2015

BAY AREA BIKE SHARING

Snehal Bhilare

INFO7250:Engineering Big-Data Systems

12/17/2015

Index

PROJECT DESCRIPTION

The “Bay Area Bike Share” is the region’s bike sharing system with 700 bikes and 70 stations across the region, with locations in San Francisco, Redwood City, Mountain View, Palo Alto, and San Jose. It provides the transportation facilities to the residents with ease of use affordable service.

To get started with this service you just need to get the annual or 3 days membership. Enter your ride code or membership key and get going with your trip.

DATASET

FILE LIST

1) 201508_status_data.csv – approx. 37 million records of bike and dock availability by minute per station.

2) 201508_station_data.csv – 70 records – station ID, name, latitude, longitude, dockcount, city, installation date

3) 201508_trip_data.csv – approx. 354,000 records of individual trips.

4) 201508_weather_data.csv – 1,825 records of daily weather by city.

Link: https://s3.amazonaws.com/babs-open-data/babs_open_data_year_2.zip

Techniques included:

- AWS EC2
- Hadoop
- HDFS
- JAVA MAPREDUCE
- CHAINING
- PIG
- UDF
- HIVE
- HBASE
- SPRING MVC
- HTML
- CSS
- BOOTSTRAP
- JAVA SCRIPT
- JSON
- D3JS
- TABLUE
- MAPS

ANALYSIS

CONFIGURATION

Psedo distributed mode AWS EC2 (Manuaaly configured):

Data inside local : /usr/local/data/ Data inside hdfs: /input/data hbase data : hbase://trip_data

JAVA MAPREDUCE:

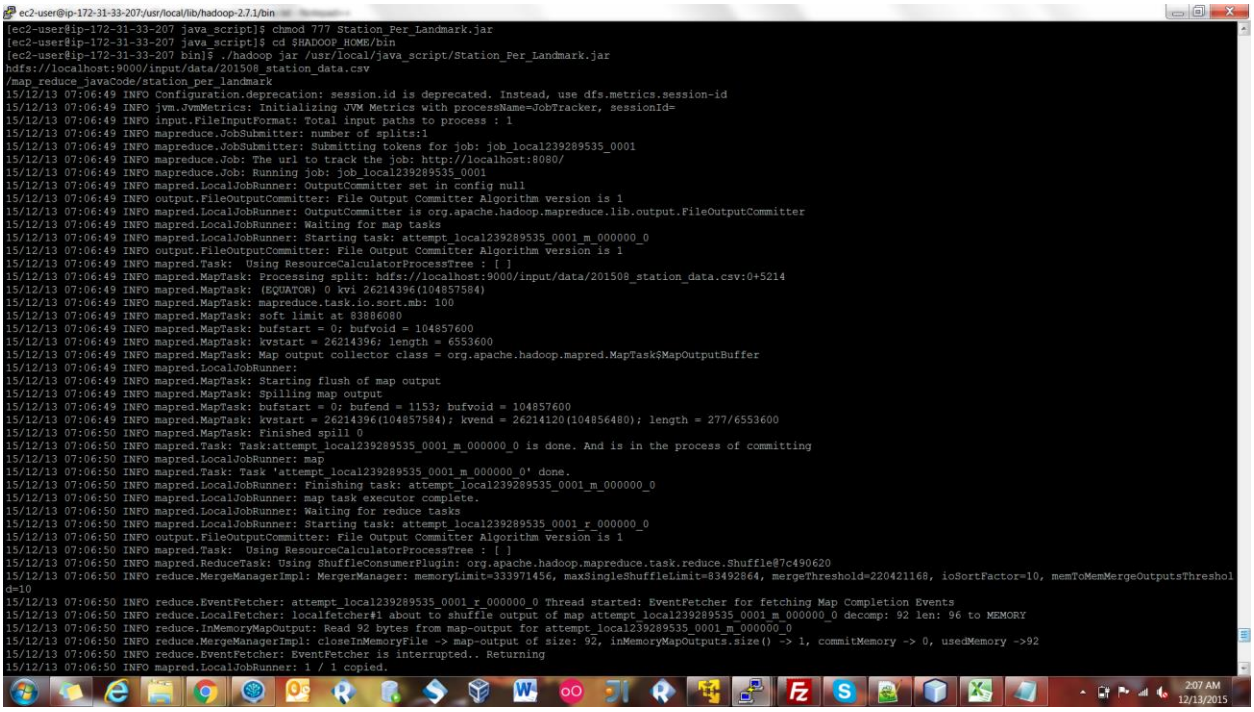
---------- **Analysis : 1** ----------

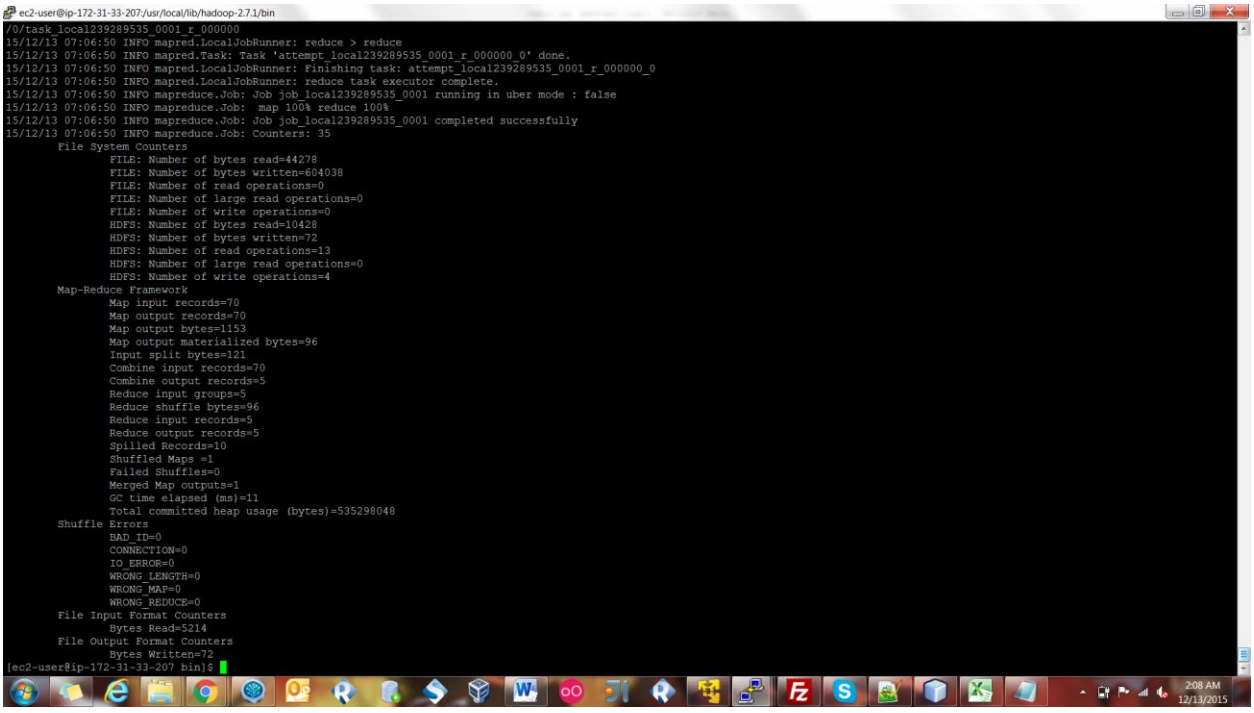
Analysis: Total no of stations per landmark from station data.

Script on ec2: /usr/local/java_script/Station_Per_Landmark.jar

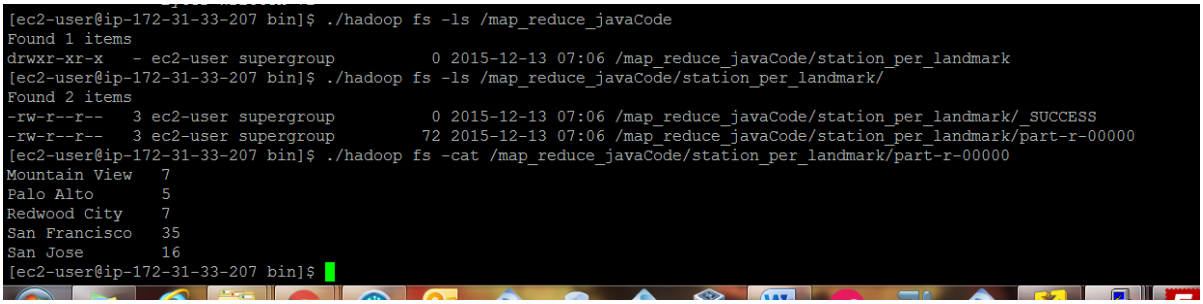
Output on ec2: /map_reduce_javaCode/station_per_landmark/part-r-00000

Run:





Output:



Mountain View 7

Palo Alto 5

Redwood City 7

San Francisco 35

San Jose 16

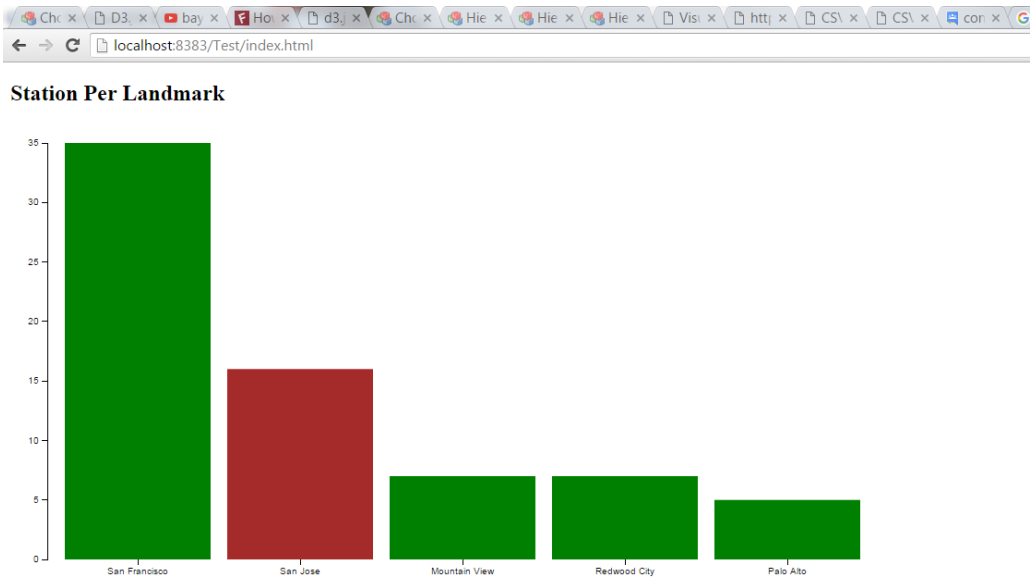
JAVA code:

Station_Per_Landmark.java

```
/*
 * To change this license header, choose License Headers in Project Properties.
 * To change this template file, choose Tools | Templates
 * and open the template in the editor.
 */
package Station;


import java.io.File;
import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.conf.Configured;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;
import org.apache.hadoop.util.Tool;
import org.apache.hadoop.util.ToolRunner;

/**
 *
 * @author ubuntu
 */
public class Station_Per_Landmark  extends Configured implements Tool{

    public static class WordCountMapper
            extends Mapper<LongWritable, Text, Text, IntWritable> {
        private static IntWritable one = new IntWritable(1);
        private Text wordText = new Text();
        public void map(LongWritable key, Text line, Context context) throws IOException, InterruptedException
{
            String []words=line.toString().split(",");//this will split the words by spaces and tabs[^\w]
instead
            wordText.set(words[5]);
```

```
            context.write(wordText, one);
        }
    }

    public static class WordCountReducer extends Reducer<Text, IntWritable, Text, IntWritable> {

        private IntWritable result = new IntWritable();

        public void reduce(Text key, Iterable<IntWritable> values,
                Context context)
                throws IOException, InterruptedException {
            int total = 0;
            for (IntWritable x : values)
            {

                total+= x.get();
            }

            result.set(total);
            context.write(key, result);
        }
    }

    public int run(String[] args) throws Exception {
        Configuration conf = getConf();
        Job job = new Job(conf, "access count");
        job.setJarByClass(Station_Per_Landmark.class);
        final File f = new
File(Station_Per_Landmark.class.getProtectionDomain().getCodeSource().getLocation().getPath());
       // String inFiles = f.getAbsolutePath().replace("/build/classes", "") + "/src/inFiles/";
        String inFiles ="hdfs://localhost:9000/input/data/201508_station_data.csv";
        String outFiles = "/map_reduce_javaCode/station_per_landmark";
        //use the arguments instead if provided.
        if (args.length > 1) {
            inFiles = args[0];
            outFiles = args[1];
        }
        System.out.println(inFiles);
        System.out.println(outFiles);
        Path in = new Path(inFiles);
        Path out = new Path(outFiles);
        FileInputFormat.setInputPaths(job, in);
        FileOutputFormat.setOutputPath(job, out);
        job.setMapperClass(WordCountMapper.class);
        job.setCombinerClass(WordCountReducer.class);
        job.setReducerClass(WordCountReducer.class);
        job.setInputFormatClass(TextInputFormat.class);
        job.setOutputFormatClass(TextOutputFormat.class);
        job.setOutputKeyClass(Text.class);
        job.setOutputValueClass(IntWritable.class);
        System.exit(job.waitForCompletion(true) ? 0 : 1);
        return 0;
    }

    public static void main(String[] args) throws Exception {
        int res = ToolRunner.run(new Configuration(), new Station_Per_Landmark(), args);
        System.exit(res);
    }

}
```

Graph:



---------- **Analysis : 2 JAVA Chaining** ----------

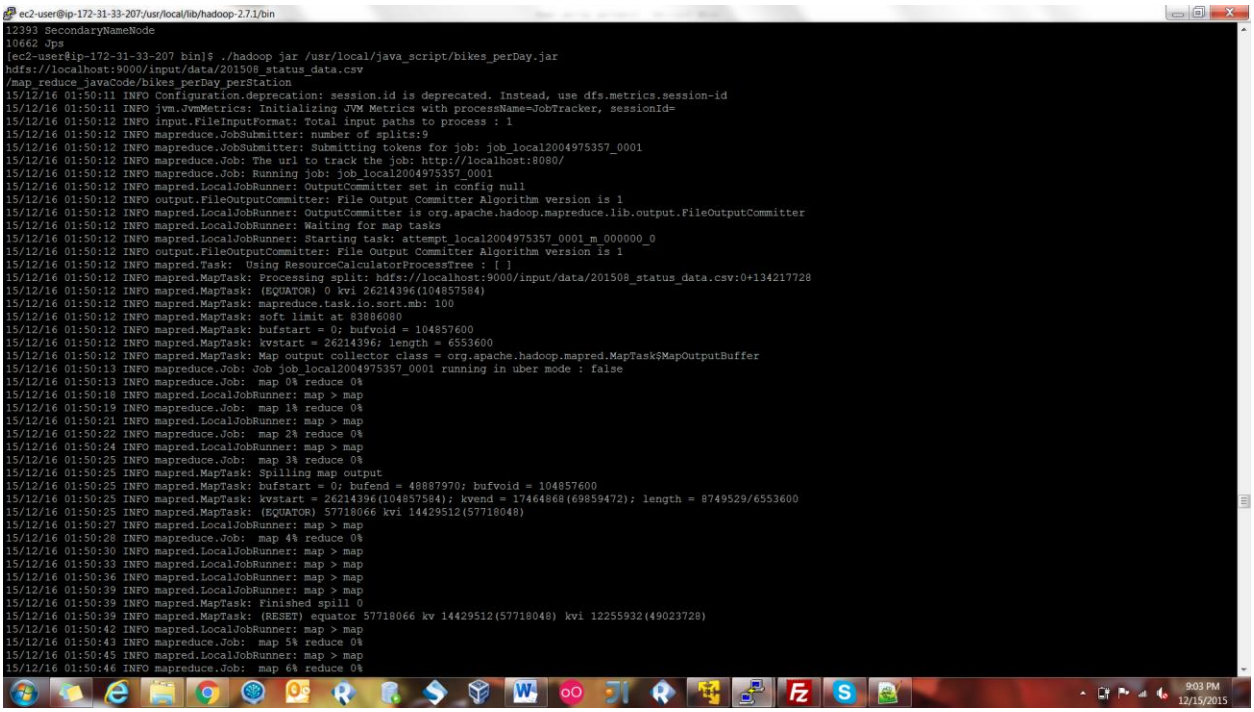
Analysis: Average bike rented per day per station. Used java chaining map reduce to generate composite key as “station_id + day of week”.

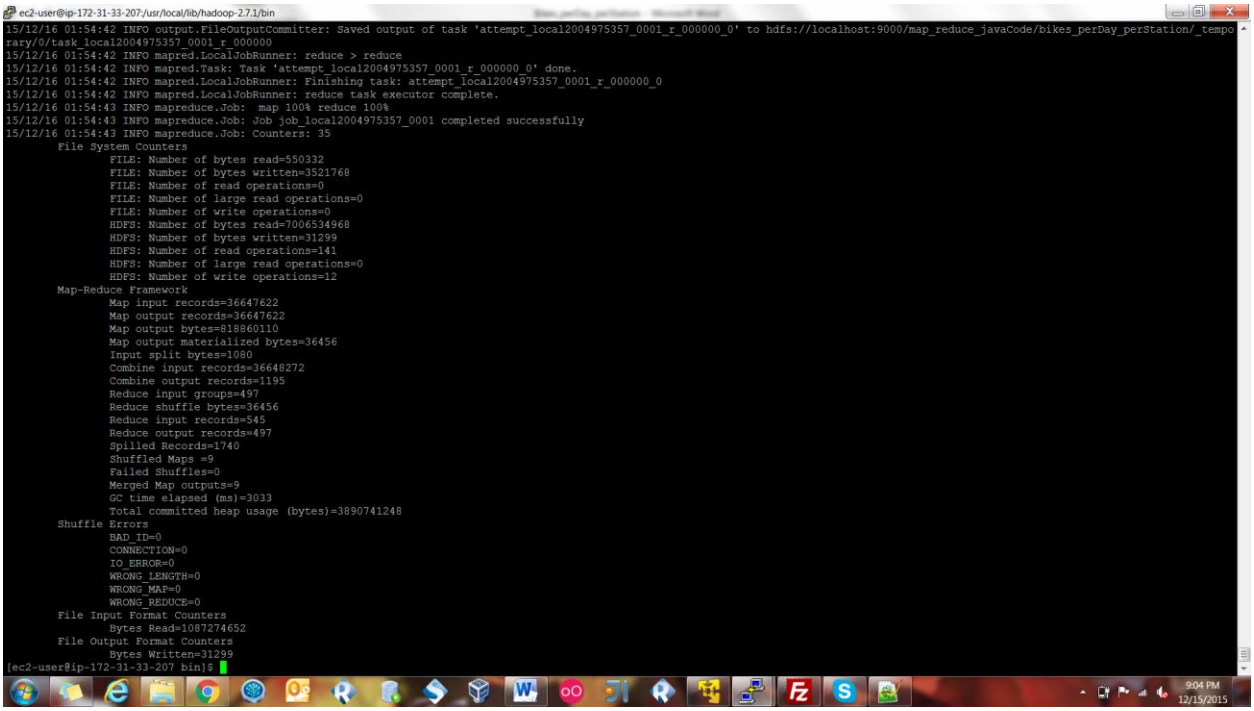
Result: Business is rarely affected by the day of the week.

