

Data Analysis of startups investments and funding trends in India

Piyush Anand Verma¹, Vikas Singhal²

¹MCA 3rd year student, Department of Computer Applications, JSS Academy of Technical Education, Noida (India)

²Professor, Department of Computer Applications, JSS Academy of Technical Education, Noida (India)

Abstract - In the present decade, India is planning a vital shift in the direction of startup welcoming policies and a business-friendly or entrepreneur's environment. India is one of the fastest growing countries in terms of entrepreneurship. Entrepreneurship is an essential input for economic development, creating new markets or opportunities. Currently, India is promoting entrepreneurship enthusiastically but it's a huge challenge for government as well as large population of India to create employment via startups. This paper intention at offering an analytical overview of the boom and potentialities of startup systems in India i.e the progress made by India so far. Therefore, this research can contribute to a better understanding of the Investment and financing strategy of entrepreneurial ventures.

Key Words: Startups, Entrepreneurs, Analytical, Business-friendly

1. INTRODUCTION

The Indian startup Ecosystem started in the late 1960s with the start of TCS, followed by Infosys in 1981 and others. These startups started as software services serving Indian software needs, and later expanding to exporting software services.

The dot-com era was a blessing for many startups, which included marketplaces, ecommerce and vendors. And finally, modern day startups comprised of the latest technology started around 2007-2008, and comprised e-commerce, logistics, marketplaces, and advertising startups. Startups like Flipkart (2007) and IXIGO (2008) were part of these. Currently, there are more than 5,200 startups in India putting it in the fourth spot just behind US, UK, and Israel (Nasscom, 2018), where Bangalore, Mumbai, and Delhi-NCR form 68% of start-up base with 44% foreign investments.

Being so startup friendly the country has attracted numerous numbers of investors, both national and international. Therefore a large amount of money is poured into the startup ecosystem. Also due to government support, technology boon and rise of tier-2 and tier-3 cities has boosted the startup ecosystem. Tier II Cities like Varanasi, Kochi and Indore and others have also attracted a few startups. Events like the launch of Startup India initiative, Digital India Initiative, US elections and the Indian banknote demonetization had a huge impact on the startup community in recent years. Military events like the

surgical strike on Pakistan border by Indian army impacted the funding in India.

This paper will analyze the data from various startups from the start of 2015(January) to mid of 2017(August). The dataset is acquired from Kaggle.com as "Indian Startup Funding. It includes columns with the date funded, the city the startup is based out of, the names of the funders, and the amount invested (in USD). This will also help to realize how big events like US elections and Indian banknote demonetization, in this case, affect the funding of different startups.

There are various questions that are answered through analyzing the available dataset are like who are the important.

2. LITERATURE REVIEW

Major people considered events of 1991 as the turning points for the Indian Startup ecosystem in India when the country was on the verge of bankruptcy that made the government to rethink its policies. Because of India's ever increasing entrepreneurial performance some analysts consider India as the next Asian miracle (Huang, 2008).

Giurca Vasilescu (2009) suggested that Investors are the main source to provide funding for the developing companies. They provide financial assistance, from the beginning till the company are ready to be on the capital market. Investors also provide managerial support to make companies survive in the competitive market.

Tracxn, a venture capital analytics firm found out in their research that a total amount of \$6.4 billion funding were given to various startups in the first nine month of 2015, out of which \$3.4 billion was invested in online market place in India. (Velayanikal, 2015)

Honorable Prime Minister Dr. Narendra Modi announced the "Start-up India, Stand-up India" on the 15th of August 2015 to promote Bank financing for startups and offer incentives to boost entrepreneurship in India. This is considered as the initial seed to the future of startups in India for the ease of doing business.

The increasing ease of doing business is also bringing in investors in some much-needed but neglected areas. For instance, the focus on the Insolvency and Bankruptcy Code (IBC) has encouraged asset reconstruction companies

(CDPQ from Canada and Encore Capital from the US) to invest in India (Hari, 2018).

Crowd funding, which involves raising small amounts of capital from a large number of individuals, is considered to be a major disruption in entrepreneurial financing (Kshetri, 2016). This is because risk taken by investors are less in crowding funding as compared other funding option where risk factor is higher.

