STARTUP IN INDIA

SUBMITTED BY

Group 11

Vanshika Sharma (2018BtechCSE120)

Samyak Bhagat (2018BtechCSE115)

Sanket Sharma (2018BtechCSE116)

Madhav Yadav (2018BtechCSE006)

as Project – III

of

Computational Data Analysis

Institute of Engineering and Technology

JK Lakshmipat University Jaipur

April 2019

ABSTRACT

India is a developing south Asian country. It is a most populous and 7th largest country by area. Large population implies a large prospective market in India and puts more pressure for employment in the country. In the present decade, India is undertaking an essential shift towards startup welcoming policies and a business-friendly environment.

India being a populated country having increasing demand which is putting a competitive environment forcing to create innovative systems. One of these systems is a Start-up ecosystem. This paper is aimed at about the growth and prospects of Startup systems in India. This is an attempt to find out the startup trends in the Indian context by considering different aspects like city wise funding received and sector wise funding received, to find which are the suitable cities for certain type of startups and which is the most funded sector in startups.

INTRODUCTION

A **Startup** is a company initiated by individual founders or entrepreneurs to search for a repeatable and scalable business model. Founders design startups to effectively develop and validate a scalable business model. Startups must not be confused with small business, as the biggest difference being is INNOVATION.

The objectives of a startup are to be one's own boss and to create employment to others which warrants lot of endurance and sacrifice. The government has introduced campaigns like 'Make in India', 'Digital India', etc., to attract foreign investments and promote private companies to come up ahead in the manufacturing sectors.

Figure 1

It is rightly said, *"Entrepreneurship is no longer being condemned as jugaad."*
India is the second most populated country of the word ranking first in the youth population with 356 million youngsters. The youth in India has bent over to developing innovative ideas and self-made job opportunities which has led to a substantial development of the startup ecosystem of India. This has made India rank second on the most no. of startups in a country, only lagging USA.

The funding in the area of startup area has also increased. The startups in India saw a 108 % growth from USD 2 billion in 2017 to USD 4.2 billion in 2018.
The government has introduced campaigns like 'Make in India' and 'Digital India' to attract foreign investments and promote private companies to come up ahead in the manufacturing sectors.

OBJECTIVES:

1. To find relation between Sector and Type of funding of Startups.
2. To find relation between City and Sector of Startups.
3. To find the average funding in the 4 Sectors of Startups – Technology, Consumer Internet, eCommerce and Other Sectors (dominated by top 4 cities).
4. To find relation between the ranking of Cities (according to number of startups in the city) in two sectors – Technology and Consumer Internet.

SUSTAINABLE DEVELOPMENT GOAL:

1. Goal 8 – Decent Work and Economic Growth
   The goal aims at decent work and growth which promotes sustainable economic growth, full and productive employment and decent work for all. It consists of 12 targets and 17 indicators.

2. Goal 9 - Industry, Innovation and Infrastructure
   The goal focuses on building resilient infrastructure, promoting inclusive and sustainable industrialization and fostering innovation. Through SDG 9, countries have determined that investing in more resilient infrastructure, cooperating across borders, and encouraging small enterprises will all be critical to ensuring sustainable industrial development.

Figure 2 – SDG Goal 8

Figure 3 – SDG Goal 9

MATHEMATICAL CONCEPT:

To reach out to results, in this project we have used mathematical concepts:

1. Linear algebra – solving linear simultaneous equations using matrix
2. Statistics – Contingency table and Spearman's Rank correlation.

LITERATURE REVIEW

The startup situation in India has gone an enormous makeover, presently individuals are not outsider with the idea of new companies. Prior individuals had no clue what this idea is about, gratitude to the ascent in media's empowering inclusion towards new businesses as of late. The idea of startup is some way or another distinctive for Indians and not all that diverse for individuals of created economies. New companies are something to do with new item/process for the whole market or portion of the market. New businesses must not be mistaken for independent company, as the greatest contrast being is INNOVATION. As of late legislature of India has propelled "Startup India" activity to cultivate/support and energize startup endeavours in India. The outcomes are exceptionally palatable with activity being acknowledged with great enthusiasm in nation, different state governments have likewise begun the comparable endeavours. India remain at a critical cross street, India remained at number three in by and large innovation driven detail ups on the planet (Top two positions are held by USA and UK individually).

The overall impact of startups is very visible initially then, only those ideas persist which are smartly implemented. In India government is constantly trying to create an environment which is both conducive and optimum for stat ups. The reason is very simple startups are necessary for the entrepreneurial and innovative growth of any nation[1]. There are nations which are smaller than ours and less naturally equipped than ours but made tremendous growth and advancements in the field of economy and overall development. The secret of their success is nothing but an appetite for innovation. If India wants to be in the front lines with developed nations in the world, innovation is the key to become so. Fortunately, India is endowed with youngest population which is primarily required for setting up startups. With the growing inclination towards “Having something of my own” attitude is also helping in bringing new ideas into successful implementation. India has produced some of the leading startups in the world, which are working as the lighthouse for the rest. The prominent example being OYO Rooms and Zomato (both catering to a very different market segment and objectives).

The government of India is also serious in promoting entrepreneurship at the startup level and has taken several initiatives to ensure appropriate support.

1) MUDRA – The Micro Units Development and Refinance Agency or MUDRA, is a flagship program by the government of India to provide funds to micro and small enterprises.
2) NABARD – The National Bank for Agriculture And Rural Development, or NABARD, for short, is primarily aimed towards providing credit benefits to agriculture as well as other cottage and village industries.

[1]Upadhyay1, C. S., & P. R. (2017, October). Start Ups; Let’s Start Them Up - An Inside View in the Indian Start Up Scenario. Retrieved from https://www.ijraset.com/fileserve.php?FID=10609

3) Credit Guarantee Scheme – The CGTMSE (Credit Guarantee Fund Trust for Micro and Small Enterprises) was set up by the Government of India to provide business loans to micro and small industries, with zero collateral.
4) Stand Up India Scheme – Launched in 2016, this scheme was implemented to cater to women entrepreneurs, as well as those from SC and ST communities.
5) NewGen IEDC – Introduced last year, the NewGen Innovation and Entrepreneurship Development Centre is applicable to industries like healthcare services, chemicals, hardware, aeronautical/defence, IT, AR/VR, construction, design, food and beverages, etc.

Apart from the above-mentioned schemes, one of the major campaigns includes the 'Make in India' campaign introduced to attract foreign investments and encourage domestic companies to participate in the manufacturing sector.

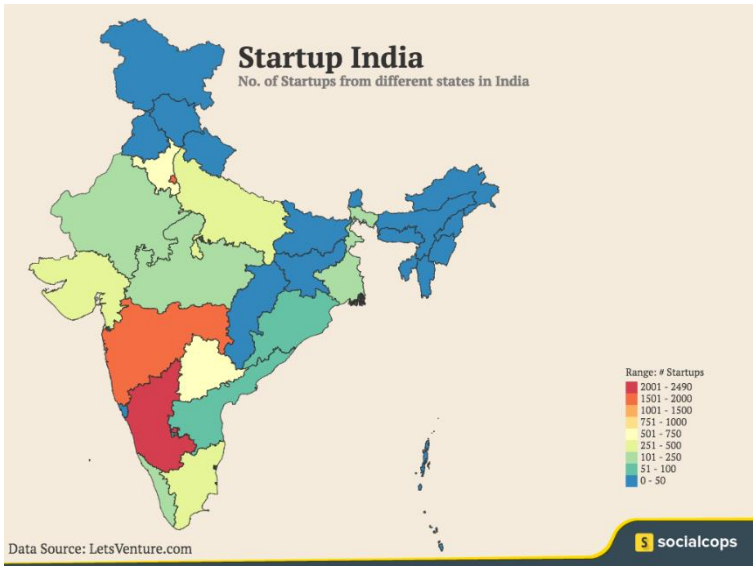


Figure 4 - Map shows the number of startups in states of India.