Jar path : /usr/local/java_script/bikes_perDay.jar

Output path: /map_reduce_javaCode/bikes_perDay_perStation/part-r-00000

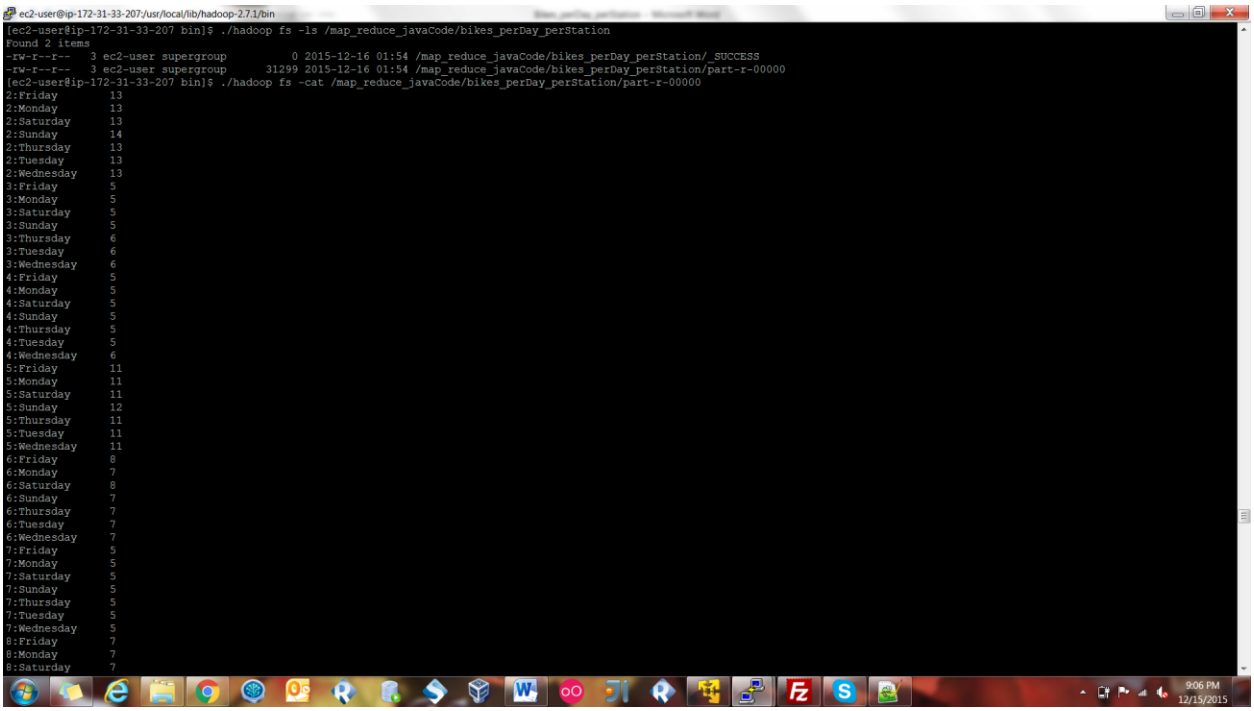
Run:

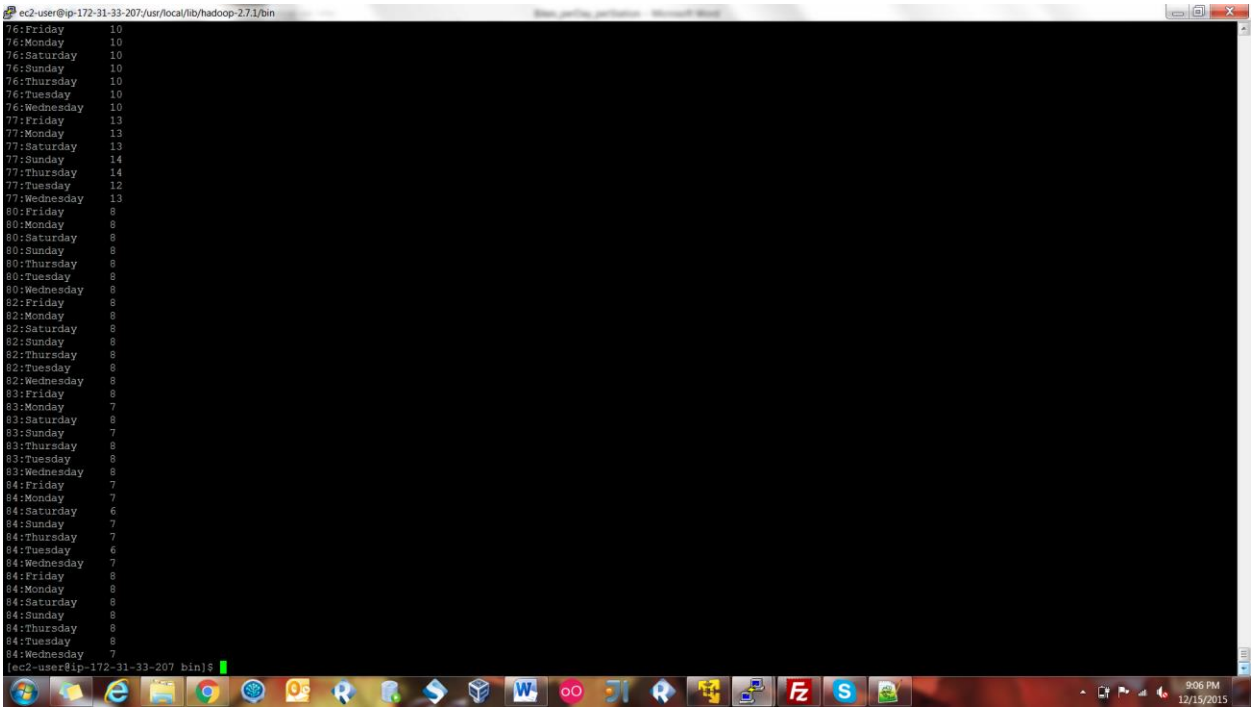




Output:

```
                Bytes Written=31299
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -ls /map_reduce_javaCode/bikes_perDay_perStation
Found 2 items
-rw-r--r--   3 ec2-user supergroup          0 2015-12-16 01:54 /map_reduce_javaCode/bikes_perDay_perStation/_SUCCESS
-rw-r--r--   3 ec2-user supergroup      31299 2015-12-16 01:54 /map_reduce_javaCode/bikes_perDay_perStation/part-r-00000
[ec2-user@ip-172-31-33-207 bin]$
```





JAVA Code:

station_perDay.java

```
/*
 * To change this license header, choose License Headers in Project Properties.
 * To change this template file, choose Tools | Templates
 * and open the template in the editor.
 */
package bikes_perDayStation;
```

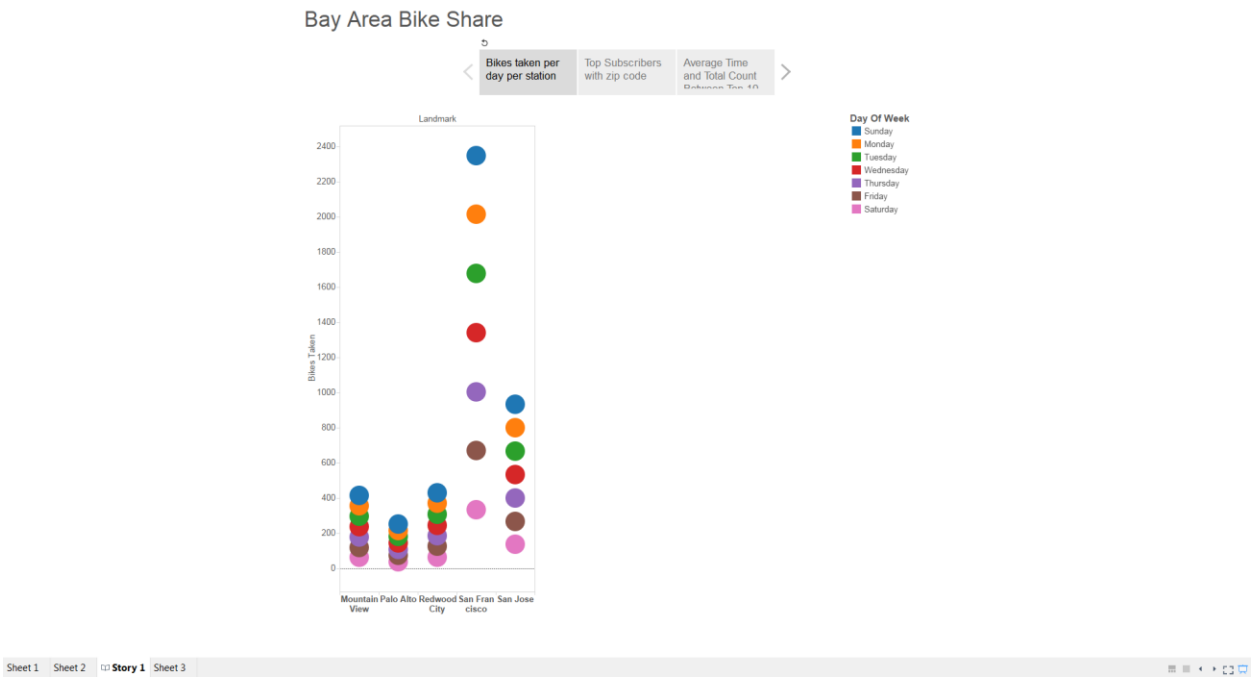
```
import java.io.DataInput;
import java.io.DataOutput;
import java.io.IOException;
import org.apache.hadoop.io.WritableComparable;

/**
 *
 * @author ubuntu
 */
public class station_perDay implements WritableComparable<station_perDay>
{


private int station_id;
private String day ;


    public String getDay() {
        return day;
    }

    public void setDay(String day) {
        this.day = day;
    }


    public int getStation_id() {
        return station_id;
    }

    public void setStation_id(int station_id) {
        this.station_id = station_id;
    }



    @Override
    public void write(DataOutput out) throws IOException {
      out.writeInt(station id);
       out.writeChars(day);


    }

    @Override
    public void readFields(DataInput in) throws IOException {


    station_id = in.readInt();
    day= in.readLine();
    }

    @Override
    public int compareTo(station_perDay t) {

       int result= (int)(this.getStation_id()-t.getStation_id());
       if (0 == result) {
            result = day.compareTo(t.day);
        }
        return result;


    }

    @Override
    public String toString() {
        return  station_id+":"+day;
    }



}
```

main.java

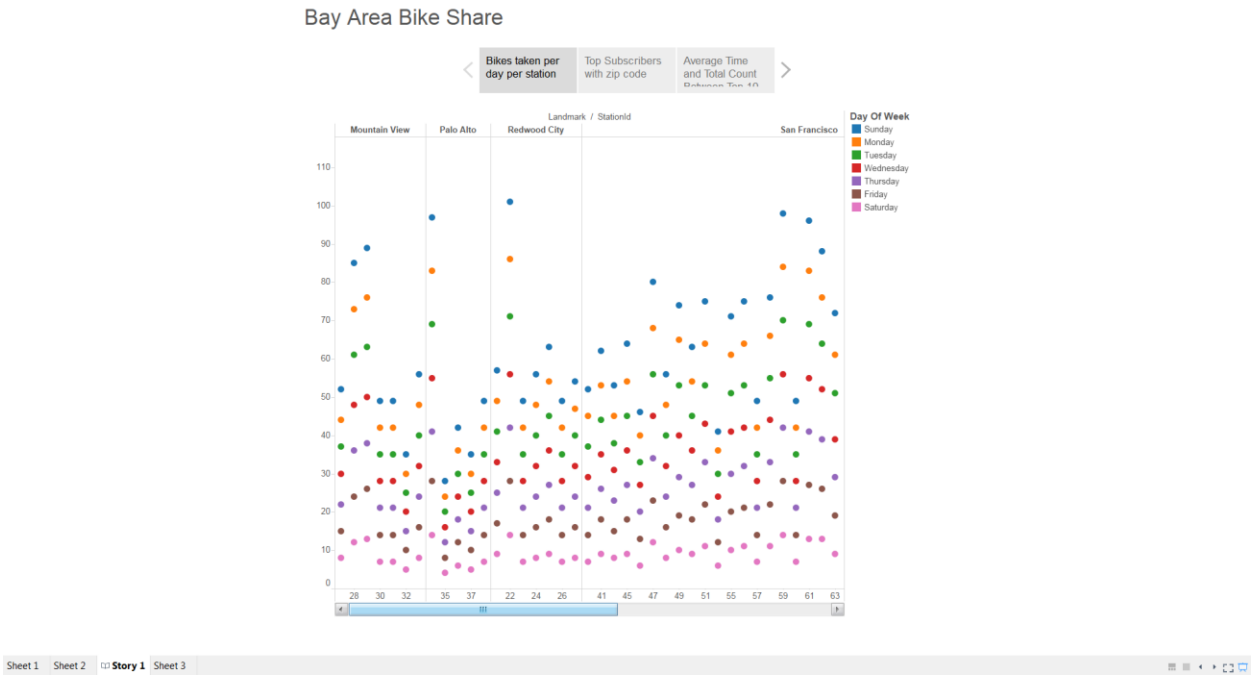
```
/*
 * To change this license header, choose License Headers in Project Properties.
 * To change this template file, choose Tools | Templates
 * and open the template in the editor.
 */
package bikes_perDayStation;


import java.io.File;
import java.io.IOException;
import java.text.DateFormat;
import java.text.ParseException;
import java.text.SimpleDateFormat;

import java.util.Date;
import java.util.logging.Level;
import java.util.logging.Logger;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.conf.Configured;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;
import org.apache.hadoop.util.Tool;
import org.apache.hadoop.util.ToolRunner;

/**
 *
 * @author ubuntu
 */
public class main  extends Configured implements Tool{

    public static class WordCountMapper
            extends Mapper<LongWritable, Text, station_perDay, IntWritable> {
        private station_perDay wordText = new station_perDay();

        public void map(LongWritable key, Text line, Context context) throws IOException, InterruptedException
{
            String []words=line.toString().split(",");//this will split the words by spaces and tabs[^\w]
instead
            if (words[0].equals("")||words[0].isEmpty()||words[0].equals("station_id")||
                    words[3].equals("")||words[3].isEmpty()||words[3].equals("time")||
                    words[2].equals("")||words[2].isEmpty()||words[2].equals("docks_available")) return;
            wordText.setStation_id(Integer.parseInt(words[0]));
            String day;

            try {
                day = getDayFromDate(words[3]);
                 wordText.setDay(day);
            } catch (ParseException ex) {
                Logger.getLogger(main.class.getName()).log(Level.SEVERE, null, ex);
            }


            IntWritable result;
            result = new IntWritable(Integer.parseInt(words[2]));

            context.write(wordText, result);
        }

        private String getDayFromDate(String word) throws ParseException {

            word = word.substring(1, word.length()-1);

            Date  yourDate= new SimpleDateFormat("yyyy-mm-dd HH:mm:ss").parse(word);
            DateFormat format2=new SimpleDateFormat("EEEE");
            String finalDay=format2.format(yourDate);

            return finalDay;
```

```
        }
    }

    public static class WordCountReducer extends Reducer<station_perDay, IntWritable, station_perDay,
IntWritable> {


        public void reduce(station_perDay key, Iterable<IntWritable> values,
                Context context)
                throws IOException, InterruptedException {


           int result =0;
           int count =0;
            for (IntWritable value : values) {

                  result += value.get();
                  count++;

            }

            IntWritable total = new IntWritable(result/count);

            context.write(key, total);

        }
    }

    public int run(String[] args) throws Exception {
        Configuration conf = getConf();
        Job job = new Job(conf, "access count");
        job.setJarByClass(main.class);
        final File f = new File(main.class.getProtectionDomain().getCodeSource().getLocation().getPath());
        //String inFiles="/usr/local/data/201508_status_data.csv";
        //String outFiles=f.getAbsolutePath().replace("/build/classes", "") + "/src/outFiles/WorWdCount";
        String inFiles ="hdfs://localhost:9000/input/data/201508_status_data.csv";
        String outFiles = "/map_reduce_javaCode/bikes_perDay_perStation";

        if (args.length > 1) {
            inFiles = args[0];
            outFiles = args[1];
        }
        System.out.println(inFiles);
        System.out.println(outFiles);
        Path in = new Path(inFiles);
        Path out = new Path(outFiles);
        FileInputFormat.setInputPaths(job, in);
        FileOutputFormat.setOutputPath(job, out);
        job.setMapperClass(WordCountMapper.class);
        job.setCombinerClass(WordCountReducer.class);
        job.setReducerClass(WordCountReducer.class);
        job.setInputFormatClass(TextInputFormat.class);
        job.setOutputFormatClass(TextOutputFormat.class);
        job.setOutputKeyClass(station_perDay.class);
        job.setOutputValueClass(IntWritable.class);
        System.exit(job.waitForCompletion(true) ? 0 : 1);
        return 0;
    }

    public static void main(String[] args) throws Exception {
        int res = ToolRunner.run(new Configuration(), new main(), args);
        System.exit(res);
    }

}
```

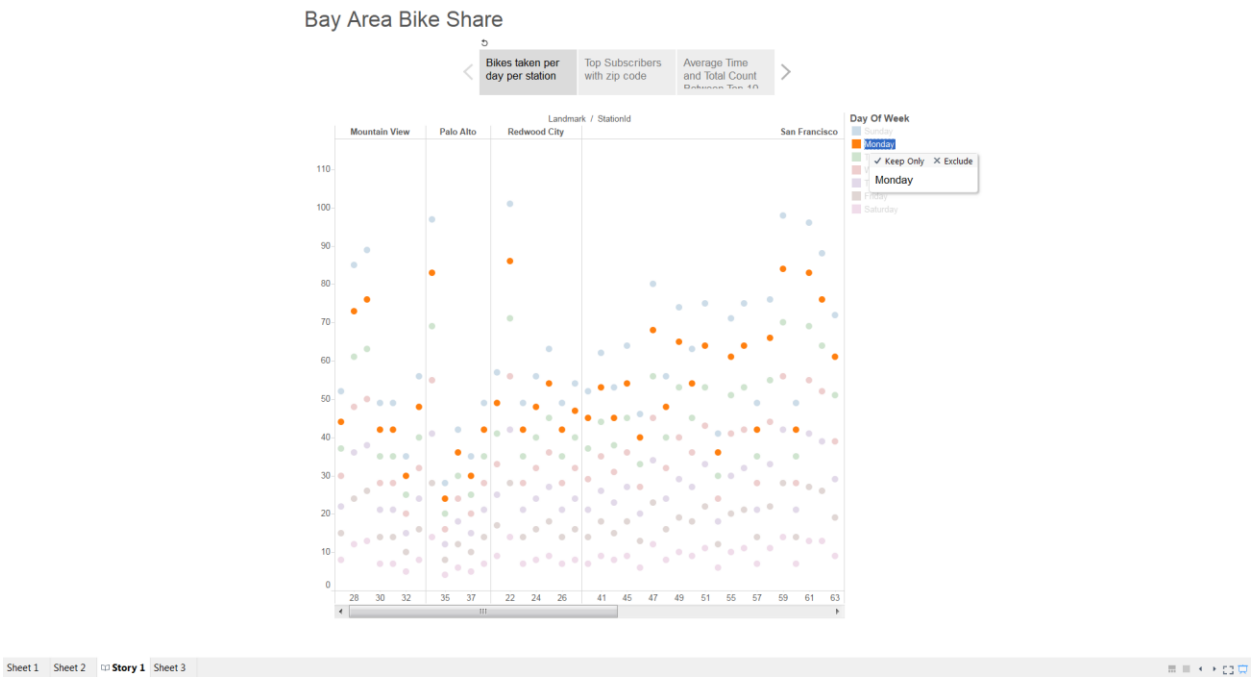
Graph:

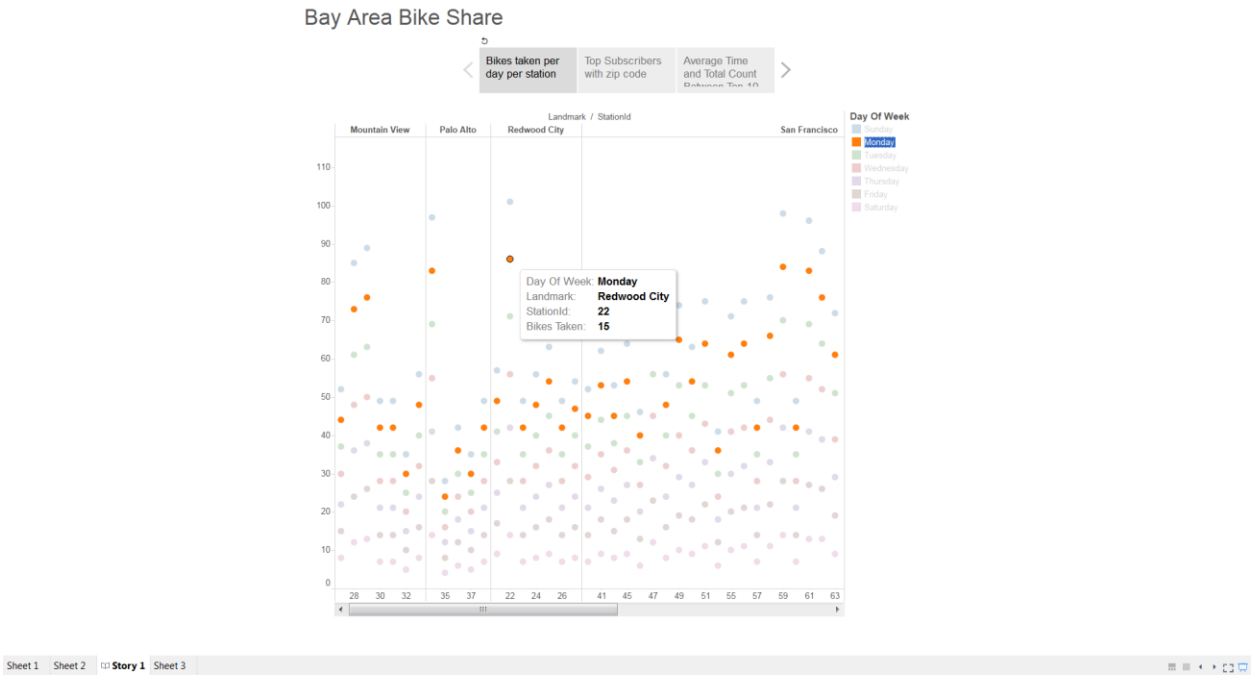


Drill down landmark



Filter by day:





PIG:

---------- **Analysis : 3** ----------

Analysis: Total no of stations per landmark from station data.

pig script: /usr/local/pig_script/station_per_landmark.pig

pig output: /pig_mapreduce/station_per_landmark/part-r-00000

Script:

```
station_data = LOAD 'hdfs:/input/data/201508_station_data.csv' USING PigStorage(',') AS
(station_id:int,name:chararray,lat:double,long:double,dockcount:int,landmark:chararray,installation:dat
etime);

B = group station_data  by landmark;

C = FOREACH B generate group , COUNT(station_data) as noOfStation;

D = ORDER C by noOfStation DESC;

Store D into 'hdfs:/pig_mapreduce/station_per_landmark';
```

Output:

```
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -ls /pig_mapreduce/station_per_landmark/
Found 2 items
-rw-r--r--   3 ec2-user supergroup          0 2015-12-13 05:54 /pig_mapreduce/station_per_landmark/_SUCCESS
-rw-r--r--   3 ec2-user supergroup         72 2015-12-13 05:54 /pig_mapreduce/station_per_landmark/part-r-00000
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -cat /pig_mapreduce/station_per_landmark/part-r-00000
San Francisco   35
San Jose        16
Mountain View   7
Redwood City    7
Palo Alto       5
[ec2-user@ip-172-31-33-207 bin]$
```

San Francisco 35

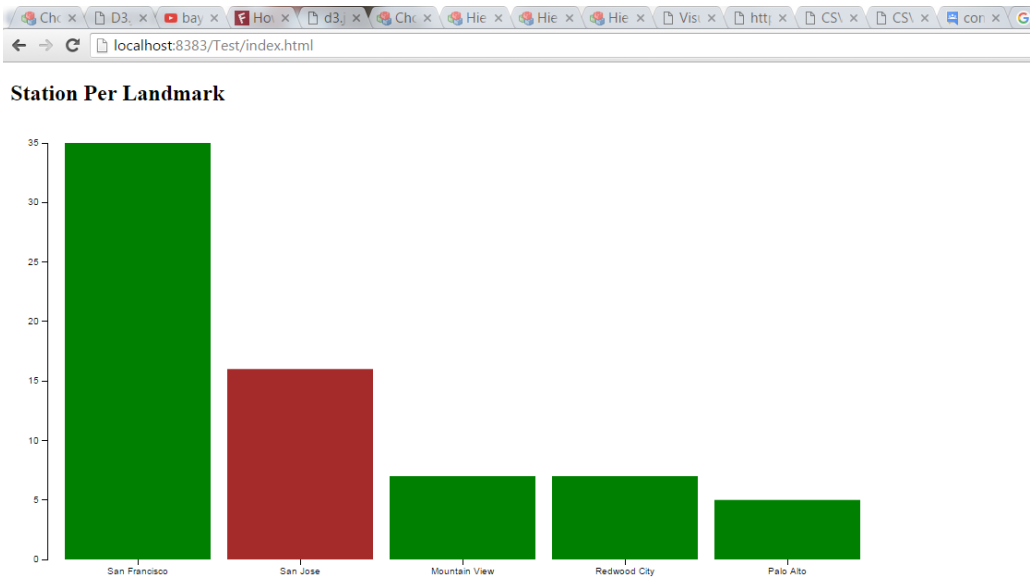
San Jose 16

Mountain View 7

Redwood City 7

Palo Alto 5

Graph:



---------- **Analysis : 4** ----------

Analysis: Estimated time between two stations (average time to reach from one station to other) and total no of trips in between them ordered by trip_count. Used start and end station as key with CONCAT function

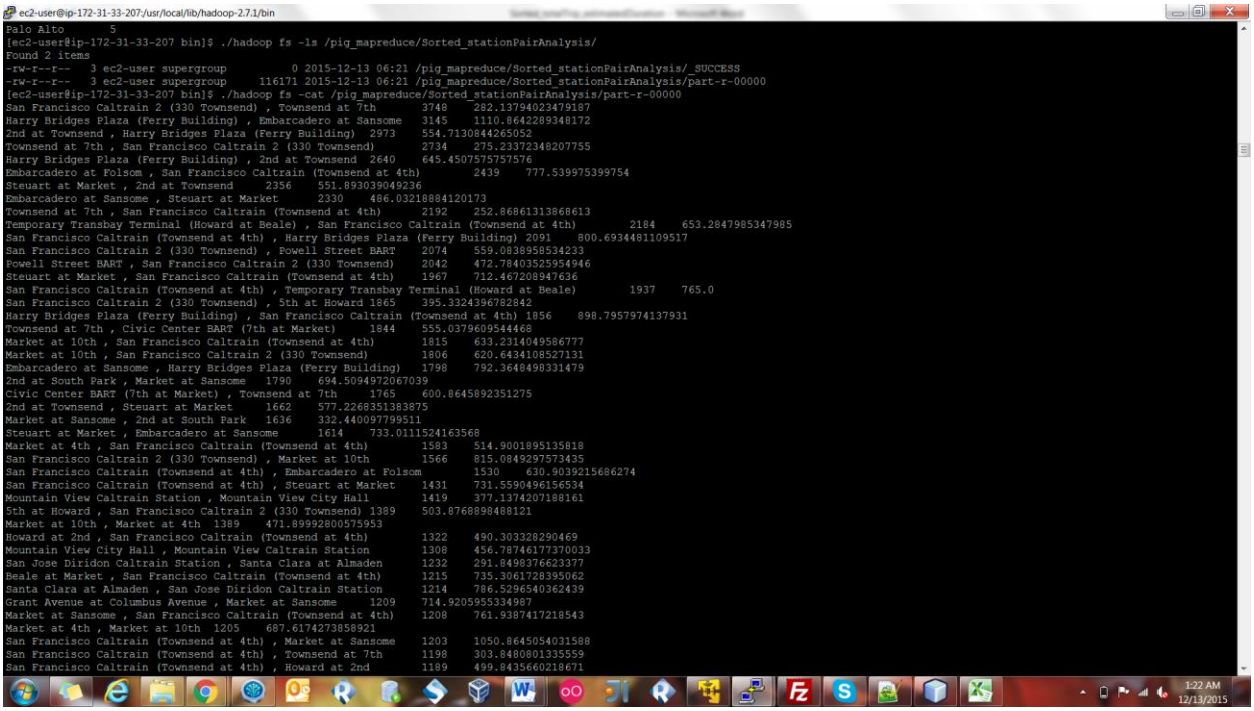
pig script: /usr/local/pig_script/Sorted_totalTrip_estimatedDuration.pig

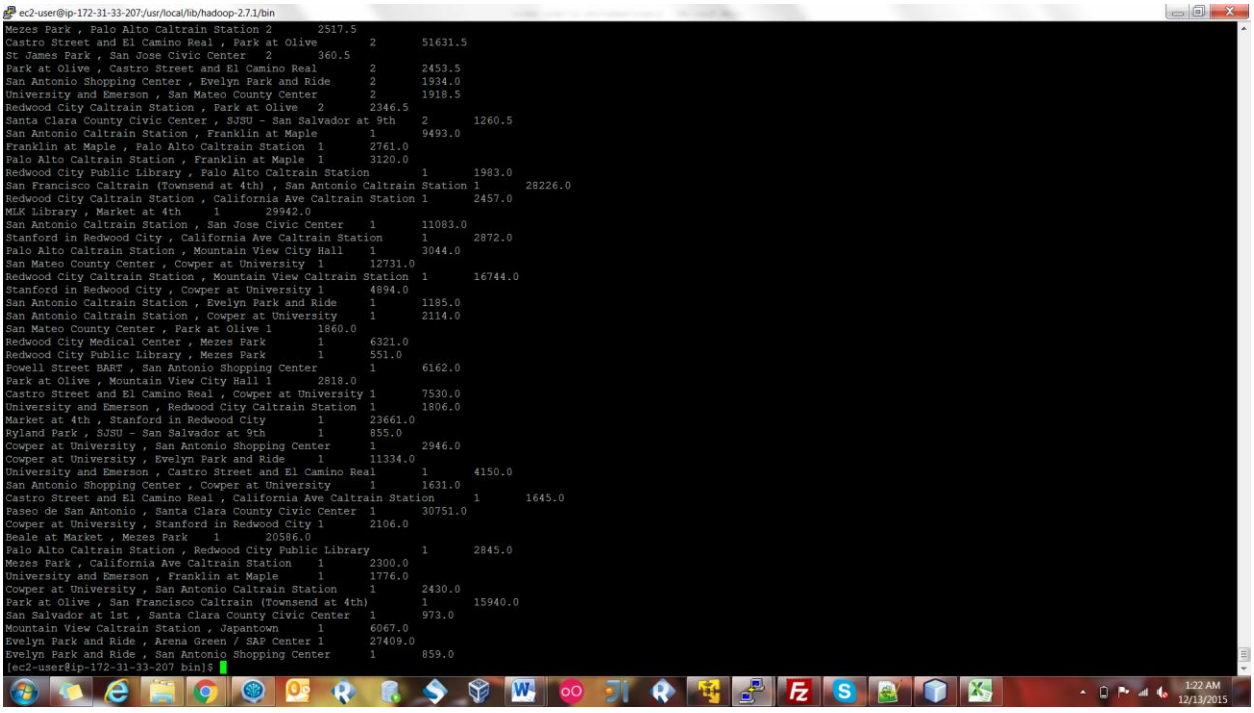
pig output: hdfs://pig_mapreduce/Sorted_stationPairAnalysis/part-r-00000

Script:

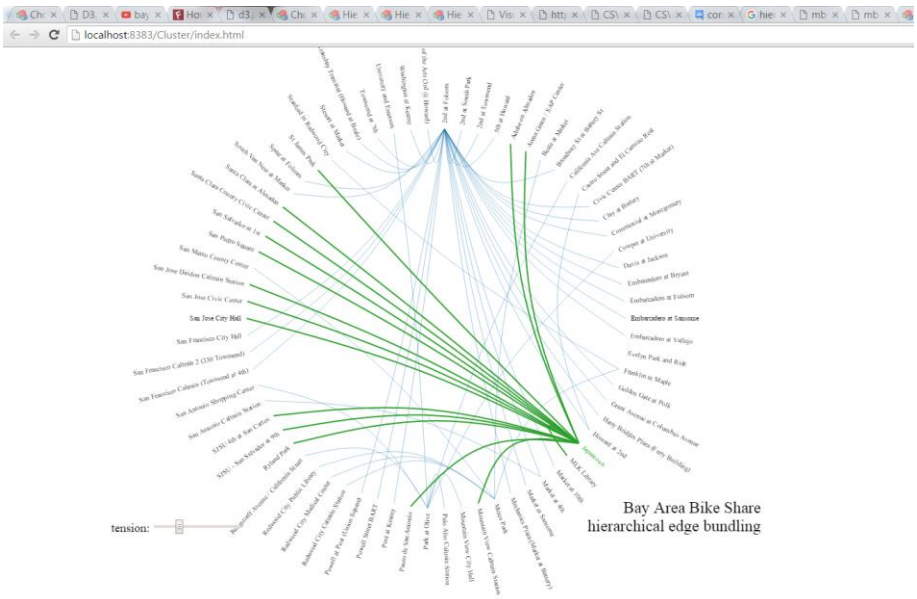
```
trip_data = LOAD 'hdfs:/input/data/201508_trip_data.csv' USING PigStorage(',') AS (trip_id:int,duration:int,start_date:chararray,start_station:chararray,start_terminal:int,end_date:chararray,end_station:chararray,end_terminal:int,bike:int,subscriber:chararray,zip:int);

group_data = foreach trip_data generate CONCAT(start_station,' , ',end_station)as stationPair,duration;

station_pairs = group group_data by stationPair;

per_pair_analysis = FOREACH station_pairs generate group , COUNT(group_data) as trip_count, AVG(group_data.duration) as estimated_time;

sorted_per_pair_analysis = ORDER per_pair_analysis by trip_count DESC;

Store sorted_per_pair_analysis into 'hdfs:/pig_mapreduce/Sorted_stationPairAnalysis';
```

Output:


ec2-user@ip-172-31-33-207:/usr/local/lib/hadoop-2.7.1/bin
Palo Alto 5
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -ls /pig_mapreduce/Sorted_stationPairAnalysis/
Found 2 items
-rw-r--r-- 3 ec2-user supergroup 0 2015-12-13 06:21 /pig_mapreduce/Sorted_stationPairAnalysis/_SUCCESS
-rw-r--r-- 3 ec2-user supergroup 116171 2015-12-13 06:21 /pig_mapreduce/Sorted_stationPairAnalysis/part-r-00000
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -cat /pig_mapreduce/Sorted_stationPairAnalysis/part-r-00000
San Francisco Caltrain 2 (330 Townsend) , Townsend at 7th 3748 282.13794023479187
Harry Bridges Plaza (Ferry Building) , Embarcadero at Sansome 3145 1110.8642289348172
2nd at Townsend , Harry Bridges Plaza (Ferry Building) 2973 554.7130844265052
Townsend at 7th , San Francisco Caltrain 2 (330 Townsend) 2734 275.23372348207755
Harry Bridges Plaza (Ferry Building) , 2nd at Townsend 2640 645.4507575757576
Embarcadero at Folsom , San Francisco Caltrain (Townsend at 4th) 2439 777.539975399754
Steuart at Market , 2nd at Townsend 2356 551.893039049236
Embarcadero at Sansome , Steuart at Market 2330 486.03218884120173
Townsend at 7th , San Francisco Caltrain (Townsend at 4th) 2192 252.86861313868613
Temporary Transbay Terminal (Howard at Beale) , San Francisco Caltrain (Townsend at 4th) 2184 653.2847985347985
San Francisco Caltrain (Townsend at 4th) , Harry Bridges Plaza (Ferry Building) 2091 800.6934481109517
San Francisco Caltrain 2 (330 Townsend) , Powell Street BART 2074 559.0838958534233
Powell Street BART , San Francisco Caltrain 2 (330 Townsend) 2042 472.78403525954946
Steuart at Market , San Francisco Caltrain (Townsend at 4th) 1967 712.467208947636
San Francisco Caltrain (Townsend at 4th) , Temporary Transbay Terminal (Howard at Beale) 1937 765.0
San Francisco Caltrain 2 (330 Townsend) , 5th at Howard 1865 395.3324396782842
Harry Bridges Plaza (Ferry Building) , San Francisco Caltrain (Townsend at 4th) 1856 898.7957974137931
Townsend at 7th , Civic Center BART (7th at Market) 1844 555.0379609544468
Market at 10th , San Francisco Caltrain (Townsend at 4th) 1815 633.2314049586777
Market at 10th , San Francisco Caltrain 2 (330 Townsend) 1806 620.6434108527131
Embarcadero at Sansome , Harry Bridges Plaza (Ferry Building) 1798 792.3648498331479
2nd at South Park , Market at Sansome 1790 694.5094972067039
Civic Center BART (7th at Market) , Townsend at 7th 1765 600.8645892351275
2nd at Townsend , Steuart at Market 1662 577.2268351383875
Market at Sansome , 2nd at South Park 1636 332.440097799511
Steuart at Market , Embarcadero at Sansome 1614 733.0111524163568
Market at 4th , San Francisco Caltrain (Townsend at 4th) 1583 514.9001895135818
San Francisco Caltrain 2 (330 Townsend) , Market at 10th 1566 815.0849297573435
San Francisco Caltrain (Townsend at 4th) , Embarcadero at Folsom 1530 630.9039215686274
San Francisco Caltrain (Townsend at 4th) , Steuart at Market 1431 731.5590496156534
Mountain View Caltrain Station , Mountain View City Hall 1419 377.1374207188161
5th at Howard , San Francisco Caltrain 2 (330 Townsend) 1389 503.8768898488121
Market at 10th , Market at 4th 1389 471.89992800575953
Howard at 2nd , San Francisco Caltrain (Townsend at 4th) 1322 490.303328290469
Mountain View City Hall , Mountain View Caltrain Station 1308 456.78746177370033
San Jose Diridon Caltrain Station , Santa Clara at Almaden 1232 291.84983766233377
Beale at Market , San Francisco Caltrain (Townsend at 4th) 1215 735.3061728395062
Santa Clara at Almaden , San Jose Diridon Caltrain Station 1214 786.5296540362439
Grant Avenue at Columbus Avenue , Market at Sansome 1209 714.9205955334987
Market at Sansome , San Francisco Caltrain (Townsend at 4th) 1208 761.9387417218543
Market at 4th , Market at 10th 1205 687.617427385892l
San Francisco Caltrain (Townsend at 4th) , Market at Sansome 1203 1050.86450540331588
San Francisco Caltrain (Townsend at 4th) , Townsend at 7th 1198 303.8480801335559
San Francisco Caltrain (Townsend at 4th) , Howard at 2nd 1189 499.8435660218671
1:22 AM
12/13/2015


ec2-user@ip-172-31-33-207:/usr/local/lib/hadoop-2.7.1/bin
Mezes Park , Palo Alto Caltrain Station 2 2517.5
Castro Street and El Camino Real , Park at Olive 2 51631.5
St James Park , San Jose Civic Center 2 360.5
Park at Olive , Castro Street and El Camino Real 2 2453.5
San Antonio Shopping Center , Evelyn Park and Ride 2 1934.0
University and Emerson , San Mateo County Center 2 1918.5
Redwood City Caltrain Station , Park at Olive 2 2346.5
Santa Clara County Civic Center , SJSU - San Salvador at 9th 2 1260.5
San Antonio Caltrain Station , Franklin at Maple 1 9493.0
Franklin at Maple , Palo Alto Caltrain Station 1 2761.0
Palo Alto Caltrain Station , Franklin at Maple 1 3120.0
Redwood City Public Library , Palo Alto Caltrain Station 1 1983.0
San Francisco Caltrain (Townsend at 4th) , San Antonio Caltrain Station 1 28226.0
Redwood City Caltrain Station , California Ave Caltrain Station 1 2457.0
MLK Library , Market at 4th 1 29942.0
San Antonio Caltrain Station , San Jose Civic Center 1 11083.0
Stanford in Redwood City , California Ave Caltrain Station 1 2872.0
Palo Alto Caltrain Station , Mountain View City Hall 1 3044.0
San Mateo County Center , Cowper at University 1 12731.0
Redwood City Caltrain Station , Mountain View Caltrain Station 1 16744.0
Stanford in Redwood City , Cowper at University 1 4894.0
San Antonio Caltrain Station , Evelyn Park and Ride 1 1185.0
San Antonio Caltrain Station , Cowper at University 1 2114.0
San Mateo County Center , Park at Olive 1 1860.0
Redwood City Medical Center , Mezes Park 1 6321.0
Redwood City Public Library , Mezes Park 1 551.0
Powell Street BART , San Antonio Shopping Center 1 6162.0
Park at Olive , Mountain View City Hall 1 2818.0
Castro Street and El Camino Real , Cowper at University 1 7530.0
University and Emerson , Redwood City Caltrain Station 1 1806.0
Market at 4th , Stanford in Redwood City 1 23661.0
Ryland Park , SJSU - San Salvador at 9th 1 855.0
Cowper at University , San Antonio Shopping Center 1 2946.0
Cowper at University , Evelyn Park and Ride 1 11334.0
University and Emerson , Castro Street and El Camino Real 1 4150.0
San Antonio Shopping Center , Cowper at University 1 1631.0
Castro Street and El Camino Real , California Ave Caltrain Station 1 1645.0
Paseo de San Antonio , Santa Clara County Civic Center 1 30751.0
Cowper at University , Stanford in Redwood City 1 2106.0
Beale at Market , Mezes Park 1 20586.0
Palo Alto Caltrain Station , Redwood City Public Library 1 2845.0
Mezes Park , California Ave Caltrain Station 1 2300.0
University and Emerson , Franklin at Maple 1 1776.0
Cowper at University , San Antonio Caltrain Station 1 2430.0
Park at Olive , San Francisco Caltrain (Townsend at 4th) 1 15940.0
San Salvador at 1st , Santa Clara County Civic Center 1 973.0
Mountain View Caltrain Station , Japantown 1 6067.0
Evelyn Park and Ride , Arena Green / SAP Center 1 27409.0
Evelyn Park and Ride , San Antonio Shopping Center 1 859.0
[ec2-user@ip-172-31-33-207 bin]$
1:22 AM
12/13/2015

Graph:



---------- **Analysis : 5 UDF** ----------

Analysis:

Timely (Hourly) analysis of total docks available (average) on particular stations (station id =2): tells how many bikes are taken from station. Can be used to analyze peak time when customer uses bike services.

Used UDF to convert date to each hour.

Result: Maximum bike taken at 9 and 10 am

UDF Jar: /usr/local/java_script/Hourly_TripStatus/sne.jar
UDF Java Class: /usr/local/java_script/Hourly_TripStatus/ConvertDate.class
UDF script For (station id =2): /usr/local/java_script/Hourly_TripStatus/bikeUsage_station2.pig

Output at hdfs:/pig_mapreduce/Hourly_TripStatus_Station2

Script:

-- myscript.pig

REGISTER /usr/local/java_script/Hourly_TripStatus/sne.jar;

status_dummy_data = LOAD 'hdfs:/input/data/201508_status_data.csv' USING PigStorage(',') AS (station_id:int,bikes_avail:int,dock_avail:int,time:chararray);

status_data = FILTER status_dummy_data BY station_id ==2;

C = FOREACH status_data GENERATE sne.ConvertDate(time) as date, dock_avail;

D = group C by date;

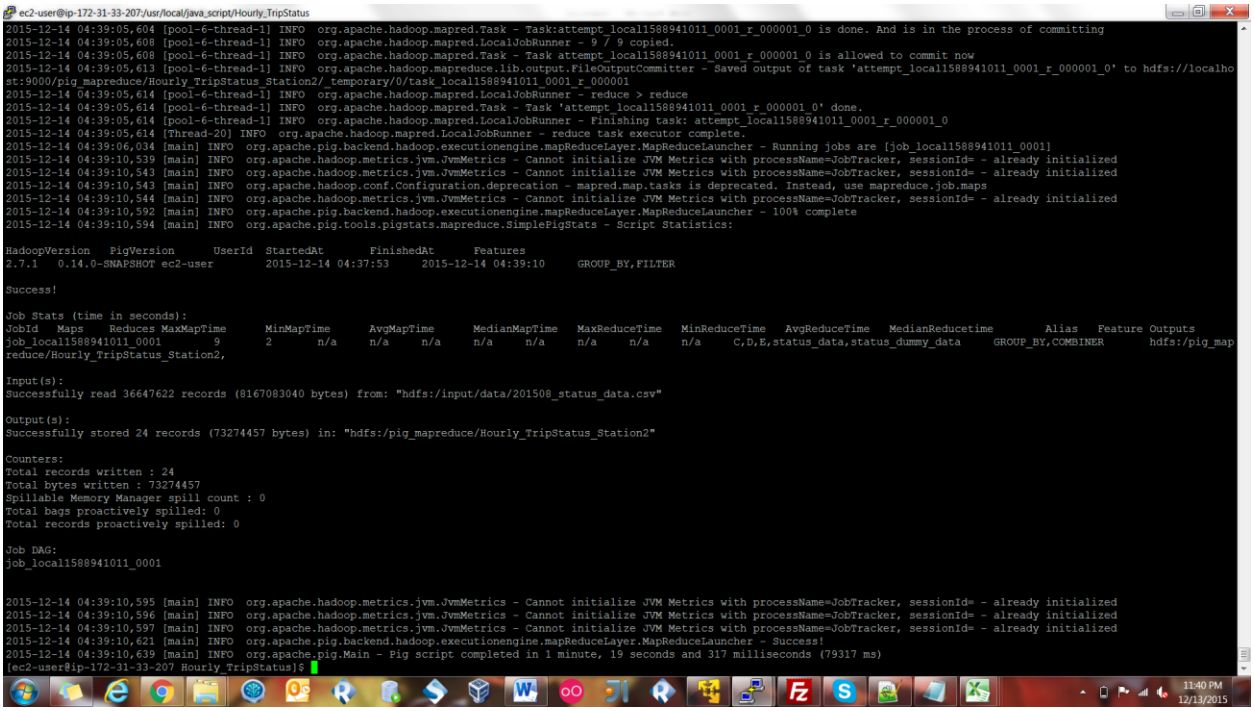
E = FOREACH D generate group, AVG(C.dock_avail) as total_rented_bikes;

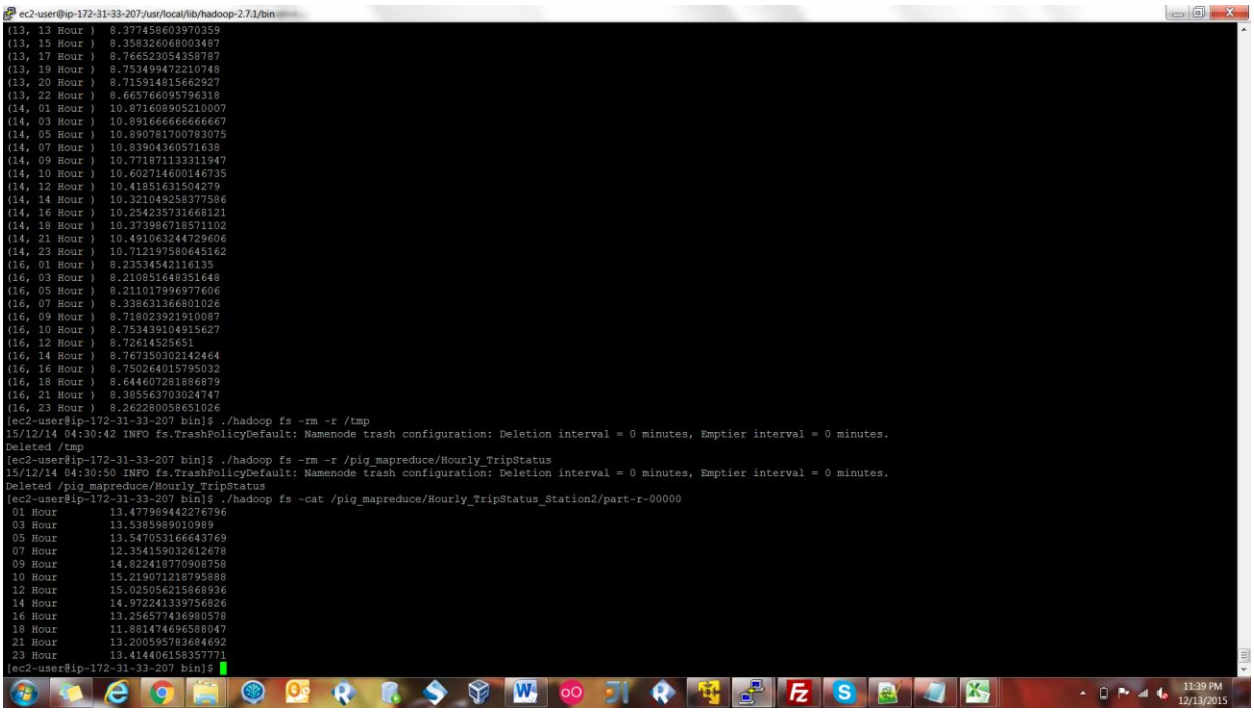
STORE E INTO 'hdfs:/pig_mapreduce/Hourly_TripStatus_Station2';

UDF class to get hours from a date:

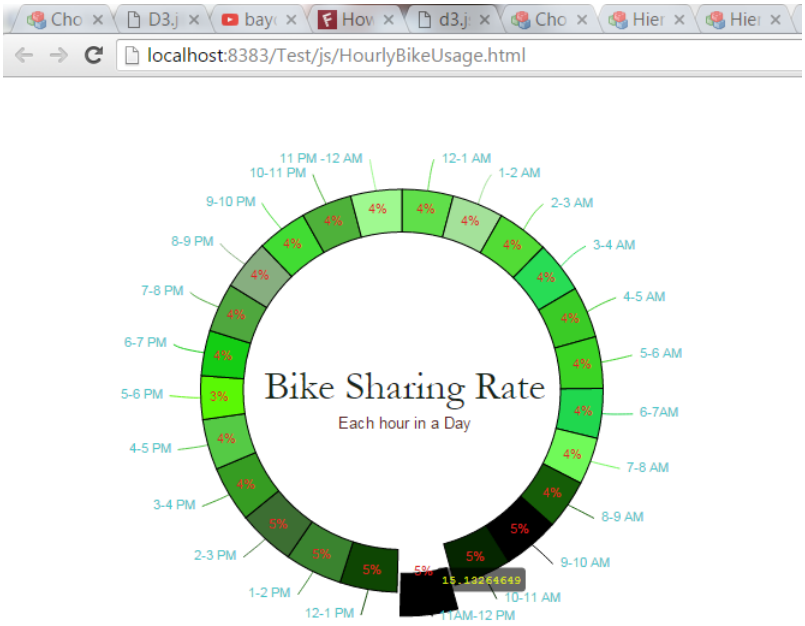
```
- ConvertDate.java-----------------------------------------------------------------------------------------

package sne;

import java.io.IOException;

import org.apache.pig.EvalFunc;

import org.apache.pig.data.Tuple;


public class ConvertDate extends EvalFunc<String>

{

    public String exec(Tuple input) throws IOException {

        if (input == null || input.size() == 0)

            return null;

        try{

            String str = (String)input.get(0);

            str = str.substring(11,str.length()-7);

            str = str+" Hour ";

           return str;

        }catch(Exception e){

            throw new IOException("Caught exception processing input row ", e);

        }

    }

  }
```

Output:


ec2-user@ip-172-31-33-207:/usr/local/java_script/Hourly_TripStatus
2015-12-14 04:39:05,604 [pool-6-thread-1] INFO org.apache.hadoop.mapred.Task - Task:attempt_local1588941011_0001_r_000001_0 is done. And is in the process of committing
2015-12-14 04:39:05,608 [pool-6-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - 9 / 9 copied.
2015-12-14 04:39:05,608 [pool-6-thread-1] INFO org.apache.hadoop.mapred.Task - Task attempt_local1588941011_0001_r_000001_0 is allowed to commit now
2015-12-14 04:39:05,613 [pool-6-thread-1] INFO org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task 'attempt_local1588941011_0001_r_000001_0' to hdfs://localhost:9000/pig_mapreduce/Hourly_TripStatus_Station2/_temporary/0/task_local1588941011_0001_r_000001
2015-12-14 04:39:05,614 [pool-6-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2015-12-14 04:39:05,614 [pool-6-thread-1] INFO org.apache.hadoop.mapred.Task - Task 'attempt_local1588941011_0001_r_000001_0' done.
2015-12-14 04:39:05,614 [pool-6-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local1588941011_0001_r_000001_0
2015-12-14 04:39:05,614 [Thread-20] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce task executor complete.
2015-12-14 04:39:06,034 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Running jobs are [job_local1588941011_0001]
2015-12-14 04:39:10,539 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-14 04:39:10,543 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-14 04:39:10,543 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
2015-12-14 04:39:10,544 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-14 04:39:10,592 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2015-12-14 04:39:10,594 [main] INFO org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.7.1 0.14.0-SNAPSHOT ec2-user 2015-12-14 04:37:53 2015-12-14 04:39:10 GROUP_BY,FILTER
Success!
Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReducetime Alias Feature Outputs
job_local1588941011_0001 9 2 n/a n/a n/a n/a n/a n/a n/a C,D,E,status_data,status_dummy_data GROUP_BY,COMBINER hdfs:/pig_mapreduce/Hourly_TripStatus_Station2,
Input(s):
Successfully read 36647622 records (8167083040 bytes) from: "hdfs:/input/data/201508_status_data.csv"
Output(s):
Successfully stored 24 records (73274457 bytes) in: "hdfs:/pig_mapreduce/Hourly_TripStatus_Station2"
Counters:
Total records written : 24
Total bytes written : 73274457
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
Job DAG:
job_local1588941011_0001
2015-12-14 04:39:10,595 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-14 04:39:10,596 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-14 04:39:10,597 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-14 04:39:10,621 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2015-12-14 04:39:10,639 [main] INFO org.apache.pig.Main - Pig script completed in 1 minute, 19 seconds and 317 milliseconds (79317 ms)
[ec2-user@ip-172-31-33-207 Hourly_TripStatus]$
11:40 PM
12/13/2015


ec2-user@ip-172-31-33-207:/usr/local/lib/hadoop-2.7.1/bin
(13, 13 Hour) 8.377458603970359
(13, 15 Hour) 8.358326068003487
(13, 17 Hour) 8.766523054358787
(13, 19 Hour) 8.753499472210748
(13, 20 Hour) 8.715914815662927
(13, 22 Hour) 8.665766095796318
(14, 01 Hour) 10.871608905210007
(14, 03 Hour) 10.891666666666667
(14, 05 Hour) 10.890781700783075
(14, 07 Hour) 10.83904360571638
(14, 09 Hour) 10.771871133311947
(14, 10 Hour) 10.602714600146735
(14, 12 Hour) 10.418516331504279
(14, 14 Hour) 10.321049258377586
(14, 16 Hour) 10.254235731668121
(14, 18 Hour) 10.373986718571102
(14, 21 Hour) 10.491063244729606
(14, 23 Hour) 10.712197580645162
(16, 01 Hour) 8.23534542116135
(16, 03 Hour) 8.210851648351648
(16, 05 Hour) 8.211017996977606
(16, 07 Hour) 8.338631366801026
(16, 09 Hour) 8.718023921910087
(16, 10 Hour) 8.753439104915627
(16, 12 Hour) 8.72614525651
(16, 14 Hour) 8.767350302142464
(16, 16 Hour) 8.750264015795032
(16, 18 Hour) 8.644607281886879
(16, 21 Hour) 8.385563703024747
(16, 23 Hour) 8.262280058651026
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -rm -r /tmp
15/12/14 04:30:42 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /tmp
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -rm -r /pig_mapreduce/Hourly_TripStatus
15/12/14 04:30:50 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /pig_mapreduce/Hourly_TripStatus
[ec2-user@ip-172-31-33-207 bin]$ ./hadoop fs -cat /pig_mapreduce/Hourly_TripStatus_Station2/part-r-00000
01 Hour 13.477989442276796
03 Hour 13.5385989010989
05 Hour 13.547053166643769
07 Hour 12.354159032612678
09 Hour 14.822418770908758
10 Hour 15.219071218795888
12 Hour 15.025056215868936
14 Hour 14.972241339756826
16 Hour 13.256577436980578
18 Hour 11.881474696588047
21 Hour 13.200595783684692
23 Hour 13.414406158357771
[ec2-user@ip-172-31-33-207 bin]$
11:39 PM
12/13/2015

Graph:



HIVE:

---------- **Analysis : 6** ----------

Creating and Loading data to HIVE:

Copy data at hdfs : ./hadoop fs -put /usr/local/data/ /input_TO_HIVE

Status table:

Create table to hive :
create table status_data(station_id int, bikes_available int, docks_available int, time String) row format delimited fields terminated by ',' stored as textfile;

Load data: LOAD DATA INPATH 'hdfs:/input_TO_HIVE/data/201508_status_data.csv' OVERWRITE INTO TABLE status_data;

Select data : select * from status_data;

```
Time taken: 0.2 seconds
hive> create table status_data(station_id int, bikes_available int, docks_available int, time timestamp) row format delimited fields terminated by ',' stored as textfile;
FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.DDLTask. AlreadyExistsException(message:Table status_data already exists)
hive> LOAD DATA INPATH 'hdfs:/input_TO_HIVE/data/201508_status_data.csv' OVERWRITE INTO TABLE status_data;
FAILED: SemanticException Line 1:17 Invalid path ''hdfs:/input_TO_HIVE/data/201508_status_data.csv'': No files matching path hdfs://localhost:9000/input_TO_HIVE/data/201508_status_data.csv
hive> LOAD DATA INPATH 'hdfs:/input_TO_HIVE/data/201508_status_data.csv' OVERWRITE INTO TABLE status_data;
Loading data to table default.status_data
Table default.status_data stats: [numFiles=1, numRows=0, totalSize=1087241884, rawDataSize=0]
OK
Time taken: 0.326 seconds
hive> select * from status_data
    > ;
Query ID = root_20151214061212_33cd63dd-1fa4-48d1-878d-bc733d4e7222
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Job running in-process (local Hadoop)
Hadoop job information for Stage-1: number of mappers: 0; number of reducers: 0
2015-12-14 06:12:03,732 Stage-1 map = 0%,  reduce = 0%
```

Weather table:

./hadoop fs -cp /input/data/201508_weather_data.csv /input_TO_HIVE

create table weather_data(PDT String,MaxTemperatureF int,MeanTemperatureF int,MinTemperatureF int,MaxDew int,MeanDew int,MinDew int,MaxHumidity int, MeanHumidity int, MinHumidity int, MaxSea int, MeanSea int, MinSea int, MaxVisibilityMiles int, MeanVisibilityMiles int, MinVisibilityMiles int, MaxWind int, MeanWind int, MaxGust int,PrecipitationIn int, CloudCover int, Events int, WindDirDegrees int,Zip int) row format delimited fields terminated by ',' stored as textfile;

LOAD DATA INPATH 'hdfs:/input_TO_HIVE/201508_weather_data.csv' OVERWRITE INTO TABLE weather_data;

```
Time taken: 2.23 seconds, Fetched: 1 row(s)
hive> create table weather_data(PDT String,MaxTemperatureF int,MeanTemperatureF int,MinTemperatureF int,MaxDew int,MeanDew int,MinDew int,MaxHumidity int, MeanHumidity int, MinHumidity int,
 MaxSea int, MeanSea int, MinSea int, MaxVisibilityMiles int, MeanVisibilityMiles int, MinVisibilityMiles int, MaxWind int, MeanWind int, MaxGust int,PrecipitationIn int, CloudCover int, Ev
ents int, WindDirDegrees int,Zip int) row format delimited fields terminated by ',' stored as textfile;
OK
Time taken: 0.253 seconds
hive> LOAD DATA INPATH 'hdfs:/input_TO_HIVE/data/201508_weather_data.csv' OVERWRITE INTO TABLE weather_data;
FAILED: SemanticException Line 1:17 Invalid path ''hdfs:/input_TO_HIVE/data/201508_weather_data.csv'': No files matching path hdfs://localhost:9000/input_TO_HIVE/data/201508_weather_data.cs
v
hive> LOAD DATA INPATH 'hdfs:/input_TO_HIVE/201508_weather_data.csv' OVERWRITE INTO TABLE weather_data;
Loading data to table default.weather_data
Table default.weather_data stats: [numFiles=1, numRows=0, totalSize=158238, rawDataSize=0]
OK
Time taken: 0.372 seconds
hive>
```

Trip_data

./hadoop fs -cp /input/data/201508_trip_data.csv /input_TO_HIV

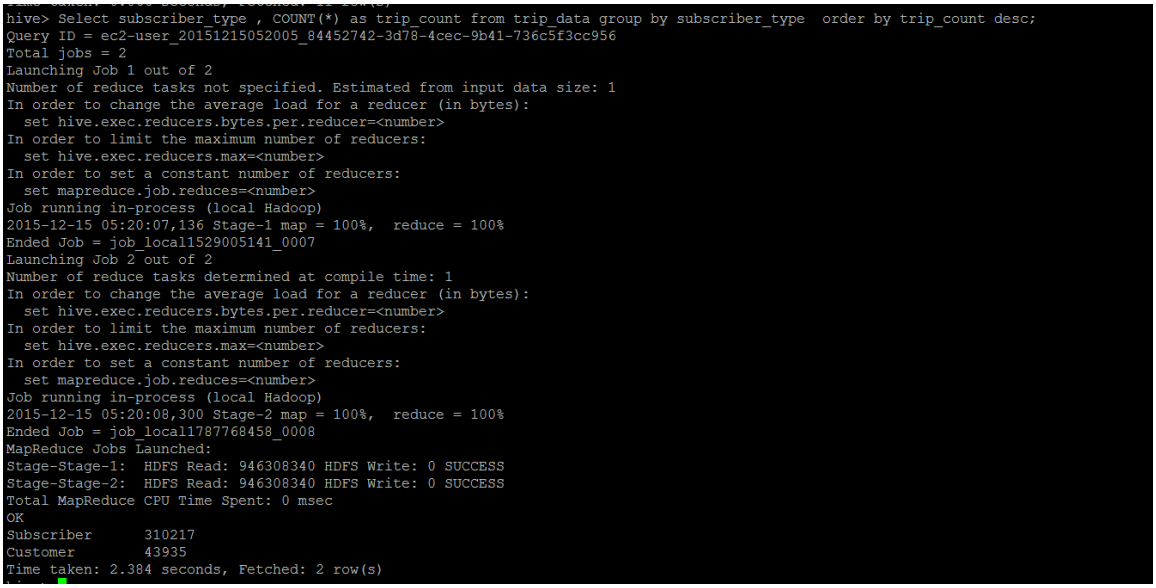
create table trip_data(Trip_ID int,Duration int,Start_Date String,Start_Station String,Start_Terminal String,End_Date String,End_Station String,End_Terminal String,Bike String,Subscriber_Type String,Zip int) row format delimited fields terminated by ',' stored as textfile;

LOAD DATA INPATH 'hdfs:/input_TO_HIVE/201508_trip_data.csv' OVERWRITE INTO TABLE trip_data;

```
hive> create table trip_data(Trip_ID int,Duration int,Start_Date String,Start_Station String,Start_Terminal String,End_Date String,End_Station String,End_Terminal String,Bike String,Subscri
ber_Type String,Zip int) row format delimited fields terminated by ',' stored as textfile;
OK
Time taken: 0.052 seconds
hive> LOAD DATA INPATH 'hdfs:/input_TO_HIVE/201508_trip_data.csv' OVERWRITE INTO TABLE trip_data;
FAILED: SemanticException Line 1:17 Invalid path ''hdfs:/input_TO_HIVE/201508_trip_data.csv'': No files matching path hdfs://localhost:9000/input_TO_HIVE/201508_trip_data.csv
hive> LOAD DATA INPATH 'hdfs:/input_TO_HIV/201508_trip_data.csv' OVERWRITE INTO TABLE trip_data;
FAILED: SemanticException Line 1:17 Invalid path ''hdfs:/input_TO_HIV/201508_trip_data.csv'': No files matching path hdfs://localhost:9000/input_TO_HIV/201508_trip_data.csv
hive> LOAD DATA INPATH 'hdfs:/input_TO_HIVE/201508_trip_data.csv' OVERWRITE INTO TABLE trip_data;
Loading data to table default.trip_data
Table default.trip_data stats: [numFiles=1, numRows=0, totalSize=43012526, rawDataSize=0]
OK
Time taken: 0.207 seconds
hive>
```

---------- **Analysis : 7** ----------

Ratio of subscriber to customer in trip data:



Query :

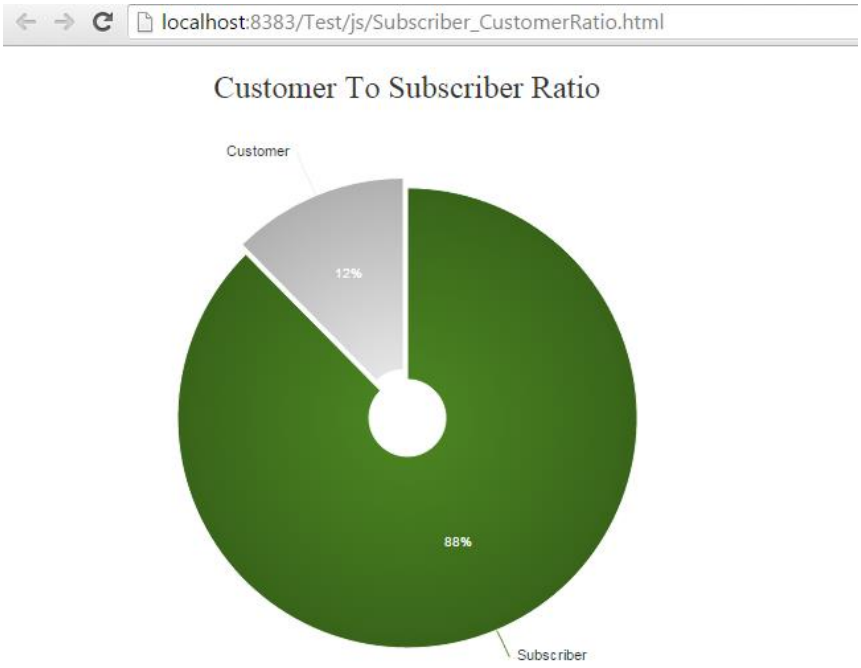
Select subscriber_type , COUNT(*) as trip_count from trip_data group by subscriber_type order by trip_count desc;

Output:

Subscriber 310217

Customer 43935

Graph:



---------- **Analysis : 8** ----------

Top 10 zip of customer type by total_trip:

```
hive> Select zip, subscriber_type ,  COUNT(*) as trip_count from trip_data group by zip,subscriber_type  order by trip_count desc limit 10;
Query ID = ec2-user_20151215052447_99ea8c0a-2329-4bfd-bcfa-3e4e9389460c
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Job running in-process (local Hadoop)
2015-12-15 05:24:48,460 Stage-1 map = 100%,  reduce = 100%
Ended Job = job_local2059511318_0011
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Job running in-process (local Hadoop)
2015-12-15 05:24:49,612 Stage-2 map = 100%,  reduce = 100%
Ended Job = job_local1343277069_0012
MapReduce Jobs Launched:
Stage-Stage-1:  HDFS Read: 1118358444 HDFS Write: 0 SUCCESS
Stage-Stage-2:  HDFS Read: 1118358444 HDFS Write: 0 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
94107   Subscriber      45819
94105   Subscriber      19915
94133   Subscriber      15868
94103   Subscriber      14446
94111   Subscriber      10697
94102   Subscriber      9696
NULL    Customer        6929
94109   Subscriber      6029
95112   Subscriber      4488
94403   Subscriber      4124
Time taken: 2.363 seconds, Fetched: 10 row(s)
```

Query:

Select zip, subscriber_type , COUNT(*) as trip_count from trip_data group by zip,subscriber_type order by trip_count desc limit 10;

Output:

94107 Subscriber 45819

94105 Subscriber 19915

94133 Subscriber 15868

94103 Subscriber 14446

94111 Subscriber 10697

94102 Subscriber 9696

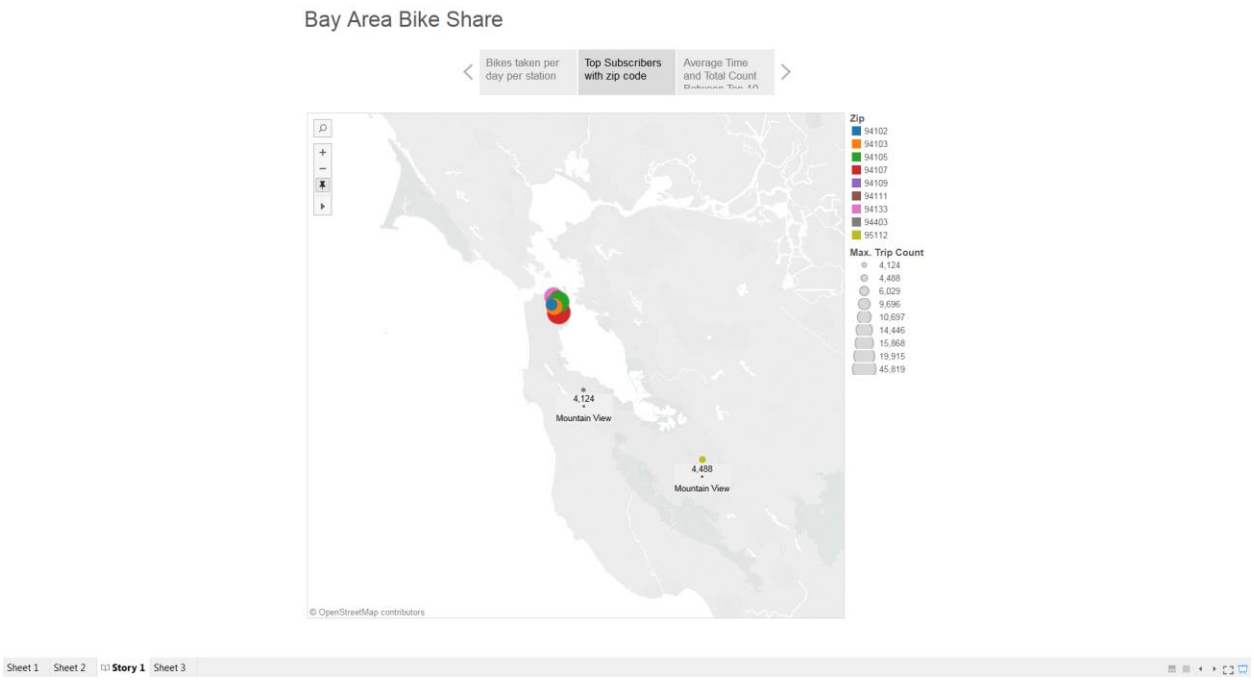
NULL Customer 6929

94109 Subscriber 6029

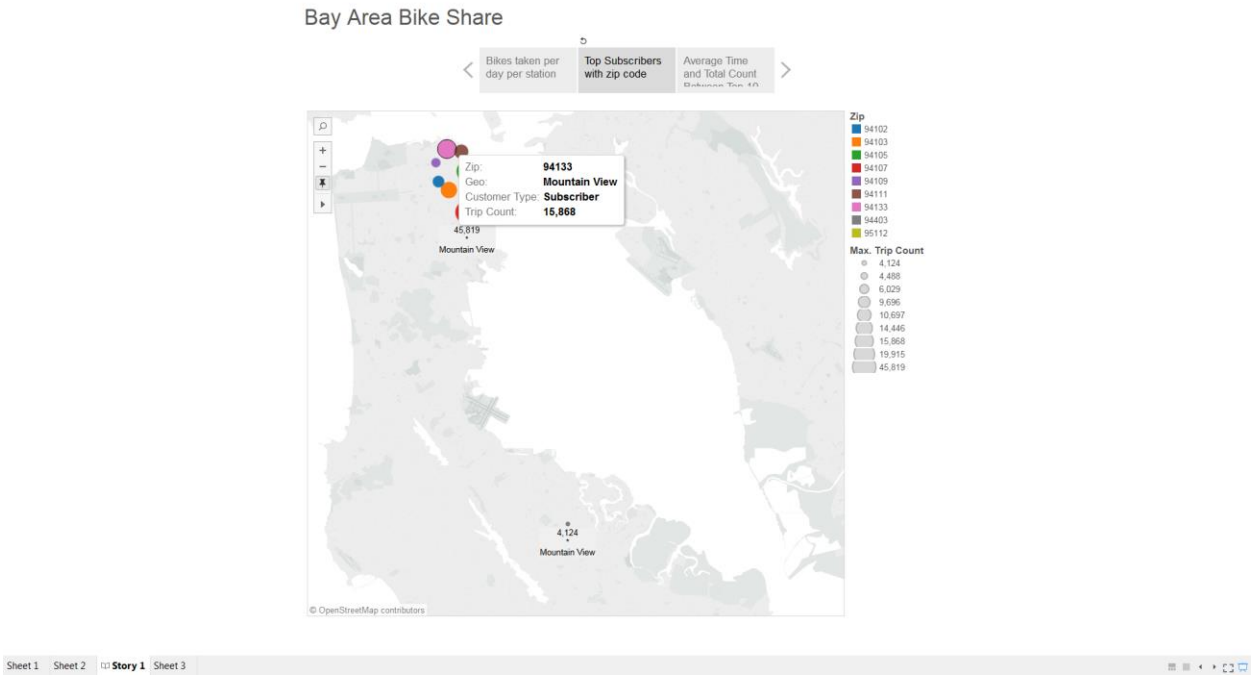
95112 Subscriber 4488

94403 Subscriber 4124

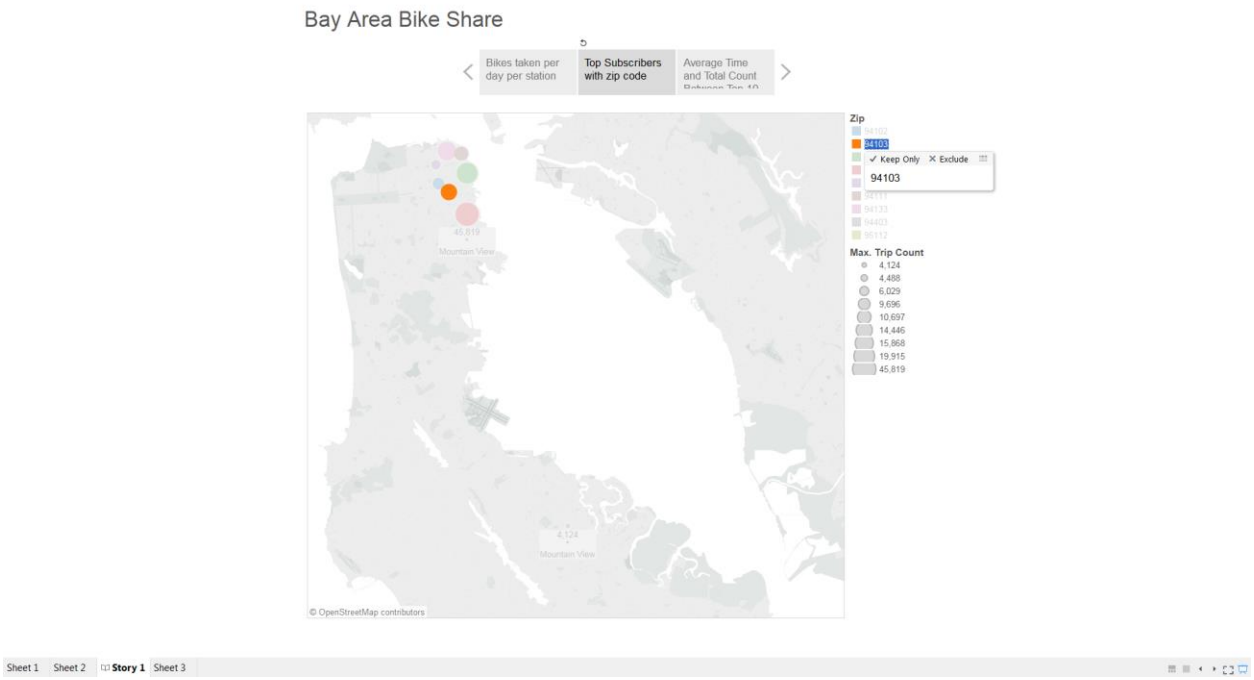
Graph:

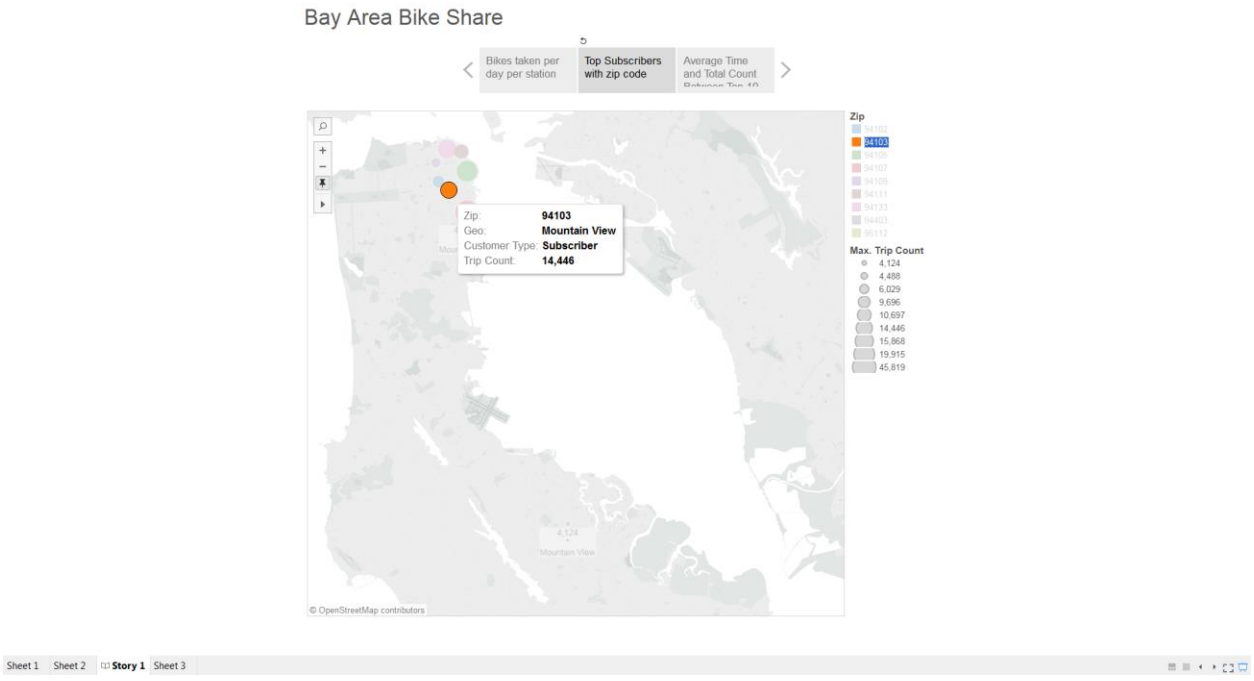

Bay Area Bike Share
Bikes taken per day per station
Top Subscribers with zip code
Average Time and Total Count
Zip
94102
94103
94105
94107
94109
94111
94133
94403
95112
Max. Trip Count
4,124
4,488
6,029
9,696
10,697
14,446
15,868
19,915
45,819
4,124
Mountain View
4,488
Mountain View
© OpenStreetMap contributors
Sheet 1
Sheet 2
Story 1
Sheet 3

Zoom

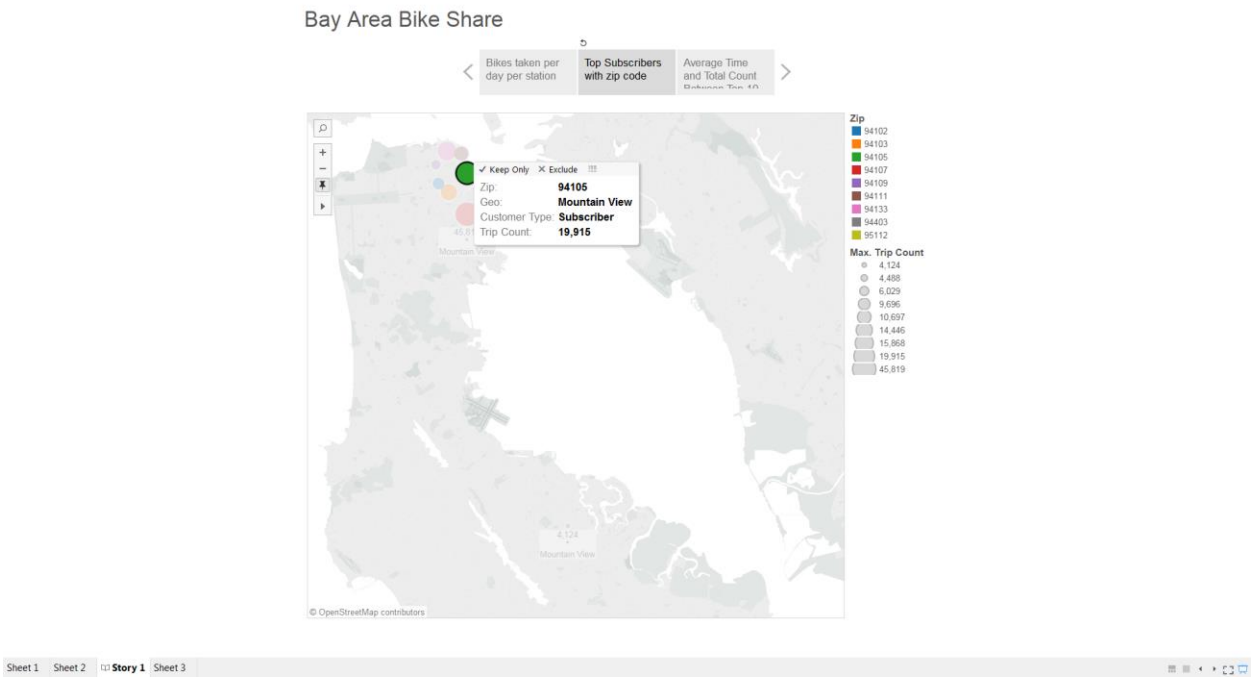

Bay Area Bike Share
Bikes taken per day per station
Top Subscribers with zip code
Average Time and Total Count
Zip: 94133
Geo: Mountain View
Customer Type: Subscriber
Trip Count: 15,868
45,819
Mountain View
Zip
94102
94103
94105
94107
94109
94111
94133
94403
95112
Max. Trip Count
4,124
4,488
6,029
9,696
10,697
14,446
15,868
19,915
45,819
4,124
Mountain View
© OpenStreetMap contributors
Sheet 1
Sheet 2
Story 1
Sheet 3

Select Zip:


Bay Area Bike Share
Bikes taken per day per station
Top Subscribers with zip code
Average Time and Total Count
Zip
94103
Keep Only
Exclude
94103
Max. Trip Count
4,124
4,488
6,029
9,696
10,697
14,446
15,868
19,915
45,819
© OpenStreetMap contributors
Sheet 1
Sheet 2
Story 1
Sheet 3


Bay Area Bike Share
Bikes taken per day per station
Top Subscribers with zip code
Average Time and Total Count
Zip
94103
Zip: 94103
Geo: Mountain View
Customer Type: Subscriber
Trip Count: 14,446
Max. Trip Count
4,124
4,488
6,029
9,696
10,697
14,446
15,868
19,915
45,819
© OpenStreetMap contributors
Sheet 1
Sheet 2
Story 1
Sheet 3

Joined with station table for landmark name:



---------- **Analysis : 9** ----------

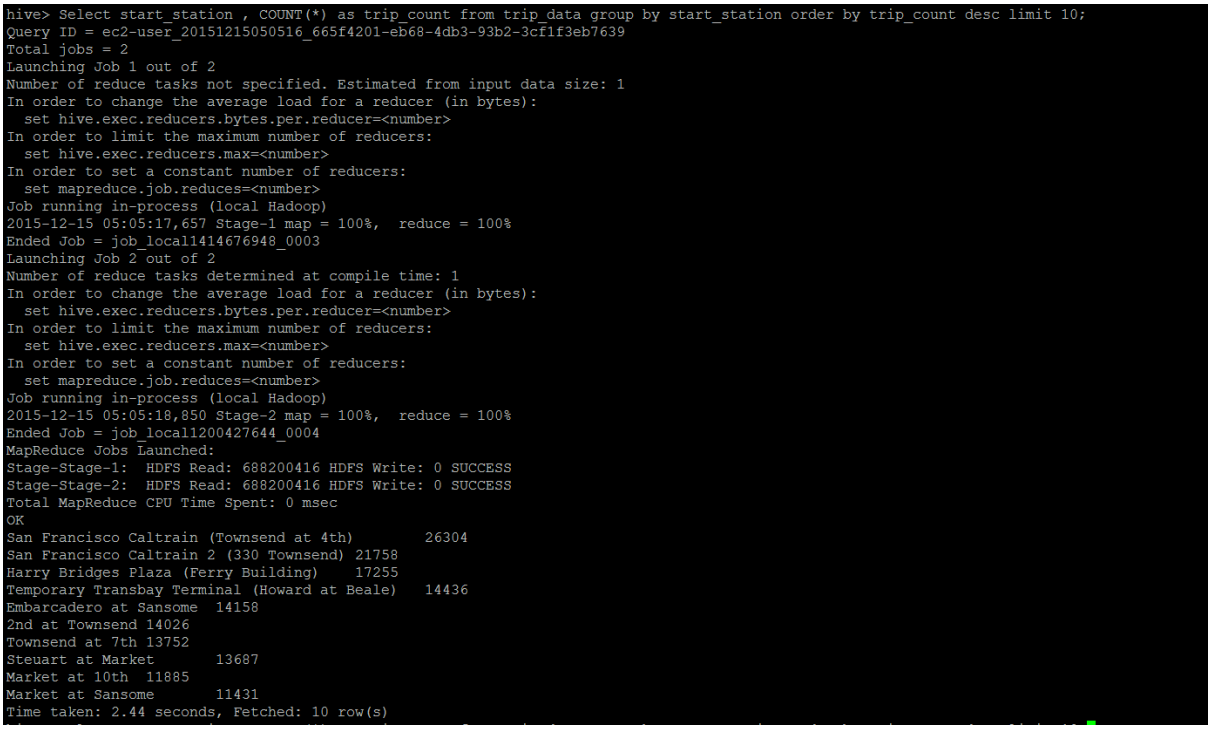
Analysis:

Most used start terminal with total trip count

Query:

Select start_station , COUNT(*) as trip_count from trip_data group by start_station order by trip_count desc limit 10;

SS:



Output:

San Francisco Caltrain (Townsend at 4th) 26304

San Francisco Caltrain 2 (330 Townsend) 21758

Harry Bridges Plaza (Ferry Building) 17255

Temporary Transbay Terminal (Howard at Beale) 14436

Embarcadero at Sansome 14158

2nd at Townsend 14026

Townsend at 7th 13752

Steuart at Market 13687

Market at 10th 11885

Market at Sansome 11431

HBASE:

---------- **Analysis : 10 Load data into HBASE using PIG script** ----------

Start HBASE:

```
[ec2-user@ip-172-31-33-207 bin]$ jps
12078 NameNode
12214 DataNode
12560 ResourceManager
11322 HQuorumPeer
11431 HMaster
13039 JobHistoryServer
12672 NodeManager
11568 HRegionServer
13364 Main
12393 SecondaryNameNode
13560 Jps
[ec2-user@ip-172-31-33-207 bin]$
```

Create table 'trip_data':

create 'trip_data','trip_info','start_details','end_details','customer_details'

```
hbase(main):004:0> create 'trip_data','trip_info','start_details','end_details','customer_details'
0 row(s) in 0.2280 seconds

=> Hbase::Table - trip_data
hbase(main):005:0> list
TABLE
trip_data
1 row(s) in 0.0070 seconds

=> ["trip_data"]
hbase(main):006:0> describe'trip_data'
Table trip_data is ENABLED
trip_data
COLUMN FAMILIES DESCRIPTION
{NAME => 'customer_details', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KE
EP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
{NAME => 'end_details', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DE
LETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
{NAME => 'start_details', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_
DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
{NAME => 'trip_info', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELE
TED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
4 row(s) in 0.0210 seconds

hbase(main):007:0>
```

Insert data to hbase using "/usr/local/pig_script/loadToHbase.pig" script:

trip table:----------

trip_data = LOAD 'hdfs:/input/data/201508_trip_data.csv' USING PigStorage(',') AS
(trip_id,duration:int,start_date:chararray,start_station:chararray,start_terminal:int,
end_date:chararray,end_station:chararray,end_terminal:int,bike:int,subscriber:chararray,zip:int);

STORE trip_data INTO 'hbase://trip_data' USING

org.apache.pig.backend.hadoop.hbase.HBaseStorage(

'trip_info:duration,

start_details:start_date,

start_details:start_station,

start_details:start_terminal,

end_details:end_date,

end_details:end_station,

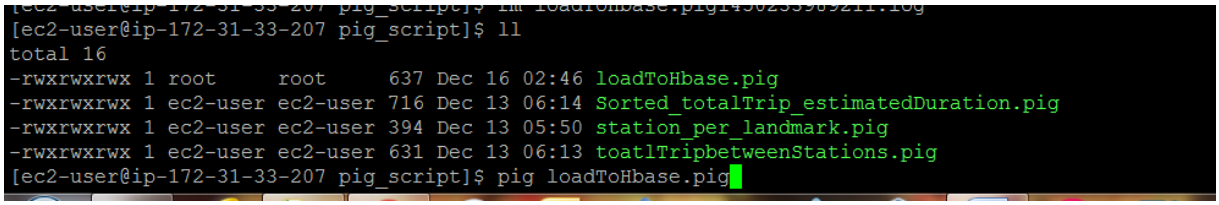
end_details:end_terminal,

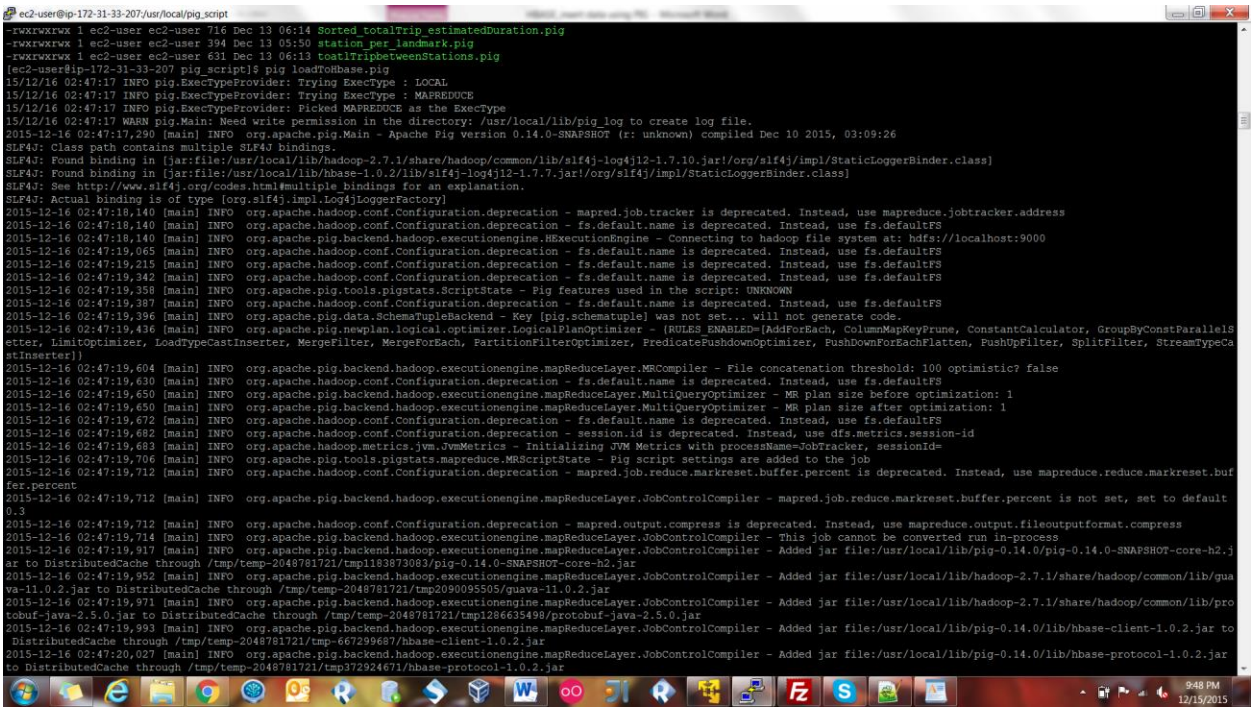
customer_details:bike,

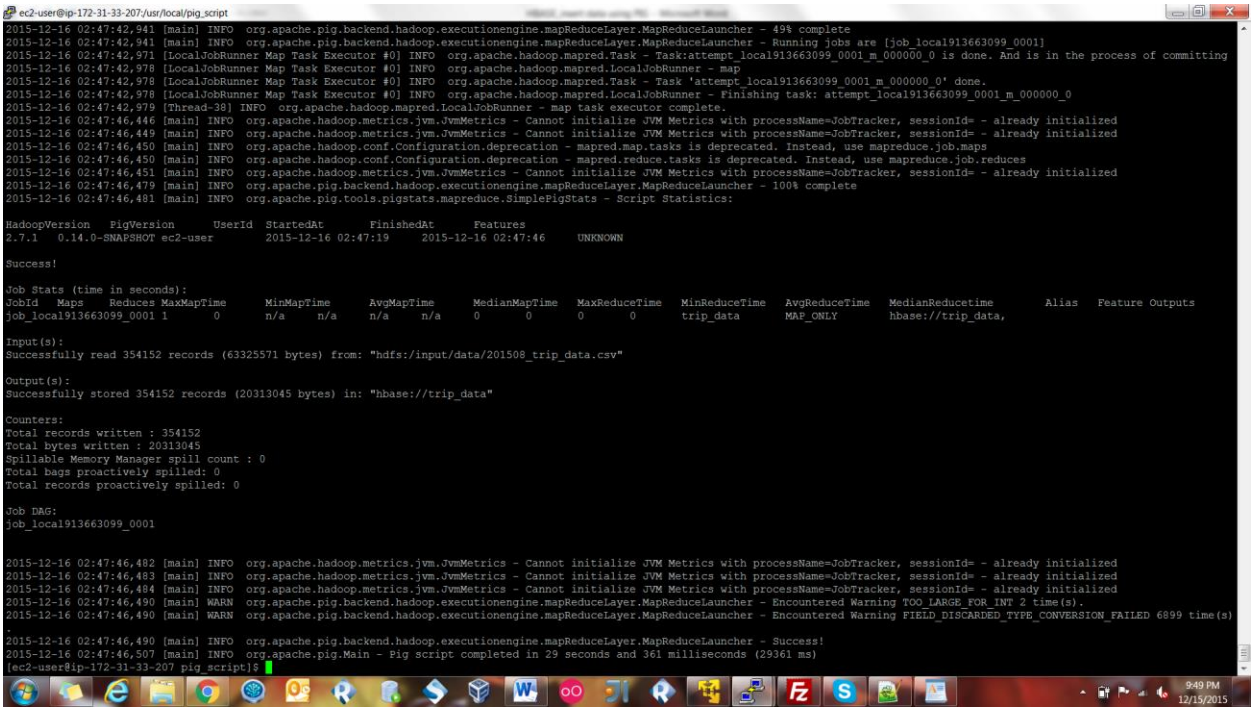
customer_details:subscriber,

customer_details:zip','loadKey true');

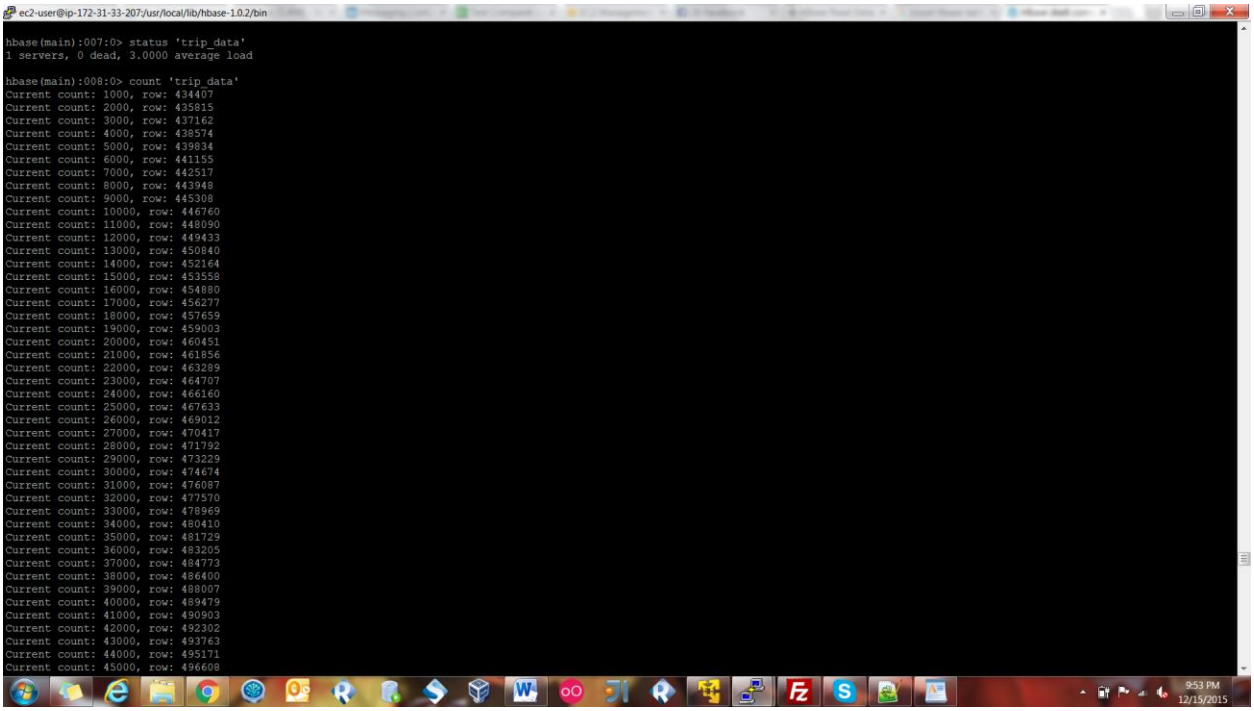
Run pig script and its output:

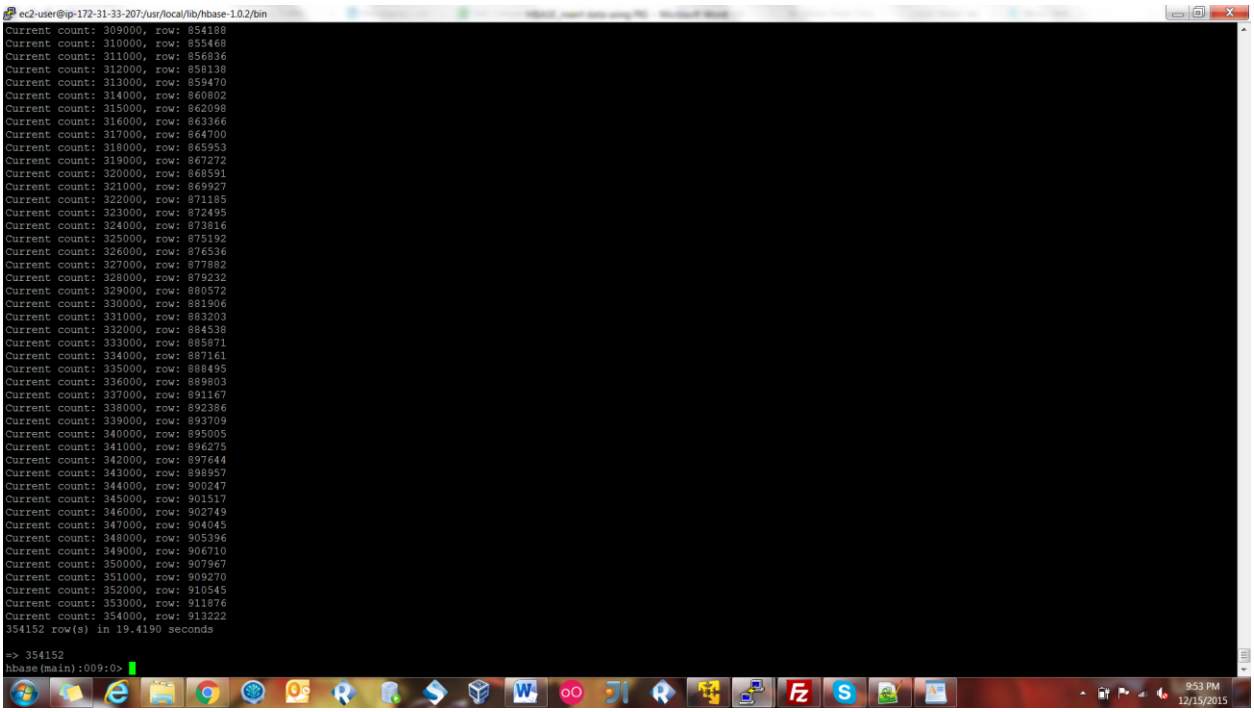



ec2-user@ip-172-31-33-207:/usr/local/pig_script
-rwxrwxrwx 1 ec2-user ec2-user 716 Dec 13 06:14 Sorted_totalTrip_estimatedDuration.pig
-rwxrwxrwx 1 ec2-user ec2-user 394 Dec 13 05:50 station_per_landmark.pig
-rwxrwxrwx 1 ec2-user ec2-user 631 Dec 13 06:13 toatlTripbetweenStations.pig
[ec2-user@ip-172-31-33-207 pig_script]$ pig loadToHbase.pig
15/12/16 02:47:17 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
15/12/16 02:47:17 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE
15/12/16 02:47:17 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType
15/12/16 02:47:17 WARN pig.Main: Need write permission in the directory: /usr/local/lib/pig_log to create log file.
2015-12-16 02:47:17,290 [main] INFO org.apache.pig.Main - Apache Pig version 0.14.0-SNAPSHOT (r: unknown) compiled Dec 10 2015, 03:09:26
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/lib/hadoop-2.7.1/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/lib/hbase-1.0.2/lib/slf4j-log4j12-1.7.7.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2015-12-16 02:47:18,140 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2015-12-16 02:47:18,140 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2015-12-16 02:47:18,140 [main] INFO org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000
2015-12-16 02:47:19,065 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2015-12-16 02:47:19,215 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2015-12-16 02:47:19,342 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2015-12-16 02:47:19,358 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2015-12-16 02:47:19,387 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2015-12-16 02:47:19,396 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
2015-12-16 02:47:19,436 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2015-12-16 02:47:19,604 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2015-12-16 02:47:19,630 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2015-12-16 02:47:19,650 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2015-12-16 02:47:19,650 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2015-12-16 02:47:19,672 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2015-12-16 02:47:19,682 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - session.id is deprecated. Instead, use dfs.metrics.session-id
2015-12-16 02:47:19,683 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Initializing JVM Metrics with processName=JobTracker, sessionId=
2015-12-16 02:47:19,706 [main] INFO org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to the job
2015-12-16 02:47:19,712 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use mapreduce.reduce.markreset.buffer.percent
2015-12-16 02:47:19,712 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2015-12-16 02:47:19,712 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.output.compress is deprecated. Instead, use mapreduce.output.fileoutputformat.compress
2015-12-16 02:47:19,714 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - This job cannot be converted run in-process
2015-12-16 02:47:19,917 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added jar file:/usr/local/lib/pig-0.14.0/pig-0.14.0-SNAPSHOT-core-h2.jar to DistributedCache through /tmp/temp-2048781721/tmp1183873083/pig-0.14.0-SNAPSHOT-core-h2.jar
2015-12-16 02:47:19,952 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added jar file:/usr/local/lib/hadoop-2.7.1/share/hadoop/common/lib/guava-11.0.2.jar to DistributedCache through /tmp/temp-2048781721/tmp2090095505/guava-11.0.2.jar
2015-12-16 02:47:19,971 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added jar file:/usr/local/lib/hadoop-2.7.1/share/hadoop/common/lib/protobuf-java-2.5.0.jar to DistributedCache through /tmp/temp-2048781721/tmp1286635498/protobuf-java-2.5.0.jar
2015-12-16 02:47:19,993 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added jar file:/usr/local/lib/pig-0.14.0/lib/hbase-client-1.0.2.jar to DistributedCache through /tmp/temp-2048781721/tmp-667299687/hbase-client-1.0.2.jar
2015-12-16 02:47:20,027 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added jar file:/usr/local/lib/pig-0.14.0/lib/hbase-protocol-1.0.2.jar to DistributedCache through /tmp/temp-2048781721/tmp372924671/hbase-protocol-1.0.2.jar
9:48 PM
12/15/2015

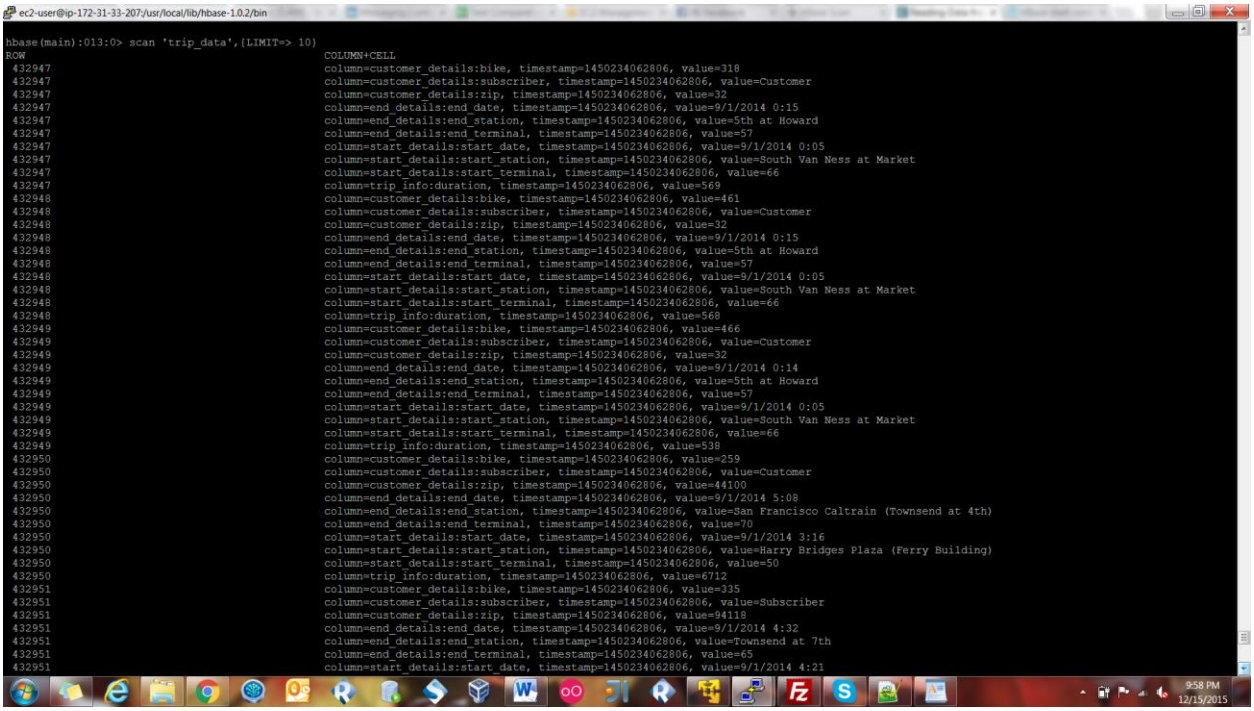

ec2-user@ip-172-31-33-207:/usr/local/pig_script
2015-12-16 02:47:42,941 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 49% complete
2015-12-16 02:47:42,941 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Running jobs are [job_local913663099_0001]
2015-12-16 02:47:42,971 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.Task - Task:attempt_local913663099_0001_m_000000_0 is done. And is in the process of committing
2015-12-16 02:47:42,978 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.LocalJobRunner - map
2015-12-16 02:47:42,978 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.Task - Task 'attempt_local913663099_0001_m_000000_0' done.
2015-12-16 02:47:42,978 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local913663099_0001_m_000000_0
2015-12-16 02:47:42,979 [Thread-38] INFO org.apache.hadoop.mapred.LocalJobRunner - map task executor complete.
2015-12-16 02:47:46,446 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-16 02:47:46,449 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-16 02:47:46,450 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
2015-12-16 02:47:46,450 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated. Instead, use mapreduce.job.reduces
2015-12-16 02:47:46,451 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-16 02:47:46,479 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2015-12-16 02:47:46,481 [main] INFO org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.7.1 0.14.0-SNAPSHOT ec2-user 2015-12-16 02:47:19 2015-12-16 02:47:46 UNKNOWN
Success!
Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReducetime Alias Feature Outputs
job_local913663099_0001 1 0 n/a n/a n/a n/a 0 0 0 0 trip_data MAP_ONLY hbase://trip_data,
Input(s):
Successfully read 354152 records (63325571 bytes) from: "hdfs:/input/data/201508_trip_data.csv"
Output(s):
Successfully stored 354152 records (20313045 bytes) in: "hbase://trip_data"
Counters:
Total records written : 354152
Total bytes written : 20313045
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
Job DAG:
job_local913663099_0001
2015-12-16 02:47:46,482 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-16 02:47:46,483 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-16 02:47:46,484 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2015-12-16 02:47:46,490 [main] WARN org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Encountered Warning TOO_LARGE_FOR_INT 2 time(s).
2015-12-16 02:47:46,490 [main] WARN org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Encountered Warning FIELD_DISCARDED_TYPE_CONVERSION_FAILED 6899 time(s)
2015-12-16 02:47:46,490 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2015-12-16 02:47:46,507 [main] INFO org.apache.pig.Main - Pig script completed in 29 seconds and 361 milliseconds (29361 ms)
[ec2-user@ip-172-31-33-207 pig_script]$
9:49 PM
12/15/2015

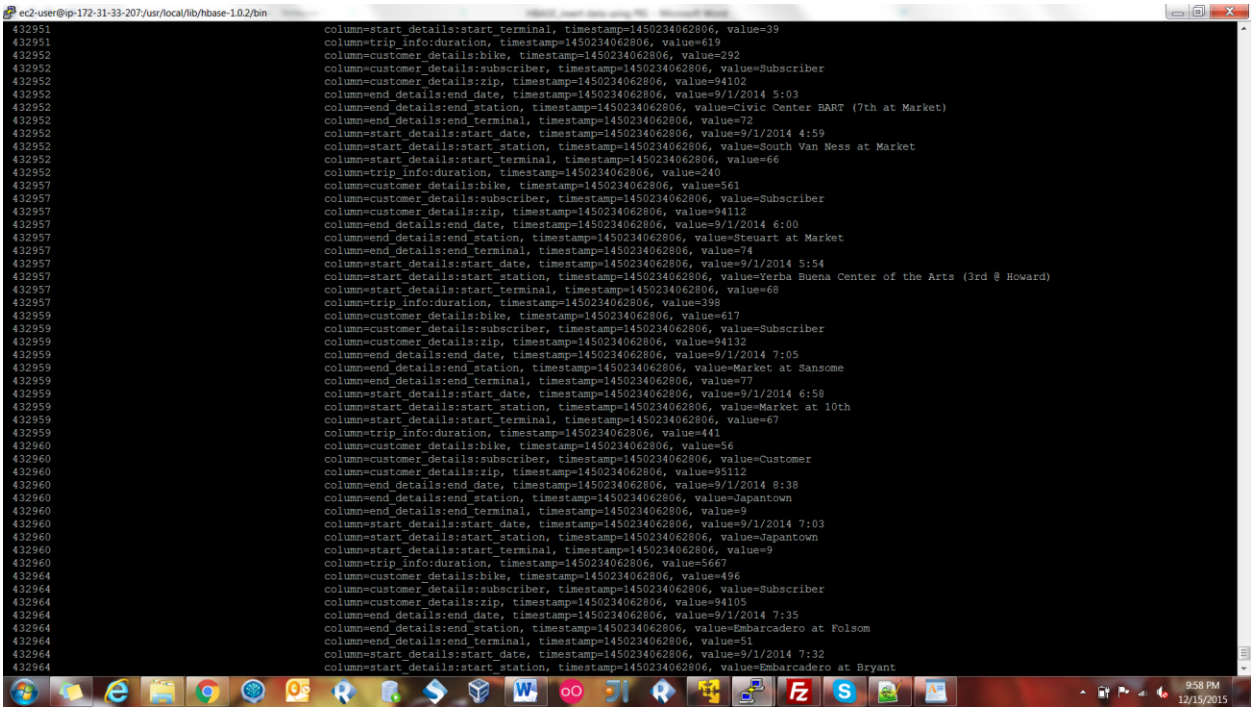
Verify 'trip_data' hbase table:





Scan table:


ec2-user@ip-172-31-33-207:/usr/local/lib/hbase-1.0.2/bin
hbase(main):013:0> scan 'trip_data',{LIMIT=> 10}
ROW COLUMN+CELL
432947 column=customer_details:bike, timestamp=1450234062806, value=318
432947 column=customer_details:subscriber, timestamp=1450234062806, value=Customer
432947 column=customer_details:zip, timestamp=1450234062806, value=32
432947 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 0:15
432947 column=end_details:end_station, timestamp=1450234062806, value=5th at Howard
432947 column=end_details:end_terminal, timestamp=1450234062806, value=57
432947 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 0:05
432947 column=start_details:start_station, timestamp=1450234062806, value=South Van Ness at Market
432947 column=start_details:start_terminal, timestamp=1450234062806, value=66
432947 column=trip_info:duration, timestamp=1450234062806, value=569
432948 column=customer_details:bike, timestamp=1450234062806, value=461
432948 column=customer_details:subscriber, timestamp=1450234062806, value=Customer
432948 column=customer_details:zip, timestamp=1450234062806, value=32
432948 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 0:15
432948 column=end_details:end_station, timestamp=1450234062806, value=5th at Howard
432948 column=end_details:end_terminal, timestamp=1450234062806, value=57
432948 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 0:05
432948 column=start_details:start_station, timestamp=1450234062806, value=South Van Ness at Market
432948 column=start_details:start_terminal, timestamp=1450234062806, value=66
432948 column=trip_info:duration, timestamp=1450234062806, value=568
432949 column=customer_details:bike, timestamp=1450234062806, value=466
432949 column=customer_details:subscriber, timestamp=1450234062806, value=Customer
432949 column=customer_details:zip, timestamp=1450234062806, value=32
432949 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 0:14
432949 column=end_details:end_station, timestamp=1450234062806, value=5th at Howard
432949 column=end_details:end_terminal, timestamp=1450234062806, value=57
432949 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 0:05
432949 column=start_details:start_station, timestamp=1450234062806, value=South Van Ness at Market
432949 column=start_details:start_terminal, timestamp=1450234062806, value=66
432949 column=trip_info:duration, timestamp=1450234062806, value=538
432950 column=customer_details:bike, timestamp=1450234062806, value=259
432950 column=customer_details:subscriber, timestamp=1450234062806, value=Customer
432950 column=customer_details:zip, timestamp=1450234062806, value=44100
432950 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 5:08
432950 column=end_details:end_station, timestamp=1450234062806, value=San Francisco Caltrain (Townsend at 4th)
432950 column=end_details:end_terminal, timestamp=1450234062806, value=70
432950 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 3:16
432950 column=start_details:start_station, timestamp=1450234062806, value=Harry Bridges Plaza (Ferry Building)
432950 column=start_details:start_terminal, timestamp=1450234062806, value=50
432950 column=trip_info:duration, timestamp=1450234062806, value=6712
432951 column=customer_details:bike, timestamp=1450234062806, value=335
432951 column=customer_details:subscriber, timestamp=1450234062806, value=Subscriber
432951 column=customer_details:zip, timestamp=1450234062806, value=94118
432951 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 4:32
432951 column=end_details:end_station, timestamp=1450234062806, value=Townsend at 7th
432951 column=end_details:end_terminal, timestamp=1450234062806, value=65
432951 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 4:21
9:58 PM
12/15/2015


ec2-user@ip-172-31-33-207:/usr/local/lib/hbase-1.0.2/bin
432951 column=start_details:start_terminal, timestamp=1450234062806, value=39
432951 column=trip_info:duration, timestamp=1450234062806, value=619
432952 column=customer_details:bike, timestamp=1450234062806, value=292
432952 column=customer_details:subscriber, timestamp=1450234062806, value=Subscriber
432952 column=customer_details:zip, timestamp=1450234062806, value=94102
432952 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 5:03
432952 column=end_details:end_station, timestamp=1450234062806, value=Civic Center BART (7th at Market)
432952 column=end_details:end_terminal, timestamp=1450234062806, value=72
432952 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 4:59
432952 column=start_details:start_station, timestamp=1450234062806, value=South Van Ness at Market
432952 column=start_details:start_terminal, timestamp=1450234062806, value=66
432952 column=trip_info:duration, timestamp=1450234062806, value=240
432957 column=customer_details:bike, timestamp=1450234062806, value=561
432957 column=customer_details:subscriber, timestamp=1450234062806, value=Subscriber
432957 column=customer_details:zip, timestamp=1450234062806, value=94112
432957 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 6:00
432957 column=end_details:end_station, timestamp=1450234062806, value=Steuart at Market
432957 column=end_details:end_terminal, timestamp=1450234062806, value=74
432957 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 5:54
432957 column=start_details:start_station, timestamp=1450234062806, value=Yerba Buena Center of the Arts (3rd @ Howard)
432957 column=start_details:start_terminal, timestamp=1450234062806, value=68
432957 column=trip_info:duration, timestamp=1450234062806, value=398
432959 column=customer_details:bike, timestamp=1450234062806, value=617
432959 column=customer_details:subscriber, timestamp=1450234062806, value=Subscriber
432959 column=customer_details:zip, timestamp=1450234062806, value=94132
432959 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 7:05
432959 column=end_details:end_station, timestamp=1450234062806, value=Market at Sansome
432959 column=end_details:end_terminal, timestamp=1450234062806, value=77
432959 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 6:58
432959 column=start_details:start_station, timestamp=1450234062806, value=Market at 10th
432959 column=start_details:start_terminal, timestamp=1450234062806, value=67
432959 column=trip_info:duration, timestamp=1450234062806, value=441
432960 column=customer_details:bike, timestamp=1450234062806, value=56
432960 column=customer_details:subscriber, timestamp=1450234062806, value=Customer
432960 column=customer_details:zip, timestamp=1450234062806, value=95112
432960 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 8:38
432960 column=end_details:end_station, timestamp=1450234062806, value=Japantown
432960 column=end_details:end_terminal, timestamp=1450234062806, value=9
432960 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 7:03
432960 column=start_details:start_station, timestamp=1450234062806, value=Japantown
432960 column=start_details:start_terminal, timestamp=1450234062806, value=9
432960 column=trip_info:duration, timestamp=1450234062806, value=5667
432964 column=customer_details:bike, timestamp=1450234062806, value=496
432964 column=customer_details:subscriber, timestamp=1450234062806, value=Subscriber
432964 column=customer_details:zip, timestamp=1450234062806, value=94105
432964 column=end_details:end_date, timestamp=1450234062806, value=9/1/2014 7:35
432964 column=end_details:end_station, timestamp=1450234062806, value=Embarcadero at Folsom
432964 column=end_details:end_terminal, timestamp=1450234062806, value=51
432964 column=start_details:start_date, timestamp=1450234062806, value=9/1/2014 7:32
432964 column=start_details:start_station, timestamp=1450234062806, value=Embarcadero at Bryant
9:58 PM
12/15/2015

---------- **Analysis : 11 Scan single data into HBASE using JAVA** ----------

Trip_data.java --------------------

```
/*
 * To change this license header, choose License Headers in Project Properties.
 * To change this template file, choose Tools | Templates
 * and open the template in the editor.
 */
package trip_package;
import java.io.FileNotFoundException;
import java.io.IOException;
import java.util.Arrays;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.hbase.HBaseConfiguration;

import org.apache.hadoop.hbase.client.Get;
import org.apache.hadoop.hbase.client.HTable;
import org.apache.hadoop.hbase.client.Result;

/**
 *
 * @author ubuntu
 */
public class Trip_data {

    /**
     * @param args the command line arguments
     * @throws java.io.IOException
     * @throws java.io.FileNotFoundException
     */

    public static void main(String[] args) throws IOException ,FileNotFoundException{

        try{
```

```
            Configuration config = HBaseConfiguration.create();
             HTable htable = new HTable(config, "trip_data");
             String rowKey ="913454";
               Get get = new Get(rowKey.getBytes());

                Result rs =htable.get(get);
            byte[] row = rs.getValue(Bytes.toBytes("trip_info"), Bytes.toBytes("duration"));
            System.out.println(row);

        }catch (IOException e){
        }
    }
    }
```

Graphs:

Top 10 stations:

From A to B:


SUM(Average Time)
525.1119691
586.84689
610.1204013
673.2582418
875.2944785
1,174.686131
1,239.125926
1,546.930851
1,822.28972
2,806.527273
SUM(No Of Trips)
107
110
137
163
182
376
418
518
540
598
Station Name
(All)
2nd at Townsend
Embarcadero at S...
Harry Bridges Pla...
Market at 10th
Market at Sansome
San Francisco Cal...
San Francisco Cal...
Steuart at Market
Temporary Transb...
Townsend at 7th
To Station
(All)
2nd at Folsom
2nd at South P...
2nd at Townsend
5th at Howard
Adobe on Alm...
Arena Green / ...
Beale at Market
Broadway St a...
California Ave ...
Castro Street ...
Civic Center B...
Clay at Battery
Commercial at...
Cowper at Uni...
Davis at Jackson
Embarcadero ...
Embarcadero ...
Embarcadero ...
Embarcadero ...
Evelyn Park a...
Franklin at Maple
Golden Gate a...
Grant Avenue ...
Harry Bridges ...
Howard at 2nd
Japantown
Market at 4th
Market at 10th
Market at San...
Mechanics Pla...
Mezes Park
MLK Library
Mountain View...
Mountain View...
Palo Alto Caltr...
Park at Olive
Paseo de San ...
Post at Kearny
Powell at Post ...
Limit
Top 10 by MAX([Trip Count])
Lat: 37.80
Long: -122.4
name: Embarcadero at Sansome
To Station: 2nd at Townsend
Average Time: 1,239.125926
No Of Trips: 540
Trip Count: 14,158
© OpenStreetMap contributors
Sheet 1
Sheet 2
Story 1
Sheet 3


SUM(Average Time)
586.84689
SUM(No Of Trips)
418
Station Name
(All)
2nd at Townsend
Embarcadero at S...
Harry Bridges Pla...
Market at 10th
Market at Sansome
San Francisco Cal...
San Francisco Cal...
Steuart at Market
Temporary Transb...
Townsend at 7th
To Station
(All)
2nd at Folsom
2nd at South P...
2nd at Townsend
5th at Howard
Adobe on Alm...
Arena Green / ...
Beale at Market
Broadway St a...
California Ave ...
Castro Street ...
Civic Center B...
Clay at Battery
Commercial at...
Cowper at Uni...
Davis at Jackson
Embarcadero ...
Embarcadero ...
Embarcadero ...
Embarcadero ...
Evelyn Park a...
Franklin at Maple
Golden Gate a...
Grant Avenue ...
Harry Bridges ...
Howard at 2nd
Japantown
Market at 4th
Market at 10th
Market at San...
Mechanics Pla...
Mezes Park
MLK Library
Mountain View...
Mountain View...
Palo Alto Caltr...
Park at Olive
Paseo de San ...
Post at Kearny
Powell at Post ...
Limit
Top 10 by MAX([Trip Count])
14,158
Embarcadero at
Davis at Jac
Lat: 37.80
Long: -122.4
name: Embarcadero at Sansome
To Station: Davis at Jackson
Average Time: 586.84689
No Of Trips: 418
Trip Count: 14,158
© OpenStreetMap contributors
Sheet 1
Sheet 2
Story 1
Sheet 3

One location multiple destinations: