javac -cp /home/student3/hadoop-common-2.6.1.jar:/home/student3/hadoop-mapreduce-client-core-2.6.1.jar:/home/student3/commons-cli-2.0.jar -d . YearTemperatureSruthi.java YearTemperatureSruthiMapper.java YearTemperatureSruthiReducer.java

jar -cvf YearTemperatureSruthi.jar ./YearTemperatureSruthi\*.class

hadoop fs -mkdir /user/student3FP

hdfs dfs -copyFromLocal CourseProjectData /user/student3FP/CourseProjectData

hadoop jar /home/student3/YearTemperatureSruthi.jar YearTemperatureSruthi /user/student3FP/CourseProjectData /user/student3FP/Project \_output/

hdfs dfs -ls /user/student3FP/Project \_output/

SCREENSHOT

hdfs dfs -cat /user/student3FP/Project \_output/part-r-00000

hdfs dfs -copyToLocal /user/student3FP/Project \_output/part-r-00000 /home/student3

#check for r-00000 in winscp

hdfs dfs -copyFromLocal Temp\_output.txt /user/student3FP/Temp\_output.txt

Screenshot AFTER LS COMMAND

Pig:

pig -x local

record = LOAD 'Temp\_output.txt' AS (year:chararray, temperature:int);

record\_filtered = FILTER record BY temperature != 9999;

DUMP record\_filtered;

#screenshot

record\_grouped = GROUP record\_filtered BY year;

DUMP record\_grouped;

maximum\_temperature = FOREACH record\_grouped GENERATE group, MAX(record\_filtered.temperature);

DUMP maximum\_temperature;

#SS

minimum\_temperature = FOREACH record\_grouped GENERATE group, MIN(record\_filtered.temperature);

DUMP minimum\_temperature;

#SS

Part 3:

ls -l | grep meta

mv metastore\_db metastore\_db.old

schematool -dbType derby -initSchema

hive

Drop table if exists records03\_student3;

CREATE TABLE records03\_student3 (year STRING, temperature INT)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t';

load data local inpath 'Temp\_output.txt' overwrite into table records03\_student3;

SELECT year, round(AVG(temperature),2) FROM records03\_student3 WHERE temperature != 9999 GROUP BY year;