SDG 8: DECENT WORK AND ECONOMIC GROWTH

Globally, labour productivity has increased and the unemployment rate has decreased. However, more progress is needed to increase employment opportunities, especially for young people, reduce informal employment and labour market inequality (particularly in terms of the gender pay gap), promote safe and secure working environments, and improve access to financial services to ensure sustained and inclusive economic growth.

1. The Challenge:
   A continued lack of decent work opportunities, insufficient investments and under-consumption has led to an erosion of the basic social contract underlying democratic societies: that all must share in progress. The startups provide an opportunity for employment and skill development.

2. Why is this important?
   While developing countries have grown at a rate faster than developed regions, sustained economic growth everywhere will be critical to fulfilling our international developmental targets over the next 15 years. Economic growth – making our world more prosperous – is inextricably linked to all our other priorities. Stronger economies will afford us more opportunities to build a more resilient and sustainable world. And economic growth must be inclusive: growth that does not improve the wellbeing of all sections of society, especially the most vulnerable, is unequal and unfair.

3. How can we address this?
   'No one left behind' is at the core of the sustainable development agenda for 2030 and if economic growth is to build a fairer world, it must be inclusive. This is the idea behind Goal 8, which aims to sustain an economic growth rate of 7% for the least developed countries by 2030, and achieve full and productive employment for all men and women everywhere in the next 15 years.

Figure 5

4. India and Goal 8:
   India can forge its own growth path, which can rely on both manufacturing and services as a growth escalator and employment generator. The challenge will be to create well-paying and productive jobs in non-farm sectors that can absorb more unskilled workers, including women and those in rural areas.
   Almost half the labour force in India still works in the agricultural sector. With low productivity, it is difficult to promote gainful employment in agriculture. Enhancing agricultural productivity through public investment and new technologies should be a priority focus area. Moreover, upgrading to high-value commodities, reforming agricultural marketing policies and market interventions, and strengthening linkages to agri-businesses are critical areas ripe for government intervention.
   Increasing the labour force participation of women is a powerful tool not only to empower women, but also to steer economic growth itself. As reported by the McKinsey Global Institute (2015), if India increases its female labour force

participation rate by 10 percentage points by 2025, its GDP could rise by as much as 16 percent as compared to the business-as-usual scenario. [2]
India's desired transition to a green economy will have a significant impact on job skill requirements within sectors, occupational profiles and business operations. Labour market and skill policies can play an important role in maximizing the benefits of economic greening for workers. Furthermore, The circular economy is gaining increasing attention as a strategy for long-term prosperity and sustainability. India's engineering workforce, its rapidly developing engineering services, R&D expertise and its geo-position in South Asia, position it as a potential global hub for both frugal manufacturing and services. Additionally, the fourth industrial revolution is both an opportunity and a challenge for India.

5. Government Schemes:
   The government's National Skill Development Mission, Deendayal Upadhyaya Antodaya Yojana, Atal Innovation Mission, as well as the National Service Scheme and the Mahatma Gandhi National Rural Employment Guarantee Scheme are some flagship programmes aimed at bringing decent work to all.

Targets:

1. Sustain per capita economic growth in accordance with national circumstances and, in particular, at least 7% gross domestic product growth per annum in the least developed countries.
2. Achieve higher levels of economic productivity through diversification, technological upgrading and innovation, including through a focus on high-value added and labour-intensive sectors.
3. Promote development-oriented policies that support productive activities, decent job creation, entrepreneurship, creativity and innovation, and encourage the formalization and growth of micro-, small- and medium-sized enterprises, including through access to financial services.
4. Improve progressively, through 2030, global resource efficiency in consumption and production and endeavour to decouple economic growth from environmental degradation, in accordance with the 10-year framework of programmes on sustainable consumption and production, with developed countries taking the lead.
5. By 2030, achieve full and productive employment and decent work for all women and men, including for young people and persons with disabilities, and equal pay for work of equal value.
6. By 2020, substantially reduce the proportion of youth not in employment, education or training.
7. Take immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labour, including recruitment and use of child soldiers, and by 2025 end child labour in all its forms.

[2] Goal 8: Decent work and economic growth. (n.d.). Retrieved from https://www.undp.org/content/undp/en/home/sustainable-development-goals/goal-8-decent-work-and-economic-growth.html

8. Protect labour rights and promote safe and secure working environments for all workers, including migrant workers, in particular women migrants, and those in precarious employment.
9. By 2030, devise and implement policies to promote sustainable tourism that creates jobs and promotes local culture and products.
10. Strengthen the capacity of domestic financial institutions to encourage and expand access to banking, insurance and financial services for all.
11. Increase aid for trade support for developing countries, in particular least developed countries, including through the Enhanced Integrated Framework for Trade-Related Technical Assistance to Least Developed Countries.
12. By 2020, develop and operationalise a global strategy for youth employment and implement the Global Jobs Pact of the International Labour Organization.

SDG 9: INDUSTRY, INNOVATION AND INFRASTRUCTURE

1. The Challenge:
   The story of industrial development has been an important determinant of the course of our history as a community of nations. From the first steam engines to the first assembly lines, to today's truly global production chains and processes, industry has changed our economies and helped drive major changes in our societies. But without sustainable practices and infrastructure in place, our growth has left vast sections of people behind. More than 937 million of the world's population in 2016 do not have access to electricity. For many lower-income countries, the existent infrastructure constraints affect firm productivity by around 40%.

Figure 6

2. Why is this important?
   Investments in transport, irrigation, energy and information and communications technology have been crucial to driving economic growth and empowering communities in many countries. The job multiplication effect of industrialisation has a positive impact on society, as every one job in manufacturing creates 2.2 jobs in other sectors. The manufacturing sector is an important employer, accounting for around 470 million jobs worldwide in 2009 – or around 16% of the world's workforce of 2.9 billion. It has long been recognised that a strong physical network of industry and communication can enhance productivity and incomes, and improve health, wellbeing and education. Technological progress similarly enhances our wellbeing as countries and can also improve the state of the planet through increased resource and energy efficiency.

3. How can we address this?

Through SDG 9, countries have determined that investing in more resilient infrastructure, cooperating across borders, and encouraging small enterprises will all be critical to ensuring sustainable industrial development. We will also have to improve our existing industrial infrastructure, and here, technological innovation will be key. Governments and businesses will have to contribute to creating a hospitable policy environment for innovation, encourage scientific research, and improve access to information technology universally.

4. India and Goal 9:
   The government's flagship interventions like Make in India and Startup India as well as Pandit Deendayal Upadhyay Shramev Jayate Karyakram are fuelling innovation and sustainable industrial and economic development.

