

A Project Report on

Market Analysis based on Social Network

Submitted in partial fulfillment of requirement
for the award of the degree

MASTER of COMPUTER APPLICATIONS

Of

Visvesvaraya Technological University

By

**Abhijith B. S. – [1PI13MCA01]
Joshua Fernandes- [1PI13MCA31]**



PES INSTITUTE OF TECHNOLOGY
100 Ft Ring Road, B.S.K 3rd Stage, Bangalore-85

2015

PES INSTITUTE OF TECHNOLOGY

Department of MCA

100 Ft Ring Road, BSK 3rd Stage

Bangalore 85

2015



C E R T I F I C A T E

This is to certify that the project entitled **“Market Analysis based on Social Network”** is a bonafide work carried out by **Abhijith B.S. [1PI13MCA01]** and **Joshua M. Fernandes [1PI13MCA31]** submitted in partial fulfillment of the requirement of Fourth semester course work of MCA during the academic session Jan-May 2015.

Ms. Deepthi S Narayan

Assistant Professor,
Department of MCA
PES Institute of Technology
Bangalore

Dr. Neelam Bawane

Professor & Head
Department of MCA
PES Institute of Technology
Bangalore

Examiners

1.

2

DECLARATION

We, **Abhijith B. S. and Joshua M. Fernandes**, students of 5th MCA, **PESIT** bearing USN **1PI13MCA01 and 1PI13MCA31** hereby declare that the project entitled "**Market Analysis based on Social Network**" has been carried out by me under the supervision of Guide (Co-guide) **Mrs. Deepthi S Narayan**, Assistant Professor, and submitted in partial fulfillment of the requirement for the award of the Degree of Master Of Computer Application by the **Visvesvaraya Technological University** during the academic year 2015. This report has not been submitted to any other Organization/University of any award of degree or certificate.

Name:

Signature:

ACKNOWLEDGEMENT

We take this opportunity to express my profound sense of sincere and deep gratitude to all those who have contributed to the knowledge and experience that I have gained during my project work.

To begin with, I express my heartfelt gratitude to **Dr. K.S. Sridhar**, Principal, PES Institute of Technology, Bangalore for providing me an opportunity to take up this project.

I would also like to express my heartfelt thanks and regards to **Dr. Neelam Bawane**, Head of the Department of MCA, PES Institute of Technology, Bangalore, for her overwhelming support in my project.

I extend my profound gratitude to the project guide, **Ms Deepthi S Narayan**, Assistant Professor, Department of MCA, PES Institute of Technology, Bangalore for her valuable guidance, co-operation and suggestions and also his encouragement to complete this project.

Finally, I would like to thank my friends for all their support and encouragement during my project work.

ABSTRACT

The **MASK project** will allow users to generate an automated report on any keyword/brand name presented based on real –time social media. It will provide in-depth analysis and cluster the information available on social media giving the users graphical and statistical data of their product/brand name of interest. The project will aggregate the information which will be provided by the users allowing them to view previous and current analysis of their keywords.

The applications of this project would be Market analysis for products, brands with the data sourced from Social network like twitter, provides real time analysis of the product and brands and presents various information which could aid in decision making.

Table of Contents

1. INTRODUCTION	1
1.1 Purpose	1
1.2 Product Perspective	1
1.3 Product Functions.....	1
1.4 User Classes and Characteristics.....	1
2. LITERATURE SURVEY	2
2.1 Existing System.....	2
2.2 Proposed System	2
3. REQUIREMENTS ANALYSIS	3
3.1 Operating Environment	3
3.2 Design and Implementation Constraints	3
3.3 User Documentation.....	3
3.4 Assumptions and Dependencies.....	3
4. SOFTWARE REQUIREMENTS SPECIFICATION.....	4
4.1 Users.....	4
4.1.1 End-user	4
4.2 Functional Requirements.....	4
4.2.1 Ability for the end-user to perform sentimental analysis on their product.	4
4.2.2 The end-user should be able to generate word cloud for his product	4
4.2.3 The system should provide user to compare two brands.	4
4.2.4 The system should display the best and the worst tweets for each sentimental report generated.	4
4.3 Non- Functional Requirements	4
4.3.1 Performance Requirements	4
4.3.2 Safety Requirements	5
4.3.3 Security Requirements	5
4.3.4 Software Quality Attributes	5
4.4 Software and Hardware Requirement	5
4.4.1 Hardware Interface.....	5
4.4.2 Software Interfaces	6
4.4.3 Communications Interfaces	6

5.	ANALYSIS AND DESIGN	7
5.1	Use Case	7
5.2	Sequence Diagram	8
5.3	Activity Diagram.....	9
5.4	ER Diagram.....	10
5.5	Data Dictionary	10
6.	IMPLEMENTATION.....	12
6.1	Screen-Shots	12
7.	SOFTWARE TESTING	16
7.1	Testing Objectives:.....	16
7.2	Testing Principles:.....	16
7.3	Unit testing:	16
7.4	Test Cases:.....	16
7.4.1	Login Page	16
7.4.2	User creation	17
7.4.3	Sentimental Search.....	17
7.4.4	Word Cloud.....	18
7.4.5	Brand comparison	18
8.	CONCLUSION.....	19
9.	FUTURE ENHANCEMENTS	20
10.	BIBLIOGRAPHY	21
11.	APPENDIX.....	22

Table of Figures

Figure 1Mask Use Case	7
Figure 2 Sequence Diagram.....	8
Figure 3Activity Diagram	9
Figure 4 ER diagram	10
Figure 5 Login Page.....	12
Figure 6 Sign Up Page	12
Figure 7 Enter Hashtag	13
Figure 8 Display sentimental analysis	13
Figure 9 Display previous sentiments.....	14
Figure 10 Display Word Cloud.....	14
Figure 11 Display previous word cloud.....	15
Figure 12 Show comparisons.....	15

1. INTRODUCTION

1.1 Purpose

The Market Analysis based on Social network serves the purpose of mining real-time information from social networks and to analyze the data to provide useful information to the end-user who is interested in getting to know the influence of his product or brand on the people and to help them make judgments or to streamline their product to achieve success in this ever changing world.

1.2 Product Perspective

The product to be developed is a first of its kind, created with a idea of providing a quick and informative insights about a product by analyzing tweets from the Twitter.com using various technologies to make it possible.

This product is a self-contained and can be extended further to provided new services.

This SRS describes the requirements of the system as a whole and each of the services provided by the product.

It also lists the interfaces to be implemented and their inter-connections.