1.1 Challenges faced by startups in India

- (i) Lack of availability of quality funds for early stage startups, especially the ones seeking seed money (initial money).
- (ii) Lack of expertise or knowledge. Government policies keep on changing day to day basis Therefore, lack of adequate mentoring/guidance.
- (iii) Hiring and retaining good quality talent is challenging, especially in the areas of product and technology.
- (iv) Availability of loans by banks and other financial institutions at right time is also challenging.
- (v) Lack of infrastructure facility in rural areas stopping startups to start over there.

- Data Cleaning
- Data Transformation
- Data Integration
- Data reduction

Here data needs to be cleaned to avoid wrong conclusions.

For this task, we have identified the null values in each column. They are as follows:

Column	Null Values
Remarks	1953
SubVertical	936
AmountInUSD	847
CityLocation	179
IndustryVertical	171
InvestorsName	8
InvestmentType	1
StartupName	0
Date	0
SNo	0

Since we can see Remarks have the most number of null values, 1953 null values, which is 82.3 % of the total remarks values. Therefore we will drop the column since it would not provide any useful information during the analysis.

We would also need to fix some date formats, spelling typos, and combine alternate names of a company into a single name.

Ex: 12/05.2015 to 12/05/2015, Seed Funding to Seed Funding, Ola to Ola Cabs, etc.

The amount of investment column is present as 'AmountInUSD' column is separated by commas which must be removed so that mathematical operations could be applied over it.

3.3 Analysis of Data

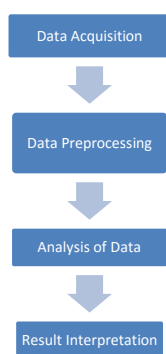
The data is analyzed using python libraries like pandas, numpy, dateutil etc. Using these and many other libraries in-depth analysis was possible.

3.4 Result Interpretation

The results are interpreted visually for a better understanding of what the data says. This visualization was possible by using python libraries like matplotlib,

3. METHODOLOGY FOR DATA ANALYSIS

The basic process of data analysis comprises only 4 steps and is a given below and is followed as well.



Methodology for Data analysis

3.1 Data acquisition

The data is acquired from kaggle.com as "Indian Startup Funding" submitted by Sudalai Rajkumar (2). The file is present in a CSV (Comma Separated Value) format.

3.2 Data preprocessing

Data preprocessing consists of the following tasks:

seaborn and squarify. Various visualization graphs and diagrams are used like bar plots and square plots. Bar plots are charts that represent categorical data with rectangular bar.

These results would help us in understand where our startup ecosystem stands and what major factors play role in the funding of a startup.

4. RESULT

The findings resulted from the data are as follows:

The dataset has 2372 rows and 10 columns. The columns are 'SNo', 'StartupName', 'IndustryVertical', 'SubVertical', 'CityLocation', 'InvestorsName', 'InvestmentType', 'AmountInUSD', and 'Remarks' out of which 1 column, i.e. Remarks was dropped because having a significantly large number of null values.

The total number of unique startups that were funded in 1997 and out of which 208 startups were funded more than 1 times with Ola Cabs topping the list with 9 funding.

The minimum investment was found out to be of \$ 16 000.0 by Hyderabad Angels. There were a total of 5 investments done by Hyderabad Angles at a startup event, Startup Heroes event, all on the same day (19/01/2015) in Hyderabad (Keshamoni, 2018). The companies that were funded were Hostel Dunia, Play your sport, Chakra Business Solutions, Enabli, and YoGrad.

The maximum investment was of \$ 1400000000.0 to Paytm and Flipkart. Softbank invested in Patym on 18/05/2017 and Microsoft, eBay, Tencent Holdings invested in Flipkart on 27/03/2017. Also, Paytm and Flipkart were funded 4 other times during 2015 to 2017.

The mean investment is of \$ 12031073.09 and the total investment is of \$ 18347386476.0 that is equal to the total money that will be needed to stretch a wall between Mexico and the US.

Now calculating the number of investments per month resulted that January 2016 had the most investment (104) followed by February (100) and June (98) of the same year.

