

NOTES:

1. I have done all this in Cloudera VM which has just 8 GB RAM and 4 CPU cores. I have just tried to save cost associated with use of EC2 instance on AWS.

Data Ingestion from the RDS to HDFS using Sqoop

1. Sqoop Import command

Listing tables first:

```
sqoop-list-tables --connect jdbc:mysql://upgradawsrds.cpclxrkdvwzm.us-east-  
1.rds.amazonaws.com:3306/indiaahs2012_13 --username upgraduser --password upgraduser
```

```
[cloudera@quickstart ~]$ sqoop-list-tables --connect jdbc:mysql://upgradawsrds.cpclxrkdvwzm.us-east-  
1.rds.amazonaws.com:3306/indiaahs2012_13 --username upgraduser --password upgraduser  
18/11/26 19:36:39 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.13.0  
18/11/26 19:36:39 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.  
18/11/26 19:36:39 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
```

Key_indicator_districtwise

```
[cloudera@quickstart ~]$
```

I am importing data from MySQL RDS into project2 directory in HDFS:

```
sqoop import --connect jdbc:mysql://upgradawsrds.cpclxrkdvwzm.us-east-  
1.rds.amazonaws.com:3306/indiaahs2012_13 --username upgraduser --password upgraduser --table  
Key_indicator_districtwise --target-dir /user/cloudera/project2
```

```
[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://upgradawsrds.cpclxrkdvwzm.us-east-  
1.rds.amazonaws.com:3306/indiaahs2012_13 --username upgraduser --password upgraduser --table Key_indicator_districtwise --target-dir  
/user/cloudera/project2  
18/11/26 19:39:37 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.13.0  
18/11/26 19:39:37 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.  
18/11/26 19:39:38 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.  
18/11/26 19:39:38 INFO tool.CodeGenTool: Beginning code generation  
18/11/26 19:39:46 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Key_indicator_districtwise` AS t LIMIT 1  
18/11/26 19:39:48 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Key_indicator_districtwise` AS t LIMIT 1  
18/11/26 19:39:50 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-mapreduce  
Note: /tmp/sqoop-cloudera/compile/151fdb601a06f2014ef4be4932419ee6/Key_indicator_districtwise.java uses or overrides a deprecated  
API.  
Note: Recompile with -Xlint:deprecation for details.  
18/11/26 19:39:56 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-  
cloudera/compile/151fdb601a06f2014ef4be4932419ee6/Key_indicator_districtwise.jar  
18/11/26 19:39:57 WARN manager.MySQLManager: It looks like you are importing from mysql.  
18/11/26 19:39:57 WARN manager.MySQLManager: This transfer can be faster! Use the --direct  
18/11/26 19:39:57 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.  
18/11/26 19:39:57 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)  
18/11/26 19:39:57 INFO mapreduce.ImportJobBase: Beginning import of Key_indicator_districtwise  
18/11/26 19:39:57 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address  
18/11/26 19:39:58 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar  
18/11/26 19:40:00 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps  
18/11/26 19:40:00 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0:8032  
18/11/26 19:40:06 INFO db.DBInputFormat: Using read committed transaction isolation  
18/11/26 19:40:06 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`ID`), MAX(`ID`) FROM `Key_indicator_districtwise`  
18/11/26 19:40:06 INFO db.IntegerSplitter: Split size: 70; Num splits: 4 from: 1 to: 284  
18/11/26 19:40:07 INFO mapreduce.JobSubmitter: number of splits:4  
18/11/26 19:40:07 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1543289208460_0001  
18/11/26 19:40:08 INFO impl.YarnClientImpl: Submitted application application_1543289208460_0001
```

```

18/11/26 19:40:08 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application_1543289208460_0001/
18/11/26 19:40:08 INFO mapreduce.Job: Running job: job_1543289208460_0001
18/11/26 19:40:20 INFO mapreduce.Job: Job job_1543289208460_0001 running in uber mode : false
18/11/26 19:40:20 INFO mapreduce.Job: map 0% reduce 0%
18/11/26 19:40:40 INFO mapreduce.Job: map 25% reduce 0%
18/11/26 19:40:41 INFO mapreduce.Job: map 50% reduce 0%
18/11/26 19:40:43 INFO mapreduce.Job: map 75% reduce 0%
18/11/26 19:40:44 INFO mapreduce.Job: map 100% reduce 0%
18/11/26 19:40:45 INFO mapreduce.Job: Job job_1543289208460_0001 completed successfully
18/11/26 19:40:45 INFO mapreduce.Job: Counters: 31
    File System Counters
        FILE: Number of bytes read=0
        FILE: Number of bytes written=822704
        FILE: Number of read operations=0
        FILE: Number of large read operations=0
        FILE: Number of write operations=0
        HDFS: Number of bytes read=405
        HDFS: Number of bytes written=1027652
        HDFS: Number of read operations=16
        HDFS: Number of large read operations=0
        HDFS: Number of write operations=8
    Job Counters
        Killed map tasks=1
        Launched map tasks=4
        Other local map tasks=4
        Total time spent by all maps in occupied slots (ms)=76401
        Total time spent by all reduces in occupied slots (ms)=0
        Total time spent by all map tasks (ms)=76401
        Total vcore-milliseconds taken by all map tasks=76401
        Total megabyte-milliseconds taken by all map tasks=78234624
    Map-Reduce Framework
        Map input records=284
        Map output records=284
        Input split bytes=405
        Spilled Records=0
        Failed Shuffles=0
        Merged Map outputs=0
        GC time elapsed (ms)=378
        CPU time spent (ms)=15590
        Physical memory (bytes) snapshot=1141731328
        Virtual memory (bytes) snapshot=6332882944
        Total committed heap usage (bytes)=1119879168
    File Input Format Counters
        Bytes Read=0
    File Output Format Counters
        Bytes Written=1027652

```

18/11/26 19:40:45 INFO mapreduce.ImportJobBase: Transferred 1,003.5664 KB in 45.399 seconds (22.1055 KB/sec)

18/11/26 19:40:45 INFO mapreduce.ImportJobBase: Retrieved 284 records.

2. Command to see the list of imported data

hadoop fs -ls /user/cloudera/project2/

```

[cloudera@quickstart ~]$ hadoop fs -ls /user/cloudera/project2/
Found 5 items
-rw-r--r-- 1 cloudera cloudera 0 2018-11-26 19:40 /user/cloudera/project2/_SUCCESS
-rw-r--r-- 1 cloudera cloudera 244283 2018-11-26 19:40 /user/cloudera/project2/part-m-00000
-rw-r--r-- 1 cloudera cloudera 249691 2018-11-26 19:40 /user/cloudera/project2/part-m-00001
-rw-r--r-- 1 cloudera cloudera 257360 2018-11-26 19:40 /user/cloudera/project2/part-m-00002
-rw-r--r-- 1 cloudera cloudera 276318 2018-11-26 19:40 /user/cloudera/project2/part-m-00003
[cloudera@quickstart ~]$

```

hadoop fs -cat /user/cloudera/project2/* (Output is - 284 rows and 645 columns so not displaying it here)

External table creation in Hive and loading the ingested data into it. Data ingestion verification.

1. Command to create the external table

First creating a separate database namely **etlproject2** and creating all objects under this database.

--- Create database etlproject2

```
create database etlproject2;
```

--- Use etlproject2 (move into etlproject2)

```
use etlproject2;
```

--- Create external table **KEY_INDICATORS_INT** using HDFS location as '**/user/cloudera/etlproject2/**'

```
CREATE EXTERNAL TABLE IF NOT EXISTS KEY_INDICATORS_EXT(
```

```
`ID` INT,
```

```
`STATE_NAME` VARCHAR(100),
```

```
`STATE_DISTRICT_NAME` VARCHAR(100),
```

```
`AA_SAMPLE_UNITS_TOTAL` DOUBLE,
```

```
`AA_SAMPLE_UNITS_RURAL` DOUBLE,
```

```
`AA_SAMPLE_UNITS_URBAN` DOUBLE,
```

```
`AA_HOUSEHOLDS_TOTAL` DOUBLE,
```

```
`AA_HOUSEHOLDS_RURAL` DOUBLE,
```

```
`AA_HOUSEHOLDS_URBAN` DOUBLE,
```

```
`AA_POPULATION_TOTAL` DOUBLE,
```

```
`AA_POPULATION_RURAL` DOUBLE,
```

```
`AA_POPULATION_URBAN` DOUBLE,
```

```
`AA_EVER_MARRIED_WOMEN_AGED_15_49_YEARS_TOTAL` DOUBLE,
```

```
`AA_EVER_MARRIED_WOMEN_AGED_15_49_YEARS_RURAL` DOUBLE,
```

```
`AA_EVER_MARRIED_WOMEN_AGED_15_49_YEARS_URBAN` DOUBLE,
```

```
`AA_CURRENTLY_MARRIED_WOMEN_AGED_15_49_YEARS_TOTAL` DOUBLE,
```

```
`AA_CURRENTLY_MARRIED_WOMEN_AGED_15_49_YEARS_RURAL` DOUBLE,
```

```
`AA_CURRENTLY_MARRIED_WOMEN_AGED_15_49_YEARS_URBAN` DOUBLE,
```

```
`AA_CHILDREN_12_23_MONTHS_TOTAL` DOUBLE,
```

```
`AA_CHILDREN_12_23_MONTHS_RURAL` DOUBLE,
```

```
`AA_CHILDREN_12_23_MONTHS_URBAN` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_SC_TOTAL` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_SC_RURAL` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_SC_URBAN` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_ST_TOTAL` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_ST_RURAL` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_ST_URBAN` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_ALL_TOTAL` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_ALL_RURAL` DOUBLE,
```

```
`BB_AVERAGE_HOUSEHOLD_SIZE_ALL_URBAN` DOUBLE,
```

```
`BB_POPULATION_BELOW AGE_15_YEARS_TOTAL` DOUBLE,
```

```
`BB_POPULATION_BELOW AGE_15_YEARS_RURAL` DOUBLE,
```

```
`BB_POPULATION_BELOW AGE_15_YEARS_URBAN` DOUBLE,
```

```
`BB_DEPENDENCY_RATIO_TOTAL` DOUBLE,
```

```
`BB_DEPENDENCY_RATIO_RURAL` DOUBLE,
```

```
`BB_DEPENDENCY_RATIO_URBAN` DOUBLE,
```

```
`BB_CURRENTLY_MARRIED_ILLITERATE_WOMEN_AGED_15_49_YEARS_TOTAL` DOUBLE,
```

`BB_CURRENTLY_MARRIED_ILLITERATE_WOMEN_AGED_15_49_YEARS_RURAL` DOUBLE,
`BB_CURRENTLY_MARRIED_ILLITERATE_WOMEN_AGED_15_49_YEARS_URBAN` DOUBLE,
`CC_SEX_RATIO_AT_BIRTH_TOTAL` DOUBLE,
`CC_SEX_RATIO_AT_BIRTH_RURAL` DOUBLE,
`CC_SEX_RATIO_AT_BIRTH_URBAN` DOUBLE,
`CC_SEX_RATIO_0_4_YEARS_TOTAL` DOUBLE,
`CC_SEX_RATIO_0_4_YEARS_RURAL` DOUBLE,
`CC_SEX_RATIO_0_4_YEARS_URBAN` DOUBLE,
`CC_SEX_RATIO_ALL_AGES_TOTAL` DOUBLE,
`CC_SEX_RATIO_ALL_AGES_RURAL` DOUBLE,
`CC_SEX_RATIO_ALL_AGES_URBAN` DOUBLE,
`DD_PERSON_TOTAL` DOUBLE,
`DD_PERSON_RURAL` DOUBLE,
`DD_PERSON_URBAN` DOUBLE,
`DD_MALE_TOTAL` DOUBLE,
`DD_MALE_RURAL` DOUBLE,
`DD_MALE_URBAN` DOUBLE,
`DD_FEMALE_TOTAL` DOUBLE,
`DD_FEMALE_RURAL` DOUBLE,
`DD_FEMALE_URBAN` DOUBLE,
`EE_MARRIAGES_AMONG_FEMALES_BELOW_LEGAL AGE_18_YEARS_TOTAL` DOUBLE,
`EE_MARRIAGES_AMONG_FEMALES_BELOW_LEGAL AGE_18_YEARS_RURAL` DOUBLE,
`EE_MARRIAGES_AMONG_FEMALES_BELOW_LEGAL AGE_18_YEARS_URBAN` DOUBLE,
`EE_MARRIAGES_AMONG_MALES_BELOW_LEGAL AGE_21_YEARS_TOTAL` DOUBLE,
`EE_MARRIAGES_AMONG_MALES_BELOW_LEGAL AGE_21_YEARS_RURAL` DOUBLE,
`EE_MARRIAGES_AMONG_MALES_BELOW_LEGAL AGE_21_YEARS_URBAN` DOUBLE,
`EE_MARRIED_WOMEN_20_24_YEARS_MARRIED_BEFORE_18_YEARS_TOTAL` DOUBLE,
`EE_MARRIED_WOMEN_20_24_YEARS_MARRIED_BEFORE_18_YEARS_RURAL` DOUBLE,
`EE_MARRIED_WOMEN_20_24_YEARS_MARRIED_BEFORE_18_YEARS_URBAN` DOUBLE,
`EE_MARRIED_MEN_25_29_YEARS_MARRIED_BEFORE_21_YEARS_TOTAL` DOUBLE,
`EE_MARRIED_MEN_25_29_YEARS_MARRIED_BEFORE_21_YEARS_RURAL` DOUBLE,
`EE_MARRIED_MEN_25_29_YEARS_MARRIED_BEFORE_21_YEARS_URBAN` DOUBLE,
`EE_MEAN AGE_AT_MARRIAGE_MALE_TOTAL` DOUBLE,
`EE_MEAN AGE_AT_MARRIAGE_MALE_RURAL` DOUBLE,
`EE_MEAN AGE_AT_MARRIAGE_MALE_URBAN` DOUBLE,
`EE_MEAN AGE_AT_MARRIAGE_FEMALE_TOTAL` DOUBLE,
`EE_MEAN AGE_AT_MARRIAGE_FEMALE_RURAL` DOUBLE,
`EE_MEAN AGE_AT_MARRIAGE_FEMALE_URBAN` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_PERSON_TOTAL` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_PERSON_RURAL` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_PERSON_URBAN` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_MALE_TOTAL` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_MALE_RURAL` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_MALE_URBAN` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_FEMALE_TOTAL` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_FEMALE_RURAL` DOUBLE,
`FF_CHILDREN_ATTENDING_SCHOOL_AGE_6_17_YEARS_FEMALE_URBAN` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_PERSON_TOTAL` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_PERSON_RURAL` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_PERSON_URBAN` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_MALE_TOTAL` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_MALE_RURAL` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_MALE_URBAN` DOUBLE,

`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_FEMALE_TOTAL` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_FEMALE_RURAL` DOUBLE,
`FF_CHILDREN_ATTENDED_BEFORE_DROP_OUT_AGE_6_17_YEARS_FEMALE_URBAN` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_PERSON_TOTAL` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_PERSON_RURAL` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_PERSON_URBAN` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_MALE_TOTAL` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_MALE_RURAL` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_MALE_URBAN` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_FEMALE_TOTAL` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_FEMALE_RURAL` DOUBLE,
`GG_CHILDREN_AGED_5_14_YEARS_ENGAGED_IN_WORK_FEMALE_URBAN` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_PERSON_TOTAL` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_PERSON_RURAL` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_PERSON_URBAN` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_MALE_TOTAL` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_MALE_RURAL` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_MALE_URBAN` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_FEMALE_TOTAL` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_FEMALE_RURAL` DOUBLE,
`GG_WORK_PARTICIPATION_RATE_15_YEARS_AND ABOVE_FEMALE_URBAN` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_PERSON_TOTAL` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_PERSON_RURAL` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_PERSON_URBAN` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_MALE_TOTAL` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_MALE_RURAL` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_MALE_URBAN` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_FEMALE_TOTAL` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_FEMALE_RURAL` DOUBLE,
`HH_PREVALENCE_DISABILITY_PER_100000_POPULATION_FEMALE_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_PERSON_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_PERSON_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_PERSON_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_MALE_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_MALE_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_MALE_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_FEMALE_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_FEMALE_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_SEVERE_FEMALE_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_PERSON_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_PERSON_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_PERSON_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_MALE_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_MALE_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_MALE_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_FEMALE_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_FEMALE_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MAJOR_FEMALE_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_PERSON_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_PERSON_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_PERSON_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_MALE_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_MALE_RURAL` DOUBLE,

`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_MALE_URBAN` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_FEMALE_TOTAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_FEMALE_RURAL` DOUBLE,
`II_INJURED_BY_TYPE_OF_TREATMENT_PER_100000_MINOR_FEMALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_PERSON_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_PERSON_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_PERSON_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_MALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_MALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_MALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_FEMALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_FEMALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_DIARRHOEA_DYSENTERY_FEMALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_PERSON_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_PERSON_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_PERSON_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_MALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_MALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_MALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_FEMALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_FEMALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_RESPIRATORY_INFECTATION_FEMALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_PERSON_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_PERSON_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_PERSON_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_MALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_MALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_MALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_FEMALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_FEMALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_FEVER_ALL_TYPES_FEMALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_PERSON_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_PERSON_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_PERSON_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_MALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_MALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_MALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_FEMALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_FEMALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_PER_100000_ANY_TYPE_OF_ACUTE_FEMALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_PERSON_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_PERSON_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_PERSON_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_MALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_MALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_MALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_FEMALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_FEMALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_FEMALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_PERSON_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_PERSON_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_PERSON_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_MALE_TOTAL` DOUBLE,