1.3 Product Functions

It will be driven by three functionalities;

1. **Sentimental Analysis:** it performs a sentimental analysis of the tweets relating to the given keyword/product/brand and display the result in a graphical charts to better understand the sentimental influence of the product.
2. **Word Cloud:** it provide a insights about other products, brand names associated with the given product, it performs a association analysis and displays the result in a Word-Cloud format.
3. **Brand Comparison:** it provides the service of comparing 2 brands side by side to help the end-use to understand where their product stands against another in gaining influence in the social media.
4. The product also stores previous analysis information and provides the same as to help the end-use to understand how their product's influence has changed.

1.4 User Classes and Characteristics

End-user:- This type of user will be able to sign-up or login and use the features provided by the product and the information accessed by this user will be stored in Database for future use

2. LITERATURE SURVEY

2.1 Existing System

The current existing system just provide you a sentimental analysis but with does not provide any info-graphics which will renders the information not much useful.

There aren't any product available in the market which provides a word-Cloud analysis and brand comparison from twitter data which are important features and useful tools for business to make better judgments.

2.2 Proposed System

The proposed system MASK provides all the tools which are missing in the current existing system along with interactive graphics to provide a better UI experience to the end-user.

It also stores information of previous session/searches to enable the user to have a timeline which gives an excellent way to understand the progress of his/her product/service.

It will be able to serve huge amount of users to enable them to analyze their product with real time data from Twitters Big-data database with just a click away.

It provides start-ups and services to analyze their product and streamline their product/service as to meet their user expectations.

It allows them to view the bad and the good tweets about their product/service and serves as a feedback system to their product.

3. REQUIREMENTS ANALYSIS

3.1 Operating Environment

Apache Server, Linux OS, MySQL DB

3.2 Design and Implementation Constraints

1. GUI is only in English.
2. Login and password is used for the identification of users.
3. Only registered users can use the service, if not registered, can sign-up in a minute.
4. Currently its Web-only accessible.
5. Limited to HTTP/HTTPS.
6. This system works on single server.

3.3 User Documentation

The Documentation of the project and the manual will be provided with the software along with the deployment.

3.4 Assumptions and Dependencies

The project assumes that the user has some knowledge of social media like twitter and how it influences the real world.

The project depends on the Apache server and the server up time to make the resource available 99.9%

4. SOFTWARE REQUIREMENTS SPECIFICATION

4.1 Users

4.1.1 End-user

This type of user will be able to sign-up or login and use the features provided by the product and the information accessed by this user will be stored in Database for future use.

4.2 Functional Requirements

4.2.1 Ability for the end-user to perform sentimental analysis on their product.

The end-user will be able to perform a sentimental analysis on their product, where the system fetches the tweets from twitter and performs analysis on each tweet and plots the graph.

4.2.2 The end-user should be able to generate word cloud for his product

The word cloud should be generated for the given product and displayed to the user and stores in database for future use, the previous stored on is overwritten .

4.2.3 The system should provide user to compare two brands.

The user should be able to enter the brand names and generate a graphical report comparing the brands with the real time twitter data .

4.2.4 The system should display the best and the worst tweets for each sentimental report generated.

It should also store the previous tweets and display along with the current one to give the user a preview about the best and worst tweets .

4.3 Non- Functional Requirements

4.3.1 Performance Requirements

(1) Static numerical requirements:

- (a) The number of terminals to be supported: 20+
- (b) The number of simultaneous users to be supported: 20+
- (c) Number of files and records to be handled: 10
- (d) Sizes of tables and files: 100GB

(2) The numbers of transactions and tasks will be 5 per second and 95% of the transactions shall be processed in less than 1 s on a normal workload conditions and the numbers of transactions and tasks will be 20 per second and 95% of the transactions shall be processed in less than 5 s on a peak workload conditions

4.3.2 Safety Requirements

Data will be transmitted with the use HTTP protocols and use inbuilt browser security measures.

4.3.3 Security Requirements

Users will be authenticated based on provided login details.

4.3.4 Software Quality Attributes

Availability

Is available anywhere throughout the world. Will display a 404 page in case of loss of connection. No data will be lost and no commits will be made until complete transaction is made to the server.

Usability

The system is easy to handle and navigates in the most expected way with no delays. The system program reacts accordingly and transverses quickly between its states.

4.4 Software and Hardware Requirement

4.4.1 Hardware Interface

Client Side			
	Processor	RAM	Disk Space
Internet Explorer – 6 or above	All Intel or AMD - 1 GHZ	256 MB	100 MB

ServerSide			
	Processor	RAM	Disk Space
Apache Server	All Intel or AMD - 2 GHZ	2GB	100 GB
MySql		512 MB	500MB(Excluding Data Size)

4.4.2 Software Interfaces

Client on Internet

Web Browser, Operating System (any)

Client on Intranet

Web Browser, Operating System (any)

Web Server

Apache (WAMP), Operating System (any)

Data Base Server

MySQL, Operating System (any)

