

### 1) Cheated version

[https://www.cloudskillsboost.google/catalog\\_lab/1307](https://www.cloudskillsboost.google/catalog_lab/1307)

```
hadoop fs -mkdir /cat/
```

```
hadoop fs -moveFromLocal Orangecat.jpeg /cat/Orangecat.jpeg
```

### 2) Docker version

#### 1. Install Docker/ Docker Desktop

<https://www.docker.com/get-started/>

#### 2. Pull Docker image (You can use other option too)

```
docker pull bde2020/hadoop-namenode:2.0.0-hadoop2.7.4-java8
```

```
docker pull bde2020/hadoop-datanode:2.0.0-hadoop2.7.4-java8
```

#### 3. Start hadoop network and run both **namenode** and **datanode**

- a. docker network create hadoop
- b. docker run -d --net hadoop --name namenode \  
-e CLUSTER\_NAME=test \  
-p 9870:9870 \  
bde2020/hadoop-namenode:2.0.0-hadoop2.7.4-java8
- c. docker run -d --net hadoop --name datanode \  
-e CORE\_CONF\_fs\_defaultFS=hdfs://namenode:8020 \  
-e CLUSTER\_NAME=test \  
bde2020/hadoop-datanode:2.0.0-hadoop2.7.4-java8

#### 4. Access hdfs

Use **docker exec -it namenode** command to access hdfs command

For example

- a. **docker exec -it namenode** hdfs dfs -ls /
- b. docker exec -it namenode hdfs dfs -mkdir /test
- c. docker exec -it namenode hdfs dfs -ls /

```
docker cp my_data.csv namenode:/tmp/my_data.csv
```

#### 5. Copy your file into docker → copy file.txt to namenode

- a. docker cp file.txt namenode:/tmp/file.txt

#### 6. Copy/move your file into hdfs → local name node

- a. docker exec -it namenode hdfs dfs -moveFromLocal tmp/file.txt /test/
- b. docker exec -it namenode hdfs dfs -ls /test/

7. Use Orangecat.jpeg

**3) Another option** → download Cloudera image (heavier version 4GB+)

1. `docker pull cloudera/quickstart:latest`
2. `docker run --hostname=quickstart.cloudera --privileged=true -t -i \`  
    `-p 8888:8888 -p 7180:7180 \`  
    `cloudera/quickstart /usr/bin/docker-quickstart`
  - a. `Hadoop fs -moveFromLocal Orangecat.jpeg /cat/Orangecat.jpeg`