`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_MALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_MALE_URBAN` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_FEMALE_TOTAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_FEMALE_RURAL` DOUBLE,
`JJ_ACUTE_ILLNESS_AND_TAKING_TREATMENT_GOVERNMENT_FEMALE_URBAN` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_PERSON_TOTAL` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_PERSON_RURAL` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_PERSON_URBAN` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_MALE_TOTAL` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_MALE_RURAL` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_MALE_URBAN` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_FEMALE_TOTAL` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_FEMALE_RURAL` DOUBLE,
`KK_SYMPTOMS_OF_CHRONIC_ILLNESS_PER_100000_FEMALE_URBAN` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_PERSON_TOTAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_PERSON_RURAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_PERSON_URBAN` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_MALE_TOTAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_MALE_RURAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_MALE_URBAN` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_FEMALE_TOTAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_FEMALE_RURAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_SOUGHT_MEDICAL_CARE_FEMALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_PERSON_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_PERSON_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_PERSON_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_MALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_MALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_MALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_FEMALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_FEMALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_DIABETES_FEMALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_PERSON_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_PERSON_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_PERSON_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_MALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_MALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_MALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_FEMALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_FEMALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_HYPERTENSION_FEMALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_TB_PERSON_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_TB_PERSON_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_TB_PERSON_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_TB_MALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_TB_MALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_TB_MALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILL_PER_100000_TB_FEMALE_TOTAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_TB_FEMALE_RURAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_TB_FEMALE_URBAN` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_PERSON_TOTAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_PERSON_RURAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_PERSON_URBAN` DOUBLE,

`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_MALE_TOTAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_MALE_RURAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_MALE_URBAN` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_FEMALE_TOTAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_FEMALE_RURAL` DOUBLE,
`KK_DIAGNOSED_FOR_CHRONIC_ILLNESS_PER_100000_ASTHMA_FEMALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_PERSON_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_PERSON_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_PERSON_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_MALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_MALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_MALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_FEMALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_FEMALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ARTHRITIS_FEMALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_PERSON_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_PERSON_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_OF_PERSON_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_OF_MALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_OF_MALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_OF_MALE_URBAN` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_OF_FEMALE_TOTAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_OF_FEMALE_RURAL` DOUBLE,
`KK_DIAG_FOR_CHRONIC_ILLNESS_PER_100000_ANY_KIND_OF_FEMALE_URBAN` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_PERSON_TOTAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_PERSON_RURAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_PERSON_URBAN` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_MALE_TOTAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_MALE_RURAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_MALE_URBAN` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_FEMALE_TOTAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_FEMALE_RURAL` DOUBLE,
`KK_CHRONIC_ILLNESS_AND_GETTING_REGULAR_TREATMENT_FEMALE_URBAN` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_PERSON_TOTAL` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_PERSON_RURAL` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_PERSON_URBAN` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_MALE_TOTAL` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_MALE_RURAL` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_MALE_URBAN` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_FEMALE_TOTAL` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_FEMALE_RURAL` DOUBLE,
`KK_CHRONIC_ILL_AND_GETTING_REGULAR_TREATMENT_GOVt_FEMALE_URBAN` DOUBLE,
`LL_CRUDE_BIRTH_RATE_CBR_TOTAL` DOUBLE,
`LL_CRUDE_BIRTH_RATE_CBR_RURAL` DOUBLE,
`LL_CRUDE_BIRTH_RATE_CBR_URBAN` DOUBLE,
`LL_NATURAL_GROWTH_RATE_TOTAL` DOUBLE,
`LL_NATURAL_GROWTH_RATE_RURAL` DOUBLE,
`LL_NATURAL_GROWTH_RATE_URBAN` DOUBLE,
`LL_TOTAL_FERTILITY_RATE_TOTAL` DOUBLE,
`LL_TOTAL_FERTILITY_RATE_RURAL` DOUBLE,
`LL_TOTAL_FERTILITY_RATE_URBAN` DOUBLE,
`LL_WOMEN_20_24_REPORTING_BIRTH_OF_ORDER_2__ABOVE_TOTAL` DOUBLE,
`LL_WOMEN_20_24_REPORTING_BIRTH_OF_ORDER_2__ABOVE_RURAL` DOUBLE,

`LL_WOMEN_20_24_REPORTING_BIRTH_OF_ORDER_2_ABOVE_URBAN` DOUBLE,
`LL_WOMEN_REPORTING_BIRTH_OF_ORDER_3_ABOVE_TOTAL` DOUBLE,
`LL_WOMEN_REPORTING_BIRTH_OF_ORDER_3_ABOVE_RURAL` DOUBLE,
`LL_WOMEN_REPORTING_BIRTH_OF_ORDER_3_ABOVE_URBAN` DOUBLE,
`LL_WOMEN_WITH_TWO_CHILDREN_WANTING_NO_MORE_CHILDREN_TOTAL` DOUBLE,
`LL_WOMEN_WITH_TWO_CHILDREN_WANTING_NO_MORE_CHILDREN_RURAL` DOUBLE,
`LL_WOMEN_WITH_TWO_CHILDREN_WANTING_NO_MORE_CHILDREN_URBAN` DOUBLE,
`LL_WOMEN_15_19_YEARS_WHO_WERE_ALREADY_MOTHERS_OR_PREGNANT_TOTAL` DOUBLE,
`LL_WOMEN_15_19_YEARS_WHO_WERE_ALREADY_MOTHERS_OR_PREGNANT_RURAL` DOUBLE,
`LL_WOMEN_15_19_YEARS_WHO_WERE_ALREADY_MOTHERS_OR_PREGNANT_URBAN` DOUBLE,
`LL_MEDIAN AGE_AT_FIRST_LIVE_BIRTH_OF_WOMEN_15_49_YEARS_TOTAL` DOUBLE,
`LL_MEDIAN AGE_AT_FIRST_LIVE_BIRTH_OF_WOMEN_15_49_YEARS_RURAL` DOUBLE,
`LL_MEDIAN AGE_AT_FIRST_LIVE_BIRTH_OF_WOMEN_15_49_YEARS_URBAN` DOUBLE,
`LL_MEDIAN AGE_AT_FIRST_LIVE_BIRTH_OF_WOMEN_25_49_YEARS_TOTAL` DOUBLE,
`LL_MEDIAN AGE_AT_FIRST_LIVE_BIRTH_OF_WOMEN_25_49_YEARS_RURAL` DOUBLE,
`LL_MEDIAN AGE_AT_FIRST_LIVE_BIRTH_OF_WOMEN_25_49_YEARS_URBAN` DOUBLE,
`LL_LIVE_BIRTHS_TAKING_PLACE_AFTER_AN_INTERVAL_OF_36_MONTHS_TOTAL` DOUBLE,
`LL_LIVE_BIRTHS_TAKING_PLACE_AFTER_AN_INTERVAL_OF_36_MONTHS_RURAL` DOUBLE,
`LL_LIVE_BIRTHS_TAKING_PLACE_AFTER_AN_INTERVAL_OF_36_MONTHS_URBAN` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_EVER_BORN_TO_WOMEN_15_49_YEARS_TOTAL` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_EVER_BORN_TO_WOMEN_15_49_YEARS_RURAL` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_EVER_BORN_TO_WOMEN_15_49_YEARS_URBAN` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_SURVIVING_TO_WOMEN_15_49_YEARS_TOTAL` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_SURVIVING_TO_WOMEN_15_49_YEARS_RURAL` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_SURVIVING_TO_WOMEN_15_49_YEARS_URBAN` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_EVER_BORN_TO_WOMEN_45_49_YEARS_TOTAL` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_EVER_BORN_TO_WOMEN_45_49_YEARS_RURAL` DOUBLE,
`LL_MEAN_NUMBER_OF_CHILDREN_EVER_BORN_TO_WOMEN_45_49_YEARS_URBAN` DOUBLE,
`MM_PREGNANCY_TO_WOMEN_15_49_YEARS_RESULTING_IN_ABORTION_TOTAL` DOUBLE,
`MM_PREGNANCY_TO_WOMEN_15_49_YEARS_RESULTING_IN_ABORTION_RURAL` DOUBLE,
`MM_PREGNANCY_TO_WOMEN_15_49_YEARS_RESULTING_IN_ABORTION_URBAN` DOUBLE,
`MM_WOMEN_WHO RECEIVED_ANY ANC BEFORE_ABORTION_TOTAL` DOUBLE,
`MM_WOMEN_WHO RECEIVED_ANY ANC BEFORE_ABORTION_RURAL` DOUBLE,
`MM_WOMEN_WHO RECEIVED_ANY ANC BEFORE_ABORTION_URBAN` DOUBLE,
`MM_WOMEN_WHO_WENT_FOR_ULTRASOUND_BEFORE_ABORTION_TOTAL` DOUBLE,
`MM_WOMEN_WHO_WENT_FOR_ULTRASOUND_BEFORE_ABORTION_RURAL` DOUBLE,
`MM_WOMEN_WHO_WENT_FOR_ULTRASOUND_BEFORE_ABORTION_URBAN` DOUBLE,
`MM_AVERAGE_MONTH_OF_PREGNANCY_AT_THE_TIME_OF_ABORTION_TOTAL` DOUBLE,
`MM_AVERAGE_MONTH_OF_PREGNANCY_AT_THE_TIME_OF_ABORTION_RURAL` DOUBLE,
`MM_AVERAGE_MONTH_OF_PREGNANCY_AT_THE_TIME_OF_ABORTION_URBAN` DOUBLE,
`MM_ABORTION_PERFORMED_BY_SKILLED_HEALTH_PERSONNEL_TOTAL` DOUBLE,
`MM_ABORTION_PERFORMED_BY_SKILLED_HEALTH_PERSONNEL_RURAL` DOUBLE,
`MM_ABORTION_PERFORMED_BY_SKILLED_HEALTH_PERSONNEL_URBAN` DOUBLE,
`MM_ABORTION_TAKING_PLACE_IN_INSTITUTION_TOTAL` DOUBLE,
`MM_ABORTION_TAKING_PLACE_IN_INSTITUTION_RURAL` DOUBLE,
`MM_ABORTION_TAKING_PLACE_IN_INSTITUTION_URBAN` DOUBLE,
`NN_CURRENT_USAGE_ANY_METHOD_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_ANY_METHOD_RURAL` DOUBLE,
`NN_CURRENT_USAGE_ANY_METHOD_URBAN` DOUBLE,
`NN_CURRENT_USAGE_ANY_MODERN_METHOD_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_ANY_MODERN_METHOD_RURAL` DOUBLE,
`NN_CURRENT_USAGE_ANY_MODERN_METHOD_URBAN` DOUBLE,
`NN_CURRENT_USAGE_FEMALE_STERILIZATION_TOTAL` DOUBLE,

`NN_CURRENT_USAGE_FEMALE_STERILIZATION_RURAL` DOUBLE,
`NN_CURRENT_USAGE_FEMALE_STERILIZATION_URBAN` DOUBLE,
`NN_CURRENT_USAGE_MALE_STERILIZATION_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_MALE_STERILIZATION_RURAL` DOUBLE,
`NN_CURRENT_USAGE_MALE_STERILIZATION_URBAN` DOUBLE,
`NN_CURRENT_USAGE_COPPER_T_IUD_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_COPPER_T_IUD_RURAL` DOUBLE,
`NN_CURRENT_USAGE_COPPER_T_IUD_URBAN` DOUBLE,
`NN_CURRENT_USAGE_PILLS_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_PILLS_RURAL` DOUBLE,
`NN_CURRENT_USAGE_PILLS_URBAN` DOUBLE,
`NN_CURRENT_USAGE_CONDOM_NIRODH_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_CONDOM_NIRODH_RURAL` DOUBLE,
`NN_CURRENT_USAGE_CONDOM_NIRODH_URBAN` DOUBLE,
`NN_CURRENT_USAGE_EMERGENCY CONTRACEPTIVE PILLS_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_EMERGENCY CONTRACEPTIVE PILLS_RURAL` DOUBLE,
`NN_CURRENT_USAGE_EMERGENCY CONTRACEPTIVE PILLS_URBAN` DOUBLE,
`NN_CURRENT_USAGE_ANY_TRADITIONAL_METHOD_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_ANY_TRADITIONAL_METHOD_RURAL` DOUBLE,
`NN_CURRENT_USAGE_ANY_TRADITIONAL_METHOD_URBAN` DOUBLE,
`NN_CURRENT_USAGE_PERIODIC_ABSTINENCE_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_PERIODIC_ABSTINENCE_RURAL` DOUBLE,
`NN_CURRENT_USAGE_PERIODIC_ABSTINENCE_URBAN` DOUBLE,
`NN_CURRENT_USAGE_WITHDRAWAL_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_WITHDRAWAL_RURAL` DOUBLE,
`NN_CURRENT_USAGE_WITHDRAWAL_URBAN` DOUBLE,
`NN_CURRENT_USAGE_LAM_TOTAL` DOUBLE,
`NN_CURRENT_USAGE_LAM_RURAL` DOUBLE,
`NN_CURRENT_USAGE_LAM_URBAN` DOUBLE,
`OO_UNMET_NEED_FOR_SPACING_TOTAL` DOUBLE,
`OO_UNMET_NEED_FOR_SPACING_RURAL` DOUBLE,
`OO_UNMET_NEED_FOR_SPACING_URBAN` DOUBLE,
`OO_UNMET_NEED_FOR_LIMITING_TOTAL` DOUBLE,
`OO_UNMET_NEED_FOR_LIMITING_RURAL` DOUBLE,
`OO_UNMET_NEED_FOR_LIMITING_URBAN` DOUBLE,
`OO_TOTAL_UNMET_NEED_TOTAL` DOUBLE,
`OO_TOTAL_UNMET_NEED_RURAL` DOUBLE,
`OO_TOTAL_UNMET_NEED_URBAN` DOUBLE,
`PP_MARRIED_PREGNANT_WOMEN_15_49_YEARS_REGISTERED_FOR_ANC_TOTAL` DOUBLE,
`PP_MARRIED_PREGNANT_WOMEN_15_49_YEARS_REGISTERED_FOR_ANC_RURAL` DOUBLE,
`PP_MARRIED_PREGNANT_WOMEN_15_49_YEARS_REGISTERED_FOR_ANC_URBAN` DOUBLE,
`PP_MOTHERS WHO RECEIVED ANY ANTEPARTUM CHECK_UP_TOTAL` DOUBLE,
`PP_MOTHERS WHO RECEIVED ANY ANTEPARTUM CHECK_UP_RURAL` DOUBLE,
`PP_MOTHERS WHO RECEIVED ANY ANTEPARTUM CHECK_UP_URBAN` DOUBLE,
`PP_MOTHERS WHO HAD ANTEPARTUM CHECK_UP_IN_FIRST_TRIMESTER_TOTAL` DOUBLE,
`PP_MOTHERS WHO HAD ANTEPARTUM CHECK_UP_IN_FIRST_TRIMESTER_RURAL` DOUBLE,
`PP_MOTHERS WHO HAD ANTEPARTUM CHECK_UP_IN_FIRST_TRIMESTER_URBAN` DOUBLE,
`PP_MOTHERS WHO RECEIVED 3 OR MORE ANTEPARTUM CARE_TOTAL` DOUBLE,
`PP_MOTHERS WHO RECEIVED 3 OR MORE ANTEPARTUM CARE_RURAL` DOUBLE,
`PP_MOTHERS WHO RECEIVED 3 OR MORE ANTEPARTUM CARE_URBAN` DOUBLE,
`PP_MOTHERS WHO RECEIVED_AT LEAST_ONE TT INJECTION_TOTAL` DOUBLE,
`PP_MOTHERS WHO RECEIVED_AT LEAST_ONE TT INJECTION_RURAL` DOUBLE,
`PP_MOTHERS WHO RECEIVED_AT LEAST_ONE TT INJECTION_URBAN` DOUBLE,