Development End

LAMP (Linux, Apache, HTML, XML, PHP), MySQL, OS (Windows),

4.4.3 Communications Interfaces

1. It uses the any Modern Web browser capable of running Java-script on client side.
2. It makes use of HTTP Protocol.

5. ANALYSIS AND DESIGN

5.1 Use Case

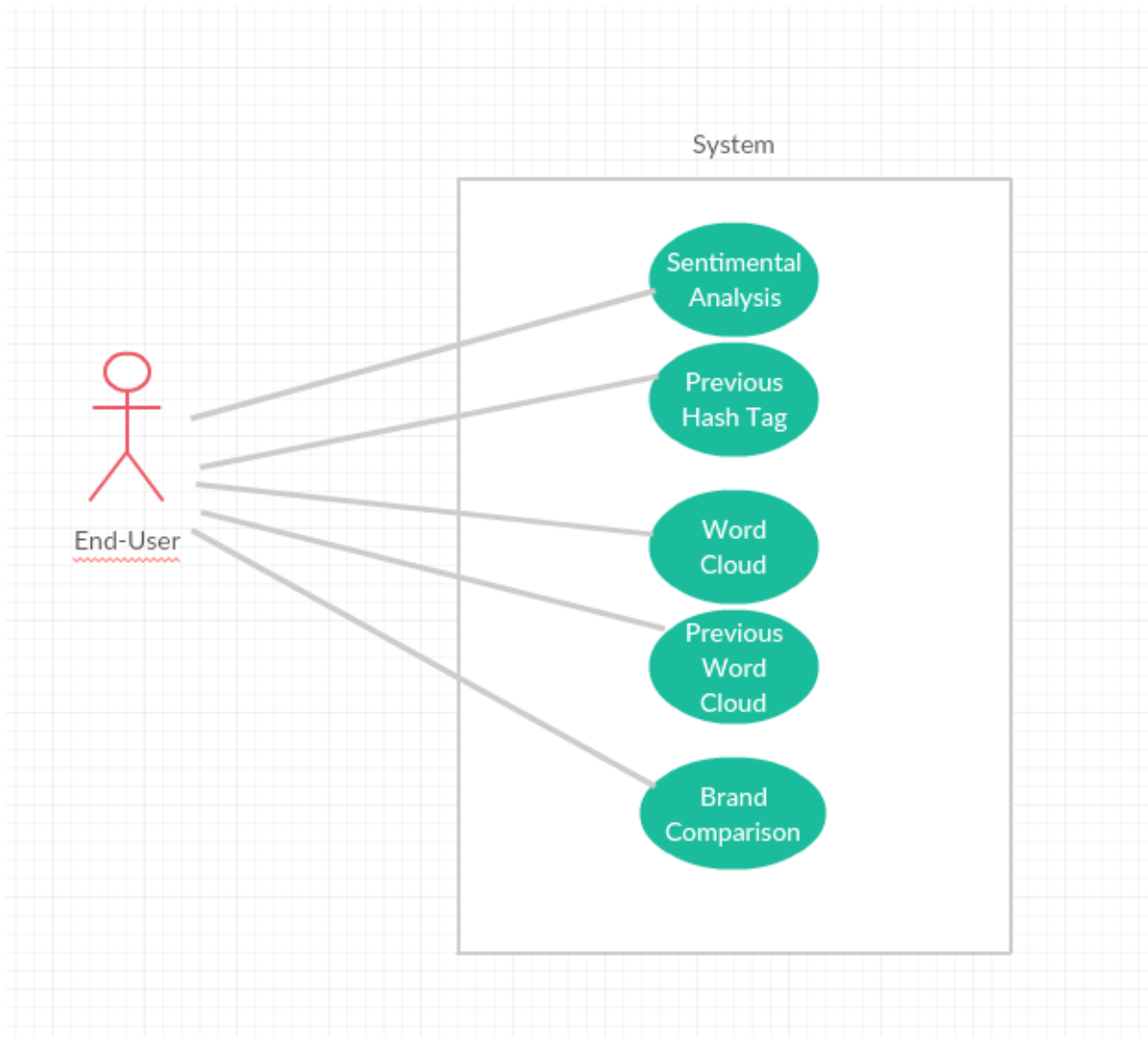


Figure 1Mask Use Case

5.2 Sequence Diagram

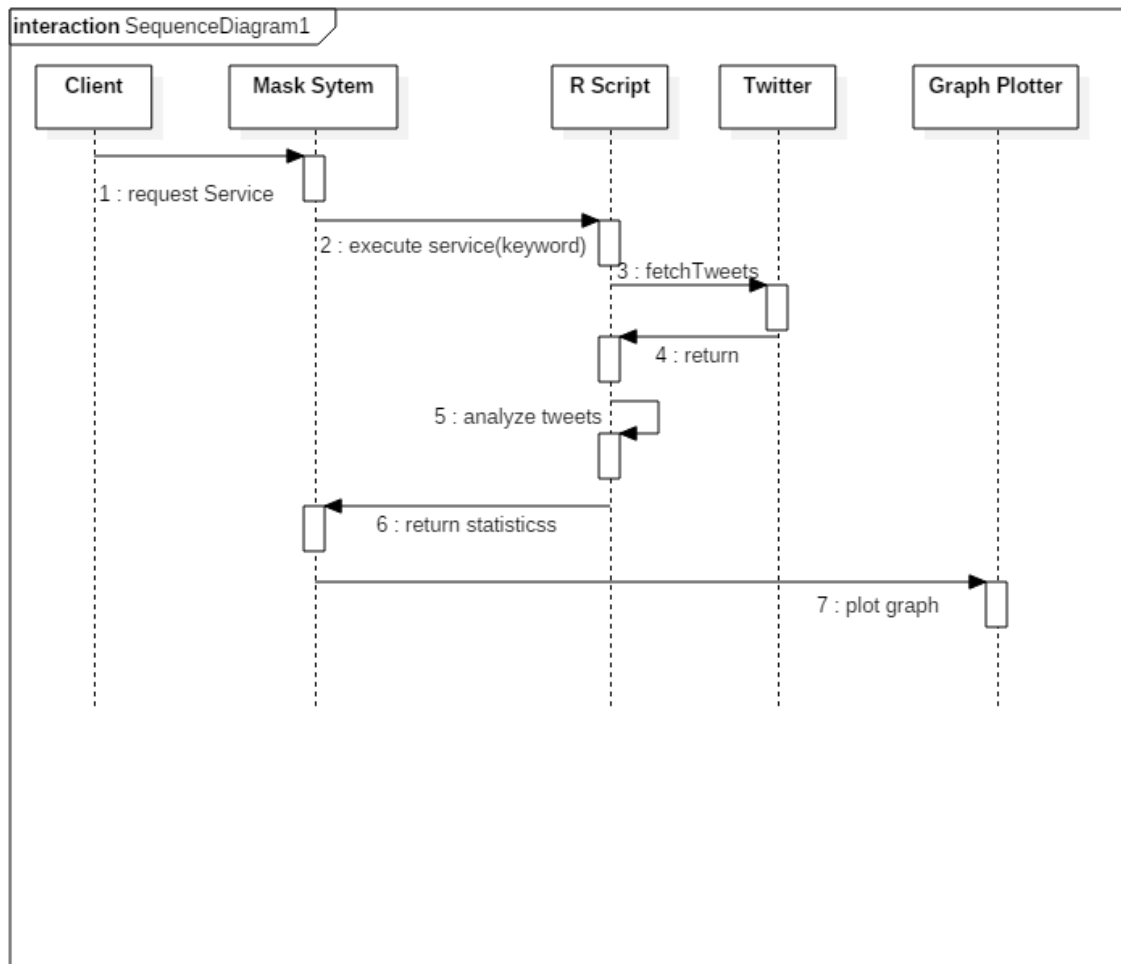


Figure 2 Sequence Diagram

5.3 Activity Diagram

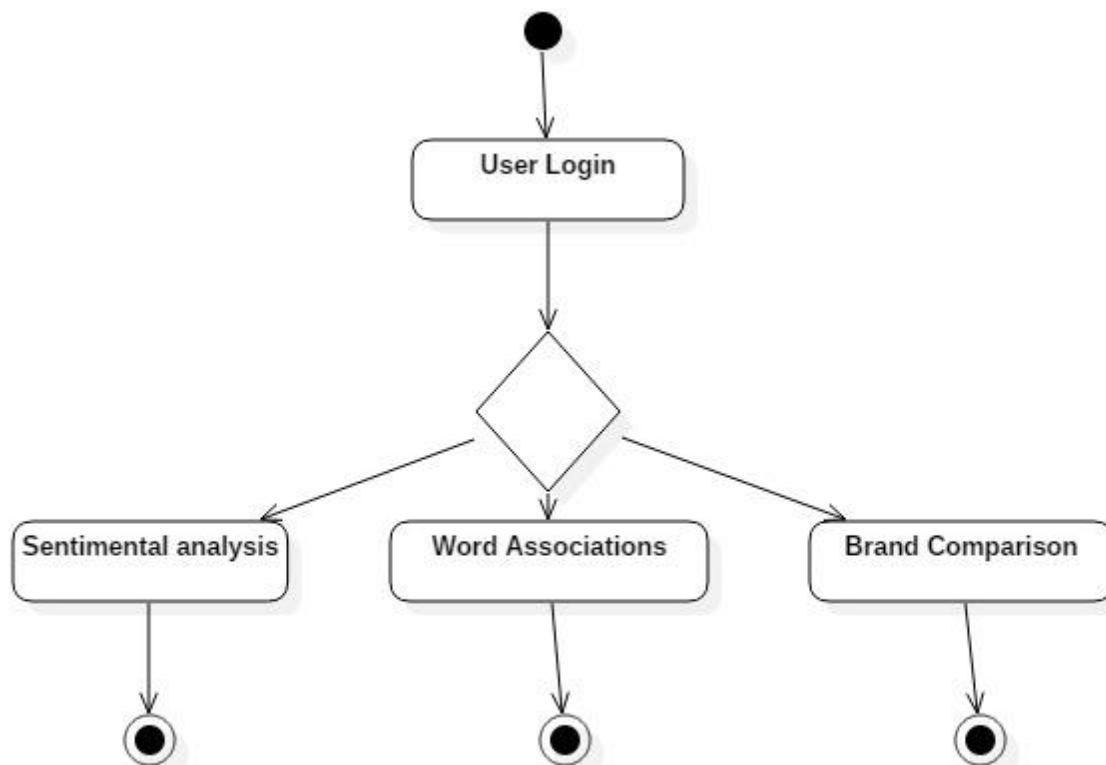


Figure 3Activity Diagram

5.4 ER Diagram

