

// Lab Execution Steps //

1. Copy the flight data files to another folder to decompress them.

```
# cd /home/labuser/  
# mkdir Data/  
# mkdir Data/flights/  
# cp /home/labuser/Project/Data/DataExpo2009/*.csv.bz2  
/home/labuser/Data/flights/  
# cp /home/labuser/Project/Data/*.csv /home/labuser/Data/  
# cd /home/labuser/Data/flights/  
# bzip2 -d *.csv.bz2# docker cp /home/labuser/Data hdp_hive-  
server:/opt/
```

2. Copy the data files to the Hive docker container and to HDFS for creating Hive external tables on top of them.

```
# docker exec -i -t hdp_hive-server bash  
# hdfs dfs -mkdir /opt  
# hdfs dfs -mkdir /opt/Data/  
# hdfs dfs -mkdir /opt/Data/airports/  
# hdfs dfs -mkdir /opt/Data/flights/  
# hdfs dfs -mkdir /opt/Data/planes/  
# hdfs dfs -mkdir /opt/Data/carriers/  
# hdfs dfs -put /opt/Data/airports.csv /opt/Data/airports/  
# hdfs dfs -put /opt/Data/plane-data.csv /opt/Data/planes/  
# hdfs dfs -put /opt/Data/carriers.csv /opt/Data/carriers/  
# hdfs dfs -put /opt/Data/flights/2005.csv /opt/Data/flights/  
# hdfs dfs -put /opt/Data/flights/2006.csv /opt/Data/flights/  
# hdfs dfs -put /opt/Data/flights/2007.csv /opt/Data/flights/  
# hdfs dfs -put /opt/Data/flights/2008.csv /opt/Data/flights/
```

3. Please refer to the Project videos and code files for step-wise execution.

Note - Code files are provided in the project folder in Lab environment.