Visualizing the data for the number of investment per month gives us an in-depth analysis of the funding trend. One advantage of big data visualization is that it provides a ready means to tell stories from the data.

In this age, the volume of data that companies are able to gather about customers and market conditions can provide business leaders with insights into new revenue and business opportunities, presuming they can spot the opportunities in the mountain of data. Using data visualization, decision-makers are able to grasp shifts in customer behaviors and market conditions across multiple dataset much more quickly.

4.1 Number of funding over months

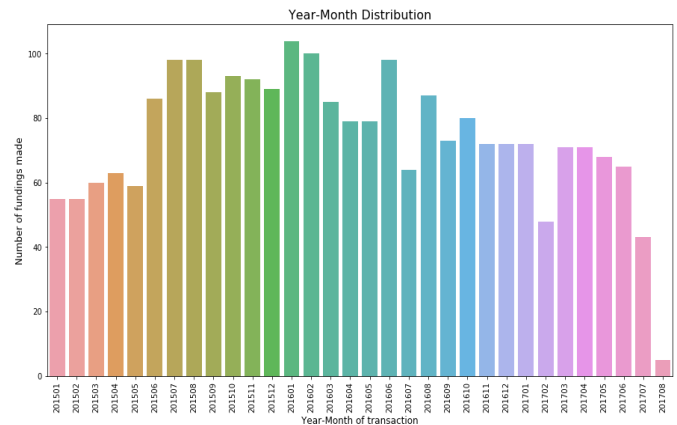


Figure 1

The above graph shows a distribution of month of transaction and number of funding made. Here there was a hike in the number of funding made from June 2015 to June 2016. Further conclusion shows:

- July and August of 2015 had the highest investment that year because of the Digital India Campaign
- January 2016 had the highest investment of 2016 and also in the dataset because of the Startup India Initiative
- As expected Demonetization lowered the investment per month from Nov '16 to July '17
- Interesting note: July '16 saw a decrease in investment and the lowest in the year, maybe because of the surgical strike that happened that month. (Wikipedia, India-Pakistan border skirmishes (2016-present), 2018).

4.2 Amount of funding over months

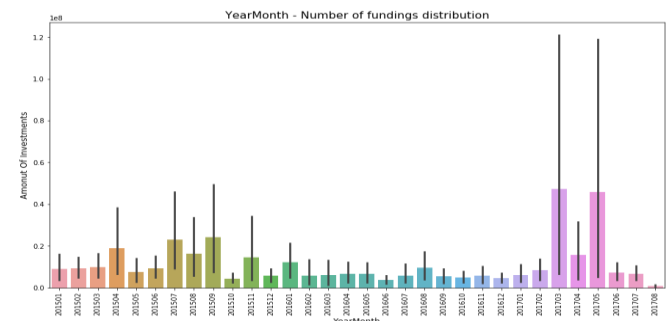


Figure 2

The above graph shows the distribution of Month to Amount of funding made. March 2017 and May 2017 show the highest money invested in the available distribution because that was the time when Flipkart and Paytm were funded then with an investment of \$ 1,400,000,000.

4.3 Startups most funded

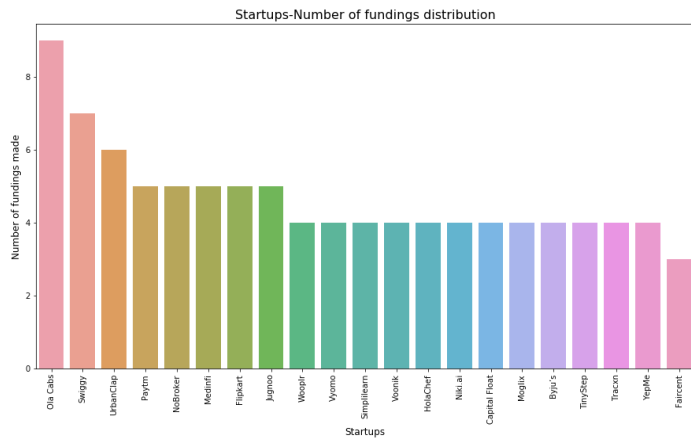


Figure 3

From Figure 3 graph we can conclude that out of the 208 startups that were funded more than 1 times 20 startups were funded 4 or more times with Ola Cabs getting funded 9 times, Swiggy 7 times, and UrbanClap 6 times.

Most of the companies that were funded 4 or more than 4 times were Consumer Internet companies with some exceptions.

There are 741 unique Industry verticals present in the dataset.

4.4 Industry verticals & sub verticals & Number of funding distribution

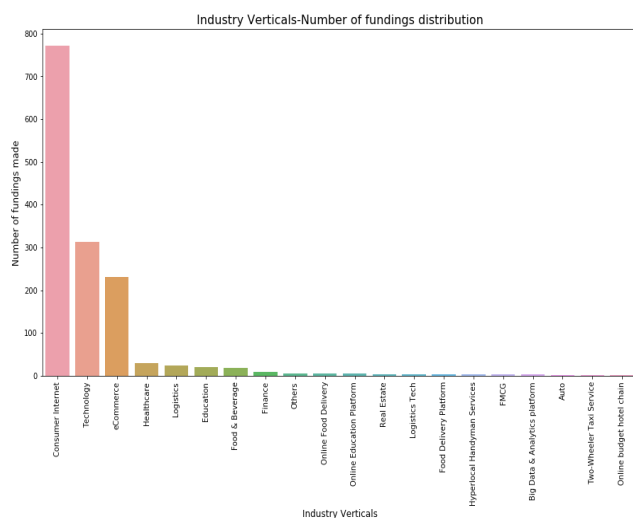


Figure 4

Consumer Internet Companies acquired the most number of funding 772. Whereas Technology and e-commerce got a total number of 313 and 230 funding respectively. The distribution is as follows

- Consumer Internet startups are the most with 32.5 % in total.

- Technology with the second most of 13.2%
- And E-commerce with 9 %

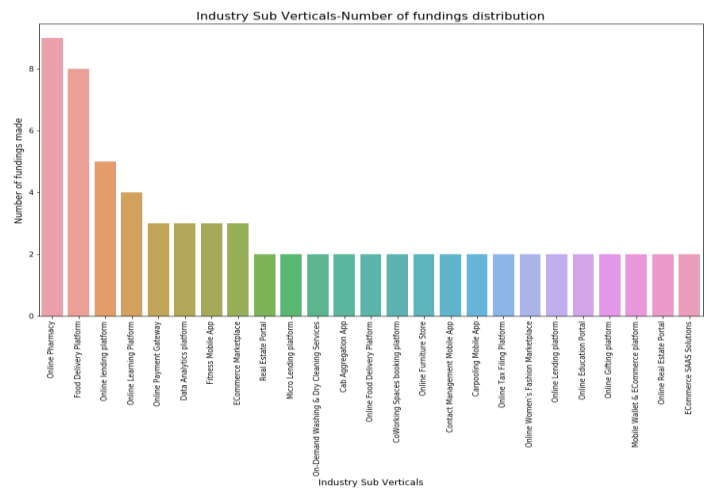


Figure 5

In figure 5 we come to know about Industry Sub Verticals, in this, we see that 'Online pharmacy' is the highest sub-vertical chosen for funding. A total number of 1365 sub-verticals are present. Online Pharmacy, Food Delivery, and Online lending platforms are the top 3 sub-verticals.

Seed Funding and Private Equity are the most preferred way of investment since investors want to bag their money on new startups by seed funding it off by providing private equity for already established ones.

Seed funding is what we require to launch the startup. Most of this money is raised to fund the marketing & promotional costs to acquire customers so that the startup generates some heat & traction in the industry to which it is catering to. A lot of the money also goes in recruiting qualitative potential resources.

Private Equity is a type of capital instrument which is bought by private investors (as supposed to the general public) in exchange for money which is generally calculated on the basis on the valuation of the startup.

