Project Report

Analysis of Meetup.com

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

Master of Technology
in
Information Technology

Submitted by

Roll No	Names of Students
MT2015082	Priya Ranjan Sinha
MT2015110	Soumyadip Saha

Under the guidance of
Prof. Shrisha Rao

April 2016

Abstract

In this project, we worked upon the data collected from meetup.com through their API and we have tried to address two problems. First, we have introduced a ranking system among groups of similar interest based on their various characteristics, which in turn would help a user to make wise decision before joining the groups. Second, we also analyzed the data generated by meetup.com to understand when and where new technology communities emerge and evolve. This helped us to understand how new technology is rising in different tech cities of India.

Contents

Chapter 1

Group score Calculation

1.1 Background

Let us say, we want to find out and join some of the best meetups in our area. We can obviously do this task manually, but there could be several groups with similar sounding names and purpose. It becomes difficult to find the right ones by just reading names. Assume that we are in a locality with more than 200 groups in area of our interest. How would we find the best ones?

1.2 Solution

We have introduced ranking system among groups with similar area of interest in a particular city, and also remove the irrelevant groups. This will enable a user to make wise decision before joining a group.
The rankings will be based on normalised group score calculated from data gathered about each group. Parameters considered for calculation of group score are as follows:

- Frequency with which the group conducts meetings(meetups) for its members
- Average rating of the group
- Number of active members in the group (A member is considered active if he has attended at least two of the last five meetups that the group conducted)
- Number of leaders of the group (A leader is a person who delivers lectures and/or is in charge of organising meetings)

- Total number of members of a group

1.3 Steps undertaken for calculation of group score

In order to introduce a ranking system based on the above mentioned parameters we have undertaken the following steps :

1. Take input from user, in which country and in which city he/she wants to search meetup groups

2. Fetch the details of the group associated with the topic by the using meetup.com API method : https://api.meetup.com/2/groups

3. Removing irrelevant groups:
   The relevancy of the group is determined by the number of times the topic (user inputs) appears in the Description and Title of the group.

4. Fetch the past events for each Group, by using the meetup.com API method : https://api.meetup.com/2/events

5. Calculate the frequency with which a particular group creates events: (No. of meetups conducted) / (Age of the group).

6. Calculate the average rating of the group: Sum of rating of all events/ Total no. of events

7. Fetch the active members for each Group, the users who have attended at least 2 of the last 5 events of the groups are termed as active members.The following meetup.com API method is used:
   https://api.meetup.com/2/rsvps

8. Fetch the leads for each Group, by using the meetup.com API method https://api.meetup.com/2/profiles

9. Now we standardized the data using Z-Score :

$$Z = ((X - \mu)/\sigma) \quad where \quad mean \quad is \quad denoted \quad by \quad \mu$$

10. We now calculate the Group Score using the following formula :

$$Z = ((w1 * Z_f) + (w2 * Z_m) + (w3 * Z_a) + (w4 * Z_r) + (w5 * Z_l))$$
$$where,\ Z_f\ denotes\ frequency\ of\ meetups,$$
$$Z_m\ denotes\ total\ number\ of\ members,$$
$$Z_a\ denotes\ number\ of\ active\ members,$$
$$Z_r\ denotes\ average\ rating,$$
$$Z_l\ denotes\ number\ of leads$$
$$and\ w_i\ are\ the\ weightages\ of\ the\ corresponding\ parameters.$$

11. Finally we convert the calculate Group Score to the scale of 10.

12. Display all the groups according to their scores to the user.

1.4 Technologies Used

- meetup.com api for fetching data
- Python (various libraries like pandas, json, etc.) for handling data coming from api.
- Flask Web framework (a python library) for dynamically generating HTML pages.
- HTML, CSS, JavaScript, BootStrap for web pages.

Chapter 2

Analysis of Indian Tech Landscape

2.1 Background

Tech meet-ups have become an important feature of the digital innovation landscape. In these events, coders, designers, hackers and entrepreneurs (among others) come together to learn from each other and network. Meetups can help participants keep their technology skills fresh in fields that move too fast for universities and training providers, and facilitate collaboration and job mobility, increasing the connectivity and efficiency of local innovation ecosystems.
Websites like Meetup has emerged to make it easier for people to create and manage meet-ups. The data generated by these platforms could help us understand when and where new technology communities emerge and evolve, and how they are connected to each other. It could also tell us something about the rise of new technologies. These are questions of obvious interest for policymakers, entrepreneurs, businesses and investors who want to identify the right communities of innovators to work with, and the right technologies to target.

2.2 Solution

We have undertaken a preliminary exploration of India tech meet-ups from Meetup.com to assess its potential as a source of information about the structure, geography and evolution of digital tech in India. The analysis is done on the following major tech cities of our country: Bangalore, Chennai, Delhi, Hyderabad, Mumbai and Pune.

2.3 Data Set

Our data set consists of total 1763 groups in tech category, with the following distribution along the cities:

- Bangalore - 586
- Chennai - 167
- Delhi - 313
- Hyderabad - 258
- Mumbai - 189
- Pune - 250

2.4 Steps undertaken for analysis

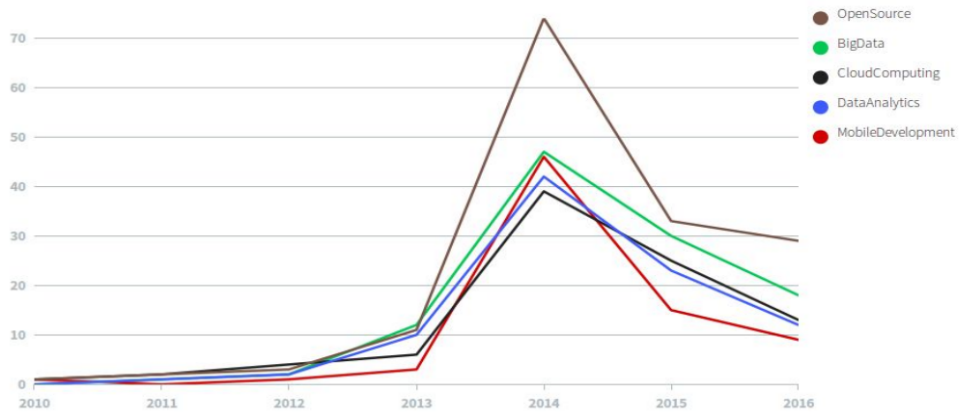
1. Using the meetup.com API we fetched all the meetup groups created in tech category in India between 2010 and 2016.
2. For each of the above groups, fetch all topics.
3. Determine the top most trending topics (which have highest no. of groups) in each city.
4. For each such topic show on graph (using D3.js) city wise, how that topic got evolved over time.

2.5 Technologies Used

- meetup.com api for fetching data
- Python (various libraries like pandas, json, etc.) for handling data coming from api.
- D3.js (a javascript library) for visualizing graphs.
- HTML, CSS, JavaScript, BootStrap for web pages.

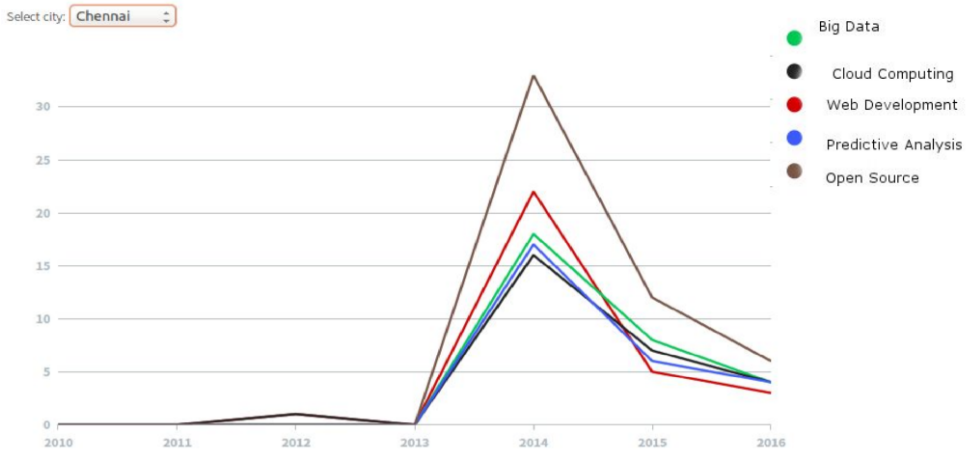
2.6 Analysis

Bangalore :



From the graph, we can conclude that Open Source is the most trending topic of meetup creation in Bangalore for the last 5 years, followed by Big Data and others.

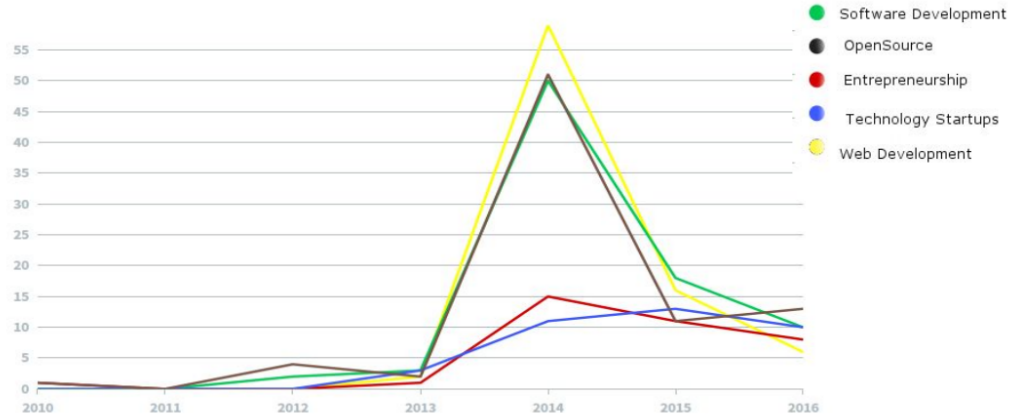
Chennai :



From the graph, we can conclude that Open Source is the most trending topic of meetup creation in Chennai for the last 5 years, followed by Web Development and others.

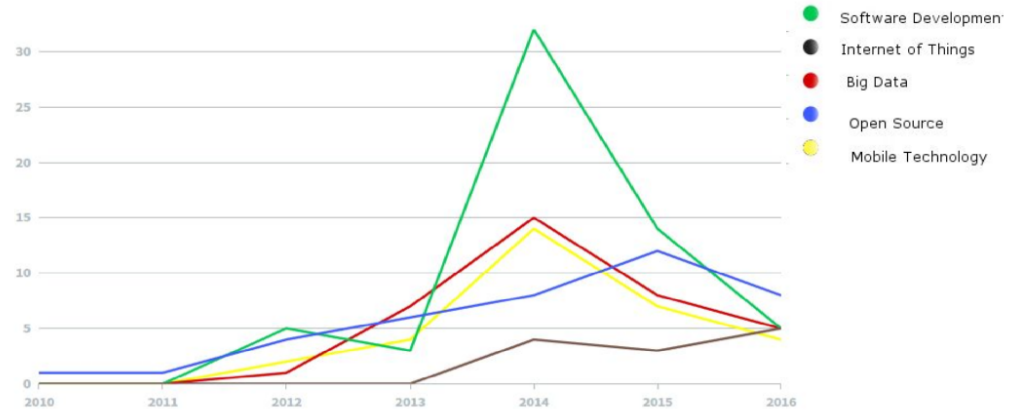
In both these cities Cloud Computing meetup groups are not created frequently.

Delhi :



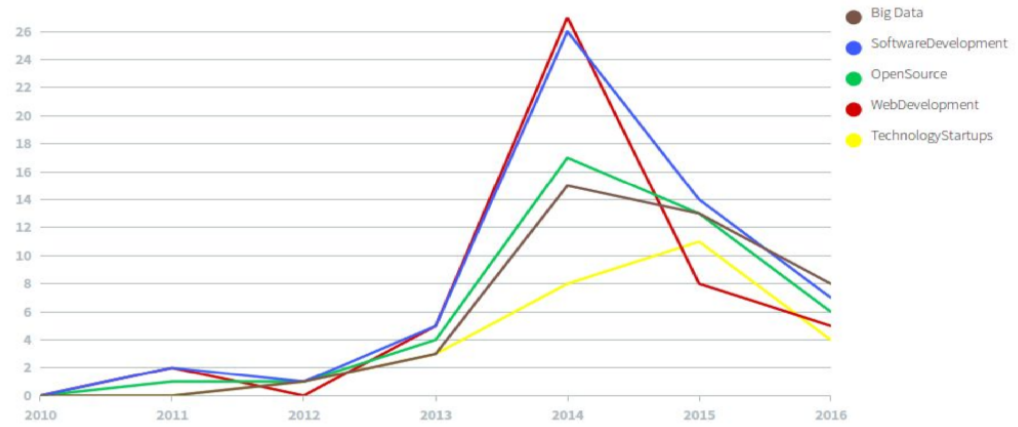
From the graph, we can conclude that Web Development is the most trending topic of meetup creation in Delhi for the last 5 years, followed by Software Development, Open Source and others. From the graph it can be told that Entrepreneurship is gaining popularity in Delhi.

Hyderabad :



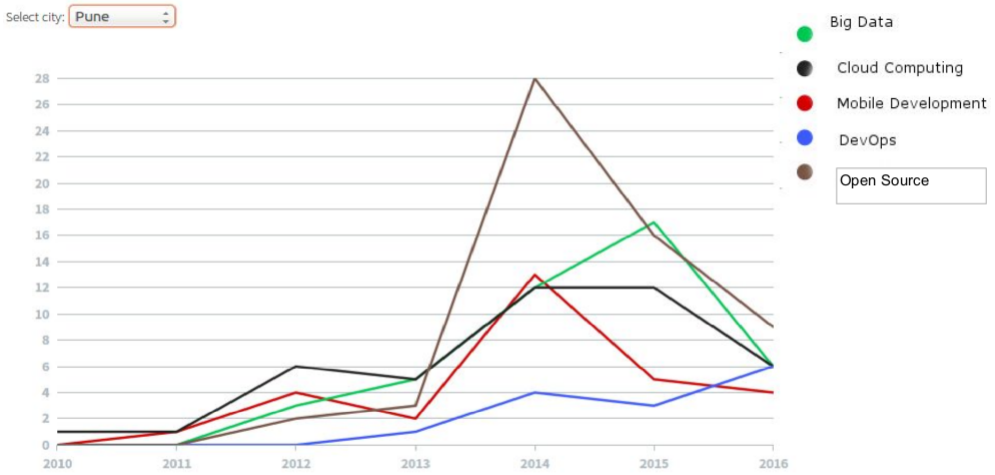
From the graph, we can conclude that Software Development is the most trending topic of meetup creation in Hyderabad for the last 5 years, followed by Big Data, Mobile Technology and others. From the graph it can be told that Internet of Things is gaining popularity in this city.

Mumbai :



From the graph, we can conclude that Web Development and Software Development are the most trending topics of meetup creation in Mumbai for the last 5 years, followed by Open Source, Big Data and others. From the graph it can be told that Start Up culture is gaining popularity in this city.

Pune :



From the graph, we can conclude that Open Scource is the most trending topic of meetup creation in Pune for the last 5 years, followed by Big Data, Cloud Computing and others. From the graph it can be told that Dev Ops is gaining popularity in this city.

Chapter 3

Future Work

Websites like meetup.com generates lots of data which could be used for improving our understanding of the emergence and evolution of new technology fields and communities, their interrelationships and their geography.
Meetup.com also provides streaming interfaces using WebSockets which could be used for real time analysis. The streaming data based on RSVP, Events can be used to provide real-time updates of :

- Popular events,topics by RSVP
- Top cities with most events
- Number of events created in a Country and visualize them on maps.

Sentimental analysis can also be done on Event comments to classify event comments into good, neutral, bad which can help event organizers and users to judge event.

Chapter 4

Conclusion

This project helped us to learn new technologies, and gave us an idea of data analysis. We have successfully implemented Group Score, which ranks the group according different parameters like frequency of meetups, total no. of members etc. This would enable a user to judge the quality of the group and make a wise decision before joining a group. We also enhanced the search feature so that groups which are irrelevant to the topic searched are not shown in search results.
We have also done a detail analysis of Indian Tech landscape using meetup.com data. We found out the trending topics in six major tech cities of India. We analysed how communities related to these topics emerge and evolve with time. Our analysis shows that not the same set of topics are popular throughout the country, in fact different topics are popular in different cities at same period of time.

References

[1] Meetup.com API docs http://www.meetup.com/meetup_api/

[2] Python Docs https://docs.python.org/2/tutorial/

[3] Data Normalization and Standardization
http://www.benetzkorn.com/2011/11/data-normalization-and-standardization/

[4] Flask Docs http://flask.pocoo.org/docs/0.10/tutorial/

[5] D3 Docs https://github.com/mbostock/d3/wiki