01 Hive Environment -KirkYagami 👨‍💻🕵️

Installation

Step 1

```
mkdir c:/hive_data
cd c:/hive_data

wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-BD0225EN-SkillsNetwork/data/emp.csv
```

Step 2

```
docker pull apache/hive:4.0.0-alpha-1
```

Step 3

-> Now, you will run the hive server on port `10002`. You will name the server instance `myhiveserver`. We will mount the local `data` folder in the hive server as `hive_custom_data`. This would mean that the whole `data` folder that you created locally, along with anything you add in the data folder, is copied into the container under the directory `hive_custom_data`.

```
docker run -d -p 10000:10000 -p 10002:10002 --env SERVICE_NAME=hiveserver2 -v C:/Hive_Data:/hive_custom_data --name nikshiveserver apache/hive:4.0.0-alpha-1
```

- `-d` = detached mode: runs the container in the background, docker runs the container and returns the control of terminal immediately
- `-p` = port mapping. `host_port:container_port`.
- `--env SERVICE_NAME=hiveserver2` : sets an environment variable in the container
- `-v C:/Hive_Data:/hive_custom_data` = This flag is for mounting a volume. The format is `host_path:container_path`
- `--name myhiveserver` = sets a name for this container
- `apache/hive:4.0.0-alpha-1` = specifies the image to be used

Step 4

1. **Launch beeline to write HiveQL**

```
docker exec -it nikshiveserver beeline -u 'jdbc:hive2://localhost:10000/'
# To execute queries
```

2. **To exit from the beeline**

```
!quit
```

3. **Launch bash to read the directories and other stuff**

```
docker exec -it nikshiveserver /bin/bash

cd /hive_custom_data
ls # files and folder will be listed which are stored in the directory mounted to
hive
```