`PP_MOTHERS_WHO_CONSUMED_IFA_FOR_100_DAYS_OR_MORE_TOTAL` DOUBLE,
`PP_MOTHERS_WHO_CONSUMED_IFA_FOR_100_DAYS_OR_MORE_RURAL` DOUBLE,
`PP_MOTHERS_WHO_CONSUMED_IFA_FOR_100_DAYS_OR_MORE_URBAN` DOUBLE,
`PP_MOTHERS_WHO_HAD_FULL_ANTENATAL_CHECK_UP_TOTAL` DOUBLE,
`PP_MOTHERS_WHO_HAD_FULL_ANTENATAL_CHECK_UP_RURAL` DOUBLE,
`PP_MOTHERS_WHO_HAD_FULL_ANTENATAL_CHECK_UP_URBAN` DOUBLE,
`PP_MOTHERS_WHO_RECEIVED_ANC_FROM_GOVTSOURCE_TOTAL` DOUBLE,
`PP_MOTHERS_WHO_RECEIVED_ANC_FROM_GOVTSOURCE_RURAL` DOUBLE,
`PP_MOTHERS_WHO_RECEIVED_ANC_FROM_GOVTSOURCE_URBAN` DOUBLE,
`PP_MOTHERS_WHOSE_BLOOD_PRESSURE_BP_TAKEN_TOTAL` DOUBLE,
`PP_MOTHERS_WHOSE_BLOOD_PRESSURE_BP_TAKEN_RURAL` DOUBLE,
`PP_MOTHERS_WHOSE_BLOOD_PRESSURE_BP_TAKEN_URBAN` DOUBLE,
`PP_MOTHERS_WHOSE_BLOOD_TAKEN_FOR_HB_TOTAL` DOUBLE,
`PP_MOTHERS_WHOSE_BLOOD_TAKEN_FOR_HB_RURAL` DOUBLE,
`PP_MOTHERS_WHOSE_BLOOD_TAKEN_FOR_HB_URBAN` DOUBLE,
`PP_MOTHERS_WHO_UNDERWENT_ULTRASOUND_TOTAL` DOUBLE,
`PP_MOTHERS_WHO_UNDERWENT_ULTRASOUND_RURAL` DOUBLE,
`PP_MOTHERS_WHO_UNDERWENT_ULTRASOUND_URBAN` DOUBLE,
`QQ_INSTITUTIONAL_DELIVERY_TOTAL` DOUBLE,
`QQ_INSTITUTIONAL_DELIVERY_RURAL` DOUBLE,
`QQ_INSTITUTIONAL_DELIVERY_URBAN` DOUBLE,
`QQ_DELIVERY_AT_GOVERNMENT_INSTITUTION_TOTAL` DOUBLE,
`QQ_DELIVERY_AT_GOVERNMENT_INSTITUTION_RURAL` DOUBLE,
`QQ_DELIVERY_AT_GOVERNMENT_INSTITUTION_URBAN` DOUBLE,
`QQ_DELIVERY_AT_PRIVATE_INSTITUTION_TOTAL` DOUBLE,
`QQ_DELIVERY_AT_PRIVATE_INSTITUTION_RURAL` DOUBLE,
`QQ_DELIVERY_AT_PRIVATE_INSTITUTION_URBAN` DOUBLE,
`QQ_DELIVERY_AT_HOME_TOTAL` DOUBLE,
`QQ_DELIVERY_AT_HOME_RURAL` DOUBLE,
`QQ_DELIVERY_AT_HOME_URBAN` DOUBLE,
`QQ_DELIVERY_AT_HOME_CONDUCTED_BY_SKILLED_HEALTH_PERSONNEL_TOTAL` DOUBLE,
`QQ_DELIVERY_AT_HOME_CONDUCTED_BY_SKILLED_HEALTH_PERSONNEL_RURAL` DOUBLE,
`QQ_DELIVERY_AT_HOME_CONDUCTED_BY_SKILLED_HEALTH_PERSONNEL_URBAN` DOUBLE,
`QQ_SAFE_DELIVERY_TOTAL` DOUBLE,
`QQ_SAFE_DELIVERY_RURAL` DOUBLE,
`QQ_SAFE_DELIVERY_URBAN` DOUBLE,
`QQ_CAESAREAN_OUT_OF_TOTAL_DELIVERY_IN_GOVERNMENT_TOTAL` DOUBLE,
`QQ_CAESAREAN_OUT_OF_TOTAL_DELIVERY_IN_GOVERNMENT_RURAL` DOUBLE,
`QQ_CAESAREAN_OUT_OF_TOTAL_DELIVERY_IN_GOVERNMENT_URBAN` DOUBLE,
`QQ_CAESAREAN_OUT_OF_TOTAL_DELIVERY_IN_PRIVATE_TOTAL` DOUBLE,
`QQ_CAESAREAN_OUT_OF_TOTAL_DELIVERY_IN_PRIVATE_RURAL` DOUBLE,
`QQ_CAESAREAN_OUT_OF_TOTAL_DELIVERY_IN_PRIVATE_URBAN` DOUBLE,
`RR_LESS_THAN_24_HRS_STAY_IN_INSTITUTION_AFTER_DELIVERY_TOTAL` DOUBLE,
`RR_LESS_THAN_24_HRS_STAY_IN_INSTITUTION_AFTER_DELIVERY_RURAL` DOUBLE,
`RR_LESS_THAN_24_HRS_STAY_IN_INSTITUTION_AFTER_DELIVERY_URBAN` DOUBLE,
`RR_MOTHERS_WHO RECEIVED_WITHIN_48_HRS_OF_DELIVERY_TOTAL` DOUBLE,
`RR_MOTHERS_WHO RECEIVED_WITHIN_48_HRS_OF_DELIVERY_RURAL` DOUBLE,
`RR_MOTHERS_WHO RECEIVED_WITHIN_48_HRS_OF_DELIVERY_URBAN` DOUBLE,
`RR_MOTHERS_WHO RECEIVED_WITHIN_1_WEEK_OF_DELIVERY_TOTAL` DOUBLE,
`RR_MOTHERS_WHO RECEIVED_WITHIN_1_WEEK_OF_DELIVERY_RURAL` DOUBLE,
`RR_MOTHERS_WHO RECEIVED_WITHIN_1_WEEK_OF_DELIVERY_URBAN` DOUBLE,
`RR_MOTHERS_WHO_DID_NOT_RECEIVE_ANY_POST_NATAL_CHECK_UP_TOTAL` DOUBLE,
`RR_MOTHERS_WHO_DID_NOT_RECEIVE_ANY_POST_NATAL_CHECK_UP_RURAL` DOUBLE,

`RR_MOTHERS_WHO_DID_NOT_RECEIVE_ANY_POST_NATAL_CHECK_UP_URBAN` DOUBLE,
`RR_NEW_BORNS_WHO_WERE_CHECKED_UP_WITHIN_24_HRS_OF_BIRTH_TOTAL` DOUBLE,
`RR_NEW_BORNS_WHO_WERE_CHECKED_UP_WITHIN_24_HRS_OF_BIRTH_RURAL` DOUBLE,
`RR_NEW_BORNS_WHO_WERE_CHECKED_UP_WITHIN_24_HRS_OF_BIRTH_URBAN` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_DELIVERY_UNDER_JSY_TOTAL` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_DELIVERY_UNDER_JSY_RURAL` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_DELIVERY_UNDER_JSY_URBAN` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_INST_DELIVERY_UNDER_JSY_TOTAL` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_INST_DELIVERY_UNDER_JSY_RURAL` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_INST_DELIVERY_UNDER_JSY_URBAN` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_GOVT_DELIVERY_UNDER_JSY_TOTAL` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_GOVT_DELIVERY_UNDER_JSY_RURAL` DOUBLE,
`SS_AVAILED_FINANCIAL_ASSISTANCE_FOR_GOVT_DELIVERY_UNDER_JSY_URBAN` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_HAVING_IMMUNIZATION_CARD_TOTAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_HAVING_IMMUNIZATION_CARD_RURAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_HAVING_IMMUNIZATION_CARD_URBAN` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_WHO_HAVE_RECEIVED_BCG_TOTAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_WHO_HAVE_RECEIVED_BCG_RURAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_WHO_HAVE_RECEIVED_BCG_URBAN` DOUBLE,
`TT_CHILDREN_12_23_MONTHS_RECEIVED_3_DOSES_OF_POLIO_VACCINE_TOTAL` DOUBLE,
`TT_CHILDREN_12_23_MONTHS_RECEIVED_3_DOSES_OF_POLIO_VACCINE_RURAL` DOUBLE,
`TT_CHILDREN_12_23_MONTHS_RECEIVED_3_DOSES_OF_POLIO_VACCINE_URBAN` DOUBLE,
`TT_CHILDREN_12_23_MONTHS_RECEIVED_3_DOSES_OF_DPT_VACCINE_TOTAL` DOUBLE,
`TT_CHILDREN_12_23_MONTHS_RECEIVED_3_DOSES_OF_DPT_VACCINE_RURAL` DOUBLE,
`TT_CHILDREN_12_23_MONTHS_RECEIVED_3_DOSES_OF_DPT_VACCINE_URBAN` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_RECEIVED_MEASLES_VACCINE_TOTAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_RECEIVED_MEASLES_VACCINE_RURAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_RECEIVED_MEASLES_VACCINE_URBAN` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_FULLY_IMMUNIZED_TOTAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_FULLY_IMMUNIZED_RURAL` DOUBLE,
`TT_CHILDRENAGED_12_23_MONTHS_FULLY_IMMUNIZED_URBAN` DOUBLE,
`TT_CHILDREN_WHO_HAVE_RECEIVED_POLIO_DOSE_AT_BIRTH_TOTAL` DOUBLE,
`TT_CHILDREN_WHO_HAVE_RECEIVED_POLIO_DOSE_AT_BIRTH_RURAL` DOUBLE,
`TT_CHILDREN_WHO_HAVE_RECEIVED_POLIO_DOSE_AT_BIRTH_URBAN` DOUBLE,
`TT_CHILDREN_WHO_DID_NOT_RECEIVE_ANY_VACCINATION_TOTAL` DOUBLE,
`TT_CHILDREN_WHO_DID_NOT_RECEIVE_ANY_VACCINATION_RURAL` DOUBLE,
`TT_CHILDREN_WHO_DID_NOT_RECEIVE_ANY_VACCINATION_URBAN` DOUBLE,
`TT_CHILDREN_6_35_MON_AT_LEAST_1_VIT_A_DOSE_LAST_6_MONTHS_TOTAL` DOUBLE,
`TT_CHILDREN_6_35_MON_AT_LEAST_1_VIT_A_DOSE_LAST_6_MONTHS_RURAL` DOUBLE,
`TT_CHILDREN_6_35_MON_AT_LEAST_1_VIT_A_DOSE_LAST_6_MONTHS_URBAN` DOUBLE,
`TT_CHILDREN_6_35_MON_IFA_TABLETS_SYRUP_LAST_3_MONTHS_TOTAL` DOUBLE,
`TT_CHILDREN_6_35_MON_IFA_TABLETS_SYRUP_LAST_3_MONTHS_RURAL` DOUBLE,
`TT_CHILDREN_6_35_MON_IFA_TABLETS_SYRUP_LAST_3_MONTHS_URBAN` DOUBLE,
`TT_CHILDREN_WHOSE_BIRTH_WEIGHT_WAS_TAKEN_TOTAL` DOUBLE,
`TT_CHILDREN_WHOSE_BIRTH_WEIGHT_WAS_TAKEN_RURAL` DOUBLE,
`TT_CHILDREN_WHOSE_BIRTH_WEIGHT_WAS_TAKEN_URBAN` DOUBLE,
`TT_CHILDREN_WITH_BIRTH_WEIGHT_LESS_THAN_2_5_KG_TOTAL` DOUBLE,
`TT_CHILDREN_WITH_BIRTH_WEIGHT_LESS_THAN_2_5_KG_RURAL` DOUBLE,
`TT_CHILDREN_WITH_BIRTH_WEIGHT_LESS_THAN_2_5_KG_URBAN` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_DIARRHOEA_TOTAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_DIARRHOEA_RURAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_DIARRHOEA_URBAN` DOUBLE,
`UU_CHILDREN_DIARRHOEA_WHO RECEIVED_HAF_ORS_ORT_TOTAL` DOUBLE,

`UU_CHILDREN_DIARRHOEA_WHO RECEIVED_HAF_ORS_ORT_RURAL` DOUBLE,
`UU_CHILDREN_DIARRHOEA_WHO RECEIVED_HAF_ORS_ORT_URBAN` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_ACUTE_RESPIRATORY_INFECTION_TOTAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_ACUTE_RESPIRATORY_INFECTION_RURAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_ACUTE_RESPIRATORY_INFECTION_URBAN` DOUBLE,
`UU_CHILDREN_ACUTE_RESPIRATORY_INFECTION_SOUGHT_TREATMENT_TOTAL` DOUBLE,
`UU_CHILDREN_ACUTE_RESPIRATORY_INFECTION_SOUGHT_TREATMENT_RURAL` DOUBLE,
`UU_CHILDREN_ACUTE_RESPIRATORY_INFECTION_SOUGHT_TREATMENT_URBAN` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_FEVER_TOTAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_FEVER_RURAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_FEVER_URBAN` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_FEVER_WHO_SOUGHT_TREATMENT_TOTAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_FEVER_WHO_SOUGHT_TREATMENT_RURAL` DOUBLE,
`UU_CHILDREN_SUFFERING_FROM_FEVER_WHO_SOUGHT_TREATMENT_URBAN` DOUBLE,
`VV_CHILDREN_BREASTFED_WITHIN_ONE_HOUR_OF_BIRTH_TOTAL` DOUBLE,
`VV_CHILDREN_BREASTFED_WITHIN_ONE_HOUR_OF_BIRTH_RURAL` DOUBLE,
`VV_CHILDREN_BREASTFED_WITHIN_ONE_HOUR_OF_BIRTH_URBAN` DOUBLE,
`VV_CHILDREN_6_35_MON_EXCL_BREASTFED_FOR_AT_LEAST_6_MON_TOTAL` DOUBLE,
`VV_CHILDREN_6_35_MON_EXCL_BREASTFED_FOR_AT_LEAST_6_MON_RURAL` DOUBLE,
`VV_CHILDREN_6_35_MON_EXCL_BREASTFED_FOR_AT_LEAST_6_MON_URBAN` DOUBLE,
`VV_OTHER_THAN_BREAST_MILK_DURING_FIRST_6_MONTHS_WATER_TOTAL` DOUBLE,
`VV_OTHER_THAN_BREAST_MILK_DURING_FIRST_6_MONTHS_WATER_RURAL` DOUBLE,
`VV_OTHER_THAN_BREAST_MILK_DURING_FIRST_6_MONTHS_WATER_URBAN` DOUBLE,
`VV_1ST_6_MONTHS_ANIMAL_FORMULA_MILK_TOTAL` DOUBLE,
`VV_1ST_6_MONTHS_ANIMAL_FORMULA_MILK_RURAL` DOUBLE,
`VV_1ST_6_MONTHS_ANIMAL_FORMULA_MILK_URBAN` DOUBLE,
`VV_1ST_6_MONTHS_SEMI_SOLID_MASHED_FOOD_TOTAL` DOUBLE,
`VV_1ST_6_MONTHS_SEMI_SOLID_MASHED_FOOD_RURAL` DOUBLE,
`VV_1ST_6_MONTHS_SEMI_SOLID_MASHED_FOOD_URBAN` DOUBLE,
`VV_1ST_6_MONTHS_SOLID_ADULT_FOOD_TOTAL` DOUBLE,
`VV_1ST_6_MONTHS_SOLID_ADULT_FOOD_RURAL` DOUBLE,
`VV_1ST_6_MONTHS_SOLID_ADULT_FOOD_URBAN` DOUBLE,
`VV_1ST_6_MONTHS_VEGETABLES_FRUITS_TOTAL` DOUBLE,
`VV_1ST_6_MONTHS_VEGETABLES_FRUITS_RURAL` DOUBLE,
`VV_1ST_6_MONTHS_VEGETABLES_FRUITS_URBAN` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_WATER_TOTAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_WATER_RURAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_WATER_URBAN` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_ANIMAL_FORMULA_MILK_TOTAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_ANIMAL_FORMULA_MILK_RURAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_ANIMAL_FORMULA_MILK_URBAN` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_SEMI_SOLID_MASHED_FOOD_TOTAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_SEMI_SOLID_MASHED_FOOD_RURAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_SEMI_SOLID_MASHED_FOOD_URBAN` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_SOLID_ADULT_FOOD_TOTAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_SOLID_ADULT_FOOD_RURAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_SOLID_ADULT_FOOD_URBAN` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_VEGETABLES_FRUITS_TOTAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_VEGETABLES_FRUITS_RURAL` DOUBLE,
`VV_AVG_MONTH_OTHER_THAN_BREAST_MILK_VEGETABLES_FRUITS_URBAN` DOUBLE,
`WW_BIRTH_REGISTERED_TOTAL` DOUBLE,
`WW_BIRTH_REGISTERED_RURAL` DOUBLE,
`WW_BIRTH_REGISTERED_URBAN` DOUBLE,

