**Section 3 : PARTITION**

**Set hive.exec.dynamic.partition=true;**

**Set hive.exec.dynamic.partition.mode = nonstrict;**

------------------------------------------------------------------

**CREATE TABLE Migrants (**

**Web\_ID INT,**

**Reported\_Date STRING,**

**Reported\_Year INT,**

**Reported\_Month VARCHAR(5),**

**No\_Dead INT,**

**Estd\_No\_Missing INT,**

**Tot\_Dead\_Missing INT,**

**No\_Survivors INT,**

**No\_Females INT,**

**No\_Males INT,**

**No\_Children INT,**

**Death\_Cause String,**

**Loc\_Desc String,**

**Info\_Src String,**

**Incident\_Region STRING**

**)**

**row format delimited fields terminated BY '\t'**

**lines terminated BY '\n'**

**tblproperties("skip.header.line.count"="1")**

**load data inpath '/user/hive/warehouse/Migrants\_Global.txt'**

**overwrite into table Migrants;**

------------------------------------------------------------------

**CREATE TABLE Migrants\_Partitioned (**

**Web\_ID INT,**

**Reported\_Date STRING,**

**Reported\_Year INT,**

**Reported\_Month VARCHAR(5),**

**No\_Dead INT,**

**Estd\_No\_Missing INT,**

**Tot\_Dead\_Missing INT,**

**No\_Survivors INT,**

**No\_Females INT,**

**No\_Males INT,**

**No\_Children INT,**

**Death\_Cause String,**

**Loc\_Desc String,**

**Info\_Src String**

**)Partitioned By (Incident\_Region STRING);**

**insert overwrite table migrants\_partitioned partition(incident\_region) select \* from migrants;**

------------------------------------------------------------------

**hive> select \* from migrants where incident\_region = "Caribbean"**

*Time taken: 51.785 seconds, Fetched: 8 row(s)*

**hive> select \* from migrants\_partitioned where incident\_region = "Caribbean";**

*Time taken: 34.972 seconds, Fetched: 8 row(s)*

------------------------------------------------------------------

**hive> select \* from migrants where incident\_region ="Caribbean" OR incident\_region = "US-Mexico Border";**

*Time taken: 0.819 seconds, Fetched: 387 row(s)*

**hive> select \* from migrants\_partitioned where incident\_region ="Caribbean" OR incident\_region = "US-Mexico Border";**

*Time taken: 1.15 seconds, Fetched: 387 row(s)*

------------------------------------------------------------------

**DESCRIBE FORMATTED migrants\_partitioned PARTITION(incident\_region="US-Mexico Border");**

------------------------------------------------------------------

**analyze table default.migrants compute statistics;**

*Table default.migrants stats: [numFiles=1, numRows=1510, totalSize=259851, rawDataSize=256831]*

*MapReduce Jobs Launched:*

*Stage-Stage-0: Map: 1 Cumulative CPU: 1.67 sec HDFS Read: 263782 HDFS Write: 78 SUCCESS*

*Total MapReduce CPU Time Spent: 1 seconds 670 msec*

*OK*

*Time taken: 95.702 seconds*

------------------------------------------------------------------

**analyze table default.migrants\_partitioned compute statistics;**

*Partition default.migrants\_partitioned{incident\_region=Caribbean} stats: [numFiles=1, numRows=8, totalSize=1352, rawDataSize=1344]*

*Partition default.migrants\_partitioned{incident\_region=Central America} stats: [numFiles=1, numRows=61, totalSize=9909, rawDataSize=9848]*

*Partition default.migrants\_partitioned{incident\_region=Central Asia} stats: [numFiles=1, numRows=1, totalSize=153, rawDataSize=152]*

*Partition default.migrants\_partitioned{incident\_region=East Asia} stats: [numFiles=1, numRows=1, totalSize=218, rawDataSize=217]*

*Partition default.migrants\_partitioned{incident\_region=Europe} stats: [numFiles=1, numRows=78, totalSize=11988, rawDataSize=11910]*

