

Principles of Big Data

Project Phase-2

Submitted by,

TEAM 27

Sravan Kumar Appana -16221049

Lakshmi Sai Krishna Ravilla – 16224836

Sri Vidya Surya Haritha Jyothula - 16222098

Sainath Reddy - 16232714

Aim :

In this project, we run five queries among which we run two queries using spark RDDs and we run another two queries using spark dataframes. Finally, the fifth query will be run by calling the public APIs.

Theme :

We have selected “Iphone 7” as search keyword. We selected Iphone7 as our theme since it is the trending hot topic in todays technology world. Our project was started before the release of Iphone7. Many apple lovers were excited regarding its specifications, new changes in the upcoming iphone 7 world-wide. The number of tweets tweeted have been increased because of some major changes like eliminating 3.5mm headphone jack, etc.,

Hashtag – Iphone7

Here five queries are run using Apache Spark and Java programming.

1. Two queries are run using Spark RDDs.
2. Another Two queries are run using Spark Data Frames.
3. One query is called the public APIs so as to update some part of information in the collection and retrieval of some other data which is not included originally.

Libraries:

1. Maven dependency.
2. Two jar files.
3. Spark-core 2.10
4. Spark-SQL 2.10

Programming Languages:

1. Environment: Eclipse
2. Tweets collection: Python
3. Programming: Java

API:

Twitter API.

Developing/Testing Machine's specifications:

[View basic information about your computer](#)



Windows edition

Windows 10 Home

© 2016 Microsoft Corporation. All rights reserved.



System

Processor: Intel(R) Core(TM) i5-4210U CPU @ 1.70GHz 2.40 GHz

Installed memory (RAM): 8.00 GB (7.88 GB usable)

System type: 64-bit Operating System, x64-based processor

Pen and Touch: Touch Support with 10 Touch Points



[Support Information](#)

Computer name, domain, and workgroup settings

Computer name: Varmas

Full computer name: Varmas

Computer description:

Workgroup: WORKGROUP

[Change settings](#)

Windows activation

Windows is activated [Read the Microsoft Software License Terms](#)