'WW_CHILDREN_REGISTERED_AND RECEIVED_BIRTH_CERTIFICATE_TOTAL` DOUBLE,
'WW_CHILDREN_REGISTERED_AND RECEIVED_BIRTH_CERTIFICATE_RURAL` DOUBLE,
'WW_CHILDREN_REGISTERED_AND RECEIVED_BIRTH_CERTIFICATE_URBAN` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_HIV_AIDS_TOTAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_HIV_AIDS_RURAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_HIV_AIDS_URBAN` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_RTI_STI_TOTAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_RTI_STI_RURAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_RTI_STI_URBAN` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_HAF_ORS_ORT_ZINC_TOTAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_HAF_ORS_ORT_ZINC_RURAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_HAF_ORS_ORT_ZINC_URBAN` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_DANGER_SIGNS_OF_ARI_PNEUMONIA_TOTAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_DANGER_SIGNS_OF_ARI_PNEUMONIA_RURAL` DOUBLE,
'XX_WOMEN_WHO_ARE_AWARE_OF_DANGER_SIGNS_OF_ARI_PNEUMONIA_URBAN` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_TOTAL_PERSON` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_TOTAL_MALE` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_TOTAL_FEMALE` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_RURAL_PERSON` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_RURAL_MALE` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_RURAL_FEMALE` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_URBAN_PERSON` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_URBAN_MALE` DOUBLE,
'YY_CRUDE_DEATH_RATE_CDR_URBAN_FEMALE` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_TOTAL_PERSON` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_TOTAL_MALE` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_TOTAL_FEMALE` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_RURAL_PERSON` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_RURAL_MALE` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_RURAL_FEMALE` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_URBAN_PERSON` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_URBAN_MALE` DOUBLE,
'YY_INFANT_MORTALITY_RATE_IMR_URBAN_FEMALE` DOUBLE,
'YY_NEONATAL_MORTALITY_RATE_TOTAL` DOUBLE,
'YY_NEONATAL_MORTALITY_RATE_RURAL` DOUBLE,
'YY_NEONATAL_MORTALITY_RATE_URBAN` DOUBLE,
'YY_POST_NEONATAL_MORTALITY_RATE_TOTAL` DOUBLE,
'YY_POST_NEONATAL_MORTALITY_RATE_RURAL` DOUBLE,
'YY_POST_NEONATAL_MORTALITY_RATE_URBAN` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_MALE` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_FEMALE` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_RURAL_PERSON` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_RURAL_MALE` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_RURAL_FEMALE` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_URBAN_PERSON` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_URBAN_MALE` DOUBLE,
'YY_UNDER_FIVE_MORTALITY_RATE_U5MR_URBAN_FEMALE` DOUBLE,
'ZZ_CRUDE_BIRTH_RATE_TOTAL_LOWER_LIMIT` DOUBLE,
'ZZ_CRUDE_BIRTH_RATE_TOTAL_UPPER_LIMIT` DOUBLE,
'ZZ_CRUDE_BIRTH_RATE_RURAL_LOWER_LIMIT` DOUBLE,
'ZZ_CRUDE_BIRTH_RATE_RURAL_UPPER_LIMIT` DOUBLE,
'ZZ_CRUDE_BIRTH_RATE_URBAN_LOWER_LIMIT` DOUBLE,

```

`ZZ_CRUDE_BIRTH_RATE_URBAN_UPPER_LIMIT` DOUBLE,
`ZZ_CRUDE_DEATH_RATE_TOTAL_LOWER_LIMIT` DOUBLE,
`ZZ_CRUDE_DEATH_RATE_TOTAL_UPPER_LIMIT` DOUBLE,
`ZZ_CRUDE_DEATH_RATE_RURAL_LOWER_LIMIT` DOUBLE,
`ZZ_CRUDE_DEATH_RATE_RURAL_UPPER_LIMIT` DOUBLE,
`ZZ_CRUDE_DEATH_RATE_URBAN_LOWER_LIMIT` DOUBLE,
`ZZ_CRUDE_DEATH_RATE_URBAN_UPPER_LIMIT` DOUBLE,
`ZZ_INFANT_MORTALITY_RATE_TOTAL_LOWER_LIMIT` DOUBLE,
`ZZ_INFANT_MORTALITY_RATE_TOTAL_UPPER_LIMIT` DOUBLE,
`ZZ_INFANT_MORTALITY_RATE_RURAL_LOWER_LIMIT` DOUBLE,
`ZZ_INFANT_MORTALITY_RATE_RURAL_UPPER_LIMIT` DOUBLE,
`ZZ_INFANT_MORTALITY_RATE_URBAN_LOWER_LIMIT` DOUBLE,
`ZZ_INFANT_MORTALITY_RATE_URBAN_UPPER_LIMIT` DOUBLE,
`ZZ_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_LOWER_LIMIT` DOUBLE,
`ZZ_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_UPPER_LIMIT` DOUBLE,
`ZZ_UNDER_FIVE_MORTALITY_RATE_U5MR_RURAL_LOWER_LIMIT` DOUBLE,
`ZZ_UNDER_FIVE_MORTALITY_RATE_U5MR_RURAL_UPPER_LIMIT` DOUBLE,
`ZZ_UNDER_FIVE_MORTALITY_RATE_U5MR_URBAN_LOWER_LIMIT` DOUBLE,
`ZZ_UNDER_FIVE_MORTALITY_RATE_U5MR_URBAN_UPPER_LIMIT` DOUBLE,
`ZZ_SEX_RATIO_AT_BIRTH_TOTAL_LOWER_LIMIT` DOUBLE,
`ZZ_SEX_RATIO_AT_BIRTH_TOTAL_UPPER_LIMIT` DOUBLE,
`ZZ_SEX_RATIO_AT_BIRTH_RURAL_LOWER_LIMIT` DOUBLE,
`ZZ_SEX_RATIO_AT_BIRTH_RURAL_UPPER_LIMIT` DOUBLE,
`ZZ_SEX_RATIO_AT_BIRTH_URBAN_LOWER_LIMIT` DOUBLE,
`ZZ_SEX_RATIO_AT_BIRTH_URBAN_UPPER_LIMIT` DOUBLE)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
LOCATION '/user/cloudera/etlproject2/';

```

Above 3 statements can be put in a .sql file namely `create_table.sql` and can be run like this:

```
[cloudera@quickstart project2]$ hive -f create_table.sql
```

```

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
OK
Time taken: 0.555 seconds
OK
Time taken: 0.128 seconds
OK
Time taken: 0.8 seconds
WARN: The method class org.apache.commons.logging.impl.SLF4JLogFactory#release() was invoked.
WARN: Please see http://www.slf4j.org/codes.html#release for an explanation.
[cloudera@quickstart project2]$

```

2. Command to load the ingested data into the external table

```
load data inpath '/user/cloudera/project2/part*' overwrite into table key_indicators_ext;
```

```

[cloudera@quickstart project2]$ hive
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> use etlproject2;
OK
Time taken: 0.432 seconds
hive> load data inpath '/user/cloudera/project2/part*' overwrite into table key_indicators_ext;
Loading data to table etlproject2.key_indicators_ext

```

```
Table etlproject2.key_indicators_ext stats: [numFiles=4, numRows=0, totalSize=1027652, rawDataSize=0]
```

```
OK
```

```
Time taken: 1.277 seconds
```

```
hive>
```

3. Queries to verify that the ingestion is correctly accomplished

- (a) Query to count the total number of rows along with the screenshots of the data fetched by the query on MySQL Workbench and Hue

MySQL:

```
select count(*) from Key_indicator_districtwise;
```

```
[cloudera@quickstart ~]$ mysql -h upgradawsrds.cpclxrkdvwzm.us-east-1.rds.amazonaws.com -u upgraduser -p
```

```
Enter password:
```

```
Welcome to the MySQL monitor. Commands end with ; or \g.
```

```
Your MySQL connection id is 25921
```

```
Server version: 5.6.39-log MySQL Community Server (GPL)
```

```
Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.
```

```
Oracle is a registered trademark of Oracle Corporation and/or its  
affiliates. Other names may be trademarks of their respective  
owners.
```

```
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
```

```
mysql> show databases;
```

```
+-----+  
| Database      |  
+-----+  
| information_schema |  
| cred_financials_data |  
| indiaahs2012_13 |  
| innodb       |  
| mysql        |  
| performance_schema |  
| sys          |  
| tmp          |  
+-----+  
8 rows in set (0.29 sec)
```

```
mysql> use indiaahs2012_13;
```

```
Reading table information for completion of table and column names
```

```
You can turn off this feature to get a quicker startup with -A
```

```
Database changed
```

```
mysql> show tables;
```

```
+-----+  
| Tables_in_indiaahs2012_13 |  
+-----+  
| Key_indicator_districtwise |  
+-----+  
1 row in set (0.29 sec)
```

```
mysql> select count(*) from Key_indicator_districtwise;
```

```
+-----+  
| count(*) |  
+-----+  
|    284 |  
+-----+  
1 row in set (0.29 sec)
```

The screenshot shows the MySQL Workbench interface. In the top navigation bar, the database 'awsrdsupgrad' is selected. The main area displays a query editor with the following SQL code:

```

1 show databases;
2
3 use indiaahs2012_13;
4
5 show tables;
6
7 select count(*) from Key_indicator_districtwise;
8
9
10

```

Below the code, the results are shown in a table:

	count(*)
	284

The left sidebar contains several sections: MANAGEMENT (Server Status, Client Connections, Users and Privileges, Status and System Variables, Data Export, Data Import/Restore), INSTANCE (Startup / Shutdown, Server Logs, Options File), PERFORMANCE (Dashboard, Performance Reports, Performance Schema Setup), and SCHEMAS (with 'indiaahs2012_13' expanded to show Tables, Views, Stored Procedures, Functions).

HUE:

select count(*) from key_indicators_ext;

```

hive> select count(*) from key_indicators_ext;
Query ID = cloudera_20181130211111_4e29de2a-2527-47ce-a531-8ac0d5f160db
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
set mapreduce.job.reduces=<number>
Starting Job = job_1543633035369_0036, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1543633035369_0036/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1543633035369_0036
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2018-11-30 21:11:45,545 Stage-1 map = 0%, reduce = 0%
2018-11-30 21:11:51,847 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.95 sec
2018-11-30 21:11:59,256 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.68 sec
MapReduce Total cumulative CPU time: 4 seconds 680 msec
Ended Job = job_1543633035369_0036
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.68 sec HDFS Read: 1274345 HDFS Write: 4 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 680 msec
OK
284
Time taken: 23.11 seconds, Fetched: 1 row(s)
hive>

```

```

use etlproject2;
select count(*) from key_indicators_ext;

INFO : Ended Job = job_1543641584890_0001
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 5.38 sec   HDFS Read: 1274253 HDFS Write: 4 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 380 msec
INFO : Completed executing command(queryId=hive_20181130212525_515e5ba5-ee44-412e-8388-7145f53c7e86); Time taken: 35.101 seconds
INFO : OK

```

Query History | Saved Queries | Results(1) | C | ✓

	_c0
1	284

(b) Query to select the top 10 rows and first 8 columns along with the screenshots of the data fetched by the query on MySQL Workbench and Hue

MySQL:

```

select ID, State_Name, State_District_Name, AA_Sample_Units_Total, AA_Sample_Units_Rural,
AA_Sample_Units_Urban, AA_Households_Total, AA_Households_Rural
from Key_indicator_districtwise limit 10;

```

```

mysql> select ID, State_Name, State_District_Name, AA_Sample_Units_Total, AA_Sample_Units_Rural,
-> AA_Sample_Units_Urban, AA_Households_Total, AA_Households_Rural
-> from Key_indicator_districtwise limit 10;
+----+-----+-----+-----+-----+-----+-----+
| ID | State_Name | State_District_Name | AA_Sample_Units_Total | AA_Sample_Units_Rural | AA_Sample_Units_Urban |
| AA_Households_Total | AA_Households_Rural |
+----+-----+-----+-----+-----+-----+
| 1 | Assam | Barpeta | 53 | 47 | 6 | 13711 | 12765 |
| 2 | Assam | Bongaigaon | 89 | 73 | 16 | 17384 | 14904 |
| 3 | Assam | Cachar | 105 | 84 | 21 | 27488 | 24207 |
| 4 | Assam | Darrang | 26 | 24 | 2 | 5951 | 5769 |
| 5 | Assam | Dhemaji | 121 | 108 | 13 | 14481 | 12619 |
| 6 | Assam | Dhubri | 42 | 35 | 7 | 11001 | 9954 |
| 7 | Assam | Dibrugarh | 91 | 66 | 25 | 21378 | 16514 |
| 8 | Assam | Goalpara | 64 | 56 | 8 | 15891 | 14630 |
| 9 | Assam | Golaghat | 70 | 61 | 9 | 16021 | 14183 |
| 10 | Assam | Hailakandi | 10 | 8 | 2 | 2802 | 2381 |
+----+-----+-----+-----+-----+-----+

```

10 rows in set (0.30 sec)

mysql>

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: awsrdsupgrad

Query 1 SQL File 3*

```

1 • show databases;
2
3 • use indiaahs2012_13;
4
5 • show tables;
6
7 • select ID, State_Name, State_District_Name, AA_Sample_Units_Total, AA_Sample_Units_Rural,
8 AA_Sample_Units_Urban, AA_Households_Total,AA_Households_Rural
9 from Key_Indicator_districtwise limit 10;
10

```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: |

ID	State_Name	State_District_Name	AA_Sample_Units_Total	AA_Sample_Units_Rural	AA_Sample_Units_Urban	AA_Households_Total	AA_Households_Rural
1	Assam	Barpeta	53	47	6	13711	12765
2	Assam	Bongaigaon	89	73	16	17384	14904
3	Assam	Cachar	105	84	21	27488	24207
4	Assam	Darrang	26	24	2	5951	5769
5	Assam	Dhemaji	121	108	13	14481	12619
6	Assam	Dhubri	42	35	7	11001	9954
7	Assam	Dibrugarh	91	66	25	21378	16514
8	Assam	Goalpara	64	56	8	15891	14630
9	Assam	Golaghat	70	61	9	16021	14183
10	Assam	Hailakandi	10	8	2	2802	2381
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

MANAGEMENT

- Server Status
- Client Connections
- Users and Privileges
- Status and System Variables
- Data Export
- Data Import/Restore

INSTANCE

- Startup / Shutdown
- Server Logs
- Options File

PERFORMANCE

- Dashboard
- Performance Reports
- Performance Schema Setup

SCHEMAS

Filter objects

- cred_financials_data
- indiaahs2012_13
 - Tables
 - Key_Indicator_districtwise
 - Views
 - Stored Procedures
 - Functions
- innodb
- sys
- tmp

HUE:

```

select ID, State_Name, State_District_Name, AA_Sample_Units_Total, AA_Sample_Units_Rural,
AA_Sample_Units_Urban, AA_Households_Total,AA_Households_Rural
from key_indicators_ext limit 10;

```

```

hive> select ID, State_Name, State_District_Name, AA_Sample_Units_Total, AA_Sample_Units_Rural,
> AA_Sample_Units_Urban, AA_Households_Total,AA_Households_Rural
> from key_indicators_ext limit 10;

```

OK

1	Assam	Barpeta	53.0	47.0	6.0	13711.0	12765.0
2	Assam	Bongaigaon	89.0	73.0	16.0	17384.0	14904.0
3	Assam	Cachar	105.0	84.0	21.0	27488.0	24207.0
4	Assam	Darrang	26.0	24.0	2.0	5951.0	5769.0
5	Assam	Dhemaji	121.0	108.0	13.0	14481.0	12619.0
6	Assam	Dhubri	42.0	35.0	7.0	11001.0	9954.0
7	Assam	Dibrugarh	91.0	66.0	25.0	21378.0	16514.0
8	Assam	Goalpara	64.0	56.0	8.0	15891.0	14630.0
9	Assam	Golaghat	70.0	61.0	9.0	16021.0	14183.0
10	Assam	Hailakandi	10.0	8.0	2.0	2802.0	2381.0

Time taken: 0.172 seconds, Fetched: 10 row(s)

hive>

The screenshot shows the Hue interface with a query editor and a results table.

Query Editor:

```

1 use etlproject2;
2 select ID, State_Name, State_District_Name, AA_Sample_Units_Total, AA_Sample_Units_Rural,
3 AA_Sample_Units_Urban, AA_Households_Total, AA_Households_Rural
4 from key_indicators_ext limit 10;
5

```

INFO : Completed executing command(queryId=hive_20181130212626_e67667e3-39a5-433f-af42-3e10141eb8be); Time taken: 0.001 seconds
INFO : OK

Results Table:

	id	state_name	state_district_name	aa_sample_units_total	aa_sample_units_rural	aa_sample_units_urban	aa_households_total	aa_households_rural
1	1	Assam	Barpeta	53	47	6	13711	12765
2	2	Assam	Bongaigaon	89	73	16	17384	14904
3	3	Assam	Cachar	105	84	21	27488	24207
4	4	Assam	Darrang	26	24	2	5951	5769
5	5	Assam	Dhemaji	121	108	13	14481	12619
6	6	Assam	Dhubri	42	35	7	11001	9954
7	7	Assam	Dibrugarh	91	66	25	21378	16514
8	8	Assam	Goalpara	64	56	8	15891	14630
9	9	Assam	Golaghat	70	61	9	16021	14183
10	10	Assam	Hailakandi	10	8	2	2802	2381

Subset schema creation in Hive to support the analyses

1. Columns used in the subset schema

ID,
 STATE_NAME,
 STATE_DISTRICT_NAME,
 AA_HOUSEHOLDS_TOTAL,
 AA_POPULATION_TOTAL,
 CC_SEX_RATIO_ALL_AGES_TOTAL,
 LL_TOTAL_FERTILITY_RATE_TOTAL,
 YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON

2. Storage format used

ORC with SNAPPY compression

3. Create and insert command for the default format

```

CREATE TABLE IF NOT EXISTS KEY_INDICATORS_INT(
ID INT,
STATE_NAME VARCHAR(100),
STATE_DISTRICT_NAME VARCHAR(100),
AA_HOUSEHOLDS_TOTAL DOUBLE,
AA_POPULATION_TOTAL DOUBLE,
CC_SEX_RATIO_ALL_AGES_TOTAL DOUBLE,
LL_TOTAL_FERTILITY_RATE_TOTAL DOUBLE,
YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON DOUBLE)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n';

```

```

hive> use etlproject2;
OK
Time taken: 0.667 seconds
hive> show tables;
OK
key_indicators_ext

