**Demo 1]Understanding Hive tables**

1) Download the attached file in **STAGING\_AREA**

2) Extract it

3) Create a folder **LABS\_HOME**/lab7.1

4) Put the extracted contents in above location

5) Wait for the instructions.

Dataset - Lab1/Lab7.1.rar

Reference - HDPDev-PigHive-LabGuide-Rev 6.doc

**Lab 2] Understanding partitions and skews**

1) Download the RAR in STAGING\_AREA.

2) Extract the contents

3) Pull the contents in LABS\_HOME/demos

4) Refer the lab manual

5) Wait for the instructions

Dataset - Lab2/partitionandskewtables.rar

Reference - HDPDev-PigHive-LabGuide-Rev 6.doc

**Lab 3] Understanding Bucketed table**

1) Download the RAR in STAGING\_AREA.

2) Extract the contents

3) Pull the contents in LABS\_HOME/bucketed\_table

4) Wait for the instructions

Dataset - Lab3/bucketed\_table\_CDP.rar

Reference - NA

**Lab 4] Analyzing BigData with Hive**

1) Refer the lab manual

2) Follow the instructions in lab manual

3) Translate the paths given in the lab manual to paths applicable for the cloudera sandbox.

Note -

1. No data for this lab, the data needed is already there with you through previous labs.
2. For your reference solution is provided in the folder Lab4/Solution.rar. Verify your solution by referring to the file

Dataset - NA

Reference - HDPDev-PigHive-LabGuide-Rev 6.doc

**Timeline =** 60 Minutes.

**Lab 5] Writing a Hive UDF**

1) Extract the rar in **STAGING\_AREA**

2) Create a folder **LABS\_HOME**/lab11.1

3) Put the extracted contents in above folder

4) Refer the lab manual

5) Wait for the instructions from an Instructor

Dataset - Lab5/HiveUDF.rar

Reference - HDPDev-PigHive-LabGuide-Rev 6.doc

**Lab 6] Computing ngrams**

1) Refer the lab manual

2) Wait for the instructions

Dataset - NA

Reference - HDPDev-PigHive-LabGuide-Rev 6.doc

**Lab 7] Joining data in Hive**

1) Refer lab manual

2) Follow instructions in lab manual and implement it.

Dataset - Lab7/Lab7.4.rar

Reference - HDPDev-PigHive-LabGuide-Rev 6.doc

**Time Line = 30 Minutes**

**Lab 8] Computing ngrams of Emails in Avro Format**

1) Download the rar in STAGING\_AREA

2) Move the extracted contents to LABS\_HOME/lab7.5

3) Refer the lab manual

4) Wait for the instructions

Dataset - Lab8/Lab7.5.rar

Reference - HDPDev-PigHive-LabGuide-Rev 6.doc

**Lab 9] Understanding tbplproperties**

1) Download the rar in STAGING\_AREA

2) Move the extracted contents to LABS\_HOME/tblproperties

3) Wait for the instructions

Dataset - Lab9/

**Lab10] Understanding file formats**

1. Download the .hive file from Lab10 folder
2. Put it in LABS\_HOME/hive\_file\_formats/
3. Wait for the instructions

Dataset - Lab10/