

## STATE PARTITIONING DOCUMENTATION

by Prabhat Srivastava

[srivastavaprabhat23@gmail.com](mailto:srivastavaprabhat23@gmail.com)

1. Create a file allStates.csv with values of **states, district, enrolment number**.

For example:

```
uttar_pradesh,varanasi,01
delhi,new_delhi,02
uttar_pradesh,allahabad,03
gujrat,ahmdabad,04
uttar_pradesh,lucknow,04
gujrat,surat,05
```

2. Creation of Table all states

```
hive> CREATE TABLE allStates(state STRING, district STRING, enrolments STRING)
row format delimited fields terminated by ',';
```

3. Loading data into created table all states

```
hive> LOAD DATA LOCAL INPATH '/home/prabhat/Desktop/allStates.csv' into TABLE
allStates;
```

4. Creation of partition table

```
hive> CREATE TABLE statePart(district STRING, enrolments STRING) PARTITIONED
BY(state STRING);
```

5. For partition we have to set this property

```
hive> set hive.exec.dynamic.partition.mode=nonstrict;
```

6. Loading data into partition table

```
hive> INSERT OVERWRITE TABLE statePart PARTITION(state) SELECT
district,enrolments,state from allStates;
```