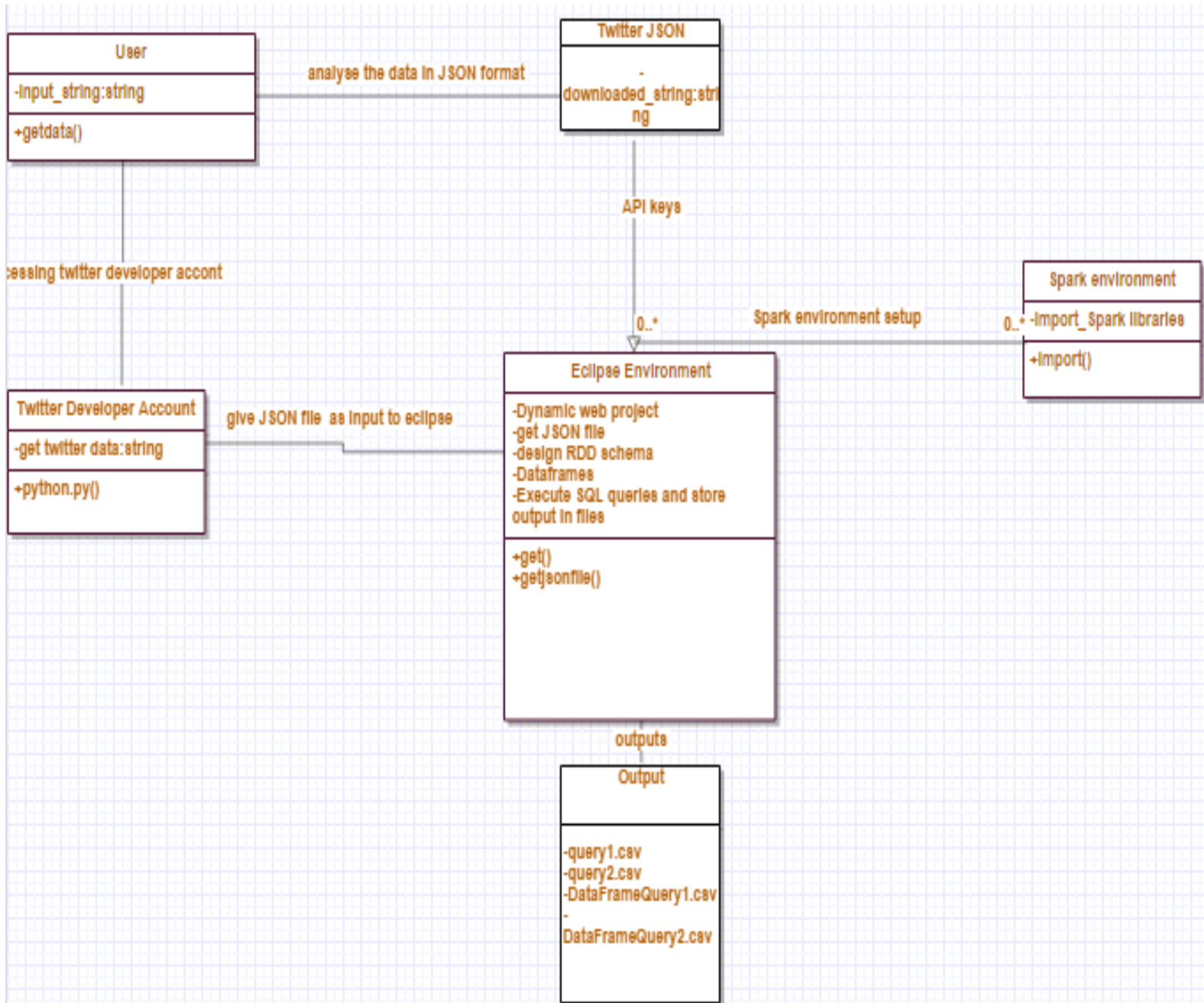
Product ID: 00326-10000-00000-AA751

[Change product key](#)

5

6

UML Diagram:



System Architecture :

1. In this project, we collect tweets using twitter streaming API using python language code with the help of filter as “Iphone7”.
2. The tweets which are collected are saved in Jason format. Those will be taken as input for java schema RDD, data frames .
3. Json file is generated which usually is reduced to tables using spark libraries such as spark_core, spark_sql.
4. Using RDD and data frames , data tables are generated.
5. From the data we will get output for the required queries by using SQL code.
6. That particular will be written in .CSV file.
7. Twitter API :: We take one of the outputs from the above 4 queries and fetch recent tweets by a particular user. The collected data is saved in .txt file.

RDD Query 1: We used RDD to retrieve the output. This query gives the count of tweets grouped by different operating systems like android,windows,ios etc., of various users.

CODE:

```
package iphone7TwitterAnalysis;

import java.net.URL;

import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.sql.api.java.JavaSQLContext;
import org.apache.spark.sql.api.java.JavaSchemaRDD;
import java.io.File;
import java.io.FileWriter;
import java.io.IOException;
import java.util.Collections;
import java.util.List;
import javax.servlet.ServletException;
import javax.servlet.annotation.WebServlet;
import javax.servlet.http.HttpServlet;
import javax.servlet.http.HttpServletRequest;
import javax.servlet.http.HttpServletResponse;
import javax.swing.JOptionPane;

import org.apache.spark.sql.api.java.Row;
import org.apache.spark.sql.api.java.JavaSQLContext;
import org.apache.spark.sql.catalyst.*;
import org.apache.spark.sql.catalyst.expressions.*;
import org.apache.spark.sql.catalyst.analysis.*;
//import org.apache.spark.sql.catalyst.*;
import org.apache.spark.sql.execution.*;
import org.apache.spark.sql.types.util.*;

public class query1 {
```

```

public static void main(String[] args) {
// TODO Auto-generated method stub
String pathToFile =
"file:/C:/Users/sravan/Downloads/SMT_EclipseWorkspace/SMT_EclipseWorkspace/org.apache.spark/src/iPhone7.json";
SparkConf conf = new SparkConf().setAppName("User mining").setMaster("local[*]");

    JavaSparkContext sc = new JavaSparkContext(conf);

    JavaSQLContext sqlContext = new JavaSQLContext(sc);

    JavaSchemaRDD tweets = sqlContext.jsonFile(pathToFile);

    JavaSchemaRDD df=sqlContext.jsonFile(pathToFile);


    tweets.registerAsTable("tweetTable");

    tweets.printSchema();

    query1(sqlContext);
    query2(sqlContext);

    sc.stop();

}
private static void query1(JavaSQLContext sqlContext) {

    try
    {

        File outputFile = new File("query1.csv");
        FileWriter fw= new FileWriter(outputFile);

        // Tweets from different kind of devices...!!

        JavaSchemaRDD count = sqlContext.sql("SELECT  COUNT(*) AS c FROM tweetTable " +
        "WHERE source LIKE '%Android%'");
        JavaSchemaRDD count1 = sqlContext.sql("SELECT  COUNT(*) AS c FROM tweetTable " +
        "WHERE source LIKE '%iPhone%'");
        JavaSchemaRDD count2 = sqlContext.sql("SELECT  COUNT(*) AS c FROM tweetTable " +
        "WHERE source LIKE '%iPad%'");
        JavaSchemaRDD count3 = sqlContext.sql("SELECT  COUNT(*) AS c FROM tweetTable " +
        "WHERE source LIKE '%Windows%'");
        JavaSchemaRDD count4 = sqlContext.sql("SELECT  COUNT(*) AS c FROM tweetTable " +
        "WHERE source LIKE '%Web Client%'");
        //JavaSchemaRDD count5 = sqlContext.sql("SELECT  COUNT(*) AS c FROM tweetTable " +
        // "WHERE source LIKE '%Golf%'");

        List<Row> Android=count.collect();
        String Android12=Android.toString();
        String Android1 = Android12.substring(Android12.indexOf("[") + 2, Android12.indexOf("]"));

        List<Row> iPhone=count1.collect();
        String iPhone12=iPhone.toString();

