

## Data Ingestion from the RDS to HDFS using Sqoop

1. Sqoop Import command

```
sqoop import --connect jdbc:mysql://upgradawsrds.cpclxrkdvwzmz.us-east-1.rds.amazonaws.com/indiaahs2012_13 --username upgraduser --password upgraduser --table Key_indicator_districtwise --target-dir /IndiaAHSCaseStudy/input/Key_indicator_districtwise
```

2. Command to see the list of imported

```
hadoop fs -cat /IndiaAHSCaseStudy/input/Key_indicator_districtwise/part*
```

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

1. Command to create the external

```
create external table if not exists IAHSInputData (
ID int,
State_Name string,
State_District_Name string,
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\_Infection\_Person\_Total double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_Person\_Rural double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_Person\_Urban double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_Male\_Total double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_Male\_Rural double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_Male\_Urban double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_Female\_Total double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_Female\_Rural double,  
JJ\_Acute\_Illness\_Per\_100000\_Respiratory\_Infection\_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\_III\_Per\_100000\_Diabetes\_Person\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Person\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Person\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Male\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Male\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Male\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Female\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Female\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Diabetes\_Female\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Person\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Person\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Person\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Male\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Male\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Male\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Female\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Female\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Hypertension\_Female\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Tb\_Person\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Tb\_Person\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Tb\_Person\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Tb\_Male\_Total double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Tb\_Male\_Rural double,  
KK\_Diag\_For\_Chronic\_III\_Per\_100000\_Tb\_Male\_Urban double,  
KK\_Diag\_For\_Chronic\_III\_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\_III\_And\_Getting-Regular\_Treatment\_Govt\_Person\_Total double,  
KK\_Chronic\_III\_And\_Getting-Regular\_Treatment\_Govt\_Person\_Rural double,  
KK\_Chronic\_III\_And\_Getting-Regular\_Treatment\_Govt\_Person\_Urban double,  
KK\_Chronic\_III\_And\_Getting-Regular\_Treatment\_Govt\_Male\_Total double,  
KK\_Chronic\_III\_And\_Getting-Regular\_Treatment\_Govt\_Male\_Rural double,  
KK\_Chronic\_III\_And\_Getting-Regular\_Treatment\_Govt\_Male\_Urban double,

KK\_Chronic\_III\_And\_Getting-Regular\_Treatment\_Govt\_Female\_Total double,  
KK\_Chronic\_III\_And\_Getting-Regular\_Treatment\_Govt\_Female\_Rural double,  
KK\_Chronic\_III\_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\_Reported\_Birth\_Of\_Order\_2\_Above\_Total double,  
LL\_Women\_20\_24\_Reported\_Birth\_Of\_Order\_2\_Above\_Rural double,  
LL\_Women\_20\_24\_Reported\_Birth\_Of\_Order\_2\_Above\_Urban double,  
LL\_Women\_Reported\_Birth\_Of\_Order\_3\_Above\_Total double,  
LL\_Women\_Reported\_Birth\_Of\_Order\_3\_Above\_Rural double,  
LL\_Women\_Reported\_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\_Antenatal\_Check\_Up\_Total double,  
PP\_Mothers\_Who\_Received\_Any\_Antenatal\_Check\_Up\_Rural double,  
PP\_Mothers\_Who\_Received\_Any\_Antenatal\_Check\_Up\_Urban double,  
PP\_Mothers\_Who\_Had\_Antenatal\_Check\_Up\_In\_First\_Trimester\_Total double,  
PP\_Mothers\_Who\_Had\_Antenatal\_Check\_Up\_In\_First\_Trimester\_Rural double,  
PP\_Mothers\_Who\_Had\_Antenatal\_Check\_Up\_In\_First\_Trimester\_Urban double,  
PP\_Mothers\_Who\_Received\_3\_Or\_More\_Antenatal\_Care\_Total double,  
PP\_Mothers\_Who\_Received\_3\_Or\_More\_Antenatal\_Care\_Rural double,  
PP\_Mothers\_Who\_Received\_3\_Or\_More\_Antenatal\_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\_Govt\_Source\_Total double,  
PP\_Mothers\_Who\_Received\_Anc\_From\_Govt\_Source\_Rural double,  
PP\_Mothers\_Who\_Received\_Anc\_From\_Govt\_Source\_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\_Assis\_For\_Inst\_Delivery\_Under\_Jsy\_Total double,  
SS\_Availed\_Financial\_Assis\_For\_Inst\_Delivery\_Under\_Jsy\_Rural double,  
SS\_Availed\_Financial\_Assis\_For\_Inst\_Delivery\_Under\_Jsy\_Urban double,  
SS\_Availed\_Financial\_Assis\_For\_Govt\_Delivery\_Under\_Jsy\_Total double,  
SS\_Availed\_Financial\_Assis\_For\_Govt\_Delivery\_Under\_Jsy\_Rural double,  
SS\_Availed\_Financial\_Assis\_For\_Govt\_Delivery\_Under\_Jsy\_Urban double,  
TT\_Children\_Aged\_12\_23\_Months\_Having\_Immunization\_Card\_Total double,  
TT\_Children\_Aged\_12\_23\_Months\_Having\_Immunization\_Card\_Rural double,  
TT\_Children\_Aged\_12\_23\_Months\_Having\_Immunization\_Card\_Urban double,  
TT\_Children\_Aged\_12\_23\_Months\_Who\_Have\_Received\_Bcg\_Total double,  
TT\_Children\_Aged\_12\_23\_Months\_Who\_Have\_Received\_Bcg\_Rural double,  
TT\_Children\_Aged\_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\_Children\_Aged\_12\_23\_Months\_Received\_Measles\_Vaccine\_Total double,  
TT\_Children\_Aged\_12\_23\_Months\_Received\_Measles\_Vaccine\_Rural double,  
TT\_Children\_Aged\_12\_23\_Months\_Received\_Measles\_Vaccine\_Urban double,  
TT\_Children\_Aged\_12\_23\_Months\_Fully\_Immunized\_Total double,  
TT\_Children\_Aged\_12\_23\_Months\_Fully\_Immunized\_Rural double,  
TT\_Children\_Aged\_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_Neo_Natal_Mortality_Rate_Total double,
YY_Neo_Natal_Mortality_Rate_Rural double,
YY_Neo_Natal_Mortality_Rate_Urban double,
YY_Post_Neo_Natal_Mortality_Rate_Total double,
YY_Post_Neo_Natal_Mortality_Rate_Rural double,
YY_Post_Neo_Natal_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 ','
location '/IndiaAHSCaseStudy/Hive/inputdata';

```

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

```

load data inpath '/IndiaAHSCaseStudy/input/Key_indicator_districtwise/part*' overwrite
into table IAHSInputData;

```

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

Query:

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

MySQL Workbench:

The screenshot shows the MySQL Workbench interface with the following details:

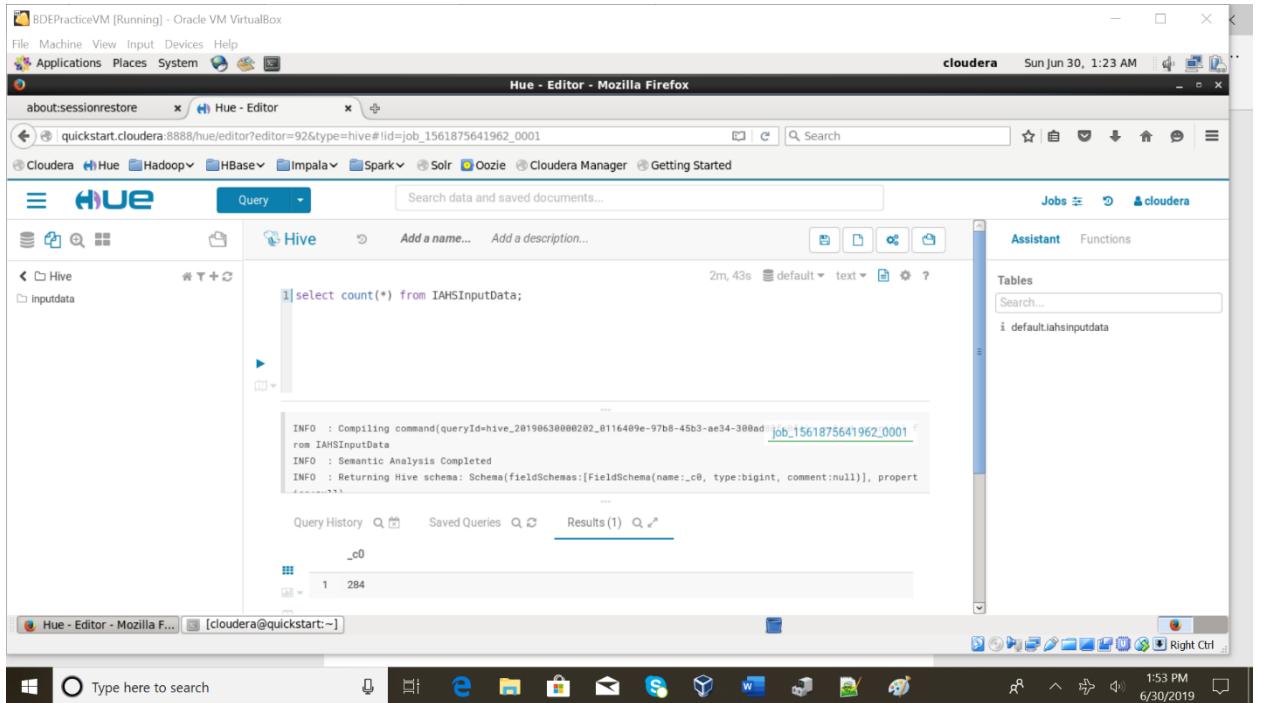
- Query Editor:** Contains the SQL command: `select count(*) from Key_indicator_districtwise;`
- Result Grid:** Displays the result of the query:
 

count(*)
284
- Output Panel:** Shows the execution log:
 

#	Time	Action
1	14.02.10	select count(*) from Key_indicator_districtwise LIMIT 0, 1000

 Message: 1 row(s) returned
 Duration / Fetch: 0.672 sec / 0.000 sec
- Navigator:** Shows the database structure, including MANAGEMENT, INSTANCE, PERFORMANCE, and Administration sections.
- SQLAdditions:** A note stating: "Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help."

## Screenshot of data:



- 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

Query:

```
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 IAHSInputData limit 10;
```

## MySQL Workbench:

The screenshot shows the MySQL Workbench interface. A query window titled "Query 1" displays the following SQL code:

```

1 • select ID,State_Name,State_District_Name,AA_Sample_Units_Total,AA_Sample_Units_Rural,AA_Sample_Units_Urban,
2 AA_Households_Total,AA_Households_Rural from Key_indicator_districtwise limit 0,10;
3

```

The results grid shows data for 10 rows, corresponding to the 10 districts listed in the query. The columns are:

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

The right side of the interface includes a "SQLAdditions" panel and a status bar indicating "Automatic context help is disabled". The output pane shows the execution log:

```

Key_indicator_districtwise 8 < />
Output:
Action Output
# Time Action
1 14:13:29 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 0,10
Message: 10 row(s) returned
Duration / Fetch: 0.500 sec / 0.000 sec

```

## Screenshot of data:

The screenshot shows the Hue interface running on a Cloudera VM. A browser window titled "Hue - Editor - Mozilla Firefox" displays the following query results:

```

about:sessionrestore < /> Hue - Editor < />
quickstart.cloudera:8888/hue/editor?editor=94#lid=job_1561875641962_0001
Cloudera Sun Jun 30, 1:54 AM
Hue - Editor - Mozilla Firefox
about:sessionrestore < /> Hue - Editor < />
quickstart.cloudera:8888/hue/editor?editor=94#lid=job_1561875641962_0001
Search... Cloudera Getting Started
Hive Add a name... Add a description...
Search data and saved documents...
1|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 IAHSInputData limit 10;
INFO : Executing command(queryId:hive_20190630014949_0fd4855-6334-4758-aef9-97f95f41698c); Time taken: 0.048 seconds
INFO : OK
Query History Q Saved Queries Q Results (10) Q v*
id state_name state_district_name aa_sample_units_total aa_sample_units_rural aa_sample_units_urba...
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

```

The results grid is identical to the one shown in the MySQL Workbench screenshot. The interface includes a navigation bar, a search bar, and various tabs for different databases and components.

## Subset schema creation in Hive to support the analyses

1. Columns used in the subset schema

```
ID , State_Name , State_District_Name, BB_Average_Household_Size_All_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
```

3. Create and insert command for the default

```
Create external table if not exists IAHSInputSubset_DefaultFormat (  
ID int,  
State_Name string,  
State_District_Name string,  
BB_Average_Household_Size_All_Total double,  
CC_Sex_Ratio_All_Ages_Total double,  
LL_Total_Fertility_Rate_Total double,  
YY_Under_Five_Mortality_Rate_U5MR_Total_Person double)  
location '/IndiaAHSCaseStudy/Hive/SubsetSchema_DefaultFormat/';
```

```
insert into IAHSInputSubset_DefaultFormat  
select ID , State_Name , State_District_Name, BB_Average_Household_Size_All_Total ,  
CC_Sex_Ratio_All_Ages_Total , LL_Total_Fertility_Rate_Total ,  
YY_Under_Five_Mortality_Rate_U5MR_Total_Person from IAHSInputData;
```

4. Create and insert command for the formats such as

```
Create external table if not exists IAHSInputSubset_ORCFormat (  
ID int,  
State_Name string,  
State_District_Name string,  
BB_Average_Household_Size_All_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  
location '/IndiaAHSCaseStudy/Hive/SubsetSchema_ORCFormat/'  
tblproperties ("orc.compress"="SNAPPY");
```

```
insert into IAHSInputSubset_ORCFormat
```

```
select ID, State_Name, State_District_Name, BB_Average_Household_Size_All_Total,  
CC_Sex_Ratio_All_Ages_Total, LL_Total_Fertility_Rate_Total,  
YY_Under_Five_Mortality_Rate_U5MR_Total_Person from IAHSInputData;
```

5. Create and insert command for the Hive-HBase integrated

```
create external table IAHS_hbase ( ID int, State_Name string, State_District_Name  
string, BB_Average_Household_Size_All_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 BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'  
WITH SERDEPROPERTIES ("hbase.columns.mapping" = ":key, location:state,  
location:district, counts:houseHoldSize, counts:sexRatio, counts:fertilityRate,  
counts:childMortalityRate")  
TBLPROPERTIES ("hbase.table.name" = "IAHS_hive");
```

```
insert overwrite table IAHS_hbase  
select ID, State_Name, State_District_Name, BB_Average_Household_Size_All_Total,  
CC_Sex_Ratio_All_Ages_Total, LL_Total_Fertility_Rate_Total,  
YY_Under_Five_Mortality_Rate_U5MR_Total_Person from  
IAHSInputSubset_ORCFormat;
```

6. Screenshot of runtimes against each query given above for the default format, formats such as ORC format as well as Hive-Hbase integration

- a. For default format:

Screenshot of run time for query 1

2m, 13s default text ?

```
1|select count(*) from IAHSInputSubset_DefaultFormat;
2|
```

INFO : Launching Job 1 out of 1  
INFO : Starting task [Stage-1:MAPRED] in serial mode  
INFO : Number of reduce tasks determined at compile time: 1  
INFO : In order to change the average load for a reducer (in bytes):  
INFO : set hive.exec.reducers.bytes.per.reducer=<number>  
INFO : In order to limit the maximum number of reducers:  
INFO : set hive.exec.reducers.max=<number>  
INFO : In order to set a constant number of reducers:  
INFO : set mapreduce.job.reduces=<number>  
INFO : number of splits:  
INFO : Submitting tokens for job: job\_1562409531409\_0003  
INFO : The url to track the job: http://quickstart.cloudera:8080/proxy/application\_1562409531409\_0003/  
INFO : Starting Job = job\_1562409531409\_0003, Tracking URL = http://quickstart.cloudera:8080/proxy/application\_1562409531409\_0003/  
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job\_1562409531409\_0003  
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
INFO : 2019-07-08 04:08:04,389 Stage-1 map = 0%, reduce = 0%  
INFO : 2019-07-08 04:08:04,389 Stage-1 map = 0%, reduce = 0%  
INFO : 2019-07-08 04:08:19,438 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.02 sec  
INFO : 2019-07-08 04:08:47,188 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.37 sec  
INFO : MapReduce Total cumulative CPU time: 8 seconds 370 msec  
INFO : Ended Job = job\_1562409531409\_0003  
INFO : Ended Job = job\_1562409531409\_0003  
INFO : MapReduce Jobs Launched:  
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.37 sec HDFS Read: 24892 HDFS Write: 4 SUCCESS  
INFO : Total MapReduce CPU Time Spent: 8 seconds 370 msec  
INFO : Completed executing command(queryId=hive\_20190708040707\_51clec97-dae6-485f-90ca-fcbdd1951d4e); Time taken: 134.324 seconds  
INFO : OK

## Screenshot of run time for query 2

1m, 50s default text ?

HUE Query Search data and saved documents...

```
1|select State_Name, count(*) from IAHSInputSubset_DefaultFormat group by State_Name;
```

INFO : Total jobs = 1  
INFO : Launching Job 1 out of 1  
INFO : Starting task [Stage-1:MAPRED] in serial mode  
INFO : Number of reduce tasks not specified. Estimated from input data size: 1  
INFO : In order to change the average load for a reducer (in bytes):  
INFO : set hive.exec.reducers.bytes.per.reducer=<number>  
INFO : In order to limit the maximum number of reducers:  
INFO : set hive.exec.reducers.max=<number>  
INFO : In order to set a constant number of reducers:  
INFO : set mapreduce.job.reduces=<number>  
INFO : number of splits:  
INFO : Submitting tokens for job: job\_1562409531409\_0005  
INFO : The url to track the job: http://quickstart.cloudera:8080/proxy/application\_1562409531409\_0005/  
INFO : Starting Job = job\_1562409531409\_0005, Tracking URL = http://quickstart.cloudera:8080/proxy/application\_1562409531409\_0005/  
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job\_1562409531409\_0005  
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
INFO : 2019-07-08 04:28:33,894 Stage-1 map = 0%, reduce = 0%  
INFO : 2019-07-08 04:29:15,408 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.04 sec  
INFO : 2019-07-08 04:29:45,220 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 0.78 sec  
INFO : MapReduce Total cumulative CPU time: 0 seconds 780 msec  
INFO : Ended Job = job\_1562409531409\_0005  
INFO : MapReduce Jobs Launched:  
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 0.78 sec HDFS Read: 25331 HDFS Write: 120 SUCCESS  
INFO : Total MapReduce CPU Time Spent: 0 seconds 780 msec  
INFO : Completed executing command(queryId=hive\_20190708042727\_3cf47779-4f08-43a9-a558-537c2b14908d); Time taken: 112.4 seconds  
INFO : OK

## Screenshot of run time for query 3

```
1|select * from IAHSInputSubset_DefaultFormat where State_Name = 'Uttar Pradesh';
```

```
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:iahsinputsubset_defaultformat.id, type:int, comment:null), FieldSchema(name:iahsinputsubset_defaultformat.state_name, type:string, comment:null), FieldSchema(name:iahsinputsubset_defaultformat.state_district_name, type:string, comment:null), FieldSchema(name:iahsinputsubset_defaultformat.as_households_total, type:double, comment:null), FieldSchema(name:iahsinputsubset_defaultformat.as_population_total, type:double, comment:null), FieldSchema(name:iahsinputsubset_defaultformat.cc_sex_ratio_all_ages_total1, type:double, comment:null), FieldSchema(name:iahsinputsubset_defaultformat.ll_total1_fertility_rate_total, type:double, comment:null), FieldSchema(name:iahsinputsubset_defaultformat.yy_infant_mortality_rate_lmr_total_person, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId:hive_20190700043232_cb7c0c40-f031-4081-8452-7303bd14ed7a); Time taken: 0.513 seconds
INFO : Executing command(queryId:hive_20190700043232_cb7c0c40-f031-4081-8452-7303bd14ed7a): select * from IAHSInputSubset_DefaultFormat where State_Name = 'Uttar Pradesh'
INFO : Query ID = hive_20190700043232_c07c0c40-f031-4081-8452-7303bd14ed7a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks is set to 0 since there's no reduce operator
INFO : number of splits:1
INFO : Submitting tokens for job: job_1562409531409_0000
INFO : The url to track the job: http://quickstart.cloudera:8088/proxy/application_1562409531409_0000/
INFO : Starting Job = job_1562409531409_0000. Tracking URL = http://quickstart.cloudera:8088/proxy/application_1562409531409_0000/
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -xkill job_1562409531409_0000
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
INFO : 2019-07-00 04:32:28,487 Stage-1 map = 0%, reduce = 0%
INFO : 2019-07-00 04:32:51,702 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.79 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 790 msec
INFO : Ended Job = job_1562409531409_0000
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Cumulative CPU: 4.79 sec  HDFS Read: 21748 HDFS Write: 4387 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 790 msec
INFO : Completed executing command(queryId:hive_20190700043232_cb7c0c40-f031-4081-8452-7303bd14ed7a); Time taken: 48.591 seconds
INFO : OK
```

b. For formats such as ORC:

Screenshot of run time for query 1

The screenshot shows the Hue interface with the 'Hive' tab selected. A search bar at the top right contains the placeholder 'Search data and saved documents...'. Below the search bar, there are buttons for 'Query' and 'Add a name...' and 'Add a description...'. On the far right, there are icons for file operations and a help menu. The main area displays a command-line interface (CLI) window. The command entered is '1| select count(\*) from IAHSInputSubset\_ORCFormat;'. The output of the query is a log of Hadoop/Hive job execution. The log includes various INFO messages about the job's progress, such as launching tasks, starting the job, and tracking URLs. It also shows the kill command being issued to stop the job. The log concludes with a message indicating the job has completed successfully.

```
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1502409531409_0007
INFO : The url to track the job: http://quickstart.cloudera:8088/proxy/application_1502409531409_0007/
INFO : Starting Job = job_1502409531409_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1502409531409_0007/
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1502409531409_0007
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-07-08 04:37:19,493 Stage-1 map = 0%, reduce = 0%
INFO : 2019-07-08 04:37:44,573 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.93 sec
INFO : 2019-07-08 04:38:09,235 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.55 sec
INFO : MapReduce Total cumulative CPU time: 7 seconds 550 msec
INFO : Ended Job = job_1502409531409_0007
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.55 sec HDFS Read: 30140 HDFS Write: 4 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 seconds 550 msec
INFO : Completed executing command(queryId=hive_20190706043036_a05ca639-99c5-4988-9e0a-cf39598d112b); Time taken: 75.84 seconds
TMRN : OK
```

Screenshot of run time for query 2

HUE    Query    Search data and saved documents...

Hive    Add a name...    Add a description...    ?

```
1|select State_Name, count(*) from IAHSInputSubset_ORCFormat group by State_Name;

INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks not specified. Estimated from input data size: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   see hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1562409531409_0008
INFO : The url to track the job: http://quickstart.cloudera:8080/proxy/application_1562409531409_0008/
INFO : Starting Job = job_1562409531409_0008, Tracking URL = http://quickstart.cloudera:8080/proxy/application_1562409531409_0008/
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1562409531409_0008
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-07-08 04:40:52,524 Stage-1 map = 0%, reduce = 0%
INFO : 2019-07-08 04:41:11,791 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.96 sec
INFO : 2019-07-08 04:41:44,487 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.62 sec
INFO : MapReduce Total cumulative CPU time: 8 seconds 028 msec
INFO : Ended Job = job_1562409531409_0008
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.62 sec HDFS Read: 31273 HDFS Write: 121 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 028 msec
INFO : Completed executing command(queryId=hive_20190708044408_a9892ec-b05d-4b03-92fa-ae155d70fc07); Time taken: 79.147 seconds
INFO : OK
```

### Screenshot of run time for query 3

HUE    Query    Search data and saved documents...

49.58s    default    text    ?

```
1|select * from IAHSInputSubset_ORCFormat where State_Name = 'Uttar Pradesh';

INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:iahsinputsubset_orcformat.id, type:int, comment:null), FieldSchema(name:iahsinputsubset_orcformat.state_name, type:string, comment:null), FieldSchema(name:iahsinputsubset_orcformat.households_total, type:double, comment:null), FieldSchema(name:iahsinputsubset_orcformat.population_total, type:double, comment:null), FieldSchema(name:iahsinputsubset_orcformat.ll_total_fertility_rate_total, type:double, comment:null), FieldSchema(name:iahsinputsubset_orcformat.yy_infant_mortality_rate, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20190708044343_32c09cfe-5f53-495c-9c89-0c179a96112c); Time taken: 0.27 seconds
INFO : Executing command(queryId=hive_20190708044343_32c09cfe-5f53-495c-9c89-0c179a96112c): select * from IAHSInputSubset_ORCFormat where State_Name = 'Uttar Pradesh'
INFO : Query ID = hive_20190708044343_32c09cfe-5f53-495c-9c89-0c179a96112c
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks is set to 0 since there's no reduce operator
INFO : number of splits:1
INFO : Submitting tokens for job: job_1562409531409_0009
INFO : The url to track the job: http://quickstart.cloudera:8080/proxy/application_1562409531409_0009/
INFO : Starting Job = job_1562409531409_0009, Tracking URL = http://quickstart.cloudera:8080/proxy/application_1562409531409_0009/
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1562409531409_0009
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
INFO : 2019-07-08 04:44:10,070 Stage-1 map = 0%, reduce = 0%
INFO : 2019-07-08 04:44:47,314 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.46 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 408 msec
INFO : Ended Job = job_1562409531409_0009
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Cumulative CPU: 4.46 sec HDFS Read: 35180 HDFS Write: 8614 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 408 msec
INFO : Completed executing command(queryId=hive_20190708044343_32c09cfe-5f53-495c-9c89-0c179a96112c); Time taken: 58.81 seconds
INFO : OK
```

### c. For Hive-Hbase Integrated Table:

## Screenshot of run time for query 1

The screenshot shows the Apache Hive interface. At the top, there's a navigation bar with 'HUE' and a dropdown menu labeled 'Query'. Below the navigation bar is a search bar with the placeholder 'Search data and saved documents...'. On the right side of the interface, there are some status indicators: '1m, 31s', 'default', 'text', and a few icons. The main area contains a code editor with the following SQL query:

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

Below the code editor is a large text block displaying the execution log for the job. The log starts with 'INFO : Total jobs = 1' and ends with 'INFO : OK'. It details the job's progress, including the start time (2019-07-06 04:49:41), the number of mappers (108), and reducers (1). It also shows the cumulative CPU time (0.98 sec) and the total map/reduce time (11.19 sec). The log concludes with 'INFO : Completed executing command(queryId=hive\_20190706044488\_1f533100-8cbd-4b37-9bff-f0a3d2bfe957); Time taken: 94.794 seconds'.

## Screenshot of run time for query 2

HUE      Query      Search data and saved documents...      2m, 12s default text ?

```
1|select State_Name, count(*) from iahs_hbase group by State_Name;
```

```

INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks not specified. Estimated from input data size: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1562409531409_0011
INFO : The url to track the job: http://quickstart.cloudera:8080/proxy/application_1562409531409_0011/
INFO : Starting Job = job_1562409531409_0011, Tracking URL = http://quickstart.cloudera:8080/proxy/application_1562409531409_0011
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1562409531409_0011
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-07-06 04:33:46,970 Stage-1 map = 0%,  reduce = 0%
INFO : 2019-07-06 04:34:41,338 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.98 sec
INFO : 2019-07-06 04:35:06,930 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 10.51 sec
INFO : MapReduce Total cumulative CPU time: 10 seconds 520 msec
INFO : Ended Job = job_1562409531409_0011
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 10.52 sec  HDFS Read: 17550 HDFS Write: 128 SUCCESS
INFO : Total MapReduce CPU Time Spent: 10 seconds 520 msec
INFO : Completed executing command(queryId:hive_20190708045252_3d3f454e-682c-4e77-af6-7412f6852453); Time taken: 133.434 seconds
INFO : OK

```

### Screenshot of run time for query 3

HUE      Query      Search data and saved documents...      1m, 33s default text ?

```
1|select * from iahs_hbase where State_Name = 'Uttar Pradesh';
```

```

INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema[fieldSchemas:[FieldSchema(name:iahs_hbase.id, type:int, comment:null), FieldSchema(name:iahs_hbase.state_name, type:string, comment:null), FieldSchema(name:iahs_hbase.state_district_name, type:string, comment:null), FieldSchema(name:iahs_hbase.ii, type:double, comment:null), FieldSchema(name:iahs_hbase.population_total, type:double, comment:null), FieldSchema(name:iahs_hbase.cc.sex_ratio_all_ages_total, type:double, comment:null), FieldSchema(name:iahs_hbase.ll.total_fertility_rate_total, type:double, comment:null), FieldSchema(name:iahs_hbase.yy.infant_mortality_rate_imr_total,person, type:double, comment:null)], properties:null]
INFO : Completed compiling command(queryId:hive_20190708045858_e4932280-257e-4b07-8141-30432c7087f7); Time taken: 0.364 seconds
INFO : Executing command(queryId:hive_20190708045858_e4932280-257e-4b07-8141-30432c7087f7); select * from iahs_hbase where State_Name = 'Uttar Pradesh'
INFO : Query ID = hive_20190708045858_e4932280-257e-4b07-8141-30432c7087f7
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks is set to 0 since there's no reduce operator
INFO : number of splits:1
INFO : Submitting tokens for job: job_1562409531409_0012
INFO : The url to track the job: http://quickstart.cloudera:8080/proxy/application_1562409531409_0012/
INFO : Starting Job = job_1562409531409_0012, Tracking URL = http://quickstart.cloudera:8080/proxy/application_1562409531409_0012
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1562409531409_0012
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
INFO : 2019-07-06 04:59:19,459 Stage-1 map = 0%,  reduce = 0%
INFO : 2019-07-06 04:59:52,307 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 7.75 sec
INFO : MapReduce Total cumulative CPU time: 7 seconds 750 msec
INFO : Ended Job = job_1562409531409_0012
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Cumulative CPU: 7.75 sec  HDFS Read: 14078 HDFS Write: 4387 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 seconds 750 msec
INFO : Completed executing command(queryId:hive_20190708045858_e4932280-257e-4b07-8141-30432c7087f7); Time taken: 94.053 seconds
INFO : OK

```

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

Create external table if not exists IAHSInputSubset\_Partitioned (ID int,

```
BB_Average_Household_Size_All_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 string, State_District_Name string)  
location '/IndiaAHSCaseStudy/Hive/SubsetSchema_partitioned/';
```

```
insert overwrite table IAHSInputSubset_Partitioned  
partition(State_Name, State_District_Name)  
select ID, State_Name, State_District_Name, BB_Average_Household_Size_All_Total,  
CC_Sex_Ratio_All_Ages_Total, LL_Total_Fertility_Rate_Total,  
YY_Under_Five_Mortality_Rate_U5MR_Total_Person from  
IAHSInputSubset_ORCFormat;
```

**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. Query on table with ORC Format

```
Select AVG(YY_Under_Five_Mortality_Rate_U5MR_Total_Person) as  
Child_Mortality_Rate from IAHSInputSubset_Partitioned where State_Name = 'Uttar  
Pradesh';
```

Screenshot of result

The screenshot shows the Hue Editor interface running in Mozilla Firefox on a Cloudera system. The query being run is:

```
1|Select AVG(YY_Under_Five_Mortality_Rate_U5MR_Total_Person) as Child_Mortality_Rate from IAHSInputSubset_Partitioned where State_Name = 'Uttar Pradesh';
```

The results of the query show a single row:

child_mortality_rate
1 90.22857142857142

Below the results, the job log for the executed Hadoop job is visible:

```
INFO : set mapreduce.job.reduces=number  
INFO : number of splits:1  
INFO : Submitting tokens for job: job_1562471303383_0005  
INFO : The url to track the job: http://quickstart.cloudera:8080/proxy/application_1562471303383_0005/  
INFO : Starting Job + job_1562471303383_0005, Tracking URL = http://quickstart.cloudera:8080/proxy/application_1562471303383_0005/  
INFO : Kill Command = /var/lib/hadoop/bin/hadoop job -kill job_1562471303383_0005
```

b. Query on the Hive-Hbase integrated table

```
Select AVG(YY_Under_Five_Mortality_Rate_U5MR_Total_Person) as  
Child_Mortality_Rate from IAHS_Hbase where State_Name = 'Uttar Pradesh';
```

## Screenshot of the result

The screenshot shows the Hue Editor interface in Mozilla Firefox. The URL is `http://quickstart.cloudera:8888/hue/editor?editor=261`. The query window contains the following SQL:

```
1|Select AVG(YY_Under_Five_Mortality_Rate_USMR_Total_Person) as Child_Mortality_Rate from IAHS_Hbase where State_Name = 'Uttar Pradesh';
```

The results pane shows the output of the query:

child_mortality_rate
1 90.228571428571428

Below the results, the command-line interface shows the execution details:

```
INFO : set mapreduce.job.reduces=number
INFO : number of splits:
INFO : Submitting tokens for job: job_1562471303383_0003
INFO : The url to track the job: http://quickstart.cloudera:8088/proxy/application_1562471303383_0003/
INFO : Starting Job = job_1562471303383_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1562471303383_0003/
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1562471303383_0003
```

## 2. The fertility rate of Bihar

### a. Query on table with ORC Format

Select AVG(LL\_Total\_Fertility\_Rate\_Total) as Fertility\_Rate from IAHSInputSubset\_Partitioned where State\_Name = 'Bihar';

### Screenshot of result:

The screenshot shows the Hue Editor interface in Mozilla Firefox. The URL is `http://quickstart.cloudera:8888/hue/editor?editor=212`. The query window contains the following SQL:

```
1|Select AVG(LL_Total_Fertility_Rate_Total) as Fertility_Rate from IAHSInputSubset_Partitioned where State_Name = 'Bihar';
```

The results pane shows the output of the query:

fertility_rate
1 3.5324324324324321

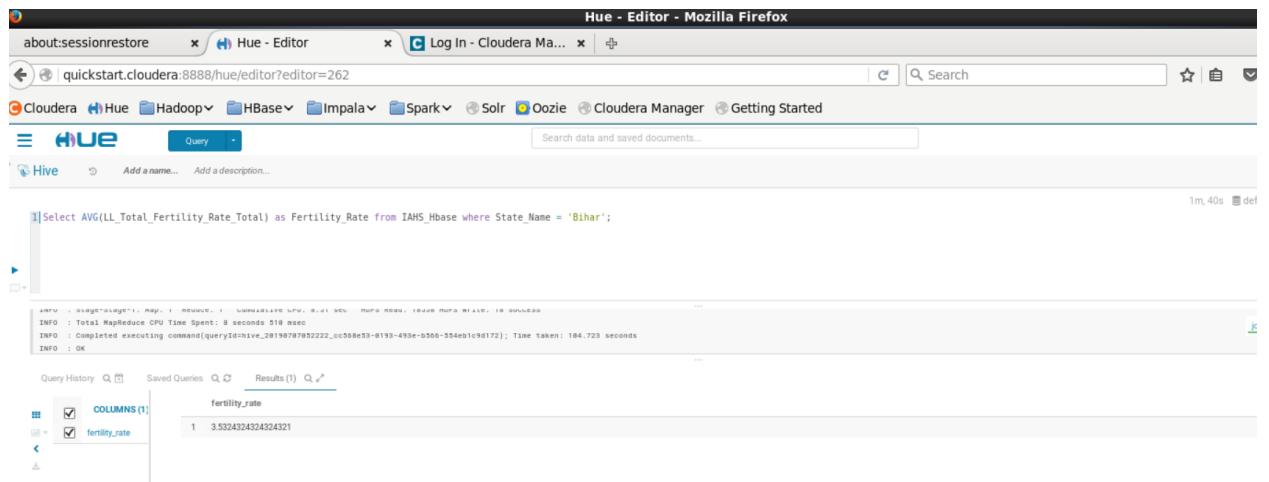
Below the results, the command-line interface shows the execution details:

```
INFO : set mapreduce.job.reduces=number
INFO : number of splits:
INFO : Submitting tokens for job: job_1562471303383_0003
INFO : The url to track the job: http://quickstart.cloudera:8088/proxy/application_1562471303383_0003/
INFO : Starting Job = job_1562471303383_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1562471303383_0003/
INFO : Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1562471303383_0003
```

b. Query on the Hive-Hbase integrated table

```
Select AVG(LL_Total_Fertility_Rate_Total) as Fertility_Rate from IAHS_Hbase where State_Name = 'Bihar'
```

Screenshot of the result



The screenshot shows the Hue interface in a Mozilla Firefox browser. The title bar reads "Hue - Editor - Mozilla Firefox". The main area displays a command-line interface for running Hive queries. A query is being run:

```
INFO : Executing command(queryId:hive_20190707052222_cc50de53-0193-493e-b506-554eb1c9d172); Time taken: 184.723 seconds
INFO : OK
```

The results pane shows the output of the query:

fertility_rate
1 3.5324324324324321

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

a. Query on the table with ORC format

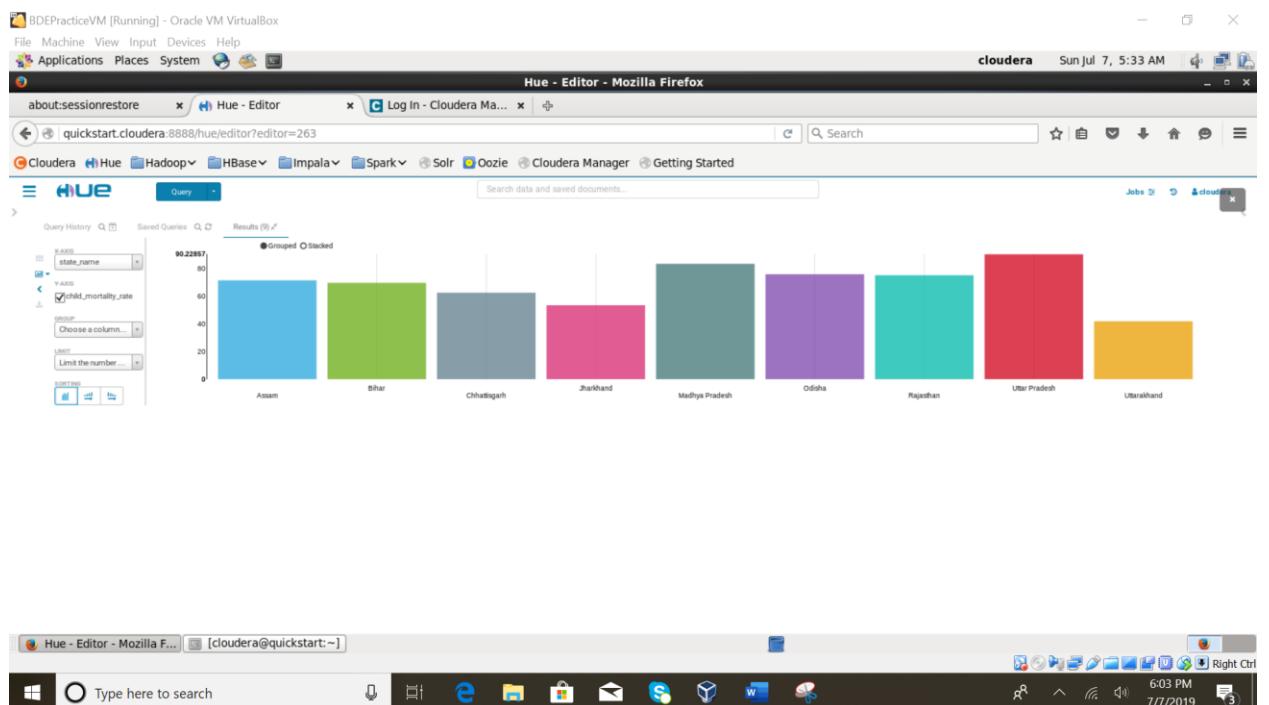
```
Select State_Name, AVG(YY_Under_Five_Mortality_Rate_U5MR_Total_Person) as Child_Mortality_Rate from IAHSInputSubset_ORCFormat GROUP BY State_Name;
```

## Screenshot of results

Query Results:

state_name	child_mortality_rate
1 Assam	71.43478290895656
2 Bihar	69.82162162162162
3 Chhattisgarh	62.5
4 Jharkhand	53.44444444444443
5 Madhya Pradesh	83.37777777777778
6 Odisha	75.79999999999997
7 Rajasthan	75.0625
8 Uttar Pradesh	90.22857142857142
9 Uttarakhand	41.846153846153847

## Chart



b. Query on the Hive-Hbase integrated table

```
Select State_Name, AVG(YY_Under_Five_Mortality_Rate_U5MR_Total_Person) as Child_Mortality_Rate from iahs_hbase GROUP BY State_Name;
```

Screenshot of the result

The screenshot shows the Hue Editor interface running on a Cloudera VM. The browser title is "Hue - Editor - Mozilla Firefox". The URL is "quickstart.cloudera:8888/hue/editor?editor=264". The page header includes links for Cloudera, Hue, Hadoop, HBase, Impala, Spark, Solr, Oozie, Cloudera Manager, and Getting Started. The main area shows a query editor with the following SQL:

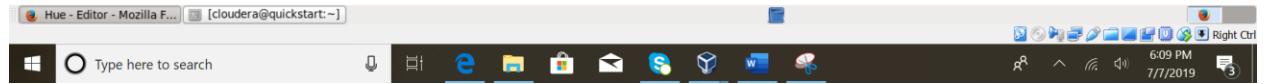
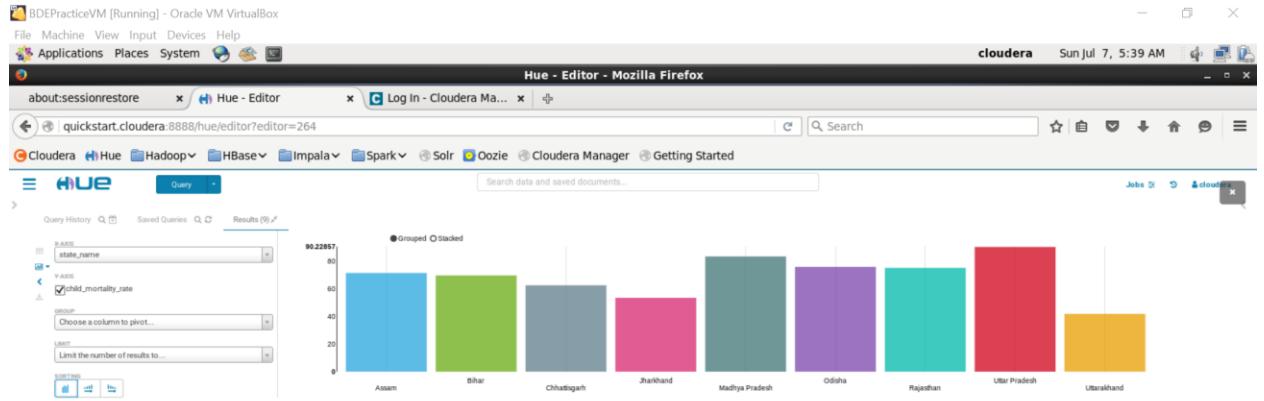
```
1 Select State_Name, AVG(YY_Under_Five_Mortality_Rate_U5MR_Total_Person) as Child_Mortality_Rate from iahs_hbase GROUP BY State_Name;
```

Below the query, the terminal output shows:

```
INFO : Total MapReduce CPU Time Spent: 7 seconds 988 msec
INFO : Completed executing command(queryId=hive_20190707053033_619c0f9-0e92-42d4-9717-9288ff99b0c1); Time taken: 100.247 seconds
INFO : OK
```

The results table has columns "state\_name" and "child\_mortality\_rate". The data is:

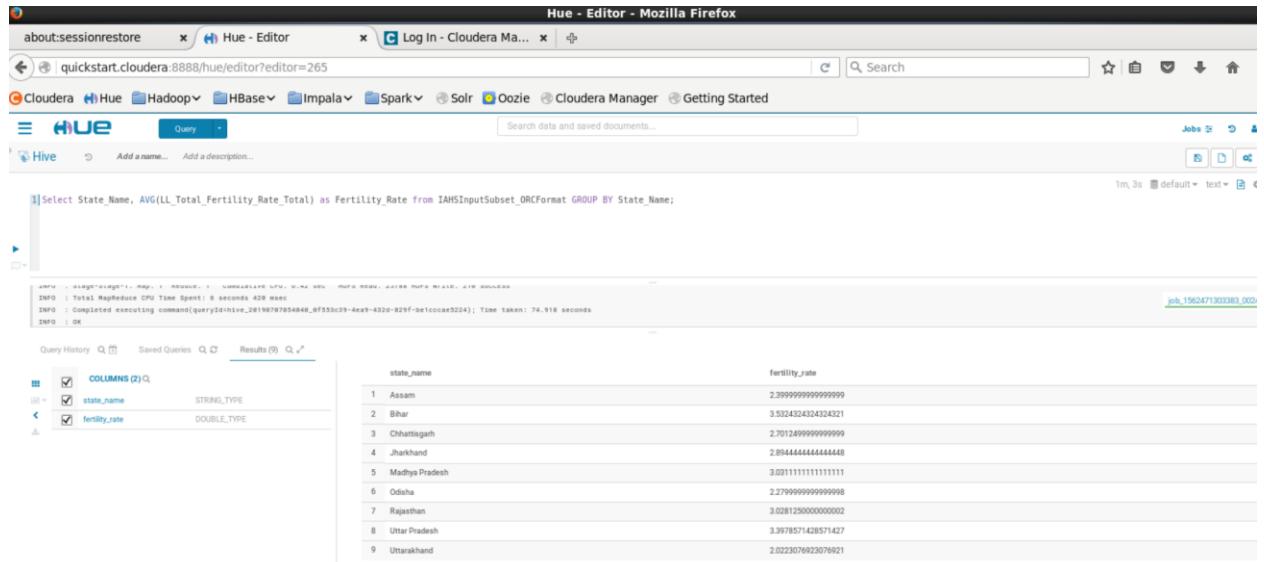
state_name	child_mortality_rate
1 Assam	71.434782608095656
2 Bihar	69.621616216216298
3 Chhattisgarh	62.5
4 Jharkhand	53.44444444444444
5 Madhya Pradesh	83.37777777777778
6 Odisha	75.79999999999999
7 Rajasthan	75.0625
8 Uttar Pradesh	90.228571428571428
9 Uttarakhand	41.846153846153847



### c. Query on the table with ORC format

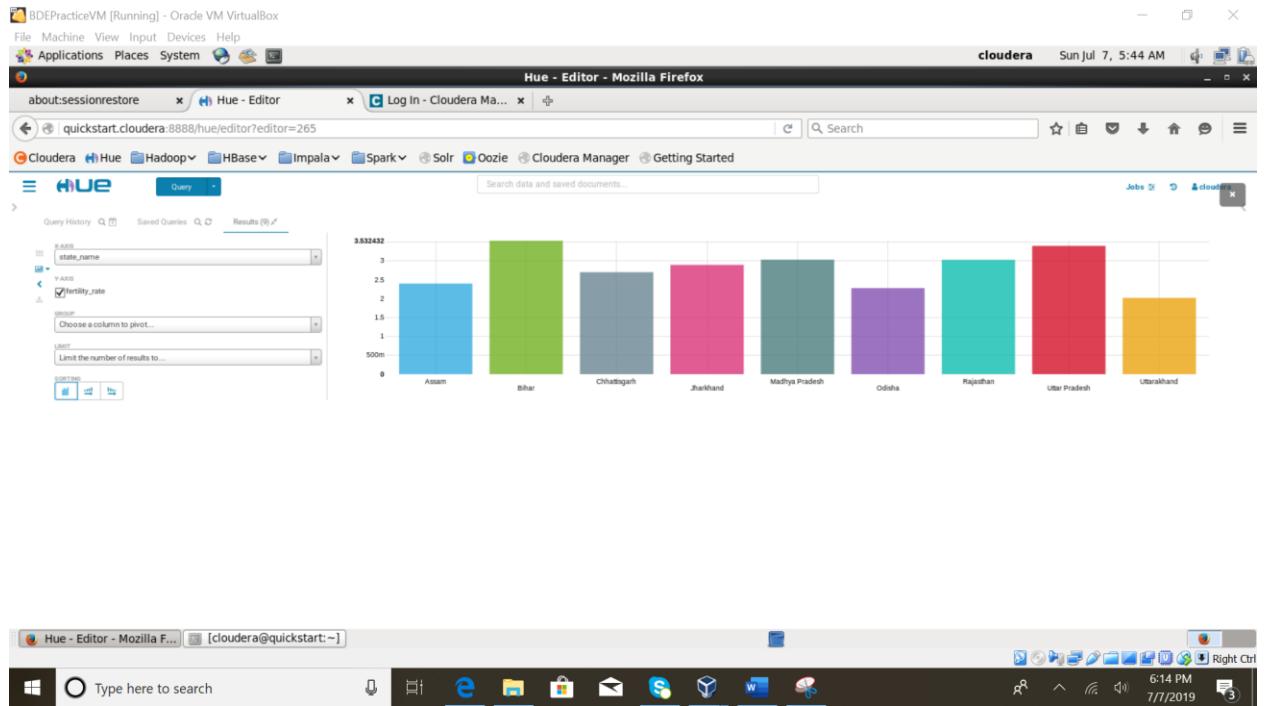
```
Select State_Name, AVG(LL_Total_Fertility_Rate_Total) as Fertility_Rate from IAHSInputSubset_ORCFormat GROUP BY State_Name;
```

### Screenshot





## Chart



### d. Query on the Hive-Hbase integrated table

```
Select State_Name, AVG(LL_Total_Fertility_Rate_Total) as Fertility_Rate from iahs_hbase GROUP BY State_Name;
```

## Screenshot of the result

The screenshot shows the Hue Editor interface on a Cloudera VM. The query window displays the following SQL:

```
1|Select State_Name, AVG(LL_Total_Fertility_Rate_Total) as Fertility_Rate from iahs_hbase GROUP BY State_Name;
```

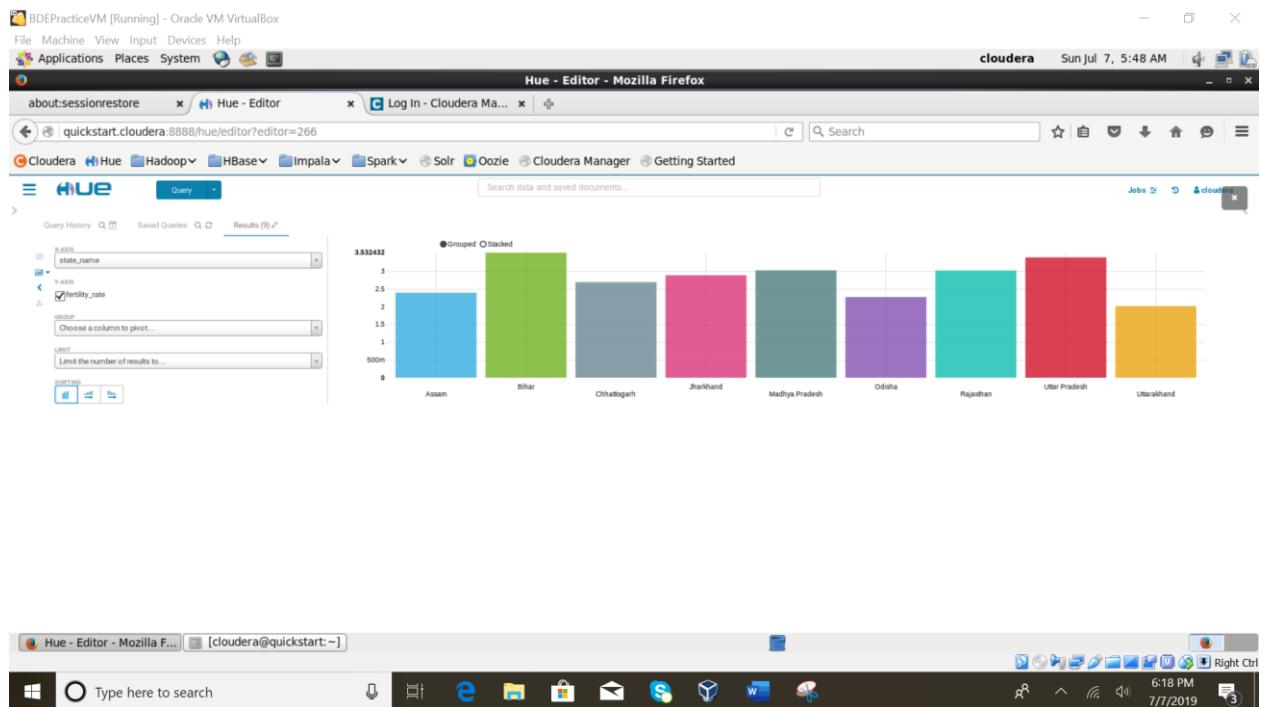
Execution details:

```
INFO : Executing query...; Row 1: RowCount: 1, ColumnCount: 2, RowSize: 1, Total RowSize: 1, Max RowSize: 1, Avg RowSize: 1
INFO : Total MapReduce CPU time Spent: 8 seconds 498 msec
INFO : Completed executing comment(queryId=hive_20190707054545_2fb44025-a8ac-4ef0-a889-082515405028); Time taken: 94.799 seconds
INFO : OK
```

Results table:

state_name	fertility_rate
Assam	2.399999999999995
Bihar	3.5224324324324321
Chhattisgarh	2.701289999999999
Jharkhand	2.8944444444444448
Madhya Pradesh	3.0311111111111112
Odisha	2.279999999999998
Rajasthan	3.0281200000000002
Uttar Pradesh	3.3978517428571427
Uttarakhand	2.0223076923076921

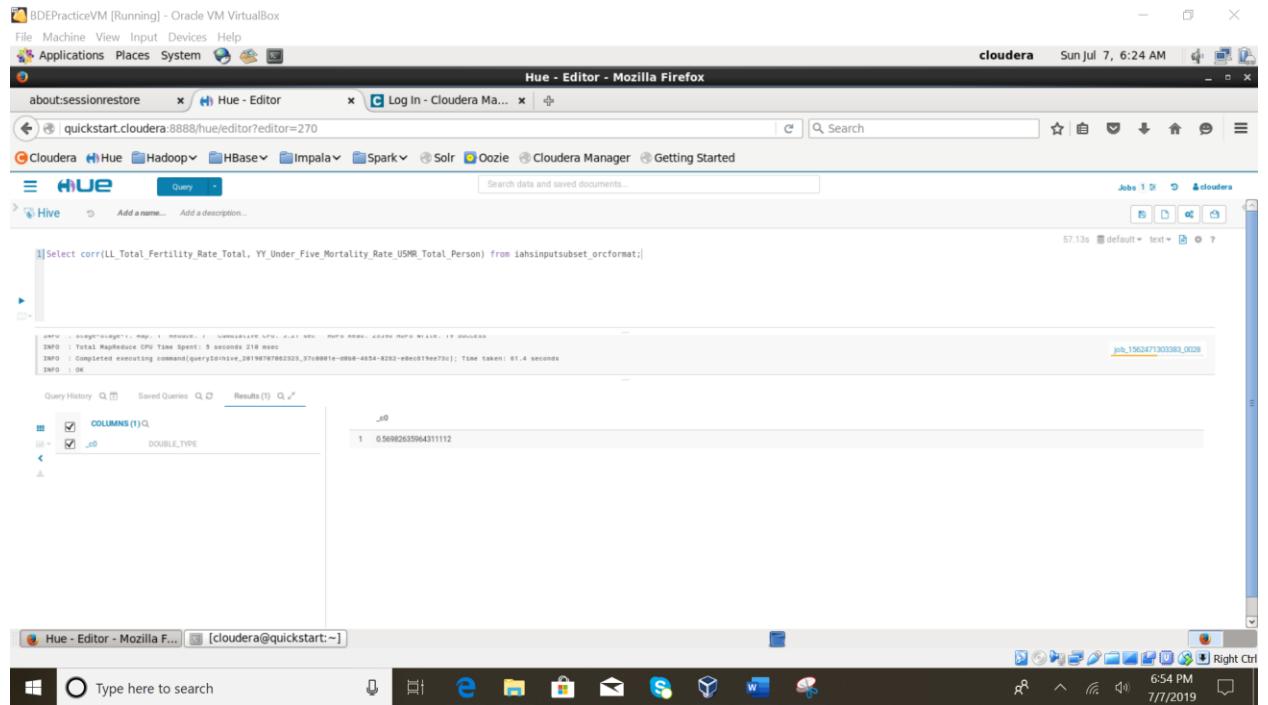
## Chart



- e. Query on the table with the chosen format such as orc

```
Select corr(LL_Total_Fertility_Rate_Total,  
YY_Under_Five_Mortality_Rate_U5MR_Total_Person) from  
iahsinputsubset_orcformat;
```

Screenshot of the result



The screenshot shows the Hue Editor interface running in a Mozilla Firefox window on a Cloudera VM. The query `Select corr(LL_Total_Fertility_Rate_Total, YY_Under_Five_Mortality_Rate_U5MR_Total_Person) from iahsinputsubset_orcformat;` has been run, and the results are displayed in the Results tab. The output shows a single row with a value of 0.56982635964311112.

Result
0.56982635964311112

- f. Query on the Hive-Hbase integrated table

```
Select corr(LL_Total_Fertility_Rate_Total,  
YY_Under_Five_Mortality_Rate_U5MR_Total_Person) from iahs_hbase;
```

## Screenshot of the result

The screenshot shows the Hue Editor interface running in a Mozilla Firefox browser window on a Cloudera VM. The URL is `quickstart.cloudera:8888/hue/editor?editor=269`. The query executed is:

```
SELECT corr(LL_Total_Fertility_Rate_Total, YY_Under_Five_Mortality_Rate_USMR_Total_Person) FROM iahs_hbase;
```

The results pane displays a single row of data:

Correlation Coefficient
0.56982635964311045

Below the results, the command history shows the execution details:

```
[INFO] : executing query: SELECT corr(LL_Total_Fertility_Rate_Total, YY_Under_Five_Mortality_Rate_USMR_Total_Person) FROM iahs_hbase;
[INFO] : Total MapReduce CPU Time Spent: 2 seconds 879 msec
[INFO] : Completed executing command(queryId=hive_201907011319_9e88-2030ff617980); Time taken: 145.899 seconds
[INFO] : OK
```

## Chart

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

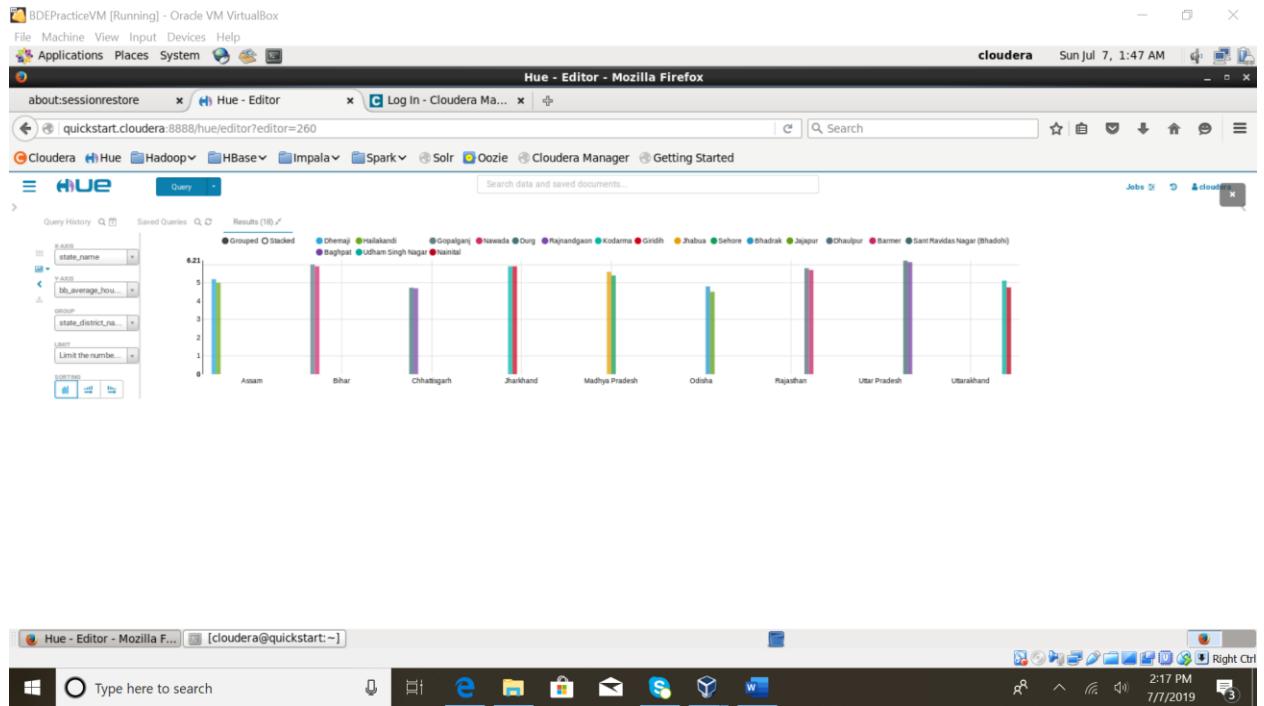
a. Query on the table with ORC Format

```
Select State_Name, state_district_name, BB_Average_Household_Size_All_Total from  
(Select State_Name, state_district_name, BB_Average_Household_Size_All_Total,  
row_number() OVER(PARTITION BY state_name ORDER BY  
BB_Average_Household_Size_All_Total desc) as rank from  
IAHSInputSubset_ORCFormat) temp where rank <= 2;
```

Screenshot of the result

state_name	state_district_name	bb_average_household_size_all_total
1 Assam	Dhemaji	5.2000000000000002
2 Assam	Halakandi	5
3 Bihar	Gopalganj	6
4 Bihar	Navada	5.9000000000000004
5 Chhattisgarh	Durg	4.7300000000000004
6 Chhattisgarh	Rajnandgaon	4.7000000000000002
7 Jharkhand	Kodama	5.9000000000000004
8 Jharkhand	Giridih	5.9000000000000004
9 Madhya Pradesh	Jhabua	5.5999999999999996
10 Madhya Pradesh	Sehore	5.4000000000000004
11 Odisha	Bhadrak	4.7999999999999998
12 Odisha	Jajpur	4.5099999999999998
13 Rajasthan	Dhaulpur	5.7999999999999998
14 Rajasthan	Bamer	5.7000000000000002
15 Uttar Pradesh	Sant Ravidas Nagar (Bhadohi)	6.21
16 Uttar Pradesh	Baghpat	6.129999999999999
17 Uttarakhand	Udham Singh Nagar	5.1200000000000001
18 Uttarakhand	Nainital	4.75

## Chart



### b. Query on the Hive-Hbase integrated table

```
Select State_Name, state_district_name, BB_Average_Household_Size_All_Total from
(Select State_Name, state_district_name, BB_Average_Household_Size_All_Total,
row_number() OVER(PARTITION BY state_name ORDER BY
BB_Average_Household_Size_All_Total desc) as rank from iahs_hbase) temp where rank
<= 2;
```

## Screenshot of the result

BDEPracticeVM [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Applications Places System

Hue - Editor - Mozilla Firefox

about:sessionrestore x Hue - Editor x C Home - Cloudera Ma... x +

quickstart.cloudera:8888/hue/editor?editor=259

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

**HUE**

Query

INFO : Total MapReduce CPU Time Spent: 9 seconds 788 msec

INFO : Completed executing command(queryId=hive\_201907071222\_382a4c44-02f2-4400-a1e9-0bf73dfa43be); Time taken: 151.831 seconds

Query History Q Saved Queries Q Results (18) Q

state\_name state\_district\_name bb\_average\_household\_size\_all\_total

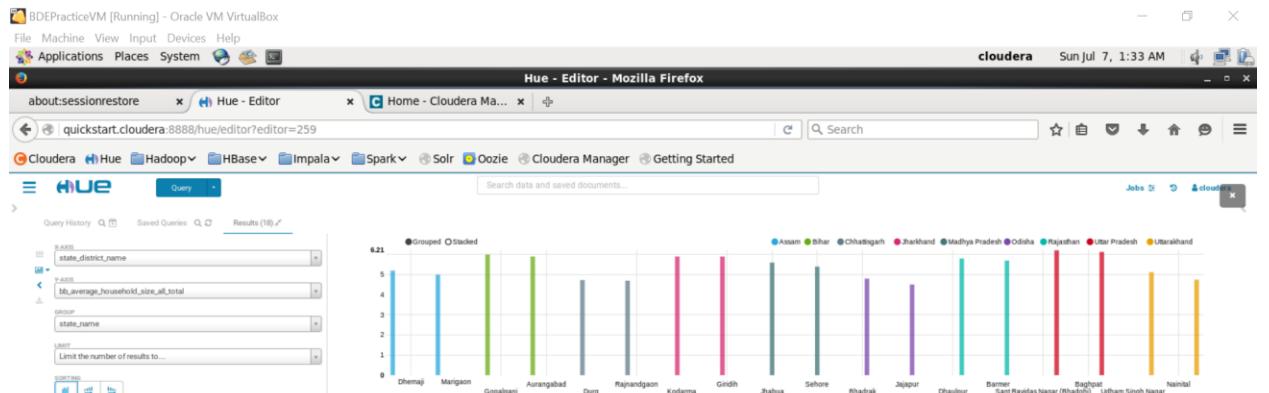
1 Assam	Dhemaj	5.2000000000000002
2 Assam	Mangon	5
3 Bihar	Gopalganj	6
4 Bihar	Aurangabad	5.9000000000000004
5 Chhattisgarh	Durg	4.7300000000000004
6 Chhattisgarh	Rajnandgaon	4.7000000000000002
7 Jharkhand	Kodama	5.9000000000000004
8 Jharkhand	Gridh	5.9000000000000004
9 Madhya Pradesh	Jhabua	5.5999999999999996
10 Madhya Pradesh	Sehore	5.4000000000000004
11 Odisha	Bhadrak	4.7999999999999998
12 Odisha	Jajpur	4.5099999999999998
13 Rajasthan	Dhadupur	5.7999999999999998
14 Rajasthan	Bamer	5.7000000000000002
15 Uttar Pradesh	Sant Ravidas Nagar (Bhadoli)	6.21
16 Uttar Pradesh	Baghpat	6.1299999999999999
17 Uttarakhand	Utham Singh Nagar	5.1200000000000001
18 Uttarakhand	Nainital	4.75

Hue - Editor - Mozilla Firefox [cloudera@quickstart:~]

Type here to search

1:56 PM 7/7/2019

## Chart



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

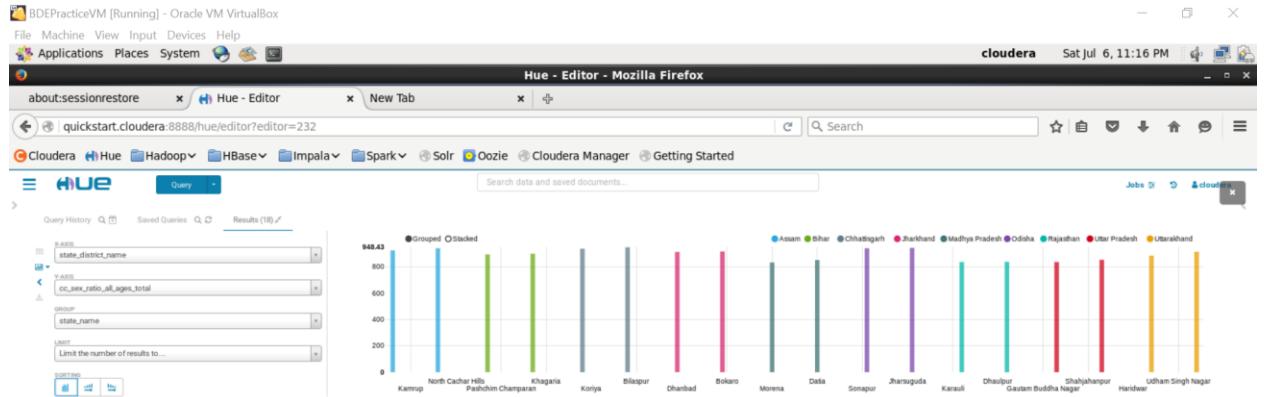
### a. Query on the table with ORC format

```
Select State_Name, state_district_name, CC_Sex_Ratio_All_Ages_Total from  
(Select State_Name, state_district_name, CC_Sex_Ratio_All_Ages_Total, row_number()  
OVER(PARTITION BY state_name ORDER BY CC_Sex_Ratio_All_Ages_Total) as rank from  
IAHSInputSubset_ORCFormat) temp where rank <= 2;
```

### Screenshot

state_name	state_district_name	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	Korba	937.299999999995
6 Chhattisgarh	Bilaspur	948.429999999995
7 Jharkhand	Dhanbad	913
8 Jharkhand	Bokaro	917
9 Madhya Pradesh	Morena	833.13
10 Madhya Pradesh	Datia	852.12
11 Odisha	Sonepur	941
12 Odisha	Jharsuguda	944
13 Rajasthan	Karauli	837
14 Rajasthan	Dhauhar	838
15 Uttar Pradesh	Gautam Buddha Nagar	836.820000000005
16 Uttar Pradesh	Shahjahanpur	853.669999999996
17 Uttarakhand	Haldwani	884.929999999995
18 Uttarakhand	Udham Singh Nagar	914.309999999995

## Chart



### b. Query on the Hive-Hbase integrated table

```
Select State_Name, state_district_name, CC_Sex_Ratio_All_Ages_Total from
(Select State_Name, state_district_name, CC_Sex_Ratio_All_Ages_Total, row_number()
OVER(PARTITION BY state_name ORDER BY CC_Sex_Ratio_All_Ages_Total) as rank from
iahs_hbase) temp where rank <= 2;
```

## Screenshot

The screenshot shows a Windows desktop environment. A Firefox browser window is open, displaying the Hue Editor interface. The URL in the address bar is `quickstart.cloudera:8888/hue/editor?editor=258`. The browser title bar says "Hue - Editor - Mozilla Firefox". The Hue interface shows a table of results with columns: state\_name, state\_district\_name, and cc\_nev\_ratio\_all\_ages\_total. The results are as follows:

state_name	state_district_name	cc_nev_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	Korba	937.299999999995
6 Chhattisgarh	Bilaspur	948.429999999995
7 Jharkhand	Dhanbad	913
8 Jharkhand	Bokaro	917
9 Madhya Pradesh	Morena	833.13
10 Madhya Pradesh	Datia	832.12
11 Odisha	Sonepur	941
12 Odisha	Jharsuguda	944
13 Rajasthan	Karauli	837
14 Rajasthan	Dholpur	838
15 Uttar Pradesh	Gautam Buddha Nagar	838.820000000005
16 Uttar Pradesh	Shahjahanpur	853.669999999995
17 Uttarakhand	Hardwar	884.929999999995
18 Uttarakhand	Udham Singh Nagar	914.309999999995

The browser title bar also shows "[cloudera@quickstart:~]". The taskbar at the bottom of the screen includes icons for File Explorer, Edge, File Manager, Mail, and others.

## Chart

