

docker compose up -d

```
PS C:\Users\nallu\Downloads\Assignment2-Document-Similarity-usingg-MapReduce-main> docker compose up
-d
time="2025-09-25T12:31:20-05:00" level=warning msg="C:\\Users\\nallu\\Downloads\\Assignment2-Document
-Similarity-usingg-MapReduce-main\\docker-compose.yml: the attribute `version` is obsolete, it will b
e ignored, please remove it to avoid potential confusion"
[+] Running 9/9
✓ Container resourcemanager Started 1.4s
✓ Container nodemanager1 Started 1.5s
✓ Container namenode Started 1.6s
✓ Container nodemanager2 Started 1.3s
✓ Container datanode1 Started 1.0s
✓ Container datanode3 Started 1.2s
✓ Container historyserver Started 1.3s
✓ Container datanode2 Started 1.1s
✓ Container nodemanager3 Started 1.2s
```

The screenshot shows the Docker Desktop application window. The top bar is blue with the Docker logo, a search bar, and a 'Ctrl+K' button. The left sidebar contains navigation options: Ask Gordon (BETA), Containers (selected), Images, Volumes, Builds, Models, MCP Toolkit (BETA), Docker Hub, Docker Scout, and Extensions. The main area displays 'Containers' with a 'Give feedback' link. It shows system usage: Container CPU usage at 37.50% / 800% (8 CPUs available) and Container memory usage at 2.53GB / 7.35GB, with a 'Show charts' link. A search bar and a toggle for 'Only show running containers' are present. Below is a table of containers:

	Name	Container ID	Image	Port(s)	Actions
<input type="checkbox"/>	postgres1	fbf9f7ded043	postgres	5432:5432	
<input checked="" type="checkbox"/>	assignment2-document-similarity-usi -		-	-	

Below the table is a 'Walkthroughs' section with two cards: 'Multi-container applications' (8 mins) and 'Containerize your application' (3 mins). A link 'View more in the Learning center' is at the bottom. The bottom status bar shows 'Engine running', system resources (RAM 4.38 GB, CPU 1.13%, Disk 4.33 GB used), and a 'New version available' notification.

```

[INFO] --- compiler:3.13.0:compile (default-compile) @ DocumentSimilarity ---
[INFO] No sources to compile
[INFO] --- resources:3.3.1:testResources (default-testResources) @ DocumentSimilarity ---
[WARNING] Using platform encoding (Cp1252 actually) to copy filtered resources, i.e. build is platform dependent!
[INFO] skip non existing resourceDirectory C:\Users\nallu\Downloads\Assignment2-Document-Similarity-usingg-MapReduce-main\src\test\resources
[INFO] --- compiler:3.13.0:testCompile (default-testCompile) @ DocumentSimilarity ---
[INFO] No sources to compile
[INFO] --- surefire:3.2.5:test (default-test) @ DocumentSimilarity ---
[INFO] No tests to run.
[INFO] --- jar:3.4.1:jar (default-jar) @ DocumentSimilarity ---
[WARNING] JAR will be empty - no content was marked for inclusion!
[INFO] Building jar: C:\Users\nallu\Downloads\Assignment2-Document-Similarity-usingg-MapReduce-main\target\DocumentSimilarity-0.0.1-SNAPSHOT.jar
[INFO] -----
[INFO] BUILD SUCCESS
[INFO] -----
[INFO] Total time: 4.419 s
[INFO] Finished at: 2025-09-25T12:35:17-05:00
[INFO] -----

```

```

2025-09-25 21:45:29,645 INFO mapreduce.Job: The url to track the job: http://resourcemanager:8088/proxy/application_1758821497360_0005/
2025-09-25 21:45:29,647 INFO mapreduce.Job: Running job: job_1758821497360_0005
2025-09-25 21:45:35,856 INFO mapreduce.Job: Job job_1758821497360_0005 running in uber mode : false
2025-09-25 21:45:35,858 INFO mapreduce.Job: map 0% reduce 0%
2025-09-25 21:45:41,975 INFO mapreduce.Job: map 100% reduce 0%
2025-09-25 21:45:48,044 INFO mapreduce.Job: map 100% reduce 100%
2025-09-25 21:45:49,067 INFO mapreduce.Job: Job job_1758821497360_0005 completed successfully
2025-09-25 21:45:49,268 INFO mapreduce.Job: Counters: 54
  File System Counters
    FILE: Number of bytes read=1134
    FILE: Number of bytes written=460801
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=3921
    HDFS: Number of bytes written=7239
    HDFS: Number of read operations=8
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Launched map tasks=1
    Launched reduce tasks=1
    Rack-local map tasks=1
    Total time spent by all maps in occupied slots (ms)=14336
    Total time spent by all reduces in occupied slots (ms)=35160
    Total time spent by all map tasks (ms)=3584
    Total time spent by all reduce tasks (ms)=4395
    Total vcore-milliseconds taken by all map tasks=3584
    Total vcore-milliseconds taken by all reduce tasks=4395
    Total megabyte-milliseconds taken by all map tasks=14680064
    Total megabyte-milliseconds taken by all reduce tasks=36003840
  Map-Reduce Framework
    Map input records=20
    Map output records=293
    Map output bytes=6089
    Map output materialized bytes=1126
    Input split bytes=116
    Combine input records=0
    Combine output records=0
    Reduce input groups=20

```

```
root@130ab40c476d:/opt/hadoop-3.2.1/share/hadoop/mapreduce# hadoop fs -cat /output/small_results/*
2025-09-25 21:46:20,412 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
Document10,Document19 Similarity: 0.42
Document20,Document3 Similarity: 0.53
Document10,Document18 Similarity: 0.50
Document20,Document4 Similarity: 0.63
Document20,Document5 Similarity: 0.61
Document10,Document15 Similarity: 0.50
Document13,Document15 Similarity: 0.47
Document7,Document8 Similarity: 0.55
Document10,Document14 Similarity: 0.40
Document13,Document16 Similarity: 0.67
Document10,Document17 Similarity: 0.50
Document7,Document9 Similarity: 0.60
Document10,Document16 Similarity: 0.61
```