```

```

String iPhone1 = iPhone12.substring(iPhone12.indexOf("[") + 2, iPhone12.indexOf("]"));

List<Row> iPad=count2.collect();
String iPad12=iPad.toString();
String iPad1 = iPad12.substring(iPad12.indexOf("[") + 2, iPad12.indexOf("]"));

List<Row> Windows=count3.collect();
String Windows12=Windows.toString();
String Windows1 = Windows12.substring(Windows12.indexOf("[") + 2, Windows12.indexOf("]"));

List<Row> Web_Client=count4.collect();
String Web_Client12=Web_Client.toString();
String Web_Client1 = Web_Client12.substring(Web_Client12.indexOf("[") + 2, Web_Client12.indexOf("]"));

//List<Row> Golf=count5.collect();
//String Golf12=Golf.toString();
//String Golf1 = Golf12.substring(Golf12.indexOf("[") + 2, Golf12.indexOf("]"));

    fw.append("DeviceName");
fw.append(',');
fw.append("Count");
fw.append("\n");
fw.append("Android");
fw.append(',');
fw.append(Android1);
fw.append("\n");
fw.append("iPhone");
fw.append(',');
fw.append(iPhone1);
fw.append("\n");
fw.append("iPad");
fw.append(',');
fw.append(iPad1);
fw.append("\n");
fw.append("Windows");
fw.append(',');
fw.append(Windows1);
fw.append("\n");
fw.append("Web Client");
fw.append(',');
fw.append(Web_Client1);
fw.append("\n");
fw.close();
String path= outputFile.getAbsolutePath();
System.out.println(path);
}
catch (Exception exp)
{
    System.out.println("Error");
}

}

```

OUTPUT:

1	DeviceName	Count
2	Android	538
3	iPhone	1122
4	iPad	46
5	Windows	12
6	Web Clier	254

RDD Query 2: We used RDD to retrieve the output. This query gives the details of the number of friends greater than given amount.

CODE:

```
package iphone7TwitterAnalysis;

import java.net.URL;

import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.sql.api.java.JavaSQLContext;
import org.apache.spark.sql.api.java.JavaSchemaRDD;
import java.io.File;
import java.io.FileWriter;
import java.io.IOException;
import java.util.Collections;
import java.util.List;
import javax.servlet.ServletException;
import javax.servlet.annotation.WebServlet;
import javax.servlet.http.HttpServlet;
import javax.servlet.http.HttpServletRequest;
import javax.servlet.http.HttpServletResponse;
import javax.swing.JOptionPane;

import org.apache.spark.sql.api.java.Row;
import org.apache.spark.sql.api.java.JavaSQLContext;
import org.apache.spark.sql.catalyst.*;
import org.apache.spark.sql.catalyst.expressions.*;
import org.apache.spark.sql.catalyst.analysis.*;
//import org.apache.spark.sql.catalyst.*;
import org.apache.spark.sql.execution.*;
import org.apache.spark.sql.types.util.*;

private static void query2(JavaSQLContext sqlContext)
{

    try
    {File outputFile = new File("query2.csv");
    FileWriter fw= new FileWriter(outputFile);
```



```
//Users with maximum friends..!!
```

```
JavaSchemaRDD count = sqlContext.sql("SELECT user.screen_name, max(user.friends_count) AS c  
FROM tweetTable " +  
    "WHERE user.friends_count>'150000'" +  
    "group by user.screen_name order by c desc limit 20" );
```

```
List<org.apache.spark.sql.api.java.Row> rows = count.collect();
```

```
    //Collections.reverse(rows);
```

```
String rows123=rows.toString();
```

```
String[] array = rows123.split(",");  
System.out.println(array.length);
```

```
System.out.println(rows123);
```

```
    fw.append("Name");  
fw.append(',');  
fw.append("Count");  
fw.append("\n");
```

```
for(int i = 0; i < array.length; i++)  
{  
    if(i==0)  
    {  
        fw.append(array[0].substring(2));  
        fw.append(',');  
        fw.append("\n");  
    }  
    else if(i == array.length-1)  
    {  
        fw.append(array[i].substring(2,array[i].length()-2));  
        fw.append(',');  
        fw.append("\n");  
    }  
    else {  
        fw.append(array[i].substring(2));  
        fw.append(',');  
        fw.append("\n");  
    }  
}  
fw.close();  
}  
catch (Exception exp)  
{  
}
```

```
}
```

```
}
```