```

```
Time taken: 0.47 seconds, Fetched: 1 row(s)
hive> CREATE TABLE IF NOT EXISTS KEY_INDICATORS_INT(
    > ID INT,
    > STATE_NAME VARCHAR(100),
    > STATE_DISTRICT_NAME VARCHAR(100),
    > AA_HOUSEHOLDS_TOTAL DOUBLE,
    > AA_POPULATION_TOTAL DOUBLE,
    > CC_SEX_RATIO_ALL_AGES_TOTAL DOUBLE,
    > LL_TOTAL_FERTILITY_RATE_TOTAL DOUBLE,
    > YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON DOUBLE)
    > ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
    > LINES TERMINATED BY '\n';
OK
```

```
Time taken: 1.592 seconds
hive>
```

```
INSERT OVERWRITE TABLE KEY_INDICATORS_INT
SELECT
ID,
STATE_NAME,
STATE_DISTRICT_NAME,
AA_HOUSEHOLDS_TOTAL,
AA_POPULATION_TOTAL,
CC_SEX_RATIO_ALL_AGES_TOTAL,
LL_TOTAL_FERTILITY_RATE_TOTAL,
YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON
FROM KEY_INDICATORS_EXT;
```

```
hive> INSERT OVERWRITE TABLE KEY_INDICATORS_INT
    > SELECT
    > ID,
    > STATE_NAME,
    > STATE_DISTRICT_NAME,
    > AA_HOUSEHOLDS_TOTAL,
    > AA_POPULATION_TOTAL,
    > CC_SEX_RATIO_ALL_AGES_TOTAL,
    > LL_TOTAL_FERTILITY_RATE_TOTAL,
    > YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON
    > FROM KEY_INDICATORS_EXT;
```

```
Query ID = cloudera_20181128194545_c1bd1cd7-5159-4b2f-8689-9c35acc77933
```

```
Total jobs = 3
```

```
Launching Job 1 out of 3
```

```
Number of reduce tasks is set to 0 since there is no reduce operator
```

```
Starting Job = job_1543462035376_0001, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1543462035376_0001/
```

```
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1543462035376_0001
```

```
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
```

```
2018-11-28 19:45:26,358 Stage-1 map = 0%, reduce = 0%
```

```
2018-11-28 19:45:33,798 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.22 sec
```

```
MapReduce Total cumulative CPU time: 2 seconds 220 msec
```

```
Ended Job = job_1543462035376_0001
```

```
Stage-4 is selected by condition resolver.
```

```
Stage-3 is filtered out by condition resolver.
```

```
Stage-5 is filtered out by condition resolver.
```

```
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/etlproject2.db/key_indicators_int/.hive-staging_hive_2018-11-28_19-45-13_184_3466385637697334553-1/ext-10000
```

```
Loading data to table etlproject2.key_indicators_int
```

```
Table etlproject2.key_indicators_int stats: [numFiles=1, numRows=284, totalSize=15799, rawDataSize=15515]
```

```
MapReduce Jobs Launched:
```

```
Stage-Stage-1: Map: 1 Cumulative CPU: 2.22 sec HDFS Read: 1181054 HDFS Write: 15889 SUCCESS
```

```
Total MapReduce CPU Time Spent: 2 seconds 220 msec
```

```
OK
```

```
Time taken: 22.288 seconds
```

4. Create and insert command for the formats such as ORC

Setting few parameters for ORC format with SNAPPY compression:

```
SET orc.compress=SNAPPY;
SET hive.exec.compress.output=true;
SET mapred.output.compression.codec=org.apache.hadoop.io.compress.SnappyCodec;
SET mapred.output.compression.type=BLOCK;
```

```
hive> SET orc.compress=SNAPPY;
hive> SET hive.exec.compress.output=true;
hive> SET mapred.output.compression.codec=org.apache.hadoop.io.compress.SnappyCodec;
hive> SET mapred.output.compression.type=BLOCK;
```

```
CREATE TABLE IF NOT EXISTS KEY_INDICATORS_INT_ORC(
ID INT,
STATE_NAME VARCHAR(100),
STATE_DISTRICT_NAME VARCHAR(100),
AA_HOUSEHOLDS_TOTAL DOUBLE,
AA_POPULATION_TOTAL DOUBLE,
CC_SEX_RATIO_ALL_AGES_TOTAL DOUBLE,
LL_TOTAL_FERTILITY_RATE_TOTAL DOUBLE,
YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON DOUBLE)
STORED AS ORC
TBLPROPERTIES("orc.compress"="SNAPPY");
```

```
hive> CREATE TABLE IF NOT EXISTS KEY_INDICATORS_INT_ORC(
> ID INT,
> STATE_NAME VARCHAR(100),
> STATE_DISTRICT_NAME VARCHAR(100),
> AA_HOUSEHOLDS_TOTAL DOUBLE,
> AA_POPULATION_TOTAL DOUBLE,
> CC_SEX_RATIO_ALL_AGES_TOTAL DOUBLE,
> LL_TOTAL_FERTILITY_RATE_TOTAL DOUBLE,
> YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON DOUBLE)
> STORED AS ORC
> TBLPROPERTIES("orc.compress"="SNAPPY");
OK
Time taken: 0.144 seconds
hive>
```

```
INSERT OVERWRITE TABLE KEY_INDICATORS_INT_ORC
SELECT
ID,
STATE_NAME,
STATE_DISTRICT_NAME,
AA_HOUSEHOLDS_TOTAL,
AA_POPULATION_TOTAL,
CC_SEX_RATIO_ALL_AGES_TOTAL,
LL_TOTAL_FERTILITY_RATE_TOTAL,
YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON
FROM KEY_INDICATORS_EXT;
```

```
hive> INSERT OVERWRITE TABLE KEY_INDICATORS_INT_ORC
> SELECT
> ID,
> STATE_NAME,
```

```

> STATE_DISTRICT_NAME,
> AA_HOUSEHOLDS_TOTAL,
> AA_POPULATION_TOTAL,
> CC_SEX_RATIO_ALL_AGES_TOTAL,
> LL_TOTAL_FERTILITY_RATE_TOTAL,
> YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON
> FROM KEY_INDICATORS_EXT;
Query ID = cloudera_20181128200303_94a8f443-bf7f-4a00-ab7b-403d1c34c8b2
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there is no reduce operator
Starting Job = job_1543462035376_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1543462035376_0006/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1543462035376_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2018-11-28 20:04:11,305 Stage-1 map = 0%, reduce = 0%
2018-11-28 20:04:23,309 Stage-1 map = 100%, reduce = 0%
MapReduce Total cumulative CPU time: 4 seconds 120 msec
Ended Job = job_1543462035376_0006
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/etlproject2.db/key_indicators_int_orc/.hive-staging_hive_2018-11-28_20-03-58_068_8785606343672482368-1/-ext-10000
Loading data to table etlproject2.key_indicators_int_orc
Table etlproject2.key_indicators_int_orc stats: [numFiles=1, numRows=284, totalSize=9392, rawDataSize=65036]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 4.12 sec HDFS Read: 1181438 HDFS Write: 9486 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 120 msec
OK
Time taken: 28.188 seconds
hive>
```

5. Screenshot of runtimes against each query given above for the default format as well as for the formats such as ORC

Note that I was not getting much difference in timings while running queries in HIVE CLI and HUE. It could be due to the fact that my cloudera VM is of just 8 GB RAM. I have used HUE to get better clarity in screenshots. For such small data, queries for both format types are taking almost same time in my cloudera VM, either I run in HIVE CLI or in HUE.

For default format:

```
select count(*) from key_indicators_int;
```

COLUMNS (1)	_c0
<input checked="" type="checkbox"/> _c0 BIGINT_TYPE	1 284

```
select state_name, count(*) from key_indicators_int group by state_name;
```

The screenshot shows the Apache Hue interface. In the top navigation bar, there are icons for Home, Query, and Jobs, along with a search bar and a cloudera link. On the left sidebar, under the 'etlproject2' section, there is a 'Tables' section containing 'key_indicators_ext', 'key_indicators_int', 'key_indicators_int_orc', and 'key_indicators_int_orc_partitioned'. Below this, there are four open queries indicated by '(4)'. The active query in the center is:

```
3| select state_name, count(*) from key_indicators_int group by state_name;
```

The query log shows the following output:

```
INFO : 2018-11-30 20:52:01,498 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.5 sec
INFO : 2018-11-30 20:52:10,026 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.04 sec
INFO : MapReduce Total cumulative CPU time: 7 seconds 40 msec
INFO : Ended Job = job_1543633035369_0031
INFO : Mapreduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.04 sec HDFS Read: 25642 HDFS Write: 120 SUCCESS
INFO : Total Mapreduce CPU Time Spent: 7 seconds 40 msec
INFO : Completed executing command(queryId=hive_20181130205151_6c309769-6e71-41c5-ad49-9655abc8b42f); Time taken: 28.826 seconds
INFO : OK
```

The results table has two columns: 'state_name' and '_c1'. The data is:

state_name	_c1
1 Assam	23
2 Bihar	37
3 Chhattisgarh	16
4 Jharkhand	18
5 Madhya Pradesh	45
6 Odisha	30
7 Rajasthan	32
8 Uttar Pradesh	70
9 Uttarakhand	13

```
select * from key_indicators_int where state_name = 'Uttar Pradesh';
```

The screenshot shows the Apache Hue interface. The layout is identical to the previous one, with the 'etlproject2' sidebar and four open queries. The active query is:

```
3| select * from key_indicators_int where state_name = 'Uttar Pradesh';
```

The query log shows the following output:

```
INFO : Concurrency mode is disabled, not creating a lock manager
INFO : Executing command(queryId=hive_20181130205454_2d8c629b-2c40-4666-b1fe-dc23b1f9675d):
select * from key_indicators_int where state_name = 'Uttar Pradesh'
INFO : Completed executing command(queryId=hive_20181130205454_2d8c629b-2c40-4666-b1fe-dc23b1f9675d); Time taken: 0.0 seconds
INFO : OK
```

The results table has eight columns: 'key_indicators_int.id', 'key_indicators_int.state_name', 'key_indicators_int.state_district_name', 'key_indicators_int.aa_households_total', 'key_indicators_int.aa_population_total', 'key_indicators_int.cc_sex_ratio_all_ages_total', 'key_indicators_int.ll_total_fertility_rate_total', and 'key_indicators_int.ly_under_five_mortality_rate_u'. The data is:

key_indicators_int.id	key_indicators_int.state_name	key_indicators_int.state_district_name	key_indicators_int.aa_households_total	key_indicators_int.aa_population_total	key_indicators_int.cc_sex_ratio_all_ages_total	key_indicators_int.ll_total_fertility_rate_total	key_indicators_int.ly_under_five_mortality_rate_u
1 202	Uttar Pradesh	Agra	20911	125614			
2 203	Uttar Pradesh	Aligarh	8844	52583			
3 204	Uttar Pradesh	Allahabad	11563	61029			
4 205	Uttar Pradesh	Ambedkar Nagar	7923	44698			
5 206	Uttar Pradesh	Auraiya	21590	107619			
6 207	Uttar Pradesh	Azamgarh	16962	103165			
7 208	Uttar Pradesh	Baghpat	15648	95759			
8 209	Uttar Pradesh	Bahraich	22906	121402			
9 210	Uttar Pradesh	Ballia	15606	87623			
10 211	Uttar Pradesh	Balrampur	7315	42016			
11 212	Uttar Pradesh	Banda	11915	59266			
12 213	Uttar Pradesh	Barabanki	11232	58722			

HUE Query Search data and saved documents... Jobs cloudera

COLUMNS (8) ▾

	key_indicators_int_id	key_indicators_int.state_name	key_indicators_int.state_district_name	key_indicators_int_aa_households_total	key_indic
13	214	Uttar Pradesh	Bareilly	13678	78492
14	215	Uttar Pradesh	Basti	8393	48055
15	216	Uttar Pradesh	Bijnor	8748	49416
16	217	Uttar Pradesh	Budhaun	8999	51993
17	218	Uttar Pradesh	Bulandshahar	10578	59473
18	219	Uttar Pradesh	Chandauli	15936	92389
19	220	Uttar Pradesh	Chitrakoot	16937	88832
20	221	Uttar Pradesh	Deoria	11641	65914
21	222	Uttar Pradesh	Etah	9054	52944
22	223	Uttar Pradesh	Etawah	16067	76793
23	224	Uttar Pradesh	Faizabad	12075	62219
24	225	Uttar Pradesh	Farrukhabad	10364	54111
25	226	Uttar Pradesh	Fatehpur	11582	60209
26	227	Uttar Pradesh	Firozabad	11053	62573
27	228	Uttar Pradesh	Gautam Buddha Nagar	16323	89498
28	229	Uttar Pradesh	Ghaziabad	20612	112985
29	230	Uttar Pradesh	Ghazipur	10337	62521
30	231	Uttar Pradesh	Gonda	14169	74324
31	232	Uttar Pradesh	Gorakhpur	17975	96497
32	233	Uttar Pradesh	Hamirpur	13042	62783
33	234	Uttar Pradesh	Hardoi	10040	52567

HUE Query Search data and saved documents... Jobs cloudera

COLUMNS (8) ▾

	key_indicators_int_id	key_indicators_int.state_name	key_indicators_int.state_district_name	key_indicators_int_aa_households_total	key_indic
34	235	Uttar Pradesh	Hathras	9779	55062
35	236	Uttar Pradesh	Jalaun	10675	53505
36	237	Uttar Pradesh	Jaunpur	7272	43285
37	238	Uttar Pradesh	Jhansi	16295	73590
38	239	Uttar Pradesh	Jyotiba Phule Nagar	6546	37927
39	240	Uttar Pradesh	Kannauj	27431	156432
40	241	Uttar Pradesh	Kanpur Dehat	10543	50626
41	242	Uttar Pradesh	Kanpur Nagar	29525	144182
42	243	Uttar Pradesh	Kaushambi	13179	67572
43	244	Uttar Pradesh	Kheri	12004	60900
44	245	Uttar Pradesh	Kushinagar	8608	48371
45	246	Uttar Pradesh	Lalitpur	8108	39529
46	247	Uttar Pradesh	Lucknow	21138	105538
47	248	Uttar Pradesh	Maharajganj	12950	68263
48	249	Uttar Pradesh	Mahoba	13461	63537
49	250	Uttar Pradesh	Mainpuri	10727	60823
50	251	Uttar Pradesh	Mathura	10406	59930
51	252	Uttar Pradesh	Mau	12606	74750
52	253	Uttar Pradesh	Meerut	12884	77688
53	254	Uttar Pradesh	Mirzapur	6709	38180
54	255	Uttar Pradesh	Moradabad	11054	66632

HUE Query Search data and saved documents... Jobs cloudera

COLUMNS (8) ▾

	key_indicators_int_id	key_indicators_int.state_name	key_indicators_int.state_district_name	key_indicators_int_aa_households_total	key_indic
55	256	Uttar Pradesh	Muzaffarnagar	12812	75749
56	257	Uttar Pradesh	Pilibhit	7773	43038
57	258	Uttar Pradesh	Pratapgarh	15695	86770
58	259	Uttar Pradesh	Rae Bareli	12981	66935
59	260	Uttar Pradesh	Rampur	11435	66460
60	261	Uttar Pradesh	Saharanpur	10259	58510
61	262	Uttar Pradesh	Sant Kabir Nagar	8028	43549
62	263	Uttar Pradesh	Sant Ravidas Nagar (Bhadoli)	8974	55736
63	264	Uttar Pradesh	Shahjahanpur	9822	55307
64	265	Uttar Pradesh	Shrawasti	7483	38131
65	266	Uttar Pradesh	Sidharthanagar	9919	56238
66	267	Uttar Pradesh	Sitapur	9323	50237
67	268	Uttar Pradesh	Sonbhadra	6838	33562
68	269	Uttar Pradesh	Sultanpur	11251	61923
69	270	Uttar Pradesh	Unnao	14128	69686
70	271	Uttar Pradesh	Varanasi	14974	86266

For formats such as ORC:

select count(*) from key_indicators_int_orc;

The screenshot shows the Hue interface with the following details:

- Project:** etlproject2
- Tables:** key_indicators_ext, key_indicators_int, key_indicators_int_orc, key_indicators_int_orc_partitioned
- Query:** use etlproject2;
3|select count(*) from key_indicators_int_orc;
- Logs:** INFO : Ended Job = job_1543633035369_0033
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.3 sec HDFS Read: 19111 HDFS Write: 4 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 300 msec
INFO : Completed executing command(queryId=hive_20181130210000_fiad577a-5d91-41d6-9415-43417c14cf60); Time taken: 28.667 seconds
INFO : OK
- Results:** Results (1) |
1 _c0
1 284

select state_name, count(*) from key_indicators_int_orc group by state_name;

The screenshot shows the Hue interface with the following details:

- Project:** etlproject2
- Tables:** key_indicators_ext, key_indicators_int, key_indicators_int_orc, key_indicators_int_orc_partitioned
- Query:** use etlproject2;
3|select state_name, count(*) from key_indicators_int_orc group by state_name;
- Logs:** INFO : Ended Job = job_1543633035369_0035
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.05 sec HDFS Read: 19601 HDFS Write: 120 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 50 msec
INFO : Completed executing command(queryId=hive_20181130210303_f4315b1e-34d8-4fa5-a8b4-eaa5c3b7852c); Time taken: 28.973 seconds
INFO : OK
- Results:** Results (9) |
state_name _c1
1 Assam 23
2 Bihar 37
3 Chhattisgarh 16
4 Jharkhand 18
5 Madhya Pradesh 45
6 Odisha 30
7 Rajasthan 32
8 Uttar Pradesh 70
9 Uttarakhand 13

select * from key_indicators_int_orc where state_name = 'Uttar Pradesh';

The screenshot shows the Hue interface with the following details:

- Project:** etlproject2
- Tables:** key_indicators_ext, key_indicators_int, key_indicators_int_orc, key_indicators_int_orc_partitioned
- Query:** 3|select * from key_indicators_int_orc where state_name = 'Uttar Pradesh';
- Logs:** INFO : Concurrency mode is disabled, not creating a lock manager
INFO : Executing command(queryId=hive_20181130210505_08706ba8-d069-40fb-ba13-891bc110cb38):
select * from key_indicators_int_orc where state_name = 'Uttar Pradesh'
INFO : Completed executing command(queryId=hive_20181130210505_08706ba8-d069-40fb-ba13-891bc110cb38); Time taken: 0.0 seconds
INFO : OK
- Results:** Results (70) |
key_indicators_int_orc.id key_indicators_int_orc.state_name key_indicators_int_orc.state_district_name key_indicators_int_orc.aa_households_total
1 202 Uttar Pradesh Agra 20911
2 203 Uttar Pradesh Aligarh 8844
3 204 Uttar Pradesh Allahabad 11563
4 205 Uttar Pradesh Ambedkar Nagar 7923
5 206 Uttar Pradesh Auraiya 21590
6 207 Uttar Pradesh Azamgarh 16962
7 208 Uttar Pradesh Baghpat 15648
8 209 Uttar Pradesh Bahraich 22906
9 210 Uttar Pradesh Ballia 15606
10 211 Uttar Pradesh Balrampur 7315
11 212 Uttar Pradesh Banda 11915
12 213 Uttar Pradesh Barabanki 11232

HUE Query Search data and saved documents... Jobs cloudera

etlproject2 Tables (4) + -

COLUMNS (8) key_indicators_int_orc.id key_indicators_int_orc.state_name key_indicators_int_orc.state_district_name key_indicators_int_orc.aa_households_total

key_indicators_int_orc.id key_indicators_int_orc.state_name key_indicators_int_orc.state_district_name key_indicators_int_orc.aa_households_total

13	214	Uttar Pradesh	Bareilly	13678
14	215	Uttar Pradesh	Basti	8393
15	216	Uttar Pradesh	Bijnor	8748
16	217	Uttar Pradesh	Budaun	8999
17	218	Uttar Pradesh	Bulandshahar	10578
18	219	Uttar Pradesh	Chandauli	15936
19	220	Uttar Pradesh	Chitrakoot	16997
20	221	Uttar Pradesh	Deoria	11641
21	222	Uttar Pradesh	Etah	9054
22	223	Uttar Pradesh	Etawah	16067
23	224	Uttar Pradesh	Faizabad	12075
24	225	Uttar Pradesh	Farrukhabad	10364
25	226	Uttar Pradesh	Fatehpur	11582
26	227	Uttar Pradesh	Firozabad	11053
27	228	Uttar Pradesh	Gautam Buddha Nagar	16323
28	229	Uttar Pradesh	Ghazababad	20612
29	230	Uttar Pradesh	Ghazipur	10337
30	231	Uttar Pradesh	Gonda	14169
31	232	Uttar Pradesh	Gorakhpur	17975
32	233	Uttar Pradesh	Hamirpur	13042
33	234	Uttar Pradesh	Hardoi	10040

HUE Query Search data and saved documents... Jobs cloudera

etlproject2 Tables (4) + -

COLUMNS (8) key_indicators_int_orc.id key_indicators_int_orc.state_name key_indicators_int_orc.state_district_name key_indicators_int_orc.aa_households_total

key_indicators_int_orc.id key_indicators_int_orc.state_name key_indicators_int_orc.state_district_name key_indicators_int_orc.aa_households_total

34	235	Uttar Pradesh	Hathras	9779
35	236	Uttar Pradesh	Jalaun	10675
36	237	Uttar Pradesh	Jaunpur	7272
37	238	Uttar Pradesh	Jhansi	16295
38	239	Uttar Pradesh	Jyotiba Phule Nagar	6546
39	240	Uttar Pradesh	Kannauj	27431
40	241	Uttar Pradesh	Kanpur Dehat	10543
41	242	Uttar Pradesh	Kanpur Nagar	29525
42	243	Uttar Pradesh	Kaushambi	13179
43	244	Uttar Pradesh	Kheri	12004
44	245	Uttar Pradesh	Kushinagar	8608
45	246	Uttar Pradesh	Lalitpur	8108
46	247	Uttar Pradesh	Lucknow	21138
47	248	Uttar Pradesh	Maharajganj	12950
48	249	Uttar Pradesh	Mahoba	13461
49	250	Uttar Pradesh	Mainpuri	10727
50	251	Uttar Pradesh	Mathura	10406
51	252	Uttar Pradesh	Mau	12606
52	253	Uttar Pradesh	Meerut	12884
53	254	Uttar Pradesh	Mirzapur	6709
54	255	Uttar Pradesh	Moradabad	11054

HUE Query Search data and saved documents... Jobs cloudera

etlproject2 Tables (4) + -

COLUMNS (8) key_indicators_int_orc.id key_indicators_int_orc.state_name key_indicators_int_orc.state_district_name key_indicators_int_orc.aa_households_total

key_indicators_int_orc.id key_indicators_int_orc.state_name key_indicators_int_orc.state_district_name key_indicators_int_orc.aa_households_total

55	256	Uttar Pradesh	Muzaffarnagar	12812
56	257	Uttar Pradesh	Pilibhit	7773
57	258	Uttar Pradesh	Pratapgarh	15695
58	259	Uttar Pradesh	Rae Bareli	12981
59	260	Uttar Pradesh	Rampur	11435
60	261	Uttar Pradesh	Saharanpur	10259
61	262	Uttar Pradesh	Sant Kabir Nagar	8028
62	263	Uttar Pradesh	Sant Ravidas Nagar (Bhadohi)	8974
63	264	Uttar Pradesh	Shahjahanpur	9822
64	265	Uttar Pradesh	Shrawasti	7483
65	266	Uttar Pradesh	Siddharthnagar	9919
66	267	Uttar Pradesh	Sitapur	9323
67	268	Uttar Pradesh	Sorbhadrak	6838
68	269	Uttar Pradesh	Sultanpur	11251
69	270	Uttar Pradesh	Unnao	14128
70	271	Uttar Pradesh	Varanasi	14974

6. Create and insert command for the partition table for analyses 1 & 2.

Setting few parameters needed for partitioned tables:

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

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

```

CREATE TABLE IF NOT EXISTS KEY_INDICATORS_INT_ORC_PARTITIONED(
ID INT,
STATE_DISTRICT_NAME VARCHAR(100),
AA_HOUSEHOLDS_TOTAL DOUBLE,
AA_POPULATION_TOTAL DOUBLE,
CC_SEX_RATIO_ALL_AGES_TOTAL DOUBLE,
LL_TOTAL_FERTILITY_RATE_TOTAL DOUBLE,
YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON DOUBLE)
PARTITIONED BY (STATE_NAME VARCHAR(100))
STORED AS ORC
TBLPROPERTIES("orc.compress"="SNAPPY");

```

```

hive> CREATE TABLE IF NOT EXISTS KEY_INDICATORS_INT_ORC_PARTITIONED(
> ID INT,
> STATE_DISTRICT_NAME VARCHAR(100),
> AA_HOUSEHOLDS_TOTAL DOUBLE,
> AA_POPULATION_TOTAL DOUBLE,
> CC_SEX_RATIO_ALL_AGES_TOTAL DOUBLE,
> LL_TOTAL_FERTILITY_RATE_TOTAL DOUBLE,
> YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON DOUBLE)
> PARTITIONED BY (STATE_NAME VARCHAR(100))
> STORED AS ORC
> TBLPROPERTIES("orc.compress"="SNAPPY");
OK
Time taken: 0.103 seconds
hive>

```

```

INSERT OVERWRITE TABLE KEY_INDICATORS_INT_ORC_PARTITIONED PARTITION(STATE_NAME)
SELECT
ID,
STATE_DISTRICT_NAME,
AA_HOUSEHOLDS_TOTAL,
AA_POPULATION_TOTAL,
CC_SEX_RATIO_ALL_AGES_TOTAL,
LL_TOTAL_FERTILITY_RATE_TOTAL,
YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON,
STATE_NAME
FROM KEY_INDICATORS_INT_ORC;

```

```

hive> INSERT OVERWRITE TABLE KEY_INDICATORS_INT_ORC_PARTITIONED PARTITION(STATE_NAME)
> SELECT
> ID,
> STATE_DISTRICT_NAME,
> AA_HOUSEHOLDS_TOTAL,
> AA_POPULATION_TOTAL,
> CC_SEX_RATIO_ALL_AGES_TOTAL,
> LL_TOTAL_FERTILITY_RATE_TOTAL,
> YY_UNDER_FIVE_MORTALITY_RATE_U5MR_TOTAL_PERSON,
> STATE_NAME
> FROM KEY_INDICATORS_INT_ORC;

```

Query ID = cloudera_20181128202121_b666e30f-21b5-49b5-8e12-05b59d9ca3de

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks is set to 0 since there is no reduce operator

Starting Job = job_1543462035376_0011, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1543462035376_0011/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1543462035376_0011

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0

2018-11-28 20:21:48,587 Stage-1 map = 0%, reduce = 0%

2018-11-28 20:22:00,958 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.73 sec

```

MapReduce Total cumulative CPU time: 5 seconds 730 msec
Ended Job = job_1543462035376_0011
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/etlproject2.db/key_indicators_int_orc_partitioned/.hive-
staging_hive_2018-11-28_20-21-36_573_4824045812001266477-1/-ext-10000
Loading data to table etlproject2.key_indicators_int_orc_partitioned partition (state_name=null)
    Time taken for load dynamic partitions : 1420
    Loading partition {state_name=Chhattisgarh}
    Loading partition {state_name=Bihar}
    Loading partition {state_name=Uttar Pradesh}
    Loading partition {state_name=Odisha}
    Loading partition {state_name=Assam}
    Loading partition {state_name=Rajasthan}
    Loading partition {state_name=Jharkhand}
    Loading partition {state_name=Madhya Pradesh}
    Loading partition {state_name=Uttarakhand}
    Time taken for adding to write entity : 4
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Assam} stats: [numFiles=1, numRows=23, totalSize=1736,
rawDataSize=3128]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Bihar} stats: [numFiles=1, numRows=37, totalSize=2138,
rawDataSize=5032]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Chhattisgarh} stats: [numFiles=1, numRows=16,
totalSize=1584, rawDataSize=2160]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Jharkhand} stats: [numFiles=1, numRows=18, totalSize=1577,
rawDataSize=2430]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Madhya Pradesh} stats: [numFiles=1, numRows=45,
totalSize=2423, rawDataSize=6030]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Odisha} stats: [numFiles=1, numRows=30, totalSize=2009,
rawDataSize=4080]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Rajasthan} stats: [numFiles=1, numRows=32, totalSize=1934,
rawDataSize=4320]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Uttar Pradesh} stats: [numFiles=1, numRows=70,
totalSize=3345, rawDataSize=9520]
Partition etlproject2.key_indicators_int_orc_partitioned{state_name=Uttarakhand} stats: [numFiles=1, numRows=13,
totalSize=1541, rawDataSize=1794]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 5.73 sec HDFS Read: 24436 HDFS Write: 19137 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 730 msec
OK
Time taken: 28.563 seconds
hive>

```

The result of each analysis along with the query and the corresponding chart generated in Hue. Keep optimizations in mind

1. The child mortality rate of Uttar Pradesh

- (a) Let's take average of child mortality rate (showing query first, then screenshot of result, followed by chart):

```

--- Select state_name, average of yy_under_five_mortality_rate_u5mr_total_person where state_name
--- is 'Uttar Pradesh' group by state_name

```

```

select state_name, avg(yy_under_five_mortality_rate_u5mr_total_person) as
avg_child_mortality_rate from key_indicators_int_orc_partitioned where state_name = 'Uttar
Pradesh' group by state_name;

```

HUE Query Search data and saved documents... Jobs cloudera

etlproject2

Tables
 key_indicators_ext
 key_indicators_int
 key_indicators_int_orc
 key_indicators_int_orc_partitioned

```
1 use etlproject2;
2
3
4
5 select state_name, avg(yy_under_five_mortality_rate_u5mr_total_person) as avg_child_mortality_rate from key_indicators_int_orc_partitioned
6 where state_name = 'Uttar Pradesh' group by state_name;
```

INFO : Stage-Stage-1: map: 1 REDUCE: 1 CUMULATIVE CPU: 0.40 SEC HDFS READ: 14152 HDFS WRITE: 52 SUCCESS
 INFO : Total MapReduce CPU Time Spent: 6 seconds 480 msec
 INFO : Completed executing command(queryId=hive_20181130192525_1232c1b0-b417-482e-a416-155889f6a319); Time taken: 25.481 seconds
 INFO : OK

job_1543633035369_0006

Query History Saved Queries Results (1)

state_name	avg_child_mortality_rate
Uttar Pradesh	90.228571428571428

HUE Query Search data and saved documents... Jobs cloudera

etlproject2

Tables
 key_indicators_ext
 key_indicators_int
 key_indicators_int_orc
 key_indicators_int_orc_partitioned

```
1 use etlproject2;
2
3
4
5 select state_name, avg(yy_under_five_mortality_rate_u5mr_total_person) as avg_child_mortality_rate from key_indicators_int_orc_partitioned
6 where state_name = 'Uttar Pradesh' group by state_name;
```

INFO : Stage-Stage-1: Map: 1 REDUCE: 1 CUMULATIVE CPU: 0.48 SEC HDFS READ: 14152 HDFS WRITE: 52 SUCCESS
 INFO : Total MapReduce CPU Time Spent: 6 seconds 480 msec
 INFO : Completed executing command(queryId=hive_20181130192525_1232c1b0-b417-482e-a416-155889f6a319); Time taken: 25.481 seconds
 INFO : OK

job_1543633035369_0006

Hive Add a name... Add a description... 24.41s etlproject2 text ?

X-AXIS: state_name
 Y-AXIS: avg_child_mortality_rate
 GROUP: Choose a column to pivot...
 LIMIT: Limit the number of results to...
 SORTING:

(b) Let's look for each district in increasing order of child mortality rate (showing query first, then screenshot of result, followed by chart):

--- Select state_district_name, yy_under_five_mortality_rate_u5mr_total_person where state_name is
 --- 'Uttar Pradesh' in increasing order by yy_under_five_mortality_rate_u5mr_total_person

```
select state_district_name, yy_under_five_mortality_rate_u5mr_total_person from
key_indicators_int_orc_partitioned where state_name = 'Uttar Pradesh'
order by yy_under_five_mortality_rate_u5mr_total_person;
```

HUE Query Search data and saved documents... Jobs cloudera

Tables (4)

key_indicators_ext
key_indicators_int
key_indicators_int_orc
key_indicators_int_orc_partitioned

Query History Saved Queries Results (70)

COLUMNS (2)

state_district_name
 yy_under_five_mortality_rate_u5mr_total_person

INFO : stage->:: map: i reduce: i cumulative CPU: 4.26 sec HDFS Read: 13241 HDFS Write: 1052 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 268 msec
INFO : Completed executing command(queryId=hive_20181108019144_49736254-b453-46c3-97d0-08a63762b988); Time taken: 23.276 seconds
INFO : OK

job_1543633035369_0003

state_district_name	yy_under_five_mortality_rate_u5mr_total_person
1 Kanpur Nagar	50
2 Lucknow	58
3 Mathura	58
4 Jhansi	59
5 Ghaziabad	59
6 Meerut	59
7 Sultanpur	66
8 Hamirpur	66
9 Agra	69
10 Baghpat	70
11 Gautam Buddha Nagar	70
12 Muzaffarnagar	71
13 Mahoba	73
14 Gorakhpur	76
15 Mainpuri	78

HUE Query Search data and saved documents... Jobs cloudera

Tables (4)

key_indicators_ext
key_indicators_int
key_indicators_int_orc
key_indicators_int_orc_partitioned

Query History Saved Queries Results (70)

COLUMNS (2)

state_district_name
 yy_under_five_mortality_rate_u5mr_total_person

state_district_name	yy_under_five_mortality_rate_u5mr_total_person
16 Ambedkar Nagar	78
17 Hathras	78
18 Bijnor	79
19 Firozabad	79
20 Moradabad	80
21 Rae Bareli	80
22 Fatehpur	81
23 Ballia	82
24 Unnao	83
25 Deoria	83
26 Aurlaya	84
27 Etawah	85
28 Mau	86
29 Etah	86
30 Rampur	86
31 Azamgarh	89
32 Bulandshahar	89
33 Varanasi	90
34 Aligarh	90
35 Pilibhit	91
36 Sant Kabir Nagar	91
37 Jaunpur	91

HUE Query Search data and saved documents... Jobs cloudera

Tables (4)

key_indicators_ext
key_indicators_int
key_indicators_int_orc
key_indicators_int_orc_partitioned

Query History Saved Queries Results (70)