4.5 Investment types in funding

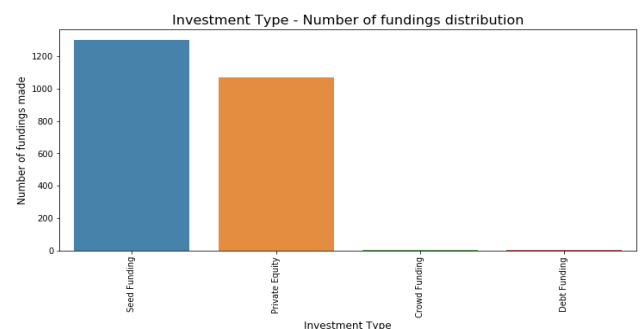


Figure 6

As we can see that Seed funding and private equity are the most preferred type of investment with 1301 and 1067 funding respectively. Whereas Crowd funding and Debt funding has only 2 and 1 investment done respectively.

4.6 Cities most favored

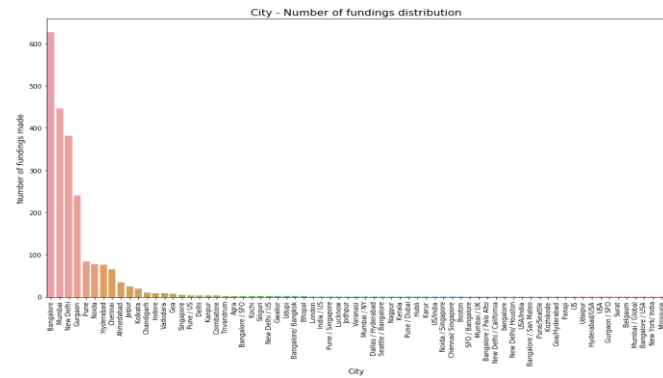


Figure 7

Figure 7 shows the distribution of City and the amount of funding made over two years of time. We could conclude the following results from it.

- Bangalore attracted the most investors with a total of 627 investors. (26.4%)
- Mumbai with 446 investors. (18.8%)
- New Delhi with 381 investors. (16.1%)

Interesting note: Few II Tier and III cities like Varanasi, Indore, Siliguri, Karur, Nagpur, Belgaum, and Kozhikode have also attracted some investors in recent years.

Now for the investors, we had to separate each cell and split individual cells since a particular cell may have more than 1 investors therefore by splitting and combining the cell we get to know the actual number of investors present in the data.

Therefore we concluded that there are 2544 unique investors, some of whom invested jointly. The below figure gives the distribution of top 50 investors to the number of investments made.

4.7 Top investors

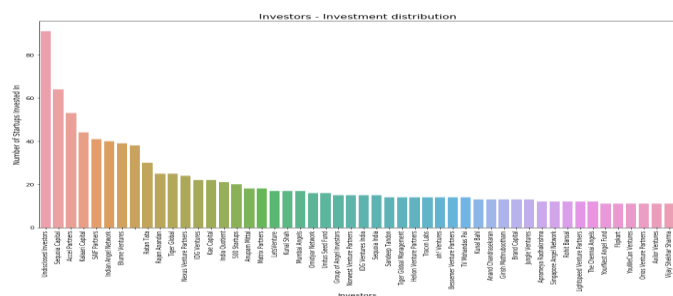


Figure 8

As we can see that most of the investors are 'Undisclosed Investors' with 91 investments, followed by Sequoia Capital with 64.

An interesting takeaway point from this is that individuals like Ratan Tata (Former chairman of Tata Sons) and Rajan Anandan (VP, Google SE Asia, and India) have invested in 30 and 25 companies respectively, most by any individuals and more than many other venture capital firms.

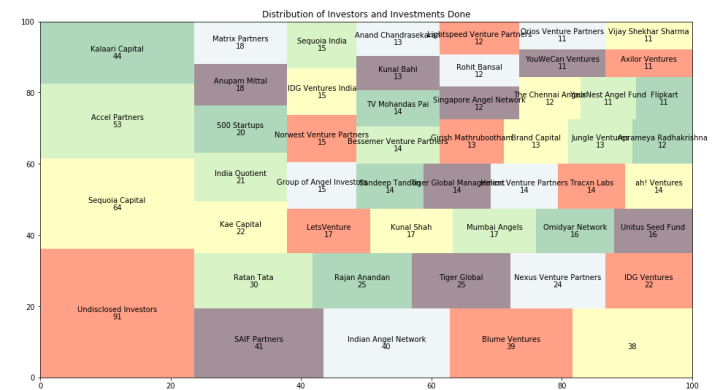


Figure 9

Above is a square plot of the same distribution as above which shows the area distribution of investors and investments. This square plot helps to visualize more how the investors fit into the picture.

