

```
CREATE database covid_db;
```

```
use covid_db;
```

```
SET hive.exec.max.dynamic.partitions=500000;
```

```
SET hive.exec.max.dynamic.partitions.pernode = 500000;
```

```
SET hive.exec.dynamic.partition=true;
```

```
SET hive.exec.dynamic.partition.mode=nonstrict;
```

```
CREATE TABLE IF NOT EXISTS covid_db.covid_staging
```

```
(  
  Country                                STRING,  
  Total_Cases                           DOUBLE,  
  New_Cases                             DOUBLE,  
  Total_Deaths                           DOUBLE,  
  New_Deaths                            DOUBLE,  
  Total_Recovered                        DOUBLE,  
  Active_Cases                          DOUBLE,  
  Serious                               DOUBLE,  
  Tot_Cases                             DOUBLE,  
  Deaths                               DOUBLE,  
  Total_Tests                           DOUBLE,  
  Tests                                DOUBLE,  
  CASES_per_Test                        DOUBLE,  
  Death_in_Closed_Cases                 STRING,  
  Rank_by_Testing_rate                   DOUBLE,  
  Rank_by_Death_rate                     DOUBLE,  
  Rank_by_Cases_rate                     DOUBLE,  
  Rank_by_Death_of_Closed_Cases          DOUBLE  
)
```

```
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
```

```
STORED as TEXTFILE
```

```
LOCATION '/user/cloudera/ds/COVID_HDFS_LZ'
```

```
tblproperties ("skip.header.line.count"="1");
```

```
CREATE EXTERNAL TABLE IF NOT EXISTS covid_db.covid_ds_partitioned
```

```
(  
  Country                                STRING,  
  Total_Cases                           DOUBLE,  
  New_Cases                             DOUBLE,  
  Total_Deaths                           DOUBLE,  
  New_Deaths                            DOUBLE,  
  Total_Recovered                        DOUBLE,  
  Active_Cases                          DOUBLE,  
  Serious                               DOUBLE,  
  Tot_Cases                             DOUBLE,  
  Deaths                               DOUBLE,  
  Total_Tests                           DOUBLE,  
  Tests                                DOUBLE,  
  CASES_per_Test                        DOUBLE,  
  Death_in_Closed_Cases                 STRING,  
  Rank_by_Testing_rate                   DOUBLE,  
  Rank_by_Death_rate                     DOUBLE,  
  Rank_by_Cases_rate                     DOUBLE,  
  Rank_by_Death_of_Closed_Cases          DOUBLE  
)
```

```
PARTITIONED BY (COUNTRY_NAME STRING)
```

```
LOCATION '/user/cloudera/ds/COVID_HDFS_PARTITIONED';
```

```
FROM
```

```
covid_db.covid_staging
INSERT INTO TABLE covid_db.covid_ds_partitioned PARTITION(COUNTRY_NAME)
SELECT *,Country WHERE Country is not null;
```

```
DROP TABLE IF EXISTS covid_final_output;
CREATE EXTERNAL TABLE covid_db.covid_final_output
(
  Country                STRING,
  TOP_DEATH               STRING,
  TOP_TEST                STRING
)
PARTITIONED BY (COUNTRY_NAME STRING)
ROW FORMAT DELIMITED FIELDS TERMINATED by ','
STORED as TEXTFILE
LOCATION '/user/cloudera/ds/COVID_FINAL_OUTPUT';
```

```
FROM
covid_db.covid_ds_partitioned
INSERT INTO TABLE covid_db.covid_final_output PARTITION(COUNTRY_NAME)
SELECT Country,Rank_by_Testing_rate,Rank_by_Death_rate,Country
WHERE Country is not null;
```