

// Lab Execution Steps //

1. Copy the PostgreSQL JAR file and hive-site.xml to spark container.

```
# docker cp  
/home/labuser/Project/Installation_and_Execution/postgresql-42.3.1.jar  
hdp_spark-master:/spark/jars/  
# docker cp hdp_hive-server:/opt/hive/conf/hive-site.xml  
/home/labuser/Downloads/  
# docker cp /home/labuser/Downloads/hive-site.xml hdp_spark-  
master:/spark/conf/
```

2. Copy the Data files to spark & hive containers.

```
# docker cp /home/labuser/Project/Data hdp_spark-master:/  
# docker cp /home/labuser/Project/Data hdp_hive-server:/opt/
```

3. Open a Scala shell.

```
# docker exec -i -t hdp_spark-master bash  
# ./spark/bin/spark-shell
```

4. Please refer the project videos for step-wise execution.

Note :

Commands to login to Hive Container and open a hive shell.

```
# docker exec -i -t hdp_hive-server bash  
# hive
```

Data files path:

Spark Container - /Data/<filename>
Hive Container - /opt/Data/<filename>