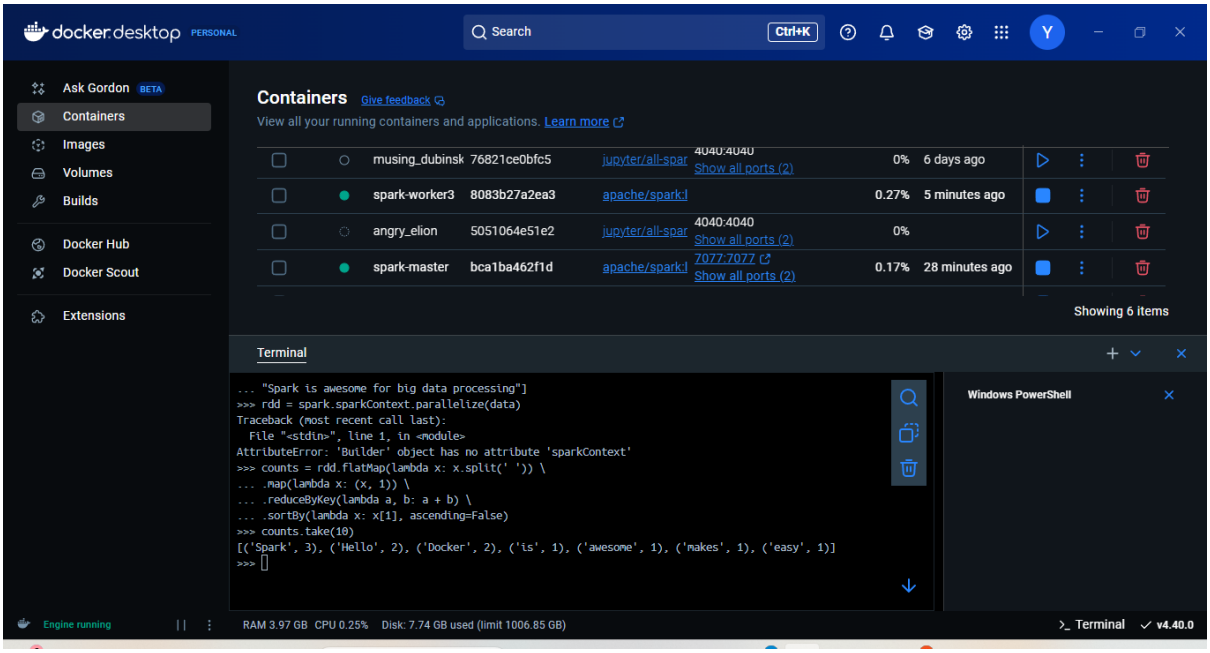


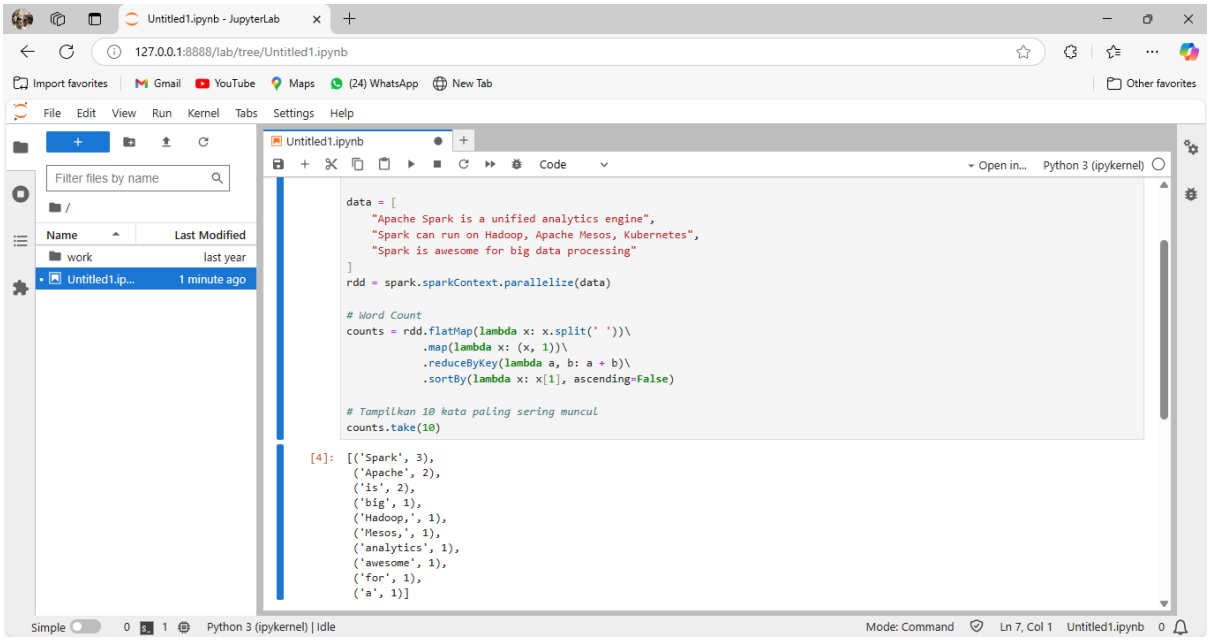
2. Menggunakan Pyspark(Python)



