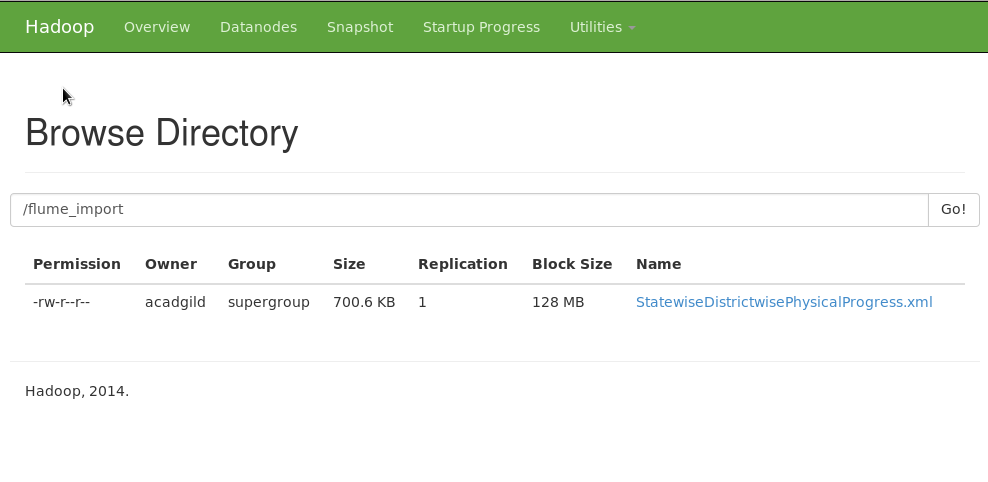
**Problem 1:**

**Step1 :**

acadgild@localhost project]$ flume-ng agent -n agent1 -f /usr/local/flume/conf/filecopy.conf



**1. Find out the districts who achieved 100 percent objective in BPL cards**

**Export the results to mysql using sqoop**

**Step 1:** lOgin to mysql and create empty table:

sql> create table results1 ( district\_name varchar(20), bpl\_objective double);

sql> desc results1

-> ;

+---------------+-------------+------+-----+---------+-------+

| Field | Type | Null | Key | Default | Extra |

+---------------+-------------+------+-----+---------+-------+

| district\_name | varchar(20) | YES | | NULL | |

| bpl\_objective | double | YES | | NULL | |

+---------------+-

**Step 2:**Run the script

Solution.sh:This script will call the pig script and uses sqoop to export results from hdfs to mysql

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

#!/bin/bash

pig project\_ans.pig

sqoop export -m 1 --connect jdbc:mysql://localhost/db1 --username root --table 'results1' --export-dir '/flume\_import/results1/part-m-00000' --input-fields-terminated-by ','

project\_ans.pig : This will filter the tables and push the results to HDFS =================using store command.

REGISTER /usr/local/pig/lib/piggybank.jar

DEFINE XPath org.apache.pig.piggybank.evaluation.xml.XPath();

DEFINE XPathAll org.apache.pig.piggybank.evaluation.xml.XPathAll();

A = LOAD '/flume\_import/StatewiseDistrictwisePhysicalProgress.xml' using org.apache.pig.piggybank.storage.XMLLoader('row') as (x:chararray);

B = FOREACH A GENERATE XPath(x,'row/District\_Name'), XPath(x,'row/Project\_Objectives\_IHHL\_BPL'),XPath(x, 'row/Project\_Performance-IHHL\_BPL');

C = foreach B generate $0, 100\*((double)$2/(double)$1);

D = filter C by ($1 == 100.0);

store D into '/flume\_import/results1' using PigStorage(',');

**Step 3:**

Go back to mysql and check if the table is populated.

sql> select \* from results1;