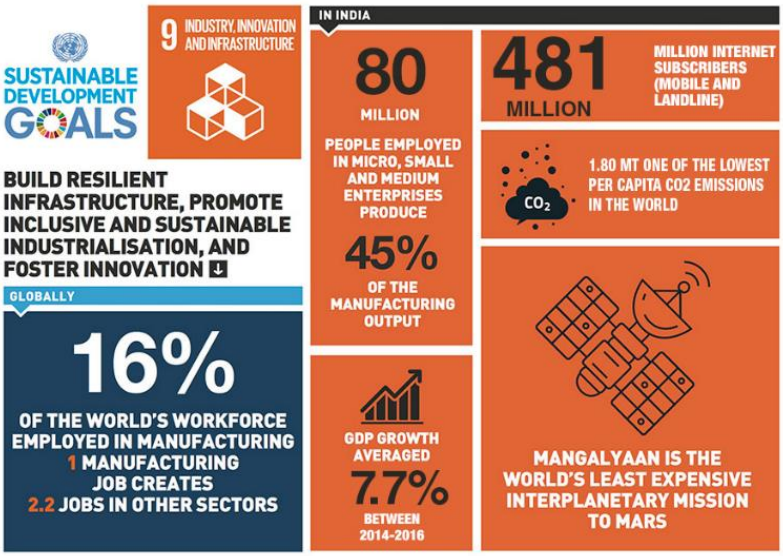


Figure 7

Targets:

1. Develop quality, reliable, sustainable and resilient infrastructure, including regional and trans-border infrastructure, to support economic development and human wellbeing, with a focus on affordable and equitable access for all.
2. Promote inclusive and sustainable industrialization and, by 2030, significantly raise industry's share of employment and gross domestic product, in line with national circumstances, and double its share in least developed countries.
3. Increase the access of small-scale industrial and other enterprises, in particular in developing countries, to financial services, including affordable credit, and their integration into value chains and markets.
4. By 2030, upgrade infrastructure and retrofit industries to make them sustainable, with increased resource-use efficiency and greater adoption of clean and environmentally sound technologies and industrial processes, with all countries taking action in accordance with their respective capabilities.
5. Enhance scientific research, upgrade the technological capabilities of industrial sectors in all countries, in particular developing countries, including, by 2030, encouraging innovation and substantially increasing the number of research

and development workers per 1 million people and public and private research and development spending.

6. Facilitate sustainable and resilient infrastructure development in developing countries through enhanced financial, technological and technical support to African countries, least developed countries, land-locked developing countries and small island developing states.
7. Support domestic technology development, research and innovation in developing countries, including by ensuring a conducive policy environment for, inter alia, industrial diversification and value addition to commodities.
8. Significantly increase access to information and communications technology and strive to provide universal and affordable access to the Internet in least developed countries by 2020.[3]

CELEBRATING Year of #startupindia

Department of Industrial Policy & Promotion
Government of India
Ministry of Commerce & Industry

ADDING **MOMENTUM** TO THE **STARTUP** SPIRIT

[3]Goal 8: Decent work and economic growth. (n.d.). Retrieved from https://www.undp.org/content/undp/en/home/sustainable-development-goals/goal-8-decent-work-and-economic-growth.html

MATHEMATICAL MODEL

1. CONTINGENCY TABLE:[4]

Contingency Tables are used in two cases:

a. In cases when we want to check whether relation exist between two categorical variables.
b. In cases when from the sample of one population we need to classify each item with respect to two.

In both the cases, we calculate the expected cell frequencies:

A chi-square test for independence shows how categorical variables are related.

$$e_{ij} = \frac{(ith\ row\ total) \times (jth\ column\ total)}{grand\ total}$$

To perform a chi square test, we then substitute into the formula:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Degrees of freedom: df=(r-1)(c-1)

Criteria of Rejection:

If χ^2 > $\chi^2_{\alpha,(r-1)(c-1)}$ reject the null hypothesis.

2. SPEARMAN'S RANK CORRELATION:

Spearman's Rank Correlation Coefficient:
The Spearman rank correlation coefficient, r_s, is the nonparametric version of the Pearson correlation coefficient. Your data must be ordinal, interval or ratio. Spearman's returns a value from -1 to 1,
where:
+1 = a perfect positive correlation between ranks
-1 = a perfect negative correlation between ranks
0 = no correlation between ranks.

[4] 4 M., & F. (n.d.). Inferences concerning proportions. In *Probability and Statistics*(p. 308).

The formula for the Spearman rank correlation coefficient when there are no tied ranks is:

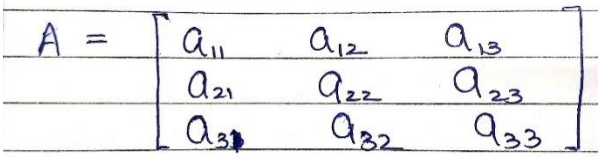
$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

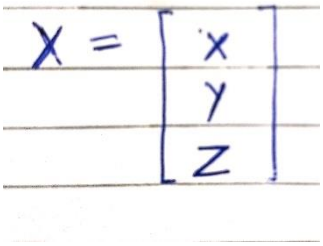
[5]

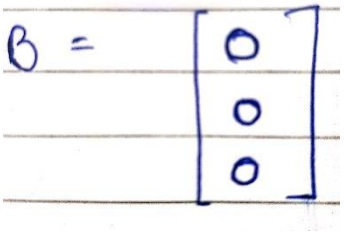
3. SOLVING LINEAR SIMULTANEOUS EQUATIONS USING MATRIX:

The system of linear equations can be easily solved using matrix by Gauss Jordan method and Gauss elimination method. If the number of variables to be solved is large iterative methods (Gauss Seidel and Gauss Jacobi method) are used.
In this project we have used Gauss Jordan method:

Consider a system of linear equations:
$a_1x + b_1y + c_1z = 0$
$a_2x + b_2y + c_2z = 0$
$a_3x + b_3y + c_3z = 0$
According to Gauss Jordan method:
AX=B;
Where







The value of $X=A^{-1}B$

[5]Spearman Rank Correlation (Spearman's Rho): Definition and How to Calculate it. (2018, September 02). Retrieved from https://www.statisticshowto.datasciencecentral.com/spearman-rank-correlation-definition-calculate/

Objective 1: To find relation between Sector and Type of funding of Startups.

MATHEMATICAL METHOD (Contingency Table) :

STEP 1:

Null Hypothesis: H_o = There is no relation between Sector and Type of funding of startups.
Alternate Hypothesis: H_a = There is relation between Sector and Type of funding of startups.

STEP 2:

Level of Significance: $\alpha = 0.05$

Criteria of rejection: If $\chi^2 > \chi^2_{\alpha,(r-1)(c-1)}$ reject the null hypothesis.

i.e $\chi^2 > 12.592$

STEP 3:

Calculations:

The observed data:

```
observation_table
```

InvestmentType	Technology	Consumer Internet	eCommerce	Others	Total_row
Seed Funding	163	473	80	500	1216
Private Equity	150	298	91	444	983
Debt Funding	0	1	0	0	1
Total_col	313	772	171	944	2200

The estimated data values calculated using the formula of e_{ij} :

```
estimation_table
```

InvestmentType	Technology	Consumer Internet	eCommerce	Others	Total_row
Seed Funding	173.00	426.71	94.52	521.77	1216.0
Private Equity	139.85	344.94	76.41	421.80	983.0
Debt funding	0.14	0.35	0.08	0.43	1.0
Total_col	312.99	772.00	171.01	944.00	2200.0

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = 21.67$$

STEP 4:

Decision:
As the calculated value χ^2 > 12.592, we REJECT the Null Hypothesis.

∴ There is relation between Sector and Type of funding of startups.

PYTHON IMPLEMENTATION:

In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
data = pd.read_csv("startup_funding.csv")
data
```

Out[2]:

	SNo	Date	StartupName	IndustryVertical	SubVertical	CityLocation	InvestorsName	InvestmentType	AmountInUSD	Remarks
0	0	01/08/2017	TouchKin	Technology	Predictive Care Platform	Bangalore	Kae Capital	Private Equity	1,300,000	NaN
1	1	02/08/2017	Ethinos	Technology	Digital Marketing Agency	Mumbai	Triton Investment Advisors	Private Equity	NaN	NaN
2	2	02/08/2017	Leverage Edu	Consumer Internet	Online platform for Higher Education Services	New Delhi	Kashyap Deorah, Anand Sankeshwar, Deepak Jain,...	Seed Funding	NaN	NaN
3	3	02/08/2017	Zepo	Consumer Internet	DIY Ecommerce platform	Mumbai	Kunal Shah, LetsVenture, Anupam Mittal, Hetal ...	Seed Funding	500,000	NaN
4	4	02/08/2017	Click2Clinic	Consumer Internet	healthcare service aggregator	Hyderabad	Narottam Thudi, Shireesh Palle	Seed Funding	850,000	NaN
5	5	01/07/2017	Billion Loans	Consumer Internet	Peer to Peer Lending platform	Bangalore	Reliance Corporate Advisory Services Ltd	Seed Funding	1,000,000	NaN
6	6	03/07/2017	Ecolibriumenergy	Technology	Energy management	Ahmedabad	Infuse Ventures, JLL	Private Equity	2,600,000	NaN

```
# -------------- Creating a contigency table between TYPE OF FUNDING & SECTOR OF STARTUP ------------------------- #
```

In [3]:
```
table1 = pd.crosstab(data.InvestmentType , data.IndustryVertical)
table1
```

Out[3]:

IndustryVertical	360-degree view creating platform	3D Printer Manufacturer	API Workflow platform	Activity Based Social Network	Advertising & Marketing Community Networking platform	Affordable Education	Affordable Hotel Booking Online	Algorithmic Match Making Platform	Alternate Mobile Monetization platform	App Analytics platform	...	online recipe sharing platform	online tiffin service aggregator
InvestmentType													
Debt Funding	0	0	0	0	0	0	0	0	0	0	...	0	0
Private Equity	1	1	0	0	0	0	1	0	1	0	...	0	0
Seed Funding	0	0	1	1	1	1	0	1	0	1	...	1	1

3 rows × 743 columns

```
#------ Extracting specific columns of sector of startup - Technology, Consumer Internet and eCommerce ------- #
```

In [4]:
```
technology_list = table1['Technology'].sort_values(ascending=False) #getting sorted no of startups in technology sector
consumerInternet_list = table1['Consumer Internet'].sort_values(ascending=False) #getting sorted no of startups in consumer intern
ecommerce_list = table1['eCommerce'].sort_values(ascending=False) #getting sorted no of startups in ecommerce sector

print("1. Technology:\n",technology_list)
print("2. Consumer Internet:\n",consumerInternet_list)
print("3. eCommerce:\n",ecommerce_list)
```

```
1. Technology:
 InvestmentType
Seed Funding      163
Private Equity    150
Debt Funding        0
Name: Technology, dtype: int64
2. Consumer Internet:
 InvestmentType
Seed Funding      473
Private Equity    298
Debt Funding        1
Name: Consumer Internet, dtype: int64
3. eCommerce:
 InvestmentType
Private Equity    91
Seed Funding      80
Debt Funding       0
Name: eCommerce, dtype: int64
```

```
#--------- Obtaining total frequency of other sectors (excluding Technology, Consumer Internet & eCommerce) ---------- #
```

In [5]:
```
table1 = table1.drop(columns = 'Technology')
table1 = table1.drop(columns = 'Consumer Internet')
table1 = table1.drop(columns = 'eCommerce')

table1 = table1.sum(axis=1) #calculating the sum of each row to obtain total frequency of other sectors

table1 = pd.DataFrame(table1,columns = ['other']) #making the list (containing the sum of each row-table1) as a dataframe
table1
```

Out[5]:

	other
InvestmentType	
Debt Funding	0
Private Equity	444
Seed Funding	500

•

#.---------------------------- Final contingency table for observed data ---------------------------------- #

In [6]:
```
observation_table = pd.DataFrame(technology_list , columns=['Technology']) #initial dataframe with 1 column
observation_table['Consumer Internet'] = consumerInternet_list
observation_table['eCommerce'] = ecommerce_list
observation_table['Others'] = table1
observation_table.loc['Total_col'] = observation_table.sum(axis=0)#sum of each column
observation_table['Total_row'] = observation_table.sum(axis=1)#sum of each row

observation_table
```

Out[6]:

InvestmentType	Technology	Consumer Internet	eCommerce	Others	Total_row
Seed Funding	163	473	80	500	1216
Private Equity	150	298	91	444	983
Debt Funding	0	1	0	0	1
Total_col	313	772	171	944	2200

#-----------------------Matrix of Estimated Values-----------------------------#

Eij = (ith row total x jth column total) / Grand Total

In [17]:
```
table2=[]
for x in range(3): #to make 3 rows in the 2D list
    row=[]
    for y in range(4): #to make 4 columns
        row.append(round((observation_table['Total_row'][x]*observation_table.loc["Total_col", : ][y])
                         /observation_table['Total_row'][3],2))
    table2.append(row)

for x in table2:
    print(x)
```

```
[173.0, 426.71, 94.52, 521.77]
[139.85, 344.94, 76.41, 421.8]
[0.14, 0.35, 0.08, 0.43]
```

In [18]:
```
#Converting the 2D list to dataframe
estimation_table = pd.DataFrame(table2, columns=['Technology','Consumer Internet','eCommerce','Others'])
estimation_table
```

Out[18]:

	Technology	Consumer Internet	eCommerce	Others
0	173.00	426.71	94.52	521.77
1	139.85	344.94	76.41	421.80
2	0.14	0.35	0.08	0.43

In [19]:
```
#changing the index from default to Investment Type
estimation_table['InvestmentType'] = ['Seed Funding' , 'Private Equity' , 'Debt funding']
estimation_table = estimation_table.set_index('InvestmentType')
estimation_table
```

•

Out[19]:

InvestmentType	Technology	Consumer Internet	eCommerce	Others
Seed Funding	173.00	426.71	94.52	521.77
Private Equity	139.85	344.94	76.41	421.80
Debt funding	0.14	0.35	0.08	0.43

In [20]:
```
estimation_table['Total_row'] = estimation_table.sum(axis=1)
estimation_table.loc['Total_col'] = estimation_table.sum() #to find the sum of each column and append sums as a row
estimation_table
```

Out[20]:

InvestmentType	Technology	Consumer Internet	eCommerce	Others	Total_row
Seed Funding	173.00	426.71	94.52	521.77	1216.0
Private Equity	139.85	344.94	76.41	421.80	983.0
Debt funding	0.14	0.35	0.08	0.43	1.0
Total_col	312.99	772.00	171.01	944.00	2200.0

Chi Square Test

In [66]:
```
#defining a dictionary to ease the access of dataframe columns in X^2 Test
sector = {0:'Technology' , 1:'Consumer Internet' , 2:'eCommerce' , 3:'Others'}
```

X^2 = sum of (Oij - Eij)^2 / Eij

In [67]:
```
numerator=0
denominator=0
sum=0 #notation for X^2
for col in range(3):
    for row in range(4):
        numerator = (observation_table[sector[row]][col] - estimation_table[sector[row]][col])**2
        denominator = estimation_table[sector[row]][col]
        sum = sum + numerator/denominator
```

In [68]:
```
sum
```

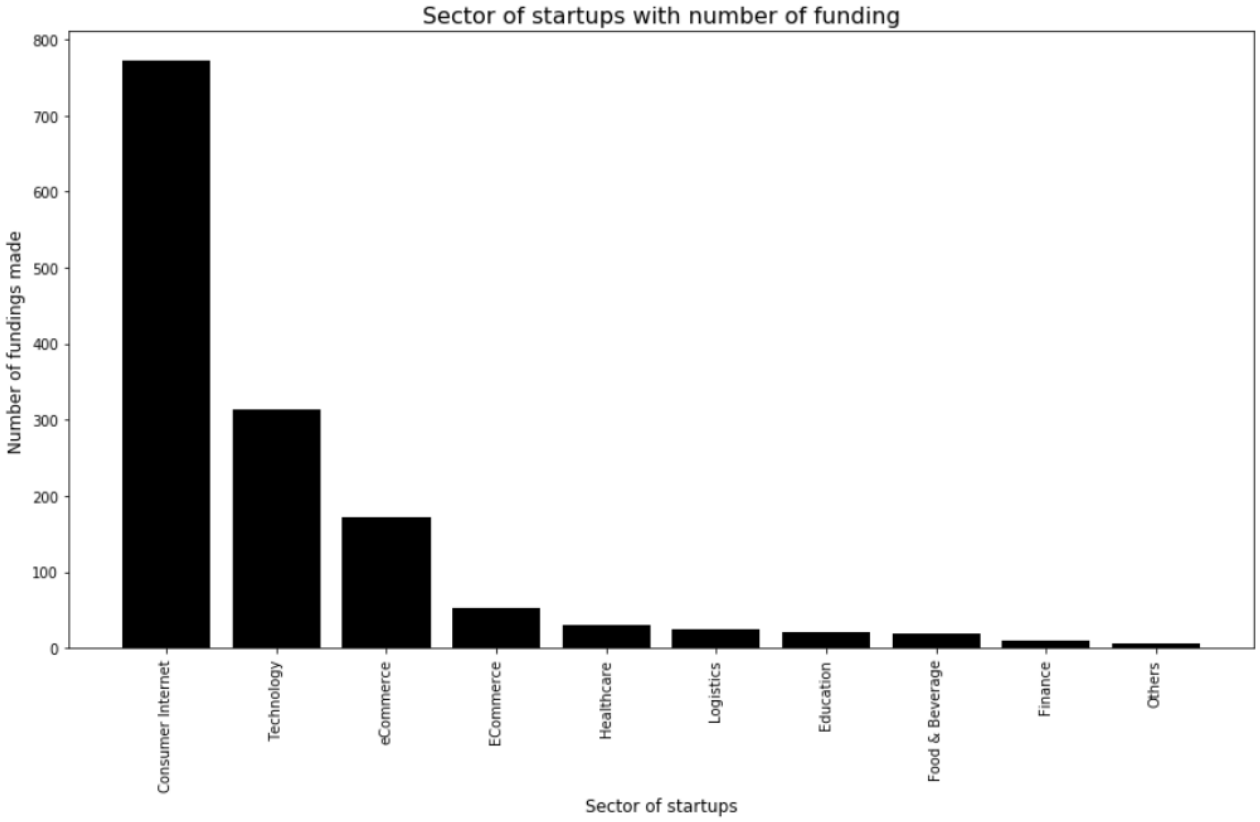
Out[68]:
```
21.670048421299196
```

```
        #------------------------ As from the X^2 table: -----------------------------#
        #------------------------ X^2 (0.05,6) = 12.592 ; and ------------------------#
        #------------------------ X^2 (calculated) = 21.67 ---------------------------#

        #------------------------ We reject the null hypothesis; ---------------------#
        #--------------This means all pi1, pi2, ... pic are not equal------------------#

    #--------The two categorical values- Sector and Type of funding of startups are not related ------#
```

.

In [11]:
```
industry = data['IndustryVertical'].value_counts().head(10)
print(industry)
plt.figure(figsize=(15,8))
plt.bar(industry.index, industry.values , color='black')
plt.xticks(rotation='vertical')
plt.xlabel('Sector of startups', fontsize=12)
plt.ylabel('Number of fundings made', fontsize=12)
plt.title("Sector of startups with number of funding", fontsize=16)
plt.show()
```

```
Consumer Internet    772
Technology           313
eCommerce            171
ECommerce             53
Healthcare            30
Logistics             24
Education             20
Food & Beverage       19
Finance                9
Others                 6
Name: IndustryVertical, dtype: int64
```



RESULT:

There is a dependence between Sector of a startup (Technology, Consumer Internet, etc.) and the type of funding it gets (Angel funding, Private Equity & Debt Funding).

CONCLUSION:

Due to dependency of type of funding and sector of the startup, an entrepreneur needs to focus on both simultaneously.
The following focus points establish relation between the study and the SDGs.

1. This study so established promote innovation and development of new start-ups in the sectors.
2. This can lead to advancement of the nation in a particular sector.
3. If startups brought up in planned manner, nation can succeed in export market.

Objective 2: To find relation between City and Sector of Startups.

MATHEMATICAL METHOD (Contingency Table) :

STEP 1:

Null Hypothesis: H_o = There is no relation between City and Sector of startups.
Alternate Hypothesis: H_a = There is relation between City and Sector of startups.

STEP 2:

Level of Significance: $\alpha = 0.05$

Criteria of rejection: If $\chi^2 > \chi^2_{\alpha,(r-1)(c-1)}$ reject the null hypothesis.

i.e $\chi^2 > 40.113$

STEP 3:

Calculations:
The observed data:

```
final_observed
```

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	106	232	41	248	627
Mumbai	47	162	34	203	446
New Delhi	39	129	34	179	381
Gurgaon	20	87	23	110	240
Pune	31	21	4	28	84
Noida	13	29	7	29	78
Hyderabad	14	25	5	32	76
Chennai	14	18	3	31	66
Ahmedabad	6	12	6	11	35
Jaipur	2	12	3	8	25
Total_col	292	727	160	879	2058

The estimated data values calculated using the formula of e_{ij} .

```
final_est
```

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	88.962099	221.491254	48.746356	267.800292	627.0
Mumbai	63.280855	157.551992	34.674441	190.492711	446.0
New Delhi	54.058309	134.590379	29.620991	162.730321	381.0
Gurgaon	34.052478	84.781341	18.658892	102.507289	240.0
Pune	11.918367	29.673469	6.530612	35.877551	84.0
Noida	11.067055	27.553936	6.064140	33.314869	78.0
Hyderabad	10.783285	26.847425	5.908649	32.460641	76.0
Chennai	9.364431	23.314869	5.131195	28.189504	66.0
Ahmedabad	4.965986	12.363946	2.721088	14.948980	35.0
Jaipur	3.547133	8.831390	1.943635	10.677843	25.0
Total_col	292.000000	727.000000	160.000000	879.000000	2058.0

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = 76.86$$

STEP 4:

Decision:
As the calculated value of χ^2 > 40.113, we REJECT the Null Hypothesis.

∴ There is a relation between City and Sector of startups.

PYTHON IMPLEMENTATION:

```
# -------------- Creating a contingency table between TYPE OF FUNDING & SECTOR OF STARTUP ------------------------- #
```

In [24]:
```
table = pd.crosstab(data.CityLocation,data.IndustryVertical) #making contingency table for city vs sectors
table
```

Out[24]:

IndustryVertical	360-degree view creating platform	3D Printer Manufacturer	API Workflow platform	Activity Based Social Network	Advertising & Marketing Community Networking platform	Affordable Education	Affordable Hotel Booking Online	Algorithmic Match Making Platform	Alternate Mobile Monetization platform	App Analytics platform	...	online recipe sharing platform	online tiffin service aggregator
CityLocation													
Agra	0	0	0	0	0	0	0	0	0	0	...	0	0
Ahmedabad	0	0	0	0	0	0	0	0	0	0	...	0	0
Bangalore	1	1	1	0	0	0	0	0	0	1	...	0	0
Bangalore / Palo Alto	0	0	0	0	0	0	0	0	0	0	...	0	0
Bangalore / SFO	0	0	0	0	0	0	0	0	0	0	...	0	0
Bangalore / San Mateo	0	0	0	0	0	0	0	0	0	0	...	0	0

```
#------ Extracting specific columns of sector of startup - Technology, Consumer Internet and eCommerce ------- #
```

In [25]:
```
technologyList = table['Technology'].sort_values(ascending=False) #getting sorted no of startups in technology sector
consumerInternetList = table['Consumer Internet'].sort_values(ascending=False) #getting sorted no of startups in consumer internet
ecommerceList = table['eCommerce'].sort_values(ascending=False) #getting sorted no of startups in ecommerce sector
```