COLUMNS (2)

state_district_name
 yy_under_five_mortality_rate_u5mr_total_person

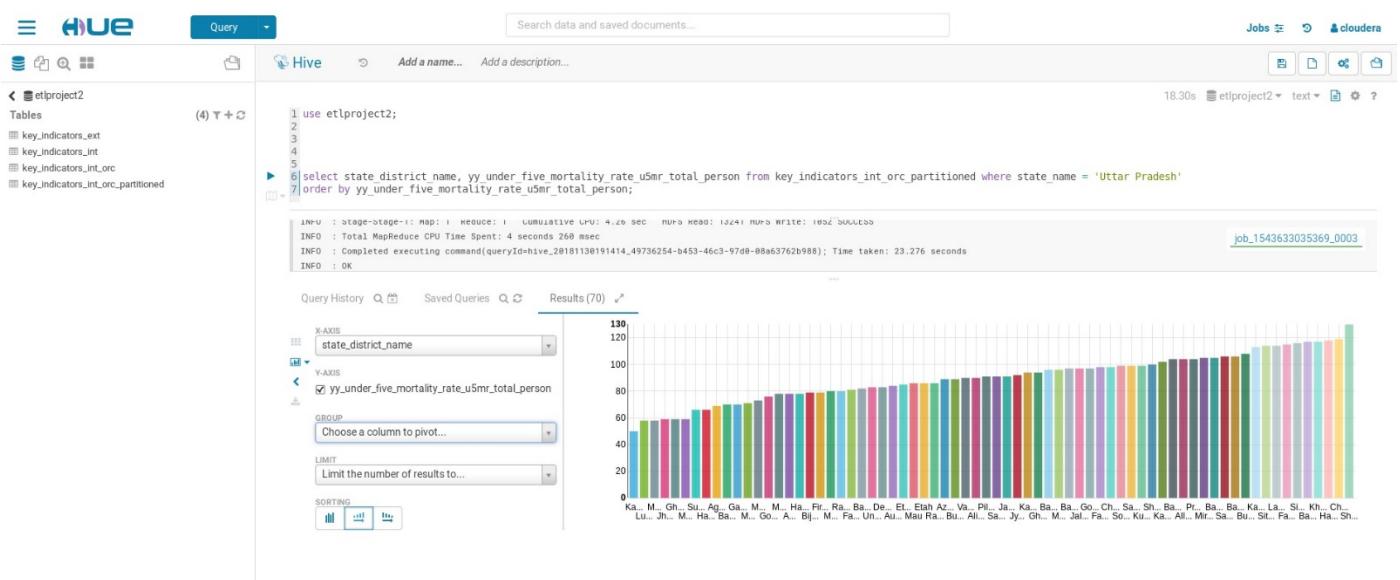
state_district_name	yy_under_five_mortality_rate_u5mr_total_person
38 Jyotiba Phule Nagar	92
39 Kanpur Dehat	94
40 Ghazipur	94
41 Banda	96
42 Maharajganj	96
43 Barabanki	97
44 Jalaun	97
45 Gonda	97
46 Farrukhabad	98
47 Chhawali	98
48 Sonbhadra	99
49 Saharanpur	99
50 Kushinagar	99
51 Shahjahanpur	100
52 Kannauj	102
53 Bareilly	104
54 Allahabad	104
55 Pratapgarh	104
56 Mirzapur	105
57 Bahraich	105
58 Sant Ravidas Nagar (Bhadoli)	106
59 Basti	106

HUE Query Search data and saved documents... Jobs cloudera

etlproject2 Tables (4) T C

state_district_name
 yy_under_five_mortality_rate_u5mr_total_person

state_district_name	yy_under_five_mortality_rate_u5mr_total_person
60 Budan	108
61 Kaushambi	113
62 Sitapur	114
63 Lalitpur	114
64 Faizabad	115
65 Siddharthnagar	116
66 Balrampur	117
67 Kheri	117
68 Hardoi	118
69 Chitrakoot	119
70 Shravasti	130



2. The fertility rate of Bihar

(a) Let's take average of fertility rate (showing query first, then screenshot of result, followed by chart):

--- Select state_name, average of ll_total_fertility_rate_total where state_name is 'Bihar' group by state_name

```
select state_name, avg(ll_total_fertility_rate_total) as avg_fertility_rate from
key_indicators_int_orc_partitioned where state_name = 'Bihar' group by state_name;
```

HUE Query Search data and saved documents... Jobs cloudera

etlproject2 Tables (4) T C

```
use etlproject2;
1
2
3
4
5 select state_name, avg(ll_total_fertility_rate_total) as avg_fertility_rate from key_indicators_int_orc_partitioned
6 where state_name = 'Bihar' group by state_name;
```

INFO : Stage->Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.91 SEC HDFS Read: 12/38 HDFS Write: 24 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 910 msec
INFO : Completed executing command(queryId=hive_20181130193131_d2f9355a-0d6d-441b-a84a-d232cdcb92f5); Time taken: 27.418 seconds
INFO : OK

job_1543633035369_0008

Query History Saved Queries Results (1)

state_name
 avg_fertility_rate

state_name	avg_fertility_rate
Bihar	3.5324324324324321

HUE

Query

Hive

Add a name... Add a description...

Search data and saved documents...

Jobs cloudera

etlproject2

Tables (4) Y + C

- key_indicators_ext
- key_indicators_int
- key_indicators_int Orc
- key_indicators_int Orc partitioned

```

1 use etlproject2;
2
3
4
5 select state_name, avg(ll_total_fertility_rate_total) as avg_fertility_rate from key_indicators_int Orc partitioned
6 where state_name = 'Bihar' group by state_name;

```

LIN-U : Stage->Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 0.91 SEC HDFS Read: 12/38 HDFS Write: 24 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 918 msec
INFO : Completed executing command(queryId=hive_20181130193131_d2f9355a-0d6d-441b-a84a-d232cdcb92f5); Time taken: 27.418 seconds
INFO : OK

job_1543633035369_0008

Query History Saved Queries Results (1)

X-AXIS state_name Y-AXIS avg_fertility_rate GROUP Choose a column to pivot... LIMIT Limit the number of results to... SORTING

3.532432

● Grouped ○ Stacked

3
2.5
2
1.5
1
0.5m
0

Bihar

- (b) Let's look for each district in increasing order of fertility rate (showing query first, then screenshot of result, followed by chart):

--- Select state_district_name, ll_total_fertility_rate_total where state_name is 'Bihar' in increasing order by ll_total_fertility_rate_total

```
select state_district_name, ll_total_fertility_rate_total from key_indicators_int Orc partitioned
where state_name = 'Bihar' order by ll_total_fertility_rate_total;
```

HUE

Query

Hive

Add a name... Add a description...

Search data and saved documents...

Jobs cloudera

etlproject2

Tables (4) Y + C

- key_indicators_ext
- key_indicators_int
- key_indicators_int Orc
- key_indicators_int Orc partitioned

```

5 select state_district_name, ll_total_fertility_rate_total from key_indicators_int Orc partitioned
6 where state_name = 'Bihar' order by ll_total_fertility_rate_total;

```

LIN-U : Stage->Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.46 sec HDFS Read: 11688 HDFS Write: 48/ SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 468 msec
INFO : Completed executing command(queryId=hive_20181130193434_e932c88f-0964-40ea-87cc-2572f8e96935); Time taken: 28.713 seconds
INFO : OK

job_1543633035369_0010

Query History Saved Queries Results (37)

COLUMNS (2)

state_district_name	ll_total_fertility_rate_total
Patna	2.6000000000000001
Bhojpur	3
Lakhisarai	3
Gaya	3
Aurangabad	3.1000000000000001
Jamui	3.1000000000000001
Jehanabad	3.1000000000000001
Navada	3.1000000000000001
Nalanda	3.1000000000000001
Kaimur (Bhabua)	3.2000000000000002
Munger	3.2000000000000002
Banka	3.2000000000000002
Saran	3.2000000000000002
Buxar	3.2000000000000002
Rohatas	3.2999999999999998

HUE Query

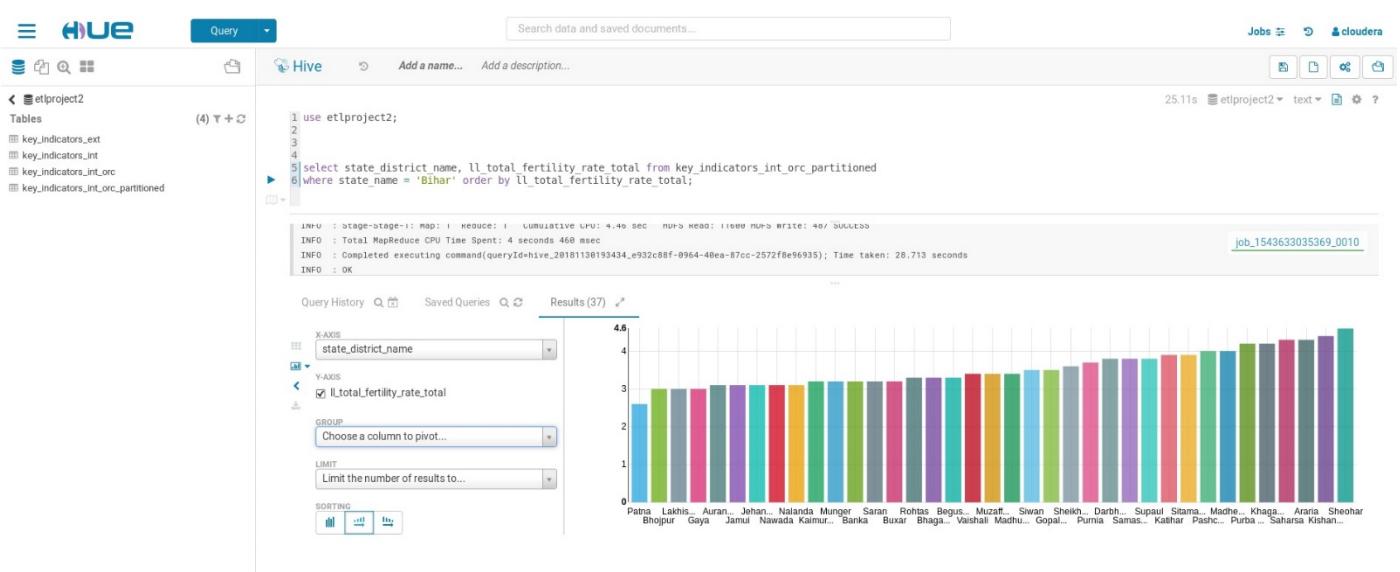
Search data and saved documents...

Jobs cloudera

etlproject2 Tables (4) ▾

COLUMNS (2) ▾

state_district_name	ll_total_fertility_rate_total
16 Bhagalpur	3.2999999999999998
17 Begusarai	3.2999999999999998
18 Vaishali	3.3999999999999999
19 Muzaffarpur	3.3999999999999999
20 Madhubani	3.3999999999999999
21 Siwan	3.5
22 Gopalganj	3.5
23 Sheikhpura	3.6000000000000001
24 Purnia	3.7000000000000002
25 Darbhanga	3.7999999999999998
26 Samastipur	3.7999999999999998
27 Supaul	3.7999999999999998
28 Kathar	3.8999999999999999
29 Sitamarhi	3.8999999999999999
30 Pashchim Champaran	4
31 Madhepura	4
32 Purba Champaran	4.2000000000000002
33 Khagaria	4.2000000000000002
34 Saharsa	4.2999999999999998
35 Araria	4.2999999999999998
36 Kishanganj	4.4000000000000004
37 Sheohar	4.5999999999999996



3. State wise child mortality rate and state wise fertility rate and does high fertility correlate with high child mortality?

(a) State wise child mortality rate (showing query first, then screenshot of result, followed by chat):

--- Select state_name, average of yy_under_five_mortality_rate_u5mr_total_person group by state_name

```
select state_name, avg(yy_under_five_mortality_rate_u5mr_total_person) as child_mortality_rate
from key_indicators_int_orc group by state_name;
```

HUE

Query

Search data and saved documents...

Jobs cloudera

Tables etlproject2

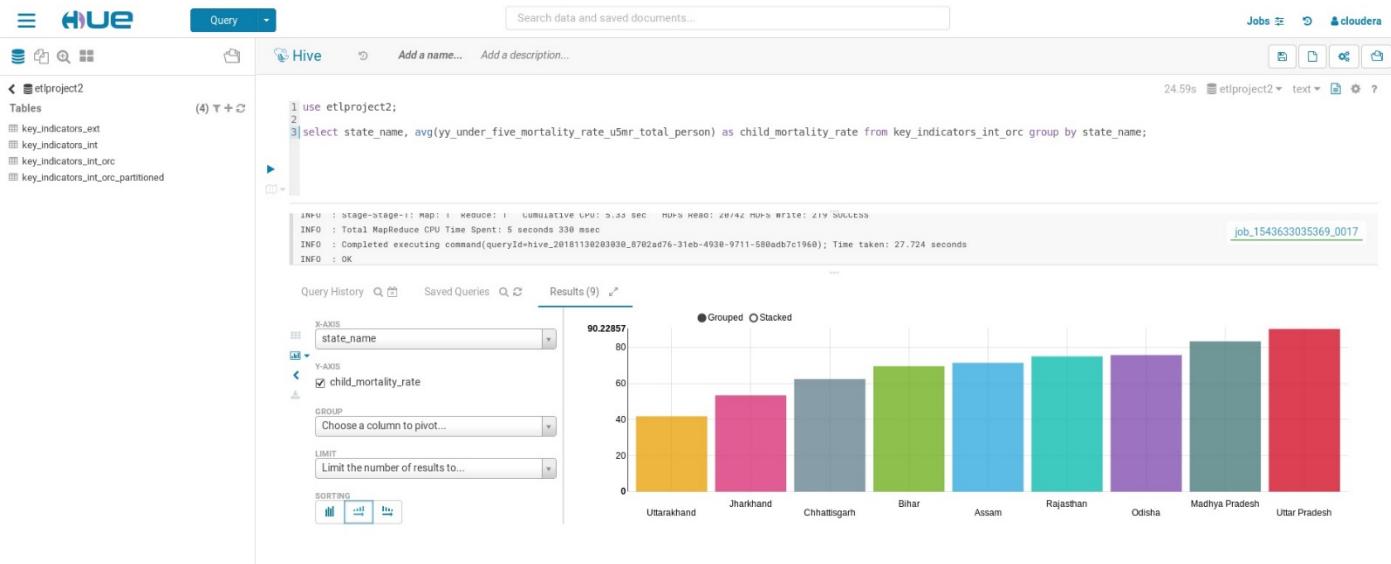
Add a name... Add a description...

```
use etlproject2;
select state_name, avg(yy_under_five_mortality_rate_u5mr_total_person) as child_mortality_rate from key_indicators_int_orc group by state_name;
```

24.59s etlproject2 text job_1543633035369_0017

Query History Saved Queries Results (9)

state_name	child_mortality_rate
Assam	71.434782608695656
Bihar	69.621621621621628
Chhattisgarh	62.5
Jharkhand	53.444444444444443
Madhya Pradesh	83.3777777777778
Odisha	75.799999999999997
Rajasthan	75.0625
Uttar Pradesh	90.228571428571428
Uttarakhand	41.846153846153847



(b) State wise fertility rate (showing query first, then screenshot of result, followed by chat):

--- Select state_name, average of ll_total_fertility_rate_total group by state_name

```
select state_name, avg(ll_total_fertility_rate_total) as fertility_rate from key_indicators_int_orc group by state_name;
```

HUE

Query

Search data and saved documents...

Jobs cloudera

etlproject2

Tables (4) Add a name... Add a description...

```

1 use etlproject2;
2
3|select state_name, avg(ll_total_fertility_rate_total) as fertility_rate from key_indicators_int_orc group by state_name;

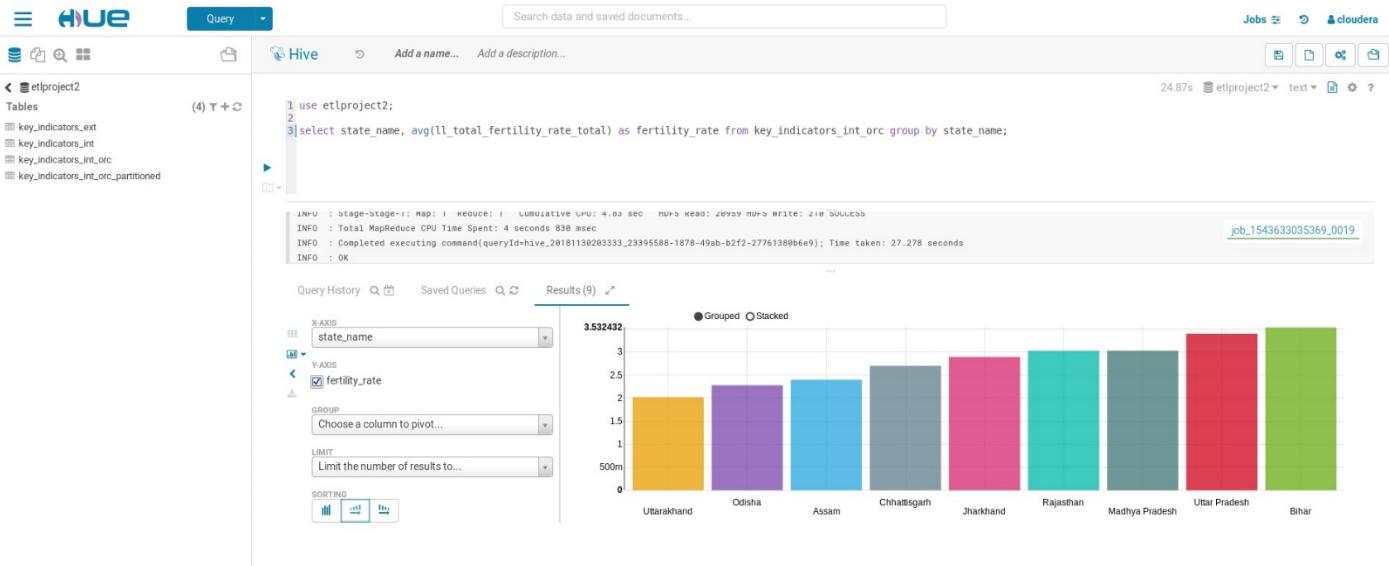
```

24.87s etlproject2 text ?