OUTPUT:

1	Name	Count
2	SalihSarik	334152
3	FresH_Bo	205108

DATAFRAMES:

Query 3:

We used data frame to retrieve the output. This query is used to count the number of tweets by each user and it will show the count of top 8 users in the descending order.

CODE:

```
public void DataframeQuery1()
{
    String pathToFile = url.toString();

    SparkConf conf = new SparkConf();
    conf.setAppName("Spark MultipleContest Test");
    conf.set("spark.driver.allowMultipleContexts", "true");
    conf.setMaster("local");

    JavaSparkContext sc = new JavaSparkContext(conf);

    SQLContext sqlContext = new SQLContext(sc);

    DataFrame tweets = sqlContext.read().json(pathToFile);

    tweets.registerTempTable("tweetTable");

    DataFrame followers = sqlContext.sql("SELECT user.name,max(user.statuses_count) AS c FROM
tweetTable " +
        "GROUP BY user.name ORDER BY c desc limit 8");

    Row[] rows = followers.collect();

    System.out.println(rows[0]);

    try
    {
        File outputFile = new
File("C:/Users/sravan/Downloads/SMT_EclipseWorkspace/SMT_EclipseWorkspace/sampleProject/WebContent/DataFrameQuery1.csv");

        FileWriter fw = new FileWriter(outputFile);

        fw.append("Name");
        fw.append(',');
        fw.append("Count");
        fw.append("\n");

        for (int i=0;i<8;i++) {
```

```

        fw.append(rows[i].get(0).toString());
        fw.append(',');
        fw.append(rows[i].get(1).toString());
        fw.append("\n");

    }

    fw.close();

}
catch(Exception e)
{

}

sc.stop();
}

```

OUTPUT:

1	Name	Count
2	John Muri	2044699
3	I Love #Fa	1639320
4	Cell Phon	1203667
5	T-Mobile I	1077218
6	Sprint Car	949032
7	Kelly	874469
8	Peggy	865265
9	amy	792143

Query 4:

We used data frame to retrieve the output. This query gives the ranking of users based on number of followers in descending order.

CODE:

```

public void DataframeQuery2()
{
    String pathToFile = url.toString();

    SparkConf conf = new SparkConf();
    conf.setAppName("Spark MultipleContext Test");
    conf.set("spark.driver.allowMultipleContexts", "true");
    conf.setMaster("local");

    JavaSparkContext sc = new JavaSparkContext(conf);

    SQLContext sqlContext = new SQLContext(sc);

    DataFrame tweets = sqlContext.read().json(pathToFile);
}

```

```

        tweets.registerTempTable("tweetTable");

        DataFrame followers = sqlContext.sql("SELECT user.name, max(user.followers_count) AS c FROM
tweetTable GROUP BY user.name ORDER BY c DESC limit 8");

        Row[] rows = followers.collect();

        System.out.println(rows[0]);

        try
        {
            File outputFile = new
File("C:/Users/sravan/Downloads/SMT_EclipseWorkspace/SMT_EclipseWorkspace/sampleProject/WebConten
t/DataFrameQuery2.csv");

            FileWriter fw = new FileWriter(outputFile);

            fw.append("Name");
            fw.append(',');
            fw.append("Count");
            fw.append("\n");

            for (int i=0;i<8;i++) {

                fw.append(rows[i].get(0).toString());
                fw.append(',');
                fw.append(rows[i].get(1).toString());
                fw.append("\n");

            }

            fw.close();

        }
        catch(Exception e)
        {

        }

        sc.stop();
    )

```

OUTPUT:

1	Name	Count
2	The Next	1686095
3	Salih SARI	428862
4	ZDNet	335574
5	T-Mobile I	294357
6	HLN	283198
7	howardlin	254664
8	MediaNar	241121
9	????? ???	233770

QUERY 5:

It requires calling public APIs . We take one of the outputs from the above 4 queries and fetch recent tweets by a particular user. The collected data is saved in .txt file.