```
#--------- Obtaining total frequency of other sectors (excluding Technology, Consumer Internet & eCommerce) ---------- #
```

In [26]:
```
other = table
other = other.drop(columns="Technology")
other = other.drop(columns="Consumer Internet")
other = other.drop(columns="eCommerce")
other = other.sum(axis=1)
other = pd.DataFrame(other,columns = ['other'])
other
```

Out[26]:

CityLocation	other
Agra	0
Ahmedabad	11
Bangalore	248
Bangalore / Palo Alto	1
Bangalore / SFO	1
Bangalore / San Mateo	1
Bangalore / USA	1

In [27]:
```
final_observed = pd.DataFrame(technologyList , columns=['Technology'])
final_observed['consumer internet'] = consumerInternetList
final_observed['Ecommerce'] = ecommerceList
final_observed['others_sector'] = other
final_observed['Total_row'] = final_observed.sum(axis=1)

final_observed = final_observed.sort_values('Total_row',ascending=False)
final_observed
```

Out[27]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	106	232	41	248	627
Mumbai	47	162	34	203	446
New Delhi	39	129	34	179	381
Gurgaon	20	87	23	110	240
Pune	31	21	4	28	84
Noida	13	29	7	29	78
Hyderabad	14	25	5	32	76

71 rows × 5 columns

In [28]:
```
final_observed = final_observed.head(10)
final_observed
```

Out[28]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	106	232	41	248	627
Mumbai	47	162	34	203	446
New Delhi	39	129	34	179	381
Gurgaon	20	87	23	110	240
Pune	31	21	4	28	84
Noida	13	29	7	29	78
Hyderabad	14	25	5	32	76
Chennai	14	18	3	31	66
Ahmedabad	6	12	6	11	35
Jaipur	2	12	3	8	25

In [29]:
```
final_observed.loc['Total_col'] = final_observed.sum() #to find the sum of each column and append sums as a row
final_observed
```

Out[29]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	106	232	41	248	627
Mumbai	47	162	34	203	446
New Delhi	39	129	34	179	381
Gurgaon	20	87	23	110	240
Pune	31	21	4	28	84
Noida	13	29	7	29	78
Hyderabad	14	25	5	32	76
Chennai	14	18	3	31	66
Ahmedabad	6	12	6	11	35
Jaipur	2	12	3	8	25
Total_col	292	727	160	879	2058

#----------Matrix of Estimated Values----------#

In [30]:
```
table2=[]
for x in range(10): #10 rows in the 2D list
    row=[]
    for y in range(4): #4 columns
        row.append((final_observed['Total_row'][x]*final_observed.loc["Total_col", : ][y]) / final_observed['Total_row'][10])
    table2.append(row)

for x in table2:
    print(x)
```

```
[88.96209912536443, 221.49125364431487, 48.74635568513119, 267.8002915451895]
[63.28085519922254, 157.55199222546162, 34.67444120505345, 190.49271137026238]
[54.05830903790088, 134.59037900874637, 29.620991253644316, 162.73032069970844]
[34.052478134110785, 84.78134110787173, 18.658892128279884, 102.50728862973762]
[11.918367346938776, 29.6734693877551, 6.530612244897959, 35.87755102040816]
[11.067055393586006, 27.55393586005831, 6.0641399416909625, 33.31486880466473]
[10.783284742468416, 26.847424684159378, 5.908649173955296, 32.46064139941691]
[9.364431486880466, 23.314868804664723, 5.131195335276968, 28.189504373177844]
[4.965986394557823, 12.363945578231293, 2.7210884353741496, 14.948979591836734]
[3.5471331389698735, 8.831389698736638, 1.9436345966958213, 10.677842565597668]
```

In [31]:
```
final_est = pd.DataFrame(table2 , columns=['Technology','consumer internet','Ecommerce','others_sector'])
final_est['CityLocation'] = ['Bangalore','Mumbai','New Delhi','Gurgaon','Pune','Noida','Hyderabad','Chennai','Ahmedabad','Jaipur']

final_est = final_est.set_index('CityLocation')
final_est
```

Out[31]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector
Bangalore	88.962099	221.491254	48.746356	267.800292
Mumbai	63.280855	157.551992	34.674441	190.492711
New Delhi	54.058309	134.590379	29.620991	162.730321
Gurgaon	34.052478	84.781341	18.658892	102.507289
Pune	11.918367	29.673469	6.530612	35.877551
Noida	11.067055	27.553936	6.064140	33.314869
Hyderabad	10.783285	26.847425	5.908649	32.460641
Chennai	9.364431	23.314869	5.131195	28.189504
Ahmedabad	4.965986	12.363946	2.721088	14.948980
Jaipur	3.547133	8.831390	1.943635	10.677843

Chi Square Test

In [32]:
```
final_observed
```

Out[32]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	106	232	41	248	627
Mumbai	47	162	34	203	446
New Delhi	39	129	34	179	381
Gurgaon	20	87	23	110	240
Pune	31	21	4	28	84
Noida	13	29	7	29	78
Hyderabad	14	25	5	32	76
Chennai	14	18	3	31	66
Ahmedabad	6	12	6	11	35
Jaipur	2	12	3	8	25
Total_col	292	727	160	879	2058

In [33]:
```
final_est['Total_row'] = final_est.sum(axis=1)
final_est.loc['Total_col'] = final_est.sum() #to find the sum of each column and append sums as a row
final_est
```

Out[34]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	88.962099	221.491254	48.746356	267.800292	627.0
Mumbai	63.280855	157.551992	34.674441	190.492711	446.0
New Delhi	54.058309	134.590379	29.620991	162.730321	381.0
Gurgaon	34.052478	84.781341	18.658892	102.507289	240.0
Pune	11.918367	29.673469	6.530612	35.877551	84.0
Noida	11.067055	27.553936	6.064140	33.314869	78.0
Hyderabad	10.783285	26.847425	5.908649	32.460641	76.0
Chennai	9.364431	23.314869	5.131195	28.189504	66.0
Ahmedabad	4.965986	12.363946	2.721088	14.948980	35.0
Jaipur	3.547133	8.831390	1.943635	10.677843	25.0
Total_col	292.000000	727.000000	160.000000	879.000000	2058.0

Chi Square = sum of (Oij - Eij)^2 / Eij

In [36]:
```
sum=0 #notation for X^2
l=['Technology' , 'consumer internet', 'Ecommerce' , 'others_sector']
for col in range(4):
    for row in range(10):
        sum+=((final_observed[l[col]][row] - final_est[l[col]][row])**2)/final_est[l[col]][row]