+----------------------+---------------+

| district\_name | bpl\_objective |

+----------------------+---------------+

| NIZAMABAD | 100 |

| TIRAP | 100 |

| HAILAKANDI | 100 |

| MADHUBANI | 100 |

| NORTH GOA | 100 |

| AHMEDABAD | 100 |

| DANGS | 100 |

| NAVSARI | 100 |

| PORBANDAR | 100 |

| SURAT | 100 |

| FARIDABAD | 100 |

| HISAR | 100 |

| JHAJJAR | 100 |

| MAHENDRAGARH | 100 |

| PANCHKULA | 100 |

| PANIPAT | 100 |

| ROHTAK | 100 |

| SIRSA | 100 |

| HAMIRPUR | 100 |

| KINNAUR | 100 |

| KULLU | 100 |

| LAHAUL & SPITI | 100 |

| SHIMLA | 100 |

| SOLAN | 100 |

| UNA | 100 |

| DEOGHAR | 100 |

| LOHARDAGA | 100 |

| HASSAN | 100 |

| MANGALORE(DAKSHINA K | 100 |

| UDUPI | 100 |

| ALAPPUZHA | 100 |

| KOLLAM | 100 |

| KOTTAYAM | 100 |

| KOZHIKODE | 100 |

| PALAKKAD | 100 |

| PATHANAMTHITTA | 100 |

| WAYANAD | 100 |

| GADCHIROLI | 100 |

| SINDHUDURG | 100 |

| WEST GARO HILLS | 100 |

| CHAMPHAI | 100 |

| LAWNGTLAI | 100 |

| HANUMANGARH | 100 |

| ERODE | 100 |

| KARUR | 100 |

| NAMAKKAL | 100 |

| TIRUCHIRAPPALLI | 100 |

| TIRUVANNAMALAI | 100 |

| DHALAI | 100 |

| SOUTH TRIPURA | 100 |

| WEST TRIPURA | 100 |

| AMBEDKAR NAGAR | 100 |

| BALRAMPUR | 100 |

| BAREILLY | 100 |

| BIJNOR | 100 |

| BUDAUN | 100 |

| ETAWAH | 100 |

| FARRUKHABAD | 100 |

| FIROZABAD | 100 |

| GHAZIABAD | 100 |

| HARDOI | 100 |

| JYOTIBA PHULE NAGAR | 100 |

| LUCKNOW | 100 |

| MAHARAJGANJ | 100 |

| MAHOBA | 100 |

| MORADABAD | 100 |

| MUZAFFARNAGAR | 100 |

| PILIBHIT | 100 |

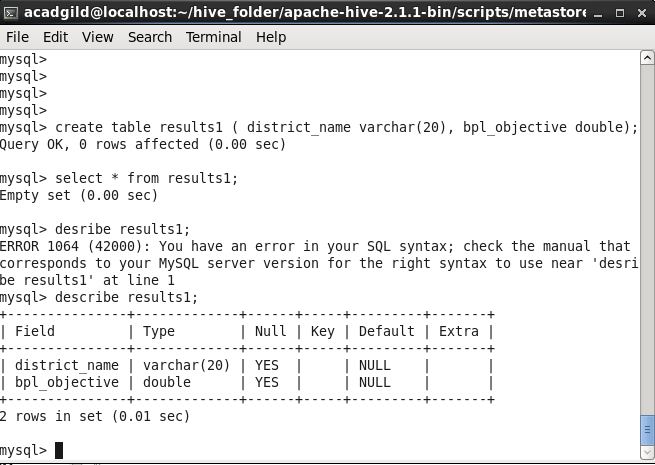
| SONBHADRA | 100 |

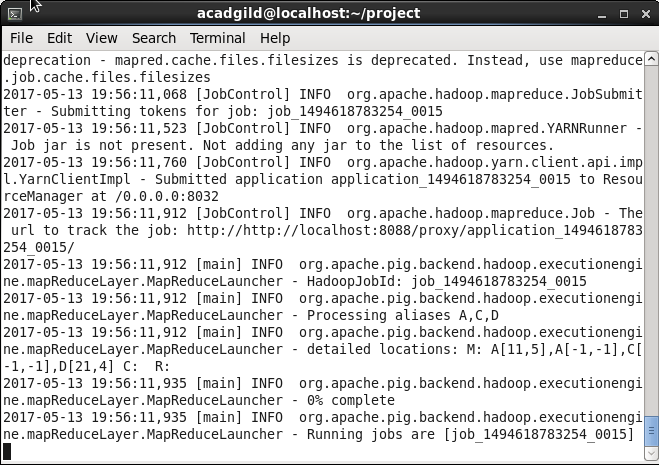
| SULTANPUR | 100 |

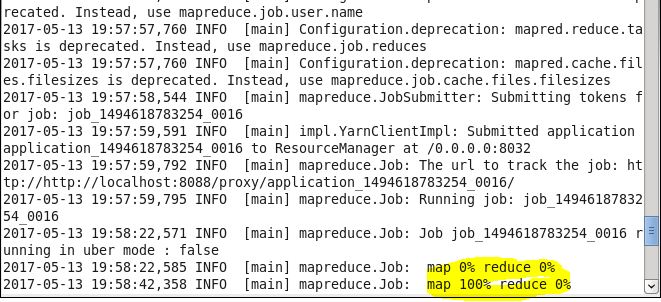
+----------------------+---------------+

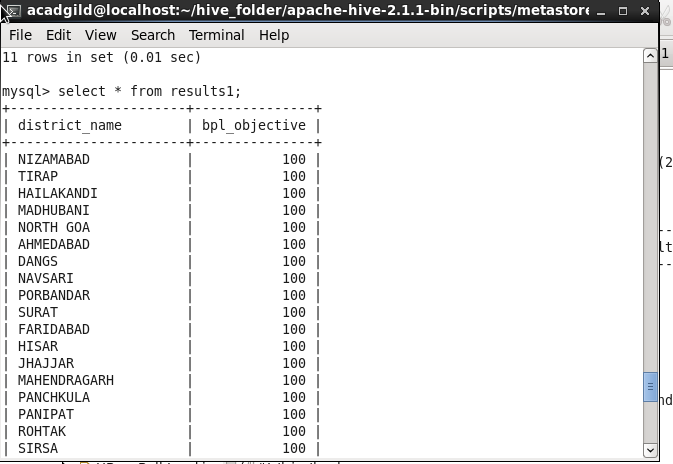
70 rows in set (0.00 sec)

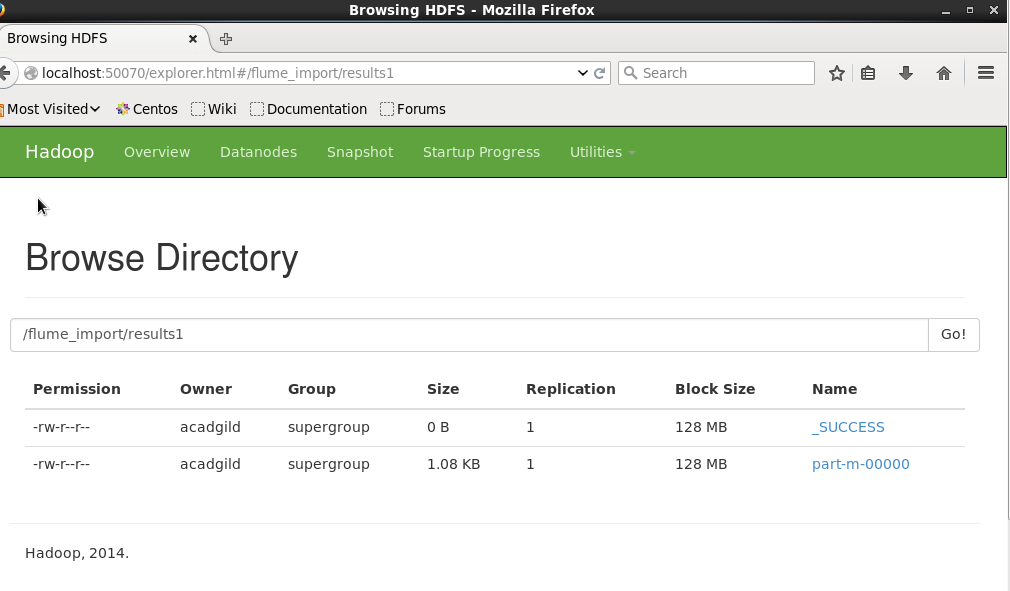
**Screenshots:**











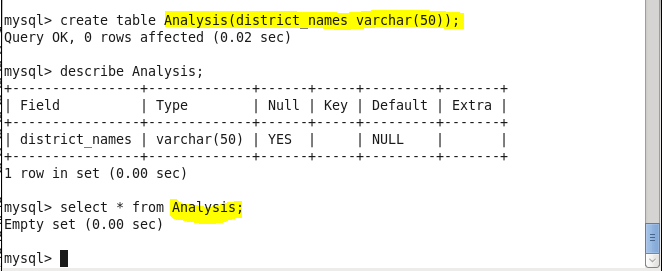
**Problem statement 2:**

**2. Write a Pig UDF to filter the districts who have reached 80% of objectives of BPL cards.**

**Export the results to mysql using sqoop.**

**Step 1:**

Create an empty table ‘Analysis’ in MySQL to store the results:

****

**Step 2:**

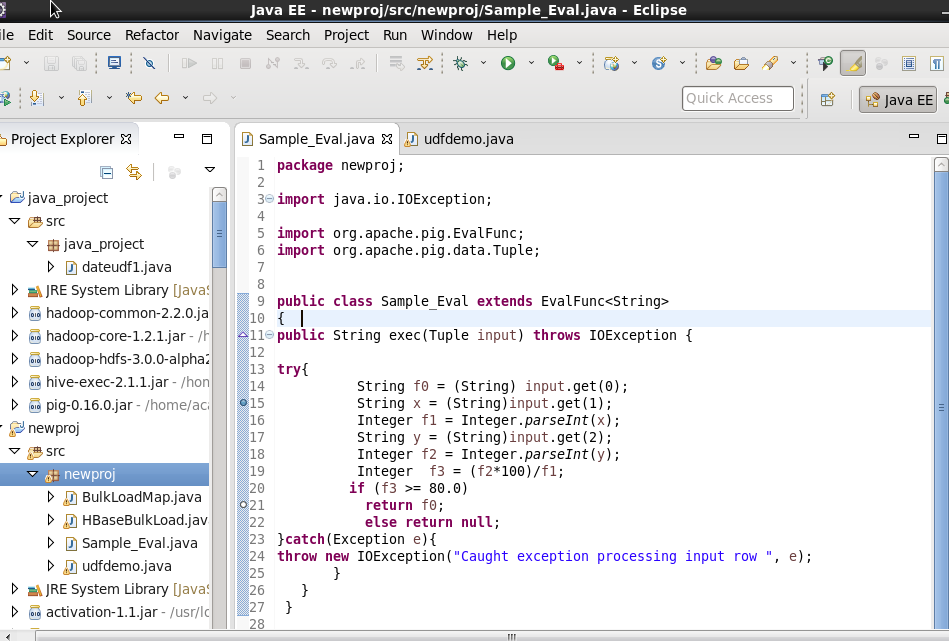
**Create an UDF in java ‘Sample\_Eval.java and export to a jar file.**

***This will compute the percentage of performance and filter those districts who have reached 80% of objectives of BPL cards.***

***Please find the jar file attached:***

***Sample\_Eval.java:***

***Export this jar to ’ /home/acadgild/project’ location.***

****

**Step 2:**Run the script

Solution.sh:This script will call the pig script and uses sqoop to export results from hdfs to mysql

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

#!/bin/bash

pig project\_ans.pig

sqoop export -m 1 --connect jdbc:mysql://localhost/db1 --username root --table 'Analysis' --export-dir '/flume\_import/results\_udf/part-m-00000' --input-fields-terminated-by ','

project\_ans.pig : This script uses the jar file which we created and gives the output.

REGISTER /usr/local/pig/lib/piggybank.jar

DEFINE XPath org.apache.pig.piggybank.evaluation.xml.XPath();

REGISTER /home/acadgild/project/pig\_xml.jar;

REGISTER /home/acadgild/project/Sample\_Eval.jar

A = LOAD '/flume\_import/StatewiseDistrictwisePhysicalProgress.xml' using org.apache.pig.piggybank.storage.XMLLoader('row') as (x:chararray);

B = FOREACH A GENERATE XPath(x,'row/District\_Name') as dist, XPath(x,'row/Project\_Objectives\_IHHL\_BPL') as obj,XPath(x, 'row/Project\_Performance-IHHL\_BPL') as per;