CODE:

```
public static void main(String args[]) throws TwitterException, IOException
{
    System.out.println("Its working");
    ConfigurationBuilder cp= new ConfigurationBuilder();
    cp.setDebugEnabled(true).setOAuthConsumerKey("BAQSR3afD856mF8NQOydbfYil").setOAuthConsumerSecret("Vt8SIImzSgM7ErdFOJDpFkKE2vggQunAJXjluVFSIJu0IptAYNj")
    .setOAuthAccessToken("766182508698435585-8SMuY4h6TtYggx4m0PyQ9UJTU31J0Xm").setOAuthAccessTokenSecret("l6rr6mQbX8vFZrNsBHkOwem6zst7L9lQZZHKvdfUGKeU9");
    TwitterFactory tf= new TwitterFactory(cp.build());
    twitter4j.Twitter twitter= tf.getInstance();
    try {
        Query query = new Query("Fresh_BoY_Will");
        QueryResult result;
        do {
            result = twitter.search(query);
            List<Status> tweets = result.getTweets();
            for (Status tweet : tweets) {
                System.out.println("@ " + tweet.getUser().getScreenName() + " - " + tweet.getText());
            }
        } while ((query = result.nextQuery()) != null);

        Query query = new Query("SalihSarikaya");
        QueryResult result;
        do {
            result = twitter.search(query);
            List<Status> tweets = result.getTweets();
            for (Status tweet : tweets) {
                System.out.println("@ " + tweet.getUser().getScreenName() + " - " + tweet.getText());
            }
        } while ((query = result.nextQuery()) != null);

        System.exit(0);
    } catch (TwitterException te) {
        te.printStackTrace();
        System.out.println("Failed to search tweets: " + te.getMessage());
        System.exit(-1);
    }
}
```

OUTPUT:

```
@lgsch_k - RT @ourlifeposts: shaved legs and fresh bed sheets are the best combination ever, no boy will
fully understand such luxuries
@your_truly1 - RT @ourlifeposts: shaved legs and fresh bed sheets are the best combination ever, no boy will
fully understand such luxuries
@_MwAh_xox - RT @ourlifeposts: shaved legs and fresh bed sheets are the best combination ever, no boy
will fully understand such luxuries
```

@xxlexybabyxx - RT @ourlifeposts: shaved legs and fresh bed sheets are the best combination ever, no boy will fully understand such luxuries
@manders32199 - RT @ourlifeposts: shaved legs and fresh bed sheets are the best combination ever, no boy will fully understand such luxuries

@RaulOcaranza1 - #Marketers #FF riend #hpy_weekend!
in

#Austin

@DialogGroup

@MargoWickersham

@rpilney

@sarahcolette_

#Houston

@SalihSarikaya

@leahfaul

@SalihSARIKAYA_ - Leonard Cohen'i kaybettik bugün, müzik dünyas?na notalar?n anahtar?n? b?rakarak ayr?ld? aram?zdan, müzik dünyas? da bugün üzdü bizdi

@yoURDATAisyours - Apparently the Internet "Tidal Wave" has passed: <https://t.co/zgZJaXyKxLhttps://t.co/1P9w8IBhnE>

@k__ristii - IE sucks - #TeamFirefox <https://t.co/Wo5zmPJJs42>

@stevenmgroves - @SalihSarikaya I think because its not "visually" smooth appealing by any means and seem to clunky.

@SalihSarikaya - 331 million users ditched Microsoft's browsers <https://t.co/PuDd5AiRhc>

@SalihSarikaya - WhatsApp enables two-step verification <https://t.co/scjR5pxlInn>

@SalihSarikaya - Instagram confirms live video is on the way <https://t.co/9rBoAe4Wwp>

@danmartin_7 - #enc4416 <https://t.co/WawTVqyGsE>

@jah_ebooks - @SalihSarikaya Those things in your twitter and Facebook fan page are run by the same person.

@SalihSarikaya - Facebook is buying CrowdTangle <https://t.co/qmv2PXCj4c>

@JOSEARM27058436 - @Gabriel61813875 @danieleduar7_ @SalihSarikaya hola como estas t gustaria conoserme x fa

@lucxsam - wohooooooooow... @befoune blame it on Social Network US version... <https://t.co/XcoPXJ43Ru>

@SalihSarikaya - Fake news on Facebook affected US election? <https://t.co/w4u6ilUNqn>

@SalihSarikaya - How will social media be used in future elections? <https://t.co/sZlxs3K8A2>

@DAM1120 - Like TV in 50's they are baby sitters and bad ones. The problems today stem from social media. Should never have be... <https://t.co/OsbyCyp2NG>

@SalihSarikaya - What are the benefits of Social Media for kids? <https://t.co/XfNunggLXw>