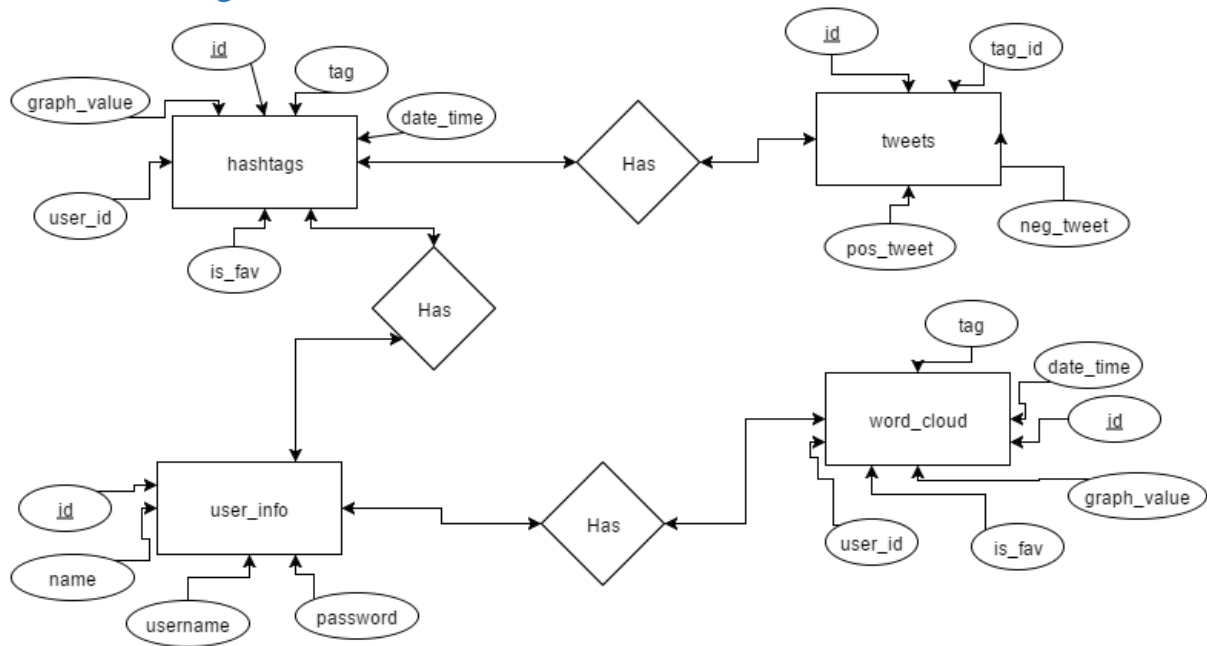


Figure 4 ER diagram

5.5 Data Dictionary

hashtags

Name	Data type	Comment
Id	int(3)	Primary Key
Tag	Varchar(50)	
Date_time	Date	
Graph_values	varchar(500)	
User_id	int(3)	
Is_fav	int(2)	

Tweet

Name	Data type	Comment
Id	int(11)	Primary Key
Neg_tweet	varchar(100)	
Pos_tweet	varchar(100)	
Tag_id	int(3)	

Word_cloud

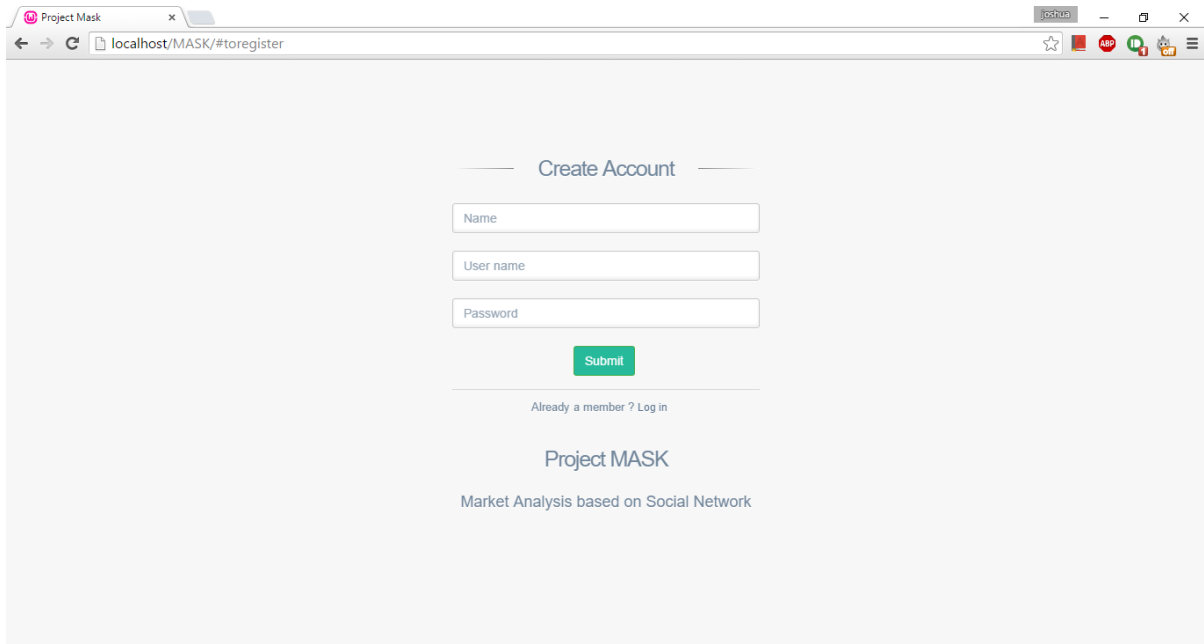
Name	Data type	Comment
Id	int(3)	Primary Key
Tag	Varchar(50)	
Date_time	Date	
Source_location	varchar(500)	
User_id	int(3)	
Is_fav	int(2)	