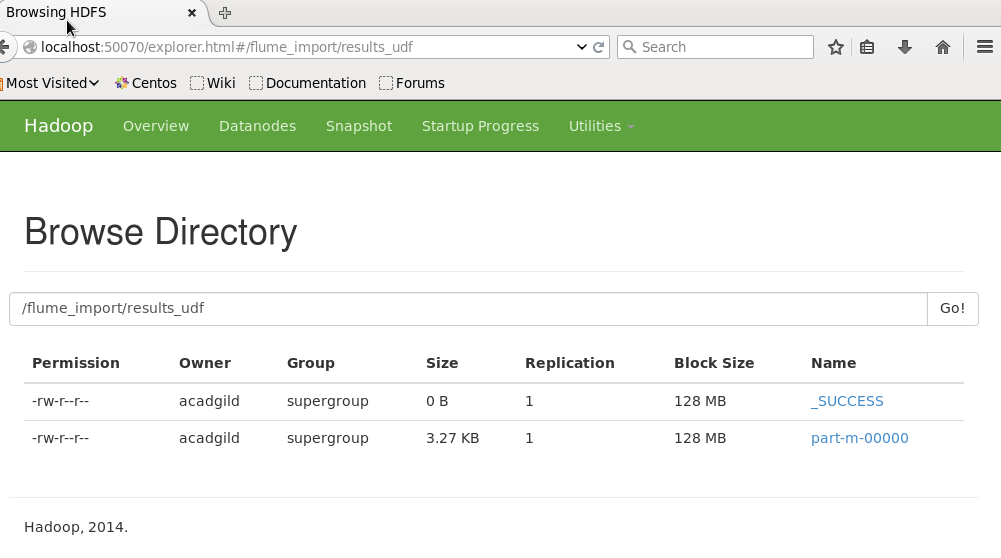
C = foreach B generate newproj.Sample\_Eval(\*);

D = filter C by $0 is not null;

store D into '/flume\_import/results\_udf' using PigStorage(',');

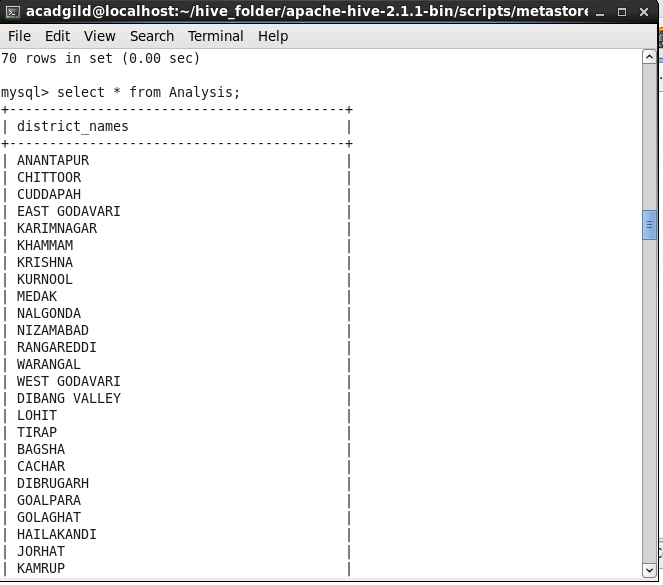
Check in the console to verify if the ‘store’ command has worked.

Then we use sqoop to move the results to MySQL.



**Step 3:**

Go back to mysql and check if the table is populated.



mysql> select \* from Analysis;

+------------------------------------------+

| district\_names |

+------------------------------------------+

| ANANTAPUR |

| CHITTOOR |

| CUDDAPAH |

| EAST GODAVARI |

| KARIMNAGAR |

| KHAMMAM |

| KRISHNA |

| KURNOOL |

| MEDAK |

| NALGONDA |

| NIZAMABAD |

| RANGAREDDI |

| WARANGAL |

| WEST GODAVARI |

| DIBANG VALLEY |

| LOHIT |

| TIRAP |

| BAGSHA |

| CACHAR |

| DIBRUGARH |

| GOALPARA |

| GOLAGHAT |

| HAILAKANDI |

| JORHAT |

| KAMRUP |

| KARIMGANJ |

| KOKRAJHAR |

| LAKHIMPUR |

| MARIGAON |

| NAGAON |

| SIBSAGAR |

| SONITPUR |

| TINSUKIA |

| BEGUSARAI |

| MADHUBANI |

| MUZAFFARPUR |

| SAHARSA |

| VAISHALI |

| DHAMTARI |

| JASHPUR |

| KANKER |

| KORBA |

| KORIYA |

| SURGUJA |

| NORTH GOA |

| AHMEDABAD |

| AMRELI |

| ANAND |

| BANAS KANTHA |

| BHARUCH |

| BHAVNAGAR |

| DAHOD |

| DANGS |

| GANDHINAGAR |

| JAMNAGAR |

| JUNAGADH |

| KACHCHH |

| KHEDA |

| MAHESANA |

| NARMADA |

| NAVSARI |

| PANCH MAHALS |

| PATAN |

| PORBANDAR |

| RAJKOT |

| SABAR KANTHA |

| SURAT |

| SURENDRANAGAR |

| VADODARA |

| VALSAD |

| AMBALA |

| BHIWANI |

| FARIDABAD |

| FATEHABAD |

| GURGAON |

| HISAR |

| JHAJJAR |

| JIND |

| KAITHAL |

| KARNAL |

| KURUKSHETRA |

| MAHENDRAGARH |

| MEWAT |

| PANCHKULA |

| PANIPAT |

| REWARI |

| ROHTAK |

| SIRSA |

| SONIPAT |

| YAMUNANAGAR |

| BILASPUR |

| CHAMBA |

| HAMIRPUR |

| KANGRA |

| KINNAUR |

| KULLU |

| LAHAUL & SPITI |

| MANDI |

| SHIMLA |

| SIRMAUR |

| SOLAN |

| UNA |

| ANANTNAG |

| LEH (LADAKH) |

| DEOGHAR |

| DUMKA |

| LATEHAR |

| LOHARDAGA |

| PAKUR |

| PURBI SINGHBHUM |

| BAGALKOT |

| BANGALORE RURAL |

| CHICKMAGALUR |

| CHITRADURGA |

| DHARWAD |

| GADAG |

| HASSAN |

| KODAGU |

| KOLAR |

| KOPPAL |

| MANDYA |

| MANGALORE(DAKSHINA KANNADA) |

| RAMANAGARA |

| SHIMOGA |

| UDUPI |

| ALAPPUZHA |

| ERNAKULAM |

| IDUKKI |

| KANNUR |

| KASARGOD |

| KOLLAM |

| KOTTAYAM |

| KOZHIKODE |

| MALAPPURAM |

| PALAKKAD |

| PATHANAMTHITTA |

| THIRUVANANTHAPURAM |

| THRISSUR |

| WAYANAD |

| ALIRAJPUR |

| ANUPPUR |

| BARWANI |

| BETUL |

| BHOPAL |

| BURHANPUR |

| DATIA |

| DEWAS |

| DHAR |

| DINDORI |

| GUNA |

| GWALIOR |

| HARDA |

| HOSHANGABAD |

| INDORE |

| JABALPUR |

| JHABUA |

| KATNI |

| KHANDWA(EAST NIMAR) |

| KHARGONE |

| MANDLA |

| MANDSAUR |

| MORENA |

| NARSINGHPUR |

| NEEMUCH |

| RAISEN |

| RAJGARH |

| RATLAM |

| REWA |

| SEHORE |

| SEONI |

| SHAHDOL |

| SHAJAPUR |

| SHEOPUR |

| SINGRAULI |

| UJJAIN |

| UMARIA |

| VIDISHA |

| AHMEDNAGAR |

| BHANDARA |

| DHULE |

| GADCHIROLI |

| GONDIA |

| HINGOLI |

| JALNA |

| KOLHAPUR |

| NAGPUR |

| OSMANABAD |

| PARBHANI |

| PUNE |

| RATNAGIRI |

| SANGLI |

| SATARA |

| SINDHUDURG |

| THANE |

| WARDHA |

| BISHNUPUR |

| IMPHAL EAST |

| TAMENGLONG |

| RI BHOI |

| SOUTH GARO HILLS |

| WEST GARO HILLS |

| CHAMPHAI |

| KOLASIB |

| LAWNGTLAI |

| LUNGLEI |

| MAMIT |

| SAIHA |

| SERCHHIP |

| KOHIMA |

| MOKOKCHUNG |

| PHEK |

| BALESWAR |

| JAGATSINGHAPUR |

| BARNALA |

| FATEHGARH SAHIB |

| HOSHIARPUR |

| JALANDHAR |

| KAPURTHALA |

| LUDHIANA |

| MANSA |

| NAWANSHAHR |

| S.A.S Nagar |

| AJMER |

| CHURU |

| DUNGARPUR |

| GANGANAGAR |

| HANUMANGARH |

| JAISALMER |

| NAGAUR |

| SIKAR |

| EAST SIKKIM |

| NORTH SIKKIM |

| SOUTH SIKKIM |

| WEST SIKKIM |

| COIMBATORE |

| CUDDALORE |

| DHARMAPURI |

| DINDIGUL |

| ERODE |

| KANCHIPURAM |

| KANYAKUMARI(NAGERCOIL) |

| KARUR |

| MADURAI |

| NAMAKKAL |

| NILGIRIS(UDHAGAMANDALAM) |

| PERAMBALUR |

| PUDUKKOTTAI |

| RAMANATHAPURAM |

| SALEM |

| SIVAGANGA |

| THENI |

| THOOTHUKUDI |

| TIRUCHIRAPPALLI |

| TIRUNELVELI |

| TIRUVANNAMALAI |

| TIRUVARUR |

| VELLORE |

| VIRUDHUNAGAR |

| DHALAI |

| NORTH TRIPURA |

| SOUTH TRIPURA |

| WEST TRIPURA |

| AGRA |

| ALIGARH |

| ALLAHABAD |

| AMBEDKAR NAGAR |

| AZAMGARH |

| BAGPAT |

| BALLIA |

| BALRAMPUR |

| BANDA |

| BARABANKI |

| BAREILLY |

| BASTI |

| BIJNOR |

| BUDAUN |

| BULANDSHAHR |

| CHANDAULI |

| CHITRAKOOT |

| DEORIA |

| ETAH |

| ETAWAH |

| FAIZABAD |

| FARRUKHABAD |

| FATEHPUR |

| FIROZABAD |

| GAUTAM BUDDHA NAGAR |

| GHAZIABAD |

| GHAZIPUR |

| GONDA |

| GORAKHPUR |

| HAMIRPUR |

| HARDOI |

| JALAUN |

| JAUNPUR |

| JHANSI |

| JYOTIBA PHULE NAGAR |

| KANNAUJ |

| KANPUR DEHAT |

| KANPUR NAGAR |

| KAUSHAMBI |

| KUSHINAGAR |

| LAKHIMPUR KHERI |

| LALITPUR |

| LUCKNOW |

| MAHAMAYA NAGAR(HATHRAS) |

| MAHARAJGANJ |

| MAHOBA |

| MAINPURI |

| MATHURA |

| MAU |

| MEERUT |

| MIRZAPUR |

| MORADABAD |

| MUZAFFARNAGAR |

| PILIBHIT |

| PRATAPGARH |

| RAE BARELI |

| RAMPUR |

| SAHARANPUR |

| SANT RAVIDAS NAGAR( BHADOHI) |

| SHAHJAHANPUR |

| SHRAVASTI |

| SIDDHARTHNAGAR |

| SITAPUR |

| SONBHADRA |

| SULTANPUR |

| UNNAO |

| VARANASI |

| BAGESHWAR |

| CHAMOLI |

| DEHRADUN |

| HARIDWAR |

| NAINITAL |

| PITHORAGARH |

| RUDRAPRAYAG |

| TEHRI GARHWAL |

| UDHAM SINGH NAGAR |

| UTTARKASHI |

| BARDHAMAN |

| DAKSHIN DINAJPUR |

| HOOGHLY |

| HOWRAH |

| JALPAIGURI |

| MIDNAPUR EAST |

| MIDNAPUR WEST |

| NADIA |

| NORTH 24 PARAGANAS |

| SOUTH 24 PARAGANAS |

+------------------------------------------+

349 rows in set (0.04 sec)