@morning_star74 - RT @SalihSarikaya: Best 10 future trends of digital media and technology

<https://t.co/WkAoPto0ev> <https://t.co/WMx2ss8ZBL>

@SalihSarikaya - RT @supersnake02: @SalihSarikaya should've better password or keep lo profile

@Diva44Es - @SalihSarikaya Hola mostrarte temas

prestashop,joomla,más.libres,gratis.Sldos <https://t.co/7jELkxLoSq>

@lisapellegrene - RT @dailynewshack: Hey @Lisa_jenifers @lisapellegrene @gazete_man @SalihSarikaya thanks for the follow. Much appreciated. via <https://t.co/7jELkxLoSq>

@SalihSarikaya - Instagram is bringing big changes to Stories <https://t.co/Y1NbdsB8qj>

@SalihSarikaya - Is that still best for your Windows laptop? <https://t.co/S1voQ5t3Cx>

@SalihSarikaya - How Much Is Facebook to Blame? <https://t.co/zlUJOUyLyL>

@supersnake02 - @SalihSarikaya should've better password or keep lo profile

@cinderellabxny - The latest Single Not Lonely Travel!! <https://t.co/03TDXRBpsX> Thanks to @SalihSarikaya

@AnnaKayThomas #ad #travel

@adebankefalade - I need to start trading in Bitcoin o. Hepp <https://t.co/8S1ZtqLZry>

@SalihSarikaya - Bitcoin is surging thanks to Trump's victory <https://t.co/cZuJ6j7S4m>

@SalihSarikaya - What to do if your Yahoo account was hacked <https://t.co/OA8M6qiFJb>

@SalihSarikaya - Snapchat is selling its Spectacles <https://t.co/HH0WQAZaor>

@Metropolis2Ltd - Hello @MakeMoneyNow21 @danglobal @cammipham @timothyjsquires @SalihSarikaya

@__abbierose thanks for following our #publicrelations agency.

@SalihSarikaya - The only way to move forward is to move forward together <https://t.co/VTPFQecWr9>

@nine_oh - #twittergoons send @MattWhite92 a #happybirthday s/o @MRCHEEKSLBFAM @TAYF3RD

@officialjaheim @flydotcomuk @SmthnglikeKites @SalihSarikaya

@BIONOVAskincare - RT @BIONOVAskincare: Personalize ur formula 2 reflect ur unique physiology <https://t.co/HlCy0eiFM1> @SalihSarikaya @factvsfitness @VerbalGo...

@CALLISTE - Not so fast @SalihSarikaya #Trump still needs to be sworn in b/4 he's Prez. Will 270 electors cast vote?... <https://t.co/9RHGEVq80K>

@ReinerKallis - RT @SalihSarikaya: Best 10 future trends of digital media and technology <https://t.co/WkAoPto0ev> <https://t.co/WMx2ss8ZBL>

@SalihSarikaya - The investor of Silicon Valley wins bet on 'Trump' <https://t.co/BbAHQr8S2s>

@SalihSarikaya - Is Facebook a potential threat to LinkedIn? <https://t.co/grrBMUuPz5>

@dailynewshack - Hey @Lisa_jenifers @lisapellegrene @gazete_man @SalihSarikaya thanks for the follow. Much appreciated. via <https://t.co/HRb8A2KA4M>

@Ravi_Bansal5 - Fantastic share .@SalihSarikaya tks <https://t.co/uJHwYvrSGm>

@ipfconline1 - RT @SocMediaMelody: 15 of the best #socialmedia #marketing blogs to follow now to get in the #smm game! <https://t.co/cPmf8kGCKv> @ipfconline1

@SocMediaMelody - 15 of the best #socialmedia #marketing blogs to follow now to get in the #smm game! <https://t.co/cPmf8kGCKv> @ipfconline1

@SalihSarikaya - How can help your customers those are using non-owned social channels? <https://t.co/n7v1NarpGa>

@Ravi_Bansal5 - .@realDonaldTrump will buy #twitter - #SocialMedia As Trend-Starter <https://t.co/Q3h36OHjMm> via .@LifeZette .@SalihSarikaya

@rodmedi12 - ??????@EdwardKoshchene @SalihSarikaya @kim_2yy1cg @RevezNexus @VeronicaD533

@danazingher @KaSaTru @JeffreyHayzlett @illwah

@SalihSARIKAYA_ - @HillaryClinton i don't know hillary, i live forever

@SalihSARIKAYA_ - ?nsan; bildikleri kadar güvende, hissettikleri kadar a??k ve vazgeçtikleri kadar cesurdur.