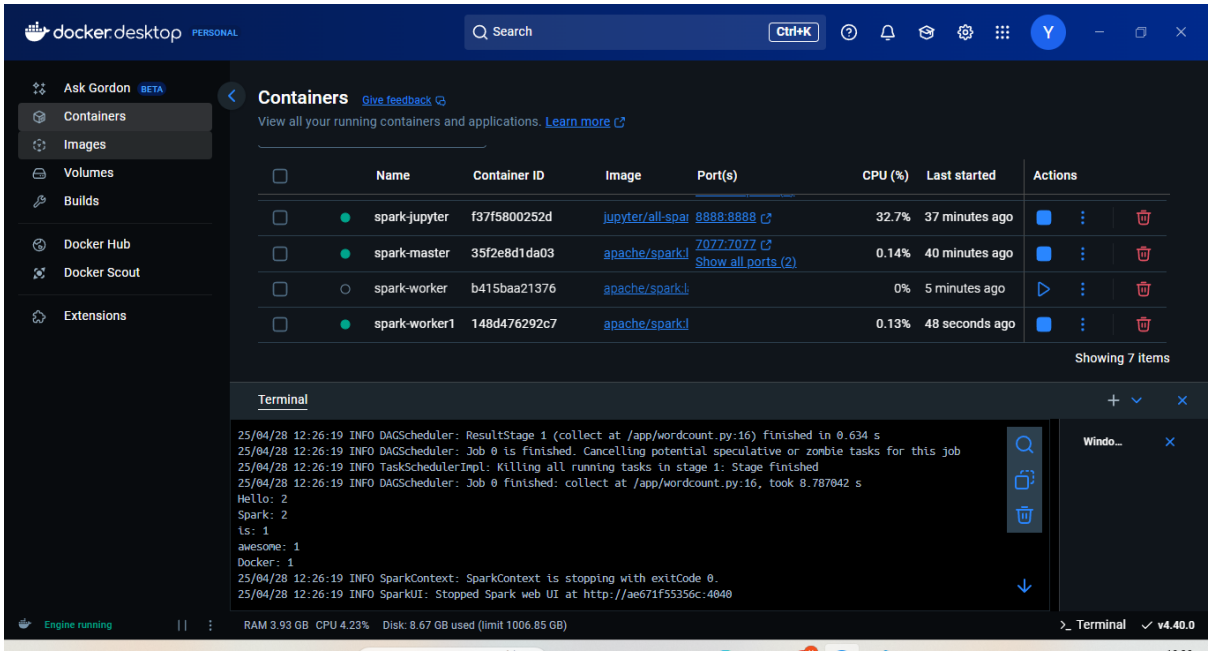


3. Menggunakan Jupyter Notebook





4. Menjalankan Program sebagai script


docker desktop PERSONAL
Search Ctrl+K
Ask Gordon BETA
Containers
Images
Volumes
Builds
Docker Hub
Docker Scout
Extensions
Containers
Give feedback
View all your running containers and applications. Learn more
Name Container ID Image Port(s) CPU (%) Last started Actions
spark-jupyter f37f5800252d jupyter/all-spar 8888:8888 32.7% 37 minutes ago
spark-master 35f2e8d1da03 apache/spark:l 7077:7077 Show all ports (2) 0.14% 40 minutes ago
spark-worker b415baa21376 apache/spark:l 0% 5 minutes ago
spark-worker1 148d476292c7 apache/spark:l 0.13% 48 seconds ago
Showing 7 items
Terminal
25/04/28 12:26:19 INFO DAGScheduler: ResultStage 1 (collect at /app/wordcount.py:16) finished in 0.634 s
25/04/28 12:26:19 INFO DAGScheduler: Job 0 is finished. Cancelling potential speculative or zombie tasks for this job
25/04/28 12:26:19 INFO TaskSchedulerImpl: Killing all running tasks in stage 1: Stage finished
25/04/28 12:26:19 INFO DAGScheduler: Job 0 finished: collect at /app/wordcount.py:16, took 8.787042 s
Hello: 2
Spark: 2
is: 1
awesome: 1
Docker: 1
25/04/28 12:26:19 INFO SparkContext: SparkContext is stopping with exitCode 0.
25/04/28 12:26:19 INFO SparkUI: Stopped Spark web UI at http://ae671f55356c:4040
Windo...
Engine running
RAM 3.93 GB CPU 4.23% Disk: 8.67 GB used (limit 1006.85 GB)
Terminal
v4.40.0