

## Intro to Docker, Hadoop, Map-Reduce and Hive

### Setting the Hadoop Environment

- Download and Install (Windows OS): <https://docs.docker.com/desktop/windows/install/>
- Test Docker installation:
  - o Open CMD
  - o Run “docker --version”
  - o Run “docker run hello-world”
    - This command will download a docker image and will eventually display some messages like the one below. This means that your docker is running.

```
Unable to find image 'hello-world:latest' locally
latest: Pulling from library/hello-world
2db29710123e: Pull complete
Digest: sha256:cc15c5b292d8525effc0f89cb299f1804f3a725c8d05e158653a563f15e4f685
Status: Downloaded newer image for hello-world:latest

Hello from Docker!
This message shows that your installation appears to be working correctly.

To generate this message, Docker took the following steps:
1. The Docker client contacted the Docker daemon.
2. The Docker daemon pulled the "hello-world" image from the Docker Hub.
   (amd64)
3. The Docker daemon created a new container from that image which runs the
   executable that produces the output you are currently reading.
4. The Docker daemon streamed that output to the Docker client, which sent it
   to your terminal.

To try something more ambitious, you can run an Ubuntu container with:
$ docker run -it ubuntu bash

Share images, automate workflows, and more with a free Docker ID:
https://hub.docker.com/

For more examples and ideas, visit:
https://docs.docker.com/get-started/
```

- Download docker images. In CMD, run the ff. commands:
  - o docker pull bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8
  - o docker pull bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8
  - o docker pull bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8
  - o docker pull bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8
  - o docker pull bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8
  - o docker pull bde2020/hive:2.3.2-postgresql-metastore

```
:\\Users\\leif\\Desktop\\docker-hadoop>docker pull bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8
```

- Place the ff. files below into a folder (To be executed during the session proper)



Map-Reduce Function Development. Download and install the ff.:

- JDK: <https://www.oracle.com/java/technologies/downloads/#jdk17-windows>
- Eclipse Environment: <https://www.eclipse.org/downloads/> (Install JDK first before installing Eclipse)
- Hadoop Libraries (Download only):
  - o <https://mvnrepository.com/artifact/org.apache.hadoop/hadoop-common/3.3.1>
  - o <https://mvnrepository.com/artifact/org.apache.hadoop/hadoop-mapreduce-client-core/3.3.1>
  - o <https://mvnrepository.com/artifact/org.apache.hadoop/hadoop-mapreduce-client-common/3.3.1>