sum
```

Out[36]: 76.81556685392074

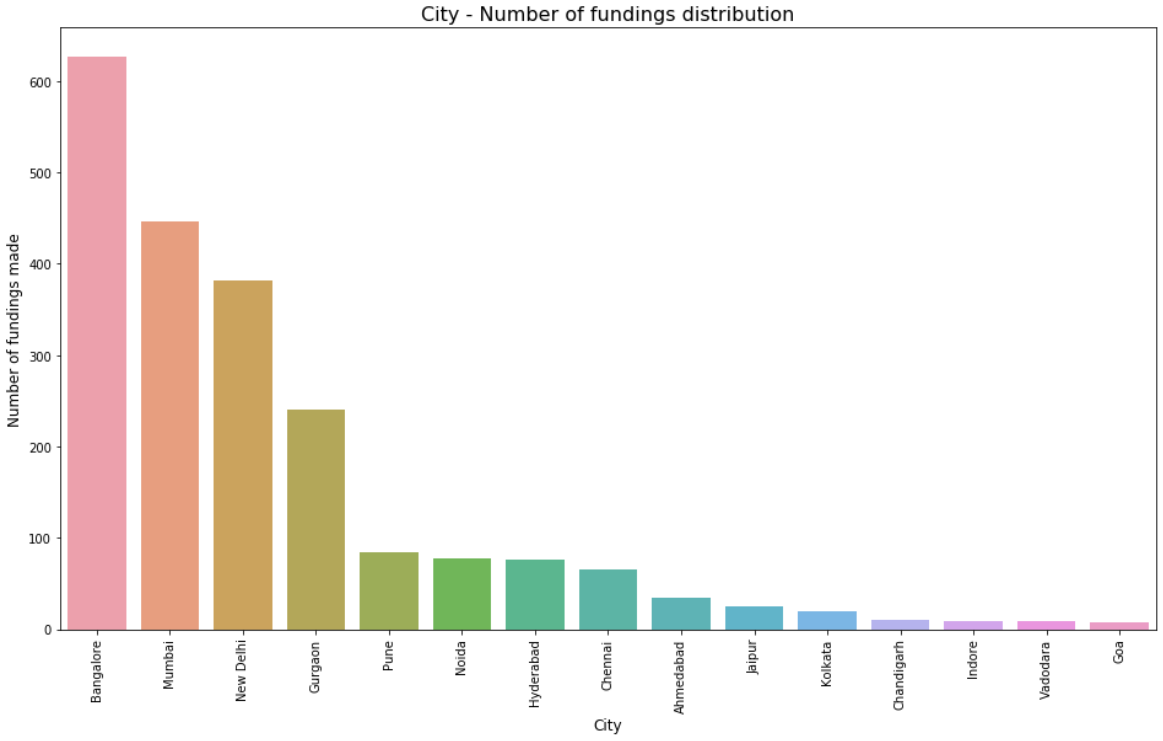
```
#------------------------ As from the X^2 table: -----------------------------#
#------------------------ X^2 (0.05,9x3=27) = 40.113 ; and -------------------------#
#------------------------ X^2 (calculated) = 76.81 ---------------------------#

#------------------------ We reject the null hypothesis; ---------------------#
#--------------This means that city and sector are related ------------------#

#--------The two categorical values- Sector and Cities of startups are related ------#
```

RESULT:

There is a dependence between Sector of a startup (Technology, Consumer Internet, etc.) and the city it is getting started in.

CONCLUSION:

Due to dependency of type of funding and city of the startup, an entrepreneur needs to focus on both simultaneously.
This means that if an entrepreneur needs funding in his startup or has just began a startup, the chances of city dependency can hinder or help.



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, Kozhikode have also attracted some investors.

Also, the type of funding is related to type of sector as established in previous objective.

Objective 3: To find the average funding in the 4 Sectors of Startups – Technology, Consumer Internet, eCommerce and Other Sectors (dominated by top 4 cities).

MATHEMATICAL METHOD (Matrix multiplication) :

The top four cities – Bangalore, Mumbai, Delhi and Gurgaon, highly contribute to the growing up of startup industry. So, this objective is fulfilled using only these 4 cities as the rest show significant differences in the number of startups coming up.

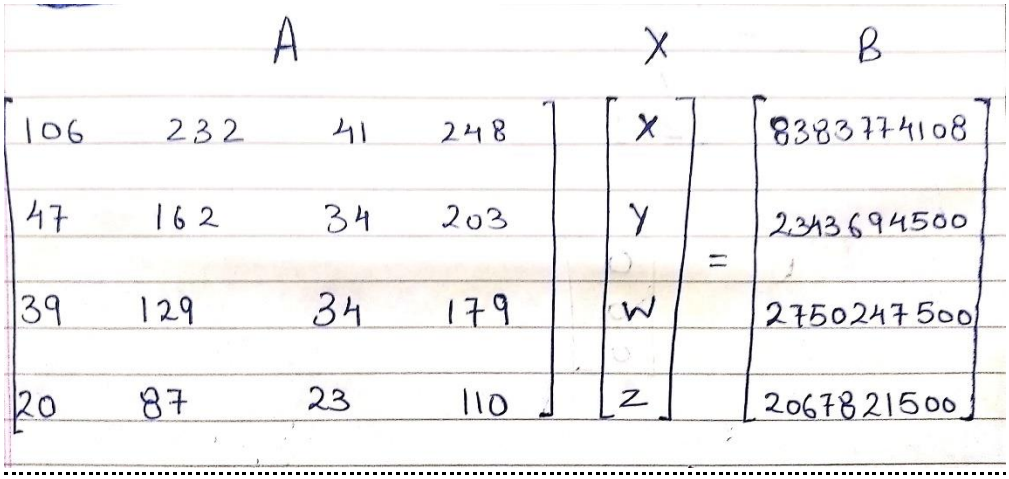
STEP 1:

Formation of equations.

106x + 232y + 41w + 248z = 8383774108.00 [Equation for Bangalore]
47x + 162y + 34w + 203z = 2343694500.00 [Equation for Mumbai]
39x + 129y + 34w + 179z = 2750247500.00 [Equation for Delhi]
20x + 87y + 23w + 110z = 2067821500.00 [Equation for Gurgaon]
Where,
x = Average funding in Technology sector
y = Average funding in Consumer Internet sector
w = Average funding in eCommerce sector
z = Average funding in Other sector

STEP 2:

Create the following matrices:



STEP 3:

The value of X will $A^{-1}B$.

PYTHON IMPLEMENTATION:

In [49]:
```
final_observed.head(4)
```

Out[49]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector	Total_row
Bangalore	106	232	41	248	627
Mumbai	47	162	34	203	446
New Delhi	39	129	34	179	381
Gurgaon	20	87	23	110	240

In [38]:
```
mat = final_observed.drop('Total_row',axis=1)
mat = mat.head(4)
mat = np.matrix(mat)
mat
```

Out[38]:
```
matrix([[106, 232,  41, 248],
        [ 47, 162,  34, 203],
        [ 39, 129,  34, 179],
        [ 20,  87,  23, 110]], dtype=int64)
```

In [39]:
```
banglore=data.loc[data['CityLocation'] == 'Bangalore']
banglore["AmountInUSD"] = banglore["AmountInUSD"].apply(lambda x: float(str(x).replace(",","")))
banglore["AmountInUSD"] = pd.to_numeric(banglore["AmountInUSD"])
banglore=np.array(banglore['AmountInUSD'])
banglore=np.nansum(banglore)
```

In [40]:
```
banglore
```

Out[40]:
```
8383774108.0
```

In [41]:
```
mumbai=data.loc[data['CityLocation'] == 'Mumbai']
mumbai["AmountInUSD"] = mumbai["AmountInUSD"].apply(lambda x: float(str(x).replace(",","")))
mumbai["AmountInUSD"] = pd.to_numeric(mumbai["AmountInUSD"])
mumbai=np.array(mumbai['AmountInUSD'])
mumbai=np.nansum(mumbai)

mumbai
```

Out[41]:
```
2343694500.0
```

In [42]:
```
delhi=data.loc[data['CityLocation'] == 'New Delhi']
delhi["AmountInUSD"] = delhi["AmountInUSD"].apply(lambda x: float(str(x).replace(",","")))
delhi["AmountInUSD"] = pd.to_numeric(delhi["AmountInUSD"])
delhi=np.array(delhi['AmountInUSD'])
delhi=np.nansum(delhi)

delhi
```

Out[42]:
```
2750247500.0
```

In [43]:
```
gurgaon=data.loc[data['CityLocation'] == 'Gurgaon']
gurgaon["AmountInUSD"] = gurgaon["AmountInUSD"].apply(lambda x: float(str(x).replace(",","")))
gurgaon["AmountInUSD"] = pd.to_numeric(gurgaon["AmountInUSD"])
gurgaon=np.array(gurgaon['AmountInUSD'])
gurgaon=np.nansum(gurgaon)

gurgaon
```

```
Out[52]: 2067821500.0
```

```
In [64]: compiled_list=np.matrix([int(banglore),int(mumbai),int(delhi),int(gurgaon)])
         compiled_list
```

