**Static Data Analysis**

**Primary Architecture of Data Pipeline**

Hadoop Tools used: Hive, Sqoop

Shell Scripting used for Automation of the entire workflow

Sqoop job

< sqoop out >

Analyzed data inserted back into Mysql Final Result Table

Table in MySql

CSV in HDFS

Sqoop job

< sqoop out >

Querying Cumulated data from all MySql tables leveraging MapReduce paradigm

From multiple RDBMS (MySQL) Tables imported into HIve

Apache Hive

Sqoop job

< sqoop in >