*Partition default.migrants\_partitioned{incident\_region=Horn of Africa} stats: [numFiles=1, numRows=66, totalSize=10326, rawDataSize=10260]*

*Partition default.migrants\_partitioned{incident\_region=Mediterranean} stats: [numFiles=1, numRows=204, totalSize=31748, rawDataSize=31544]*

*Partition default.migrants\_partitioned{incident\_region=Middle East} stats: [numFiles=1, numRows=52, totalSize=7786, rawDataSize=7734]*

*Partition default.migrants\_partitioned{incident\_region=North Africa} stats: [numFiles=1, numRows=318, totalSize=58823, rawDataSize=58505]*

*Partition default.migrants\_partitioned{incident\_region=Region of Incident} stats: [numFiles=1, numRows=1, totalSize=102, rawDataSize=101]*

*Partition default.migrants\_partitioned{incident\_region=South America} stats: [numFiles=1, numRows=11, totalSize=1867, rawDataSize=1856]*

*Partition default.migrants\_partitioned{incident\_region=South Asia} stats: [numFiles=1, numRows=13, totalSize=2076, rawDataSize=2063]*

*Partition default.migrants\_partitioned{incident\_region=Southeast Asia} stats: [numFiles=1, numRows=30, totalSize=5331, rawDataSize=5301]*

*Partition default.migrants\_partitioned{incident\_region=Sub-Saharan Africa} stats: [numFiles=1, numRows=287, totalSize=45149, rawDataSize=44862]*

*Partition default.migrants\_partitioned{incident\_region=US-Mexico Border} stats: [numFiles=1, numRows=379, totalSize=59599, rawDataSize=59220]*

*MapReduce Jobs Launched:*

*Stage-Stage-0: Map: 1 Cumulative CPU: 2.28 sec HDFS Read: 257014 HDFS Write: 1228 SUCCESS*

*Total MapReduce CPU Time Spent: 2 seconds 280 msec*

*OK*

*Time taken: 41.824 seconds*

------------------------------------------------------------------

**show partitions migrants\_partitioned;**

*incident\_region=Caribbean*

*incident\_region=Central America*

*incident\_region=Central Asia*

*incident\_region=East Asia*

*incident\_region=Europe*

*incident\_region=Horn of Africa*

*incident\_region=Mediterranean*

*incident\_region=Middle East*

*incident\_region=North Africa*

*incident\_region=Region of Incident*

*incident\_region=South America*

*incident\_region=South Asia*

*incident\_region=Southeast Asia*

*incident\_region=Sub-Saharan Africa*

*incident\_region=US-Mexico Border*

*Time taken: 4.386 seconds, Fetched: 15 row(s)*

------------------------------------------------------------------

**alter table migrants\_partitioned partition (incident\_region = "Region of Incident") rename to partition (incident\_region="WasteData");**

------------------------------------------------------------------

The following statements prevent the partition from being queried

**ALTER TABLE migrants\_partitioned PARTITION(incident\_region = "WasteData") ENABLE OFFLINE;**

**select \* FROM migrants\_partitioned where incident\_region = "WasteData";**

*Error while compiling statement: FAILED: SemanticException [Error 10113]: Query against an offline table or partition Table migrants\_partitioned Partition incident\_region=WasteData*

------------------------------------------------------------------

The following statements prevent the partition from being dropped

**ALTER TABLE migrants\_partitioned PARTITION(incident\_region = "Europe") ENABLE NO\_DROP;**

**ALTER TABLE migrants\_partitioned DROP IF EXISTS PARTITION(incident\_region = "Europe");**

*Error while compiling statement: FAILED: SemanticException [Error 30011]:*

*Partition protected from being dropped default@migrants\_partitioned@incident\_region=Europe*

------------------------------------------------------------------

Extras

**DROP TABLE migrants;**

**TRUNCATE TABLE migrants\_partitioned**

To Solve Code Error 2, see below

**set hive.exec.max.dynamic.partitions=700;**

**set hive.exec.max.dynamic.partitions.pernode=700;**

--------------------------------------- THE END --------------------------------------------------------------