User_info

Name	Data type	Comment
Id	int(11)	Primary Key
Username	varchar(30)	Unique
Password	varchar(18)	
Name	varchar(30)	

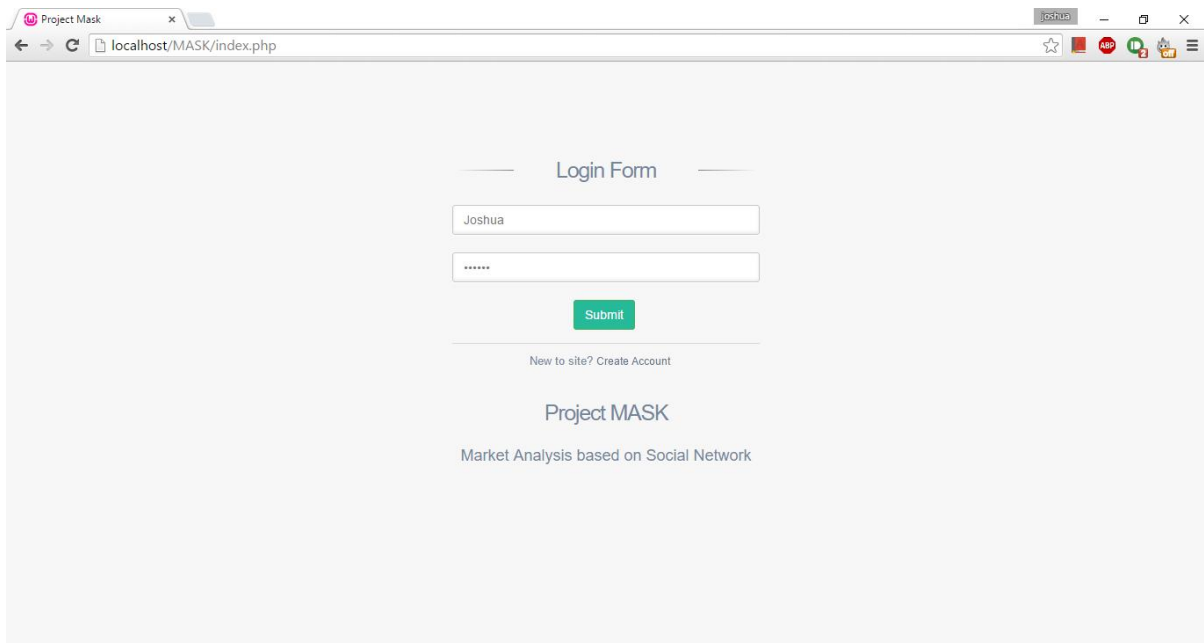
6. IMPLEMENTATION

6.1 Screen-Shots



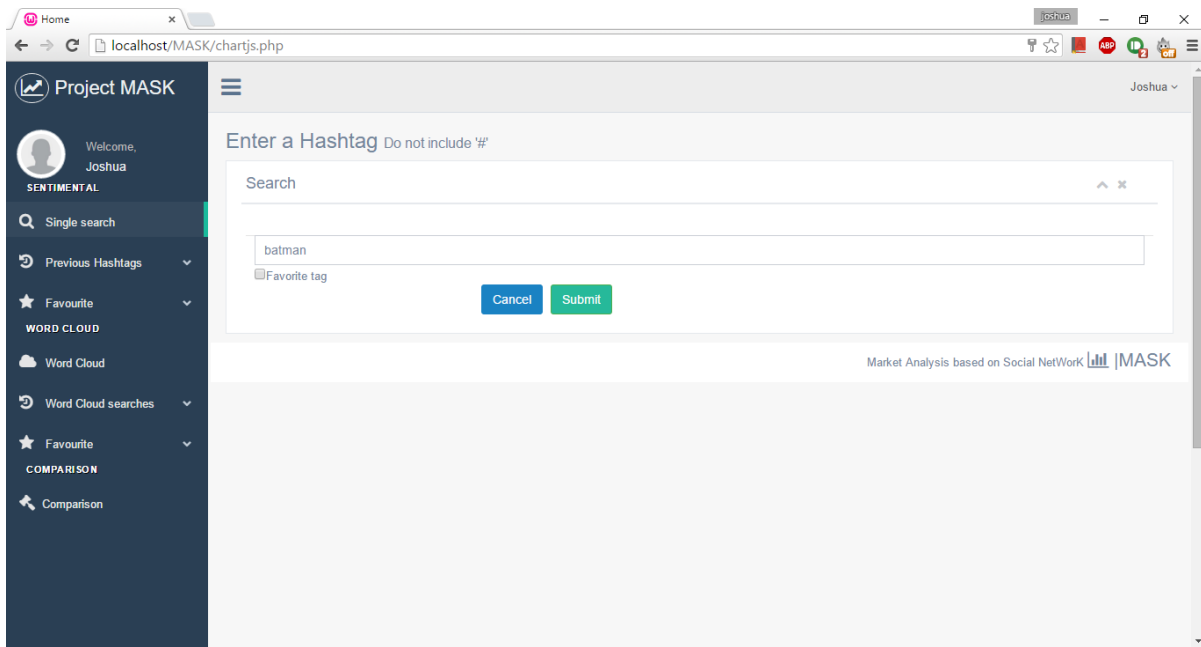
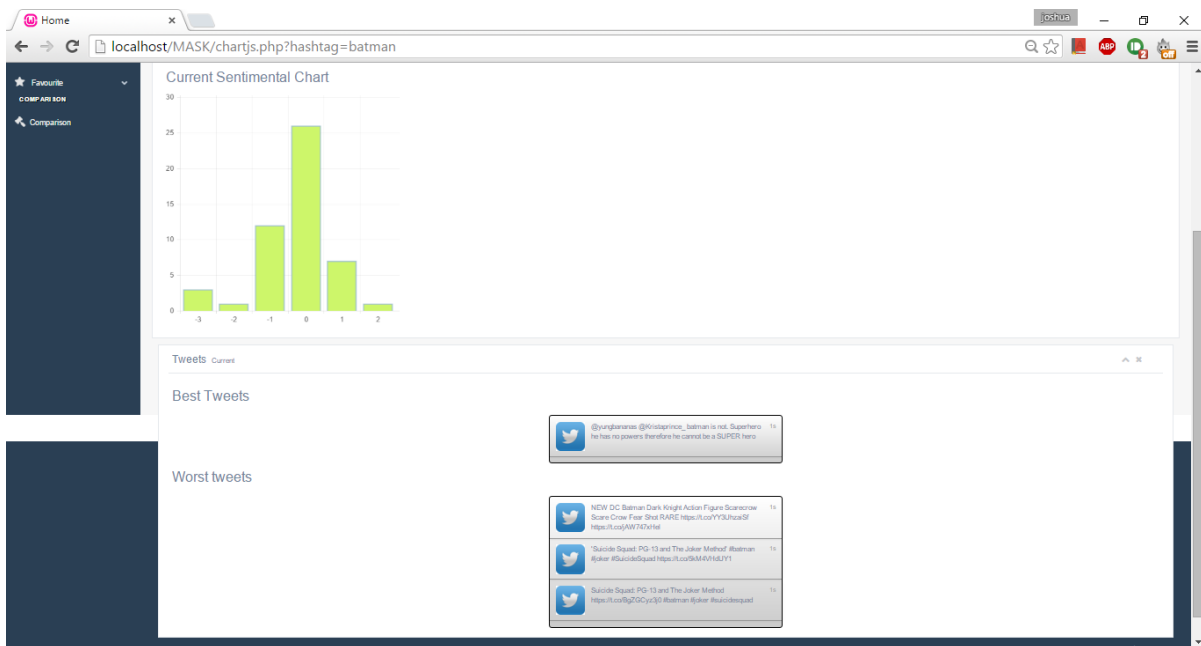
A screenshot of a web browser displaying the registration page for 'Project MASK'. The browser's address bar shows 'localhost/MASK/#toregister'. The page has a light gray background. At the top center, the text 'Create Account' is flanked by horizontal lines. Below this, there are three input fields: 'Name', 'User name', and 'Password'. A green 'Submit' button is positioned below the 'Password' field. Underneath the button, the text 'Already a member? Log in' is centered. At the bottom of the form area, the text 'Project MASK' is displayed, followed by the tagline 'Market Analysis based on Social Network'.

Figure 5 Login Page



A screenshot of a web browser displaying the login page for 'Project MASK'. The browser's address bar shows 'localhost/MASK/index.php'. The page has a light gray background. At the top center, the text 'Login Form' is flanked by horizontal lines. Below this, there are two input fields: the first contains the text 'Joshua', and the second contains a series of asterisks '*****'. A green 'Submit' button is positioned below the second input field. Underneath the button, the text 'New to site? Create Account' is centered. At the bottom of the form area, the text 'Project MASK' is displayed, followed by the tagline 'Market Analysis based on Social Network'.

Figure 6 Sign Up Page

*Figure 7 Enter Hashtag**Figure 8 Display sentimental analysis*

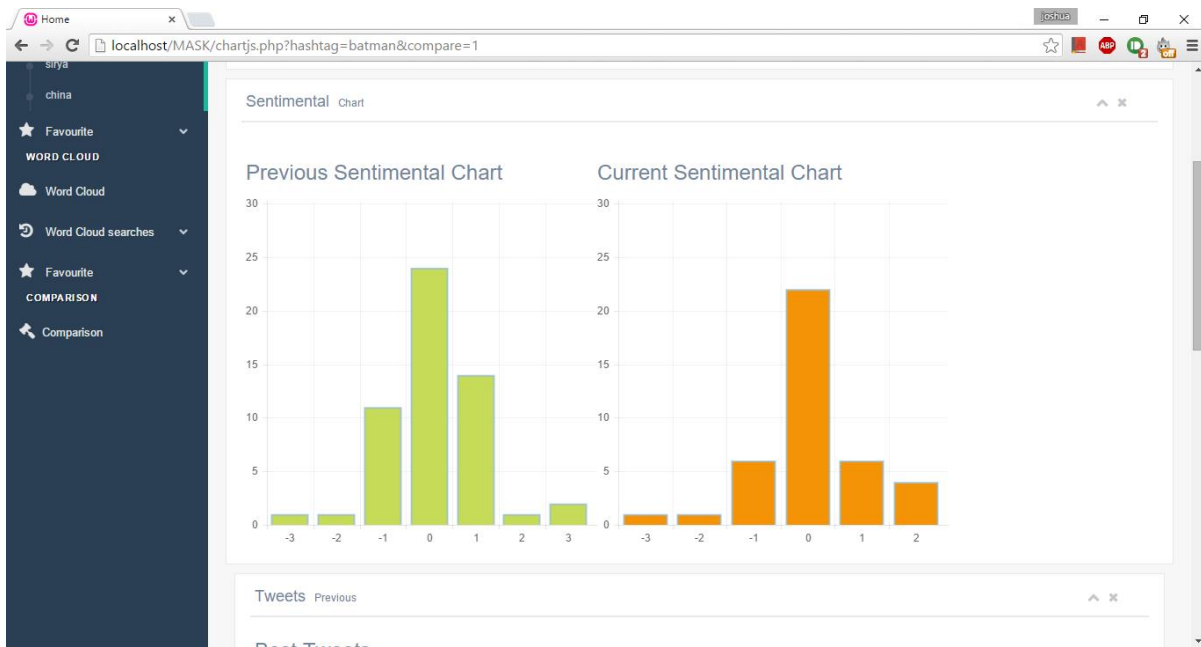


Figure 9 Display previous sentiments

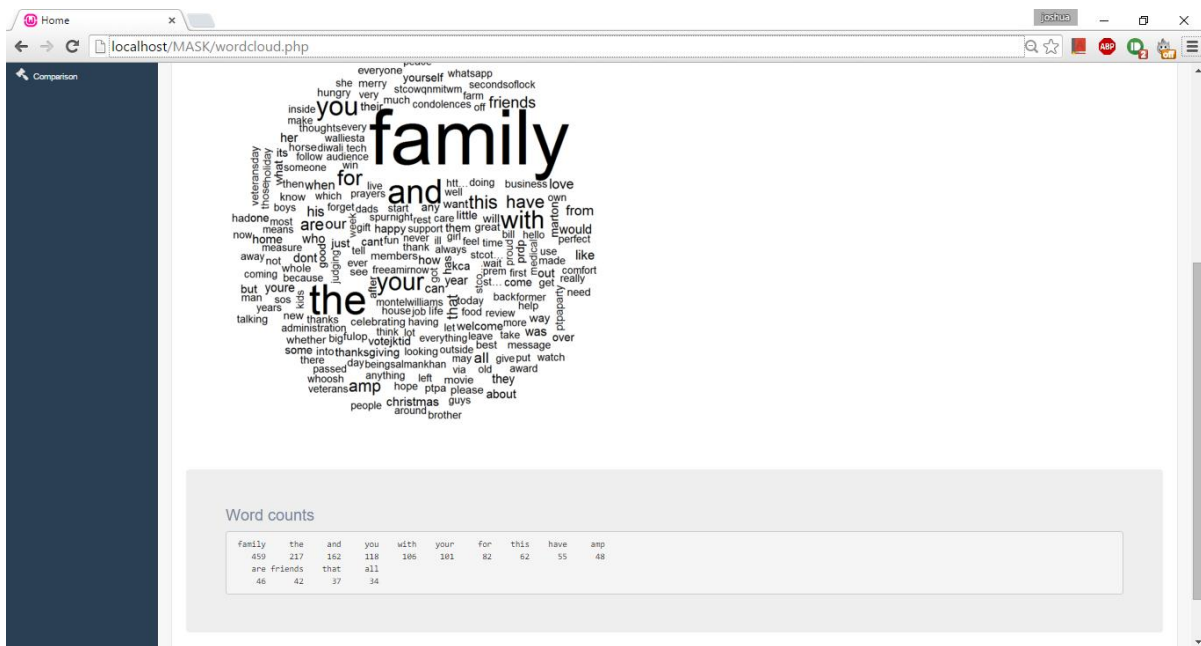


Figure 10 Display Word Cloud

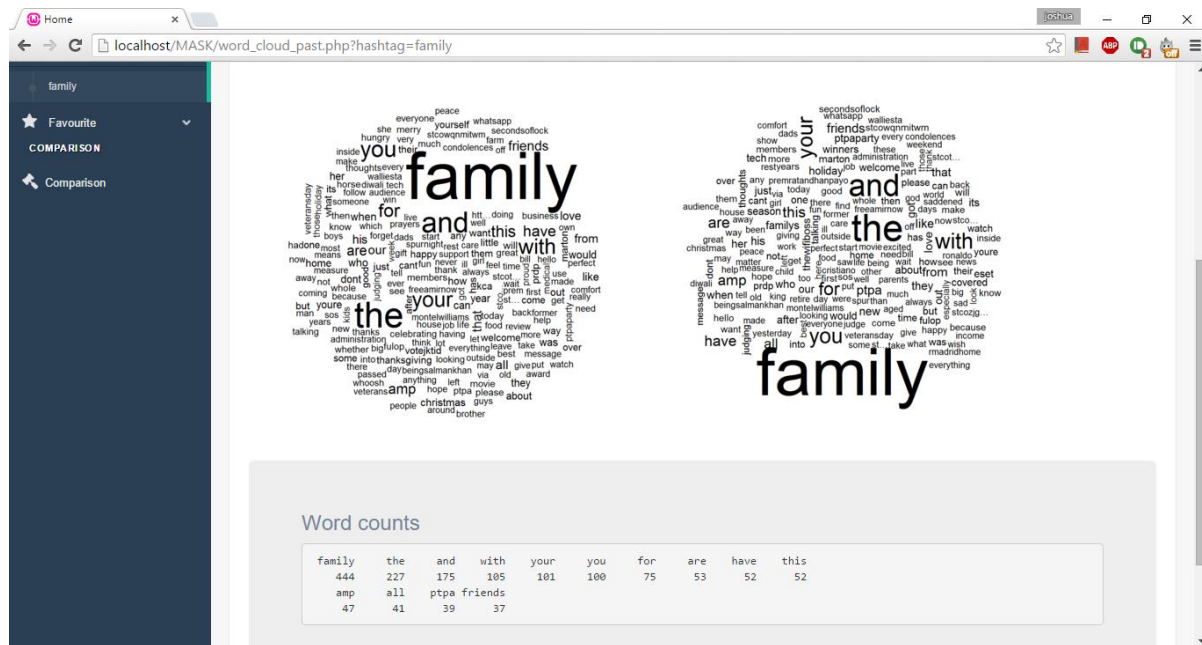


Figure 11 Display previous word cloud

