Assignment\_04 week-09

Created all the folders hierarchy

A screenshot of a computer

Description automatically generated

Let's clone the data set in our virtual box: with the command:

git clone [git@github.com:datacharmer/test\_db.git](mailto:git@github.com:datacharmer/test_db.git)

A screenshot of a computer

Description automatically generated

Done cloning the test db in virtual box.

A screenshot of a computer

Description automatically generated

How do we load that employees.sql into our vb?

With this command:

sudo mysql < employees.sql

A screenshot of a computer

Description automatically generated

Now lets login:

sudo mysql

A screenshot of a computer

Description automatically generated

Whats the command to show databases?

show databases;

A screenshot of a computer program

Description automatically generated except employees remaining 4 comes naturally

How do we inspect the employees database? Just play with some commands to explore the tables

use tab employees;

show tables;

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

Now how do we create a user?

First we need to save the sql file into our local repository

A screen shot of a computer

Description automatically generated

Then we do git pull in our vb to save it over there also

First we do quit in maria db and do git pull

A screen shot of a computer program

Description automatically generated

Lets create a user with the command:

sudo mysql < create-user.sql

A screenshot of a computer program

Description automatically generated

Now we need to get the JDBC Driver

Where do we get the connector j?

A screenshot of a computer

Description automatically generated

Download it..the zip file

A screenshot of a computer

Description automatically generated

Don’t worry abput oracle sign up ,..just go and download

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Don’t extract it here…we keep it compressed..we will move this to our local repo

A screenshot of a computer

Description automatically generated

A screen shot of a computer

Description automatically generated

Go ahead and push it and pull it

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

Now lets unzip it with the command:

unzip mysql-connector-j-8.3.0.zip

A screenshot of a computer program

Description automatically generated

A screen shot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

Copy this jar file into our spark directory.

The command is;

cp ./mysql-connector-j-8.3.0.jar ~/spark/jars

cause this is where spark keeps all the jar files

then do ls ~/spark/jars

A screen shot of a computer

Description automatically generated

Check for that jar file in the list:…yes its there

A screen shot of a computer

Description automatically generated

Now we know how to connect to our database:

Page 133

Part 1:

Writing the spark application in pyspark

mysql --password=cluster --user=worker to connect to mysql database

command to run

use employees;

describe aces;

describe left\_table;

describe left\_tempview;

describe left\_df;

spark submit command:

spark-submit --jars ~/spark/jars/mysql-connector-j-8.3.0.jar /home/vagrant/itmd-521/nannadevara/itmd-521/labs/week-09/py/assignment\_04.py