@SalihSARIKAYA_ - Köprüdeki as?l bal?kç?lar?n, denizden en büyük bal?klar? çeken de?il, en s?? derinliklerden gerçekleri çeken bal?kç?lar?n olduklar?n? gördüm

@SalihSARIKAYA_ - Trump'?n seçilmesi ile birlikte ABD'de eyaletlerin ayr?lmak istemesi ve seçim sisteminin de?iştirilmesi ile ciddi tart??malar ba?layacaktır

@chrisos44 - RT @Concrte_Jungle: @SalihSarikaya maybe it's because apple is empowering us to be linear and not unpredictable like a random musical Play...

@septamush - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>

@Rocky_R_Arnold - RT @socialmannet: 15 Best Social Media #Marketing #Blogs To Follow <https://t.co/ELi35Gdm2H>

@Ellebie - 15 Best Social Media Marketing Blogs To Follow - Digital Marketing <https://t.co/cuC7EAxZT8> #marketing

@socialmannet - 15 Best Social Media #Marketing #Blogs To Follow <https://t.co/ELi35Gdm2H>

@nueski1 - @SalihSarikaya Nueski - Play Time <https://t.co/OAP3SkpphE>

@christosval - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>

@africanbandanna - RT @SalihSarikaya: Search Engine Optimization

@360FocusDigital <https://t.co/p2ofRgrhni>

@africanbandanna - RT @SalihSarikaya: What do your potential clients see when they search your social media?

@360FocusDigital <https://t.co/HNdL3K12BI>

@AnujSaxena - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>

@katfigurepro - @fusion_xtreme Yes ! @SalihSarikaya @musclemacros

@SMJobs4u - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>

@Gary_Archibald - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>

@ghumber720 - RT @SalihSarikaya: What do your potential clients see when they search your social media?

@360FocusDigital <https://t.co/HNdL3K12BI>

@ghumber720 - RT @SalihSarikaya: Search Engine Optimization

@360FocusDigital <https://t.co/p2ofRgrhni>

@SarvOfficial - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>

@ChrisRoss2009 - @SalihSarikaya @rikwalters What do you want? (Motivational Video) watch and Retweet: <https://t.co/fqk4bTVEJ0>
@ReeteshSharma8 - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@LElkan - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@LA_Spenc3r - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@VISROXiTech - 15 Best Social Media Marketing Blogs To Follow - Digital Marketing <https://t.co/52ttoQgGyW><https://t.co/VffZjCDVvR>
@netwwork24 - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@BIONOVAskincare - Personalize ur formula 2 reflect ur unique physiology <https://t.co/HlCy0eiFM1> @SalihSarikaya @factvsfitness @VerbalGoldBlog @lukedancy
@Branding_123 - #branding

15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/9tpWQZfkdk> SalihSarikaya #SMM <https://t.co/ukAdziKWhL>

— ipfconli...
@Digital_Mkt101 - #DigitalMarketing

15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/EajB1xf0D1> SalihSarikaya #SMM <https://t.co/MlPsgzLQ75>

— ...
@KELiCHiA - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@uzzalhossain153 - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@Lindon97202483 - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@ELCavalos - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@ipfconline1 - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@pierrepinna - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@z2duqOb3i6eJzMO - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@Gagaween - @SalihSarikaya Hi, how are you?
@SalihSARIKAYA_ - @fatihportakal bildiriyorum; siyasi tutuklamalara destek vermen ve hep kar??t dü?üncelere yer vermen insani bir durum olmad???n? gösterir
@Makeover_Social - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@TrueNSocial - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@RoyalGroupVisas - RT @ipfconline1: 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@SalihSARIKAYA_ - @fatihportakal gerçekleri ama sadece gerçekleri ?imdiden kamera kar??s?nda yay?nda bildir, daha sonra s?ran?n sana gelece?i hapiste de?il
@SalihSARIKAYA_ - Cumhuriyeti birlikte kuranlar, demokratik cumhuriyete birlikte sahip ç?kabilmelidirler.
@ipfconline1 - 15 Best #SocialMedia #Marketing Blogs To Follow <https://t.co/7agF2ZiRJP> @SalihSarikaya #SMM <https://t.co/QC4KKLuUBF>
@suediamond11 - .@IAM2skilled
@SalihSarikaya
@Brodiesmith21
@KalaniBallFree
@KevinKennethLau

Meet @realDonaldTrump

VOTE #ElectionDay

<https://t.co/z5quXrpNNL>

@SalihSARIKAYA_ - @NumanKurtulmus nerden nereye, önce HAS Parti kongrende o kadar insana verdirdi?in sözlerle bak, sonra da dönüp ?imdi ki bir haline bak

@SalihSARIKAYA_ - @fatihportakal b?k b?k ancak konu?uyorsun ancak hiç bir ?ey yok, anti-demokratik uygulamalar kar??s?nda vitrine oynuyorsun, biraz kendin ol

@SalihSARIKAYA_ - Hrant Dink'in, Tahir Elçi'nin böylesine zor günlerde olmas?n? çok isterdim, ama ne yaz?k ki katledildiler.

@nine_oh - #1love @SkroogMkduk @DjTonyharder @DjQuestGh_Promo @Dj_Quest_Gh @deejayAlexM

@Deejay_ze @REALDJPAPITO730 @SmthnglikeKites @SalihSarikaya

@EmailMktgAlerts - RT @KayELacey: 15 Best Social Media Marketing Blogs To Follow - Digital

Marketing <https://t.co/dkWCORXlab>

@LeblogdeFatiha - RT @SalihSarikaya: Search Engine Optimization

@360FocusDigital <https://t.co/p2ofRgrhni>

@KayELacey - 15 Best Social Media Marketing Blogs To Follow - Digital Marketing <https://t.co/dkWCORXlab>

@BIONOVAskincare - RT @BIONOVAskincare: @Briankrzanich @StrattonGraeme @beauxbeautyblog

@SalihSarikaya @KellerMedialnc @JeffreyHayzlett @OwenFrostie @Dermstor...

@WeCT_Connect - RT @SalihSarikaya: Best 10 future trends of digital media and technology

<https://t.co/WkAoPto0ev> <https://t.co/WMx2ss8ZBL>

@SalihSARIKAYA_ - E?kiya, dünyaya hükümdar olmaz.

@TheAndrewKang - Cool tweets from @SalihSarikaya Digital technology trends: Future trends of digital medium

<https://t.co/2qz2O0Nf1b>

@EmailMktgAlerts - RT @huhuzgz: 15 Best Social Media Marketing Blogs To Follow - Digital

Marketing <https://t.co/TaMKxKWc5C>

@therealCCASH - @SalihSarikaya Listen to SHOOTAZ"FREESTYLE" by Ccash GZM #np on #SoundCloud

<https://t.co/liZNKaJh2j>

@g28_a - RT @huhuzgz: 15 Best Social Media Marketing Blogs To Follow - Digital

Marketing <https://t.co/TaMKxKWc5C>

@YouBrandInc - RT @huhuzgz: 15 Best Social Media Marketing Blogs To Follow - Digital

Marketing <https://t.co/TaMKxKWc5C>

@huhuzgz - 15 Best Social Media Marketing Blogs To Follow - Digital Marketing <https://t.co/TaMKxKWc5C>

@charismagold - #Mustfollow @SeanMaxwell @Lovi_Poe @MySexyGirlsPics @sexx_freak @LuxusGirls

@SalihSarikaya @LAJogger

#SaturdayVibes

@sabrinaonmove - Thank you to my engaged followers! @helenageorge110 @Reevoo @SalihSarikaya

@KymistryMng @FungusInstitute @christofouche @TheAirbnbExpert #FF

@thewriterben - RT @SalihSarikaya: Best 10 future trends of digital media and technology

<https://t.co/WkAoPto0ev> <https://t.co/WMx2ss8ZBL>

@SalihSARIKAYA_ - @siring korku imparatorlu?u bize uzak, soraca??n?z sorular ise umudumuz olsun

@SalihSARIKAYA_ - Yalan ate? gibidir, sadece elinde ta??yan? yakar, gerçek ise su gibidir, ate? ne kadar büyük olursa olsun su o ate?i söndürür

@akara_david - 15 Best Social Media Marketing Blogs To Follow - Digital

Marketing <https://t.co/zYx5DTcJf8><https://t.co/yNwh4GrSVX>

@midorimiller - RT @SalihSarikaya: Social media allows you to hear what your customers and prospects are actually saying. <https://t.co/2B6FboiaDw>

@ktboweswrites - @laura_fletcher9 @SalihSarikaya Got it thanks!

Run Time Measurements for each query:

RDD Query 1 :

Start time:

2016-11-13 19:43:31.929

End time:

2016-11-13 19:43:31.932

RDD Query 2 :

Start time:

2016-11-13 19:53:32.685

End time:

2016-11-13 19:53:33.178

DataFrame Query 3 :

Start time:

2016-11-13 19:15:32.824

End time:

2016-11-13 19:15:34.678

DataFrame Query 4 :

Start time:

2016-11-13 19:19:11.391

End time:

2016-11-13 19:19:11.806

API Query 5:

Start time:

2016-11-13 19:27:19.672

End time:

2016-11-13 19:27:19.685