5. CONCLUSION

The conclusion obtained from the analysis and visualization for the 'Indian Startup Funding' data is the trends of how the events like Digital India initiative, Startup India, Indian money demonetization and surgical strike affected the funding by investors into the companies.

Also, it is observed that even though tier-II and tier-III cities attracted some investors still tier-I cities like Bangalore, Mumbai, and Delhi leads the way for maximum investments.

Unicorns like Flipkart and Paytm still attract investors due to its position in the market. And after demonetization Paytm adopted the cashless India policy to boost their revenue.

Ola cabs and Swiggy attracted the most investors in the two years and both are planning to join the unicorn clubs.

Still today Consumer internet is more favored to investors than any other industry verticals, but to the due advancement of AI, Machine learning, and other advancements in technology, the technology vertical has also attracted a large number of investors.

With consumer internet leading the way, its sub-verticals like online pharmacy, online lending platform, online

payment gateway, etc have also attracted the most number of investors.

Since India has become a hub for startups, new startup emerges, therefore, seed funding is important for them, therefore, seed funding is the most preferred type of investment. The single most common reason for an investment amount remaining undisclosed is that the size of the round would be viewed by the market as derisory in comparison to competitors. This is particularly common at the seed stage. Therefore there is a number of undisclosed investors in the data. (O'Hear, 2013)

In recent years India has attracted a lot of foreign investors like Sequoia Capital, Accel Partners, and others, out of which Sequoia Capital and Accel partners lead the pack with most investment meaning foreign investors invest more than Indian investors.

An interesting point to take is that individuals like Ratan Tata and Rajan Anandan have taken a keen interest in investing in startups since they invested more than any other individuals.

With the visualizations, we came to know that various events do affect the funding. Also the cities and industry verticals play an important part in acquiring funding from investors, and foreign investors invest a lot into Indian startups.

6. REFERENCES

1. Hari, T. N. (2018). The Indian Startup Report 2018. Bengaluru: YourStory Media Pvt Ltd.
2. Huang, Y. (2008). The next Asian miracle. Foreign Policy , 32-40.
3. Giurca Vasilescu, L. (2009). Business angels: potential financial engines for startups, *Ekonomika istraživanja*, 22(3), 86-97.
4. Keshamoni, S. (2018, 05 30). 4 Start-Ups funded at TiE Startup Heroes challenge. Retrieved 05 30, 2018, from <http://startuphyderabad.com/4-startups-funded-at-tie-startup-heroes-challenge/>
5. Kshetri, N. (2016). Fostering Startup Ecosystems in India. *Asian Research Policy* , 94-103.
6. Nasscom. (2018, 05 29). NASSCOM START-UP REPORT' 2017. Retrieved 05 29, 2018, from <https://www.investindia.gov.in/sites/default/files/2018-03/Indian%20Startup%20Ecosystem%20-%20Traversing%20the%20maturity%20cycle.pdf>
7. O'Hear, S. (2013, 05 26). Anatomy Of An Undisclosed Investment Or Exit. Retrieved 08 26, 2018, from [Techcrunch.com:](https://techcrunch.com/2013/05/26/follow-the-money/)
8. Velayanikal, M. (2015, 10 23). What are the hottest sectors for venture capital in India this year? Retrieved from [Techinasia.com:](https://www.techinasia.com/hottest-sectors-venture-capital-india-infographic)
9. Wikipedia. (2018, 08 13). India Startup Ecosystem TimeLine. Retrieved 08 27, 2018, from Wikipedia: https://en.wikipedia.org/wiki/India_Startup_Ecosystem_TimeLine
10. Wikipedia. (2018, 08 24). India–Pakistan border skirmishes (2016–present). Retrieved 08 26, 2018, from Wikipedia, The Free Encyclopedia.: [https://en.wikipedia.org/wiki/India%E2%80%993Pakistan_border_skirmishes_\(2016%E2%80%933present\)](https://en.wikipedia.org/wiki/India%E2%80%993Pakistan_border_skirmishes_(2016%E2%80%933present))