INFO : Stage->stage-1: Map: 1 REDUCE: 1 CUMULATIVE CPU: 4.65 SEC HDFS READ: 29959 HDFS WRITE: 216 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 838 msec
INFO : Completed executing command(queryId=hive_20181130203333_23395588-1878-49ab-b2f2-27761388b6e9); Time taken: 27.278 seconds
INFO : OK

Query History Saved Queries Results (9)

state_name	fertility_rate
Assam	2.3999999999999999
Bihar	3.5324324324321
Chhattisgarh	2.7012499999999999
Jharkhand	2.8944444444444448
Madhya Pradesh	3.0311111111111111
Odisha	2.2799999999999998
Rajasthan	3.0281250000000002
Uttar Pradesh	3.3978571428571427
Uttarakhand	2.0223076923076921



(c) Correlation between fertility rate and child mortality rate (showing query first, then screenshot of result, followed by chat):

(a) Let's look at state wise correlation between fertility rate and child mortality rate:

---- select state_name and compute correlation between ll_total_fertility_rate_total and
---- yy_under_five_mortality_rate_u5mr_total_person using corr() function and group by state_name

```

select state_name, corr(ll_total_fertility_rate_total,
yy_under_five_mortality_rate_u5mr_total_person) as coefficient
from key_indicators_int_orc
group by state_name;

```

HUE

Query Search data and saved documents... Jobs cloudera 46.98s etlproject2 text ?

```
1 use etlproject2;
2
3 show tables;
4
5 select state_name, corr(ll_total_fertility_rate_total, yy_under_five_mortality_rate_u5mr_total_person) as coefficient
6 from key_indicators_int_orc
7 group by state_name;
```

INFO : Ended Job = job_1543989456638_0009
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.36 sec HDFS Read: 22891 HDFS Write: 264 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 368 msec
INFO : Completed executing command(queryId=hive_20181204221313_1ade765d-abc9-4dfd-b96e-2b9e3ff21f2a); Time taken: 49.636 seconds
INFO : OK

Results (9)

	state_name	coefficient
1	Assam	0.3915829744764518
2	Bihar	0.7233339655385265
3	Chhattisgarh	0.45514212030970008
4	Jharkhand	0.79369672885119091
5	Madhya Pradesh	0.70515294385635441
6	Odisha	0.31167885766913672
7	Rajasthan	0.59922095505522754
8	Uttar Pradesh	0.62975299692871178
9	Uttarakhand	0.8430606003649156

HUE

Query Search data and saved documents... Jobs cloudera 46.98s etlproject2 text ?

```
1 use etlproject2;
2
3 show tables;
4
5 select state_name, corr(ll_total_fertility_rate_total, yy_under_five_mortality_rate_u5mr_total_person) as coefficient
6 from key_indicators_int_orc
7 group by state_name;
```

INFO : Ended Job = job_1543989456638_0009
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.36 sec HDFS Read: 22891 HDFS Write: 264 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 368 msec
INFO : Completed executing command(queryId=hive_20181204221313_1ade765d-abc9-4dfd-b96e-2b9e3ff21f2a); Time taken: 49.636 seconds
INFO : OK

Results (9)

state_name	coefficient
Odisha	343.0609
Assam	400
Chhattisgarh	400
Rajasthan	600
Uttar Pradesh	600
Madhya Pradesh	600
Bihar	600
Jharkhand	600
Uttarakhand	600

Result: Correlation coefficient is different for different state and in all cases, it indicates that there is a correlation between fertility rate and child mortality rate.

Greater than 0.5 indicates STRONG (POSITIVE) i.e. at high values
Between 0.3 and 0.5 indicates MODERATE (POSITIVE)

(b) Let's look at correlation between fertility rate and child mortality rate across all states:

---- select compute correlation between ll_total_fertility_rate_total and
---- yy_under_five_mortality_rate_u5mr_total_person using corr() function

```
select corr(ll_total_fertility_rate_total, yy_under_five_mortality_rate_u5mr_total_person) as coefficient
from key_indicators_int_orc;
```

HUE

Query →

Search data and saved documents... Add a name... Add a description...

Jobs cloudera

```
1 use etlproject2;
2
3 show tables;
4
5 select corr(ll.total_fertility_rate_total, yy_under_five_mortality_rate_u5mr_total_person) as coefficient
6 from key_indicators_int_orc;
```

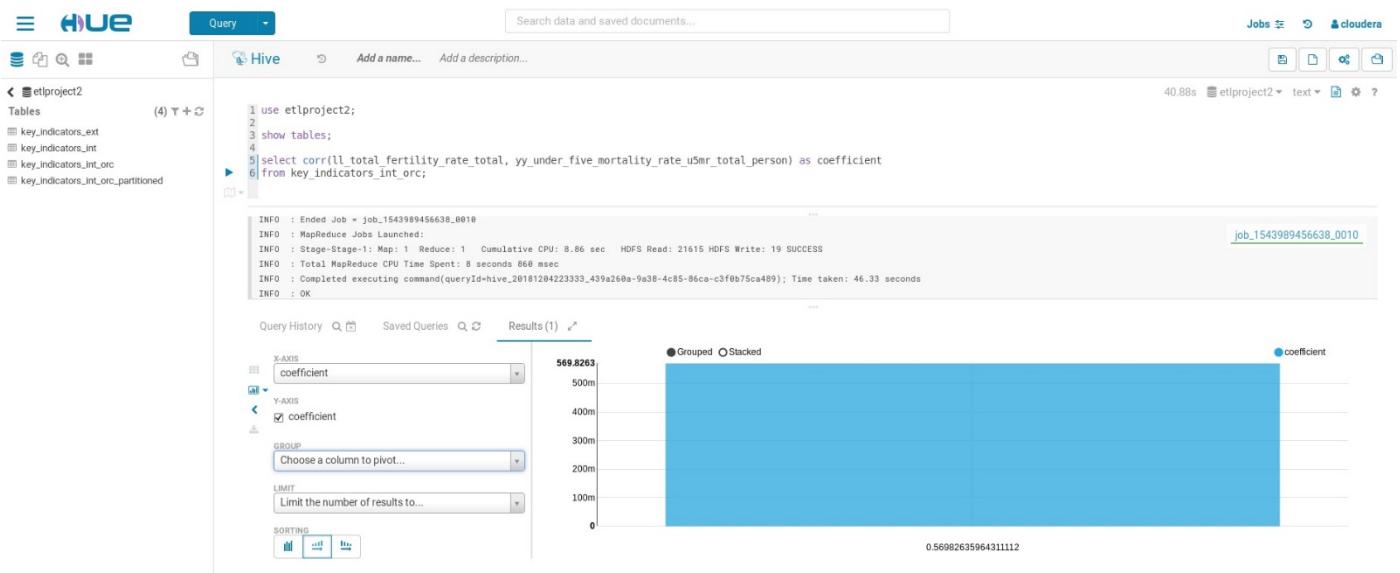
INFO : Ended Job = job_1543989456638_0010
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.86 sec HDFS Read: 21615 HDFS Write: 19 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 860 msec
INFO : Completed executing command(queryId=hive_2018120423333_439a260a-9e38-4c85-86ca-c3f0b75ca489); Time taken: 46.33 seconds
INFO : OK

Query History Saved Queries Results (1)

COLUMNS (1)	
<input checked="" type="checkbox"/> coefficient	DOUBLE_TYPE

coefficient

1 0.5698263596431112



Result: As correlation coefficient is 0.5698 which is greater than 0.5, it indicates that there is a correlation between fertility rate and child mortality rate at high values.

Greater than 0.5 indicates STRONG (POSITIVE) i.e. at high values

4. Find top 2 districts per state with the highest population per household

Showing query first, then screenshot of result, followed by chart:

--- Select state_name, state_district_name, aa_population_total divided by aa_households_total as population_per_household in the innermost query. In immediate outer query, select state_name, state_district_name, population_per_household and compute rank over partition by state_name order by population_per_household in decreasing order. In outermost query, select state_name, state_district_name, population_per_household where ranking is 1 and 2.

```

select ranking.state_name, ranking.state_district_name, ranking.population_per_household from (
select divide.state_name, divide.state_district_name, divide.population_per_household, rank()
over(partition by divide.state_name order by divide.population_per_household desc) as rk from (
select state_name, state_district_name, aa_population_total/aa_households_total as
population_per_household from key_indicators_int_orc) as divide)as ranking
where ranking.rk between 1 and 2;

```

HUE Query

Search data and saved documents...

Jobs cloudera

etlproject2 Tables (4)

```

3 select ranking.state_name, ranking.state_district_name, ranking.population_per_household from (
4 select divide.state_name, divide.state_district_name, divide.population_per_household, rank() over(partition by divide.state_name order by divide.population_per_household desc) as rk from (
5 select state_name, state_district_name, aa_population_total/aa_households_total as population_per_household from key_indicators_int_orc) as divide)as ranking
6 where ranking.rk between 1 and 2;

INFO : stage-stage-1:: map: 1 reduce: 1  cumulative CPU: 8.842 SEC  HDFS Read: 27/27  HDFS Write: 689 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 20 msec
INFO : Completed executing command(queryId=hive_20181130195858_cbf95349-8fd6-4d23-8fe7-7df4d2a95fdc), Time taken: 28.016 seconds
INFO : OK

```

job_1543633035369_0015

Query History Saved Queries Results (18)

	ranking.state_name	ranking.state_district_name	ranking.population_per_household
1	Assam	Dhemaji	5.2103445894620535
2	Assam	Marigaon	4.9784451264065472
3	Bihar	Gopalganj	5.9791953017618393
4	Bihar	Nawada	5.9449784554192906
5	Chhattisgarh	Durg	4.7164080168447322
6	Chhattisgarh	Rajnandgaon	4.6511627906976747
7	Jharkhand	Kodarma	5.8681674629524654
8	Jharkhand	Giridih	5.7871069648057656
9	Madhya Pradesh	Jhabua	5.5903925014645575
10	Madhya Pradesh	Sehore	5.366774123724636
11	Odisha	Bhadrak	4.7659507430551908
12	Odisha	Jajapur	4.4941458678393973
13	Rajasthan	Dhaulpur	5.8109722222222224
14	Rajasthan	Barmer	5.629192111322455

HUE Query

Search data and saved documents...

Jobs cloudera

etlproject2 Tables (4)

```

3 select ranking.state_name, ranking.state_district_name, ranking.population_per_household from (
4 select divide.state_name, divide.state_district_name, divide.population_per_household, rank() over(partition by divide.state_name order by divide.population_per_household desc) as rk from (
5 select state_name, state_district_name, aa_population_total/aa_households_total as population_per_household from key_indicators_int_orc) as divide)as ranking
6 where ranking.rk between 1 and 2;

INFO : stage-stage-1:: map: 1 reduce: 1  cumulative CPU: 8.842 SEC  HDFS Read: 27/27  HDFS Write: 689 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 20 msec
INFO : Completed executing command(queryId=hive_20181130195858_cbf95349-8fd6-4d23-8fe7-7df4d2a95fdc), Time taken: 28.016 seconds
INFO : OK

```

job_1543633035369_0015

Query History Saved Queries Results (18)

	ranking.state_name	ranking.state_district_name	ranking.population_per_household
1	Assam	Dhemaji	5.2103445894620535
2	Assam	Marigaon	4.9784451264065472
3	Bihar	Gopalganj	5.9791953017618393
4	Bihar	Nawada	5.9449784554192906
5	Chhattisgarh	Durg	4.7164080168447322
6	Chhattisgarh	Rajnandgaon	4.6511627906976747
7	Jharkhand	Kodarma	5.8681674629524654
8	Jharkhand	Giridih	5.7871069648057656
9	Madhya Pradesh	Jhabua	5.5903925014645575
10	Madhya Pradesh	Sehore	5.366774123724636
11	Odisha	Bhadrak	4.7659507430551908
12	Odisha	Jajapur	4.4941458678393973
13	Rajasthan	Dhaulpur	5.8109722222222224
14	Rajasthan	Barmer	5.629192111322455
15	Uttar Pradesh	Sant Ravidas Nagar (Bhadohi)	6.2108312903944727
16	Uttar Pradesh	Baghpat	6.119579959100202
17	Uttarakhand	Udham Singh Nagar	5.1164532900989546
18	Uttarakhand	Nainital	4.7489136595503494

HUE Query

Search data and saved documents...

Jobs etlproject2 text

etlproject2 Tables (4)

```

1 use etlproject2;
2
3 select ranking.state_name, ranking.state_district_name, ranking.population_per_household from (
4 select divide.state_name, divide.state_district_name, divide.population_per_household, rank() over(partition by divide.state_name order by divide.population_per_household desc) as rk from (
5 select state_name, state_district_name, aa_population_total/aa_households_total as population_per_household from key_indicators_int_orc) as divide)as ranking
6 where ranking.rk between 1 and 2;

INFO : stage-stage-1:: map: 1 reduce: 1  cumulative CPU: 0.02 SEC  HDFS Read: 27/27  HDFS Write: 689 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 20 msec
INFO : Completed executing command(queryId=hive_20181130195858_cbf95349-8fd6-4d23-8fe7-7df4d2a95fdc), Time taken: 28.016 seconds
INFO : OK

```

job_1543633035369_0015

Query History Saved Queries Results (18)

X-AXIS: ranking.state_name
Y-AXIS: ranking.population_per_household
GROUP: ranking.state_district_name
LIMIT: Limit the number of results to...
SORTING:

5. Find top 2 districts per state with the lowest sex ratios

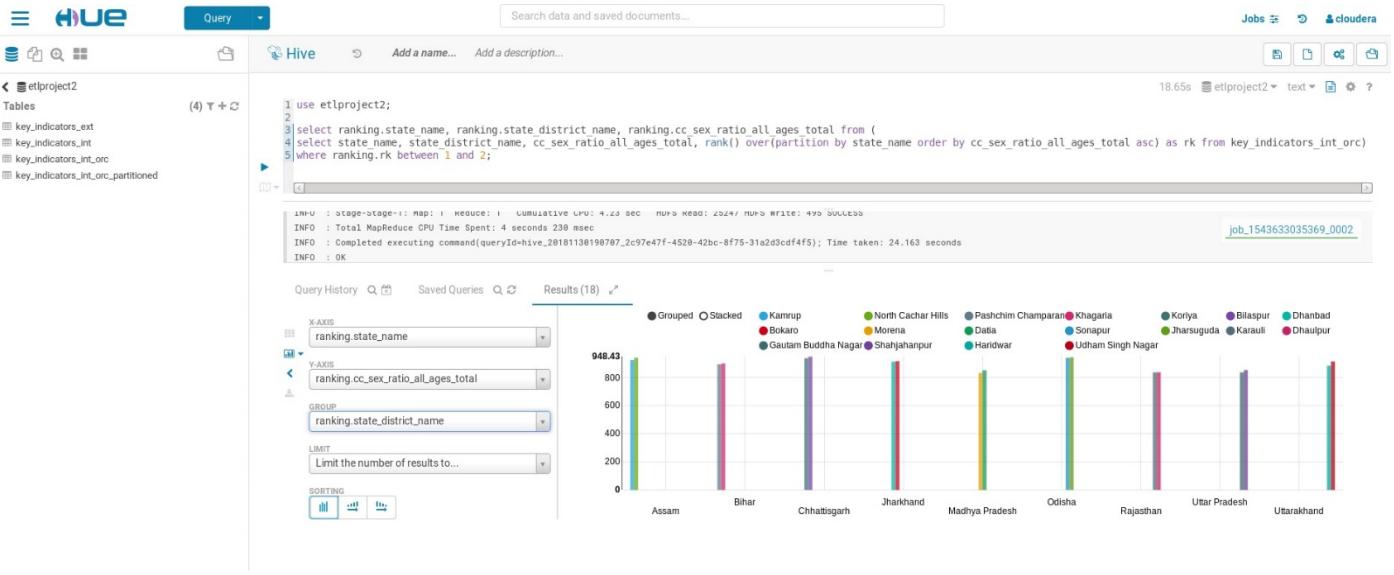
Showing query first, then screenshot of result, followed by chart:

```
--- Select state_name, state_district_name, cc_sex_ratio_all_ages_total and compute rank over partition by
--- state_name order by cc_sex_ratio_all_ages_total in increasing order, in the innermost query. In
--- outermost query, select state_name, state_district_name, cc_sex_ratio_all_ages_total where ranking is 1
--- and 2.
```

```
select ranking.state_name, ranking.state_district_name, ranking.cc_sex_ratio_all_ages_total from (
select state_name, state_district_name, cc_sex_ratio_all_ages_total, rank() over(partition by state_name
order by cc_sex_ratio_all_ages_total asc) as rk from key_indicators_int_orc) as ranking
where ranking.rk between 1 and 2;
```

ranking.state_name	ranking.state_district_name	ranking.cc_sex_ratio_all_ages_total
1 Assam	Kamrup	925
2 Assam	North Cachar Hills	941
3 Bihar	Pashchim Champaran	894
4 Bihar	Khagaria	900
5 Chhattisgarh	Korla	937.2999999999995
6 Chhattisgarh	Bilaspur	948.4299999999995
7 Jharkhand	Dhanbad	913
8 Jharkhand	Bokaro	917
9 Madhya Pradesh	Morena	833.13
10 Madhya Pradesh	Datia	852.12
11 Odisha	Sonapur	941
12 Odisha	Jharsuguda	944
13 Rajasthan	Karauli	837
14 Rajasthan	Dhaulpur	838

ranking.state_name	ranking.state_district_name	ranking.cc_sex_ratio_all_ages_total
1 Assam	Kamrup	925
2 Assam	North Cachar Hills	941
3 Bihar	Pashchim Champaran	894
4 Bihar	Khagaria	900
5 Chhattisgarh	Konya	937.2999999999995
6 Chhattisgarh	Bilaspur	948.4299999999995
7 Jharkhand	Dhanbad	913
8 Jharkhand	Bokaro	917
9 Madhya Pradesh	Morena	833.13
10 Madhya Pradesh	Datia	852.12
11 Odisha	Sonapur	941
12 Odisha	Jharsuguda	944
13 Rajasthan	Karauli	837
14 Rajasthan	Dhaulpur	838
15 Uttar Pradesh	Gautam Buddha Nagar	836.8200000000005
16 Uttar Pradesh	Shahjahanpur	851.6699999999995
17 Uttarakhand	Haridwar	884.9299999999995
18 Uttarakhand	Udham Singh Nagar	914.3099999999995



END OF DOCUMENT