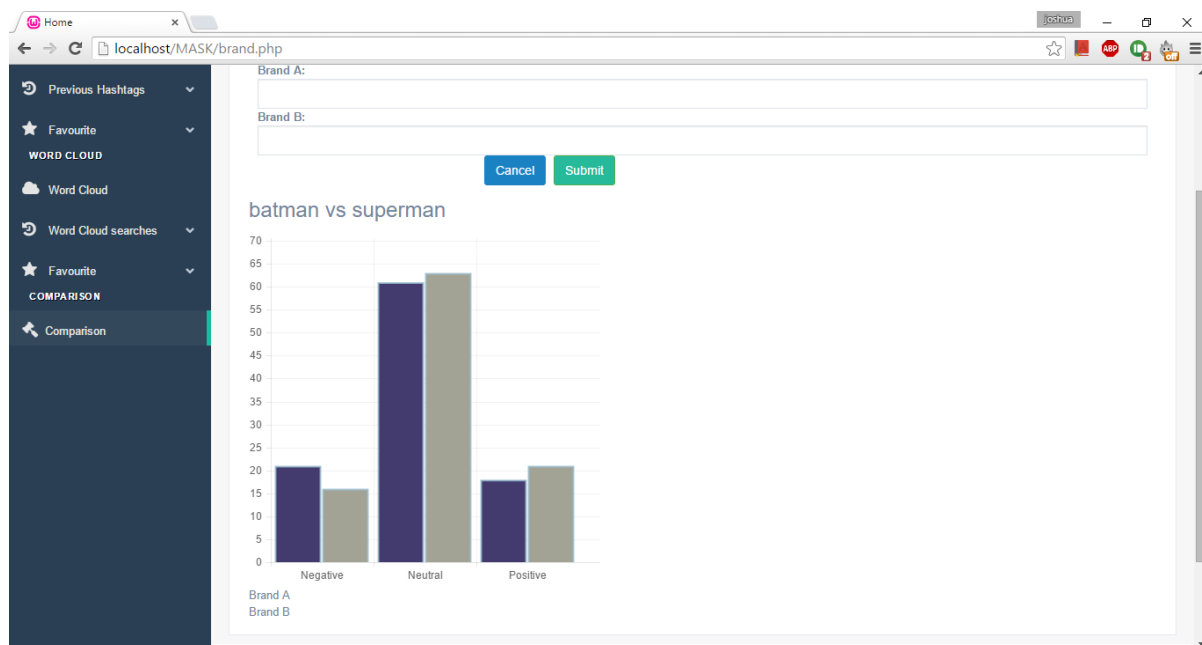


Figure 12 Show comparisons

7. SOFTWARE TESTING

7.1 Testing Objectives:

Testing is the execution an application with the goal of finding an error. A good test case runs multiple tests to find various errors and bugs in the system. A successful test is one that uncovers an unknown error.

7.2 Testing Principles:

- All tests should follow the user requirements
- We should plan long before testing begins
- Testing should start with small modules and move onto the whole application.

7.3 Unit testing:

It is the method of testing during the coding phase. It is uses the requirement document to test the modules one by one and revel an inaccuracies in the code and application.

7.4 Test Cases:

7.4.1 Login Page

S.NO.	Test	User Input	Expected Result	Actual Result	Status
1	Check login	Valid username and invalid password	Print "Incorrect Password"	Set's placeholder of password an invalid	Pass
2		Invalid username	Print "Invalid username"	Sets placeholder of username as invalid	Pass
3		Invalid Password	Print "Invalid Password"	Sets placeholder of password as invalid	Pass

7.4.2 User creation

S.NO.	Test	User Input	Expected Result	Actual Result	Status
1.	Check User sign up.	Null name or only space is provided	Print "Enter a name"	Java alert "Enter a valid name"	Pass
2		Invalid name(special characters)	Prints "Invalid name"	Java alert "Enter a Valid name"	Pass
3		Null username	Print "Enter a user name"	Java alert "Enter a valid user name"	Pass
4		Null Password	Print "Enter a Password"	Java alert "Enter a Password"	Pass
5		Password length less than 8	Print "Password is too short"	Java alert "Please Enter a password with 8 Characters"	Pass
6		Null repeat Password	Print "Enter the repeat Password"	Java alert "Enter the repeat Password"	Pass
7		Repeat password does not match	Print "Passwords do not match"	Java Alert "Passwords do not match"	Pass

7.4.3 Sentimental Search

S.NO.	Test	User Input	Expected Result	Actual Result	Status
1.	Check hashtag form input	Null hashtag	Print "Input hashtag"	Java alert "Enter a hashtag"	Pass
2		hashtag with special characters	Print "Invalid hashtag "	Java alert "Enter valid hashtag"	Pass
3		Hashtag with space in between	Print "Input valid hashtag name"	Java alert "Enter a valid hashtag name"	Pass

7.4.4 Word Cloud

S.NO.	Test	User Input	Expected Result	Actual Result	Status
1.	Check hashtag form input	Null hashtag	Print "Input hashtag"	Java alert "Enter a hashtag"	Pass
2		hashtag with special characters	Print "Invalid hashtag "	Java alert "Enter valid hashtag"	Pass
3		Hashtag with space in between	Print "Input valid hashtag name"	Java alert "Enter a valid hashtag name"	Pass

7.4.5 Brand comparison

S.NO.	Test	User Input	Expected Result	Actual Result	Status
1.	Check two hashtag form input	Null hashtag in textbox1	Print "Input hashtag"	Java alert "Enter a hashtag"	Pass
2		hashtag with special characters in textbox1	Print "Invalid hashtag "	Java alert "Enter valid hashtag"	Pass
3		Hashtag with space in between in textbox1	Print "Input valid hashtag name"	Java alert "Enter a valid hashtag name"	Pass
4		Null hashtag in textbox1	Print "Input hashtag"	Java alert "Enter a hashtag"	Pass
5		hashtag with special characters in textbox1	Print "Invalid hashtag "	Java alert "Enter valid hashtag"	Pass
6		Hashtag with space in between in textbox1	Print "Input valid hashtag name"	Java alert "Enter a valid hashtag name"	Pass
7		Hashtag in textbox1 and textbox2 are different.	Print "Please do not enter the same hashtag in both inputs"	Java alert "Please do not enter the same hashtag in both inputs"	Pass

8. CONCLUSION

The foundation idea of the project was to provide a convenient and unpretentious platform for users to analyze any keyword/product/brand name based on social media, using an online interface to share real time information about the current influence of the same. This project also identifies key words associated with said keyword/product/brand name based on the number of frequent words tweeted by various users.

This project uses data mining and data visualization techniques to provide a graphical informational tool for users using bar charts and word clouds.

This project “Market Analysis based on Social Network” uses the features provided by the various web technologies like, PHP , R scripting language, MySQL v5, Apache server v2 in the backend and HTML v5 and JavaScript in the front end.

This particular project deals with information science dealing with providing crucial analysis of their keyword/product/brand name to the users so that they may take necessary judgment calls to improve their keyword/product/brand name image.

9. FUTURE ENHANCEMENTS

This project provides a base platform and an idea which can be taken forward to involve more modules to provide users to filter out tweets based on geographical locations such as cities, states, countries and continents.

A key improvement on the system would be implementing artificial intelligence and machine learning to analyze the tweets such that a more precise and accurate semantic can be achieved.

Another feature would be to analyze images along with the tweets using OpenCV to identify possible emotions.

Additional to Twitter we could provide the ability to fetch post from other social media such as Facebook and Google Plus.

In regard of data mining and visualization, more reports and visualization technique like bar, line graphs could be generated.

10. BIBLIOGRAPHY

1. www.w3schools.com
2. www.php.net
3. en.wikipedia.org/
4. www.hotscripts.com/category/php/
5. www.apache.org/
6. www.mysql.com/
7. www.stackoverflow.com
8. www.charts-js.com – chart.js library
9. www.C-Ran.com – R scripts and libraries.

11. APPENDIX

Twitter - Social networking website used to post status along with images.

Tweets – Users personal opinion on any matter published on the website.

Hashtag – Used to highlight certain keywords in twitter tweets which can be used to track trends.

Sentimental – An emotional expression, in this case used as positive and negative emotions.

Word Cloud – Collection of words in the form of a cloud to show associations.

Graph – Plotting of values in the x and y axis to show graphical illustrations of values.

Market Analysis – The process of researching the market to find out how a particular brand name/product is doing.

Data mining – The process of fetching and scrubbing data to get valuable information.