```
Out[64]: matrix([[8383774108, 2343694500, 2750247500, 2067821500]], dtype=int64)
```

```
In [67]: compiled_list = compiled_list.reshape(4,1)
         compiled_list
```

```
Out[67]: matrix([[8383774108],
                 [2343694500],
                 [2750247500],
                 [2067821500]], dtype=int64)
```

```
In [68]: np.linalg.inv(mat)
```

```
Out[68]: matrix([[ 0.01850635, -0.04318754,  0.04781073, -0.03982385],
                 [ 0.00331828,  0.02309472, -0.06741913,  0.05960786],
                 [ 0.02267993, -0.1660561 ,  0.04558694,  0.18113366],
                 [-0.01073141,  0.02430728,  0.03509773, -0.06868619]])
```

```
In [75]: result = np.matmul(np.linalg.inv(mat) , compiled_list)
         result
```

```
Out[75]: matrix([[ 1.03077375e+08],
                 [ 1.97857545e+07],
                 [ 3.00886104e+08],
                 [ 7.85042458e+07]])
```

```
In [80]: result=np.array(result)
```

```
In [83]: print("Average funding of :")

         print("Technology         = ",result[0][0])
         print("Consumer Internet = ",result[1][0])
         print("Ecommerce          = ",result[2][0])
         print("Other Sectors      = ",result[3][0])
```

```
Average funding of :
Technology         =  103077374.97699003
Consumer Internet =  19785754.51530163
Ecommerce          =  300886104.08223027
Other Sectors      =  78504245.78420305
```

RESULT:

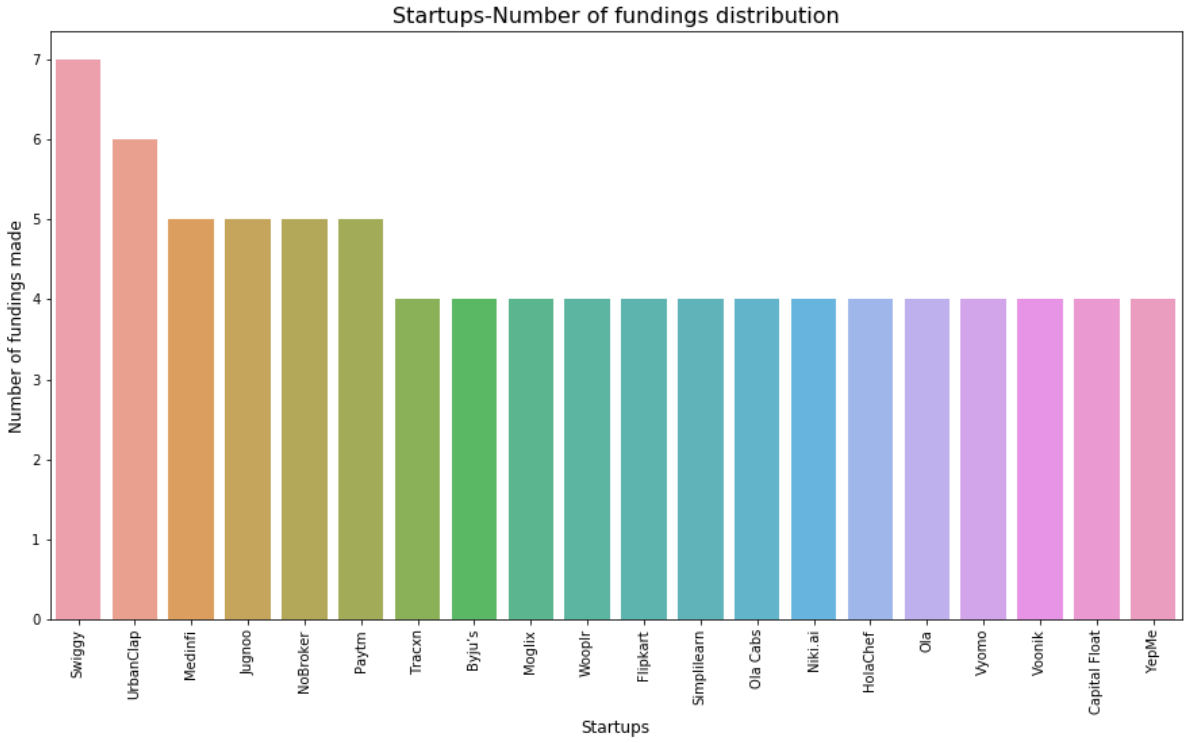
The average funding of the following sectors are as follows:

```
Average funding of :
Technology Sector         =  103077374.97699003
Consumer Internet Sector  =  19785754.51530163
Ecommerce Sector          =  300886104.08223027
Other Sectors             =   78504245.78420305
```

CONCLUSION:

1. Ecommerce sector has received the most funding across the four dominantly cities.
2. If a person is about to start a startup, he may consider it in the ecommerce sector so that in future he might get desired funding amounts.
3. This will help in the development of ecommerce sector which can also provide employment opportunities.

But if there is huge amount of funding being provided to the startups, there should be nominal revenue too. Only then we will be able to speak to the nation's growth – financially and in innovation too.



Conclusion from above:

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

Objective 4: To find relation between the ranking of Cities (according to number of startups in the city) in two sectors – Technology and Consumer Internet.

MATHEMATICAL METHOD (Spearman's rank correlation):

STEP 1:

Find the ranks for each individual subject.

S.No	City location	Tech	Cons. Int	R_1	R_2	D	D^2
1	Bangalore	106	232	1	1	0	0
2	Mumbai	47	162	2	2	0	0
3	New Delhi	39	129	3	3	0	0
4	Pune	31	21	4	7	-3	9
5	Gurgaon	20	87	5	4	1	1
6	Chennai	14	18	6.5	8	-1.5	2.25
7	Hyderabad	14	25	6.5	6	0.5	0.25
8	Noida	13	29	8	5	3	9
9	Ahmedabad	6	12	9	9.5	-0.5	0.25
10	Kolkata	4	8	10	11	-1	1
11	Singapore	2	1	11.5	12.5	-1	1
12	Jaipur	2	12	11.5	9.5	2	4
13	Kozhikode	1	0	14	14.5	-0.5	0.25
14	Missourie	1	0	14	14.5	-0.5	0.25
15	Kochi	1	1	14	12.5	1.5	2.25
N = 15							$\leq D^2 = 30.5$

STEP 2:

Add a third column, "D", to your data. The d is the difference between ranks.

STEP 3:

Sum (add up) all your D-squared values.

STEP 4:

Insert the values into the formula. These ranks are not tied, so use the first formula:

$$R = 1 - \frac{6\left\{\Sigma D^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2) + \ldots\right\}}{N^3 - N}$$

Calculation :

$$R = 1 - \frac{6\left\{30.5 + \frac{6}{12} + \frac{6}{12} + \frac{6}{12} + \frac{6}{12} + \frac{6}{12}\right\}}{3360}$$

$$R = 1 - \frac{6\{30.5 + 0.5 \times 5 + 2\}}{3360}$$

$$R = 1 - \frac{6\{35\}}{3360} = 1 - \frac{210}{336} = 1 - 0.0625$$

$$R = 0.94$$

PYTHON IMPLEMENTATION:

```
In [53]: from scipy.stats import spearmanr
```

```
In [54]: final_observed
```

Out[75]:

CityLocation	Technology	consumer internet	Ecommerce	others_sector
Bangalore	106	232	41	248
Mumbai	47	162	34	203
New Delhi	39	129	34	179
Pune	31	21	4	28
Gurgaon	20	87	23	110
Chennai	14	18	3	31
Hyderabad	14	25	5	32
Noