Problem 1

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

Export the results to mysql using sqoop

Pig Script to calculate 100% Objective

input data set

[acadgild@localhost ~]$ hadoop fs -ls /pig
17/11/07 04:01:27 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r--   1 acadgild supergroup   717414 2017-11-07 01:45 /pig/state.xml

Problem1.pig

--register piggybank lib

REGISTER /home/acadgild/piggybank-0.15.0.jar

--defining xml-xpath for importing datas

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

data = LOAD '/pig/state.xml' using org.apache.pig.piggybank.storage.XMLLoader('row') as (x:chararray);

--reading proper values for computation from xml file

State = FOREACH data GENERATE XPath(x,'row/State_Name') AS statename, XPath(x,'row/District_Name') AS disname,XPath(x,'row/Project_Objectives_IHHL_BPL') AS BPL,XPath(x,'row/Project_Objectives_IHHL_TOTAL') AS total ;

--filtering as per logic

Obj = FILTER State BY BPL == total;

--store in the hdfs

STORE Obj INTO '/vishnu/pigout' USING PigStorage(',');

Output Content

[acadgild@localhost ~]$ hadoop fs -ls /vishnu/pigout
17/11/07 04:09:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup   0 2017-11-07 04:08 /vishnu/pigout/_SUCCESS
-rw-r--r--   1 acadgild supergroup   889 2017-11-07 04:08 /vishnu/pigout/part-m-00000

[acadgild@localhost ~]$ hadoop fs -cat /vishnu/pigout/part-m-00000
17/11/07 04:10:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Arunachal Pradesh,ANJAW,3232,3232
Arunachal Pradesh,DIBANG VALLEY,1085,1085
Arunachal Pradesh,KURUNG KUMEY,22036,22036
Arunachal Pradesh,LOHIT,8800,8800
Arunachal Pradesh,WEST SIANG,11472,11472
Bihar,BANKA,82439,82439
D & N Haveli,DADRA AND NAGAR HAVELI,2480,2480
Goa,NORTH GOA,15000,15000
Jammu & Kashmir,KARGIL,8475,8475
Jammu & Kashmir,KISHTWAR,22318,22318
Jammu & Kashmir,LEH (LADAKH),6090,6090
Jammu & Kashmir,REASI   ,21500,21500
Jammu & Kashmir,SAMBA   ,9849,9849
Jammu & Kashmir,SHOPIAN   ,10196,10196
Kerala,KANNUR,34121,34121
Manipur,CHANDEL,17610,17610
Nagaland,LONGLENG,6438,6438
Nagaland,TUENSANG,13027,13027
Nagaland,ZUNHEBOTO,20570,20570
Puducherry,PONDICHERRY,18000,18000
Punjab,FARIDKOT,6000,6000
Punjab,HOSHIARPUR,11112,11112
Punjab,MOGA,37170,37170
Punjab,MUKTSAR,33148,33148

Copy the output to HDFS

Run MySQL using the following commands
 sudo service mysqld start
 mysql -u root

Create a new Database:
mysql>create database state;
mysql>use state;

Create a table similar to the structure of the output file contents:
mysql> create table stateBpl (state varchar(20),district varchar(20),
    -> bpl int,total int);
Query OK, 0 rows affected (0.01 sec)

Run the sqoop command to Export the Pig Script analysis from HDFS to MySQL:

[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/state --username 'root' -P --table stateBpl --export-dir '/vishnu/pigout/part-m-00000' --input-fields-terminated-by ',' -m 1

Output Sql

mysql> select * from stateBpl;
```
+-------------------+---------------------+-------+-------+
| state             | district            | bpl   | total |
+-------------------+---------------------+-------+-------+
| Arunachal Pradesh | ANJAW               |  3232 |  3232 |
| Arunachal Pradesh | DIBANG VALLEY       |  1085 |  1085 |
| Arunachal Pradesh | KURUNG KUMEY        | 22036 | 22036 |
| Arunachal Pradesh | LOHIT               |  8800 |  8800 |
| Arunachal Pradesh | WEST SIANG          | 11472 | 11472 |
| Bihar             | BANKA               | 82439 | 82439 |
| D & N Haveli      | DADRA AND NAGAR HAVE |  2480 |  2480 |
| Goa               | NORTH GOA           | 15000 | 15000 |
| Jammu & Kashmir   | KARGIL              |  8475 |  8475 |
| Jammu & Kashmir   | KISHTWAR            | 22318 | 22318 |
| Jammu & Kashmir   | LEH (LADAKH)        |  6090 |  6090 |
| Jammu & Kashmir   | REASI               | 21500 | 21500 |
| Jammu & Kashmir   | SAMBA               |  9849 |  9849 |
| Jammu & Kashmir   | SHOPIAN             | 10196 | 10196 |
| Kerala            | KANNUR              | 34121 | 34121 |
| Manipur           | CHANDEL             | 17610 | 17610 |
| Nagaland          | LONGLENG            |  6438 |  6438 |
| Nagaland          | TUENSANG            | 13027 | 13027 |
| Nagaland          | ZUNHEBOTO           | 20570 | 20570 |
| Puducherry        | PONDICHERRY         | 18000 | 18000 |
| Punjab            | FARIDKOT            |  6000 |  6000 |
| Punjab            | HOSHIARPUR          | 11112 | 11112 |
| Punjab            | MOGA                | 37170 | 37170 |
| Punjab            | MUKTSAR             | 33148 | 33148 |
+-------------------+---------------------+-------+-------+
```
24 rows in set (0.00 sec)

Screen Shots

```
                          acadgild@localhost:~/project2                          _ □ ×
File   Edit   View   Search   Terminal   Help

Job Stats (time in seconds):
JobId    Maps    Reduces MaxMapTime       MinMapTime      AvgMapTime       MedianMapTime   MaxReduceTime   MinReduceTime   AvgRe
duceTime        MedianReducetime       Alias   Feature Outputs
job_1509997605925_0025  1       0       23      23      23      23      0       0       0       0       Obj,State,data  MAP_O
NLY     /vishnu/pigout,

Input(s):
Successfully read 0 records from: "/pig/state.xml"

Output(s):
Successfully stored 0 records in: "/vishnu/pigout"

Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_1509997605925_0025


2017-11-07 04:08:29,382 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032
2017-11-07 04:08:29,406 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2017-11-07 04:08:29,561 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032
2017-11-07 04:08:29,579 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2017-11-07 04:08:29,683 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032
2017-11-07 04:08:29,719 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalAp
plicationStatus=SUCCEEDED. Redirecting to job history server
2017-11-07 04:08:29,851 [main] WARN  org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Unable
to retrieve job to compute warning aggregation.
2017-11-07 04:08:29,851 [main] INFO  org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success
!
2017-11-07 04:08:30,040 [main] INFO  org.apache.pig.Main - Pig script completed in 58 seconds and 252 milliseconds (58252 ms)
[acadgild@localhost project2]$
```
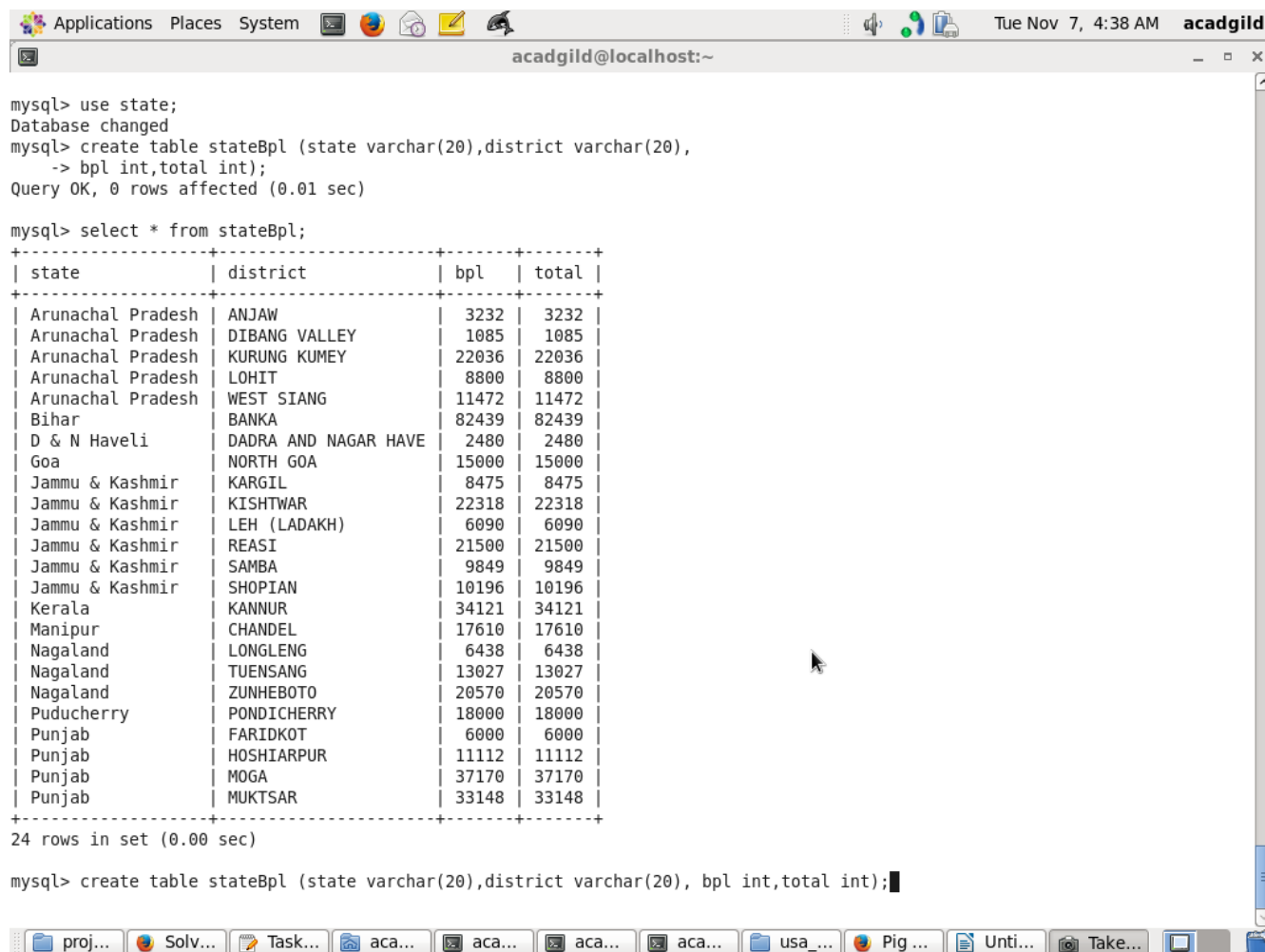
`📁 proj...` `🦊 Solv...` `📝 prob...` `📁 aca...` `✉ aca...` `✉ aca...` `✉ aca...` `📁 usa_...` `🐬 Pig ...` `📄 Unti...` `📷 Take...` `🔲`

```
[acadgild@localhost ~]$ hadoop fs -ls /vishnu/pigout
17/11/07 04:09:08 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--    1 acadgild supergroup            0 2017-11-07 04:08 /vishnu/pigout/_S
UCCESS
-rw-r--r--    1 acadgild supergroup          889 2017-11-07 04:08 /vishnu/pigout/pa
rt-m-00000
[acadgild@localhost ~]$ hadoop fs -cat /vishnu/pigout/part-m-00000
17/11/07 04:10:10 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
Arunachal Pradesh,ANJAW,3232,3232
Arunachal Pradesh,DIBANG VALLEY,1085,1085
Arunachal Pradesh,KURUNG KUMEY,22036,22036
Arunachal Pradesh,LOHIT,8800,8800
Arunachal Pradesh,WEST SIANG,11472,11472
Bihar,BANKA,82439,82439
D & N Haveli,DADRA AND NAGAR HAVELI,2480,2480
Goa,NORTH GOA,15000,15000
Jammu & Kashmir,KARGIL,8475,8475
Jammu & Kashmir,KISHTWAR,22318,22318
Jammu & Kashmir,LEH (LADAKH),6090,6090
Jammu & Kashmir,REASI                                    ,21500,21500
Jammu & Kashmir,SAMBA                                    ,9849,9849
Jammu & Kashmir,SHOPIAN                                  ,10196,10196
Kerala,KANNUR,34121,34121
Manipur,CHANDEL,17610,17610
Nagaland,LONGLENG,6438,6438
Nagaland,TUENSANG,13027,13027
Nagaland,ZUNHEBOTO,20570,20570
Puducherry,PONDICHERRY,18000,18000
Punjab,FARIDKOT,6000,6000
Punjab,HOSHIARPUR,11112,11112
Punjab,MOGA,37170,37170
Punjab,MUKTSAR,33148,33148
```

Screen Shots SQL Table


```
mysql> use state;
Database changed
mysql> create table stateBpl (state varchar(20),district varchar(20),
    -> bpl int,total int);
Query OK, 0 rows affected (0.01 sec)

mysql> select * from stateBpl;
+-------------------+----------------------+-------+-------+
| state             | district             | bpl   | total |
+-------------------+----------------------+-------+-------+
| Arunachal Pradesh | ANJAW                | 3232  | 3232  |
| Arunachal Pradesh | DIBANG VALLEY        | 1085  | 1085  |
| Arunachal Pradesh | KURUNG KUMEY         | 22036 | 22036 |
| Arunachal Pradesh | LOHIT                | 8800  | 8800  |
| Arunachal Pradesh | WEST SIANG           | 11472 | 11472 |
| Bihar             | BANKA                | 82439 | 82439 |
| D & N Haveli      | DADRA AND NAGAR HAVE | 2480  | 2480  |
| Goa               | NORTH GOA            | 15000 | 15000 |
| Jammu & Kashmir   | KARGIL               | 8475  | 8475  |
| Jammu & Kashmir   | KISHTWAR             | 22318 | 22318 |
| Jammu & Kashmir   | LEH (LADAKH)         | 6090  | 6090  |
| Jammu & Kashmir   | REASI                | 21500 | 21500 |
| Jammu & Kashmir   | SAMBA                | 9849  | 9849  |
| Jammu & Kashmir   | SHOPIAN              | 10196 | 10196 |
| Kerala            | KANNUR               | 34121 | 34121 |
| Manipur           | CHANDEL              | 17610 | 17610 |
| Nagaland          | LONGLENG             | 6438  | 6438  |
| Nagaland          | TUENSANG             | 13027 | 13027 |
| Nagaland          | ZUNHEBOTO            | 20570 | 20570 |
| Puducherry        | PONDICHERRY          | 18000 | 18000 |
| Punjab            | FARIDKOT             | 6000  | 6000  |
| Punjab            | HOSHIARPUR           | 11112 | 11112 |
| Punjab            | MOGA                 | 37170 | 37170 |
| Punjab            | MUKTSAR              | 33148 | 33148 |
+-------------------+----------------------+-------+-------+
24 rows in set (0.00 sec)

mysql> create table stateBpl (state varchar(20),district varchar(20), bpl int,total int);
```