**ITIS 6190 Cloud Computing for Data Analysis & Comp Advt**

**Fall 2016**

**Group 5 - Project Part 2**

**Table Creation in Hive**

Query: hadoop fs -put ./projectPart2/patentdata.txt /user/lsridhar/projectPart2\_input/

CREATE TABLE IF NOT EXISTS Group5\_PatentData (doc\_number INT, invention\_title STRING, classification\_national STRING, year INT, org\_name STRING, word1 STRING, word2 STRING, word3 STRING) ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';

LOAD DATA INPATH '/user/lsridhar/projectPart2\_input/patentData.txt' INTO TABLE Group5\_PatentData;

**2.1. All Patents from year 2015**

Query: INSERT OVERWRITE DIRECTORY 'patentdata\_output/query1' ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE SELECT \* FROM Group5\_PatentData where year = 2015;

**2.2. All Patents owned by the company ' Apple '**

Query: INSERT OVERWRITE DIRECTORY 'patentdata\_output/query2' ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE SELECT \* FROM Group5\_PatentData WHERE LOWER(org\_name) LIKE '%apple %';

**2.3. All patents in class electronics, that contain the frequent word ' Samsung ' AND ' Apple '**

Query: INSERT OVERWRITE DIRECTORY 'patentdata\_output/query3' ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE SELECT COUNT(\*) FROM Group5\_PatentData WHERE LOWER(classification\_national) LIKE '%electronic%' AND (((LOWER(word1) LIKE '%samsung%' OR LOWER(word1) LIKE '%apple%') AND (LOWER(word2) LIKE '%samsung%' OR LOWER(word2) LIKE '%apple%')) OR ((LOWER(word2) LIKE '%samsung%' OR LOWER(word2) LIKE '%apple%') AND (LOWER(word3) LIKE '%samsung%' OR LOWER(word3) LIKE '%apple%')) OR ((LOWER(word1) LIKE '%samsung%' OR LOWER(word1) LIKE '%apple%') AND (LOWER(word3) LIKE '%samsung%' OR LOWER(word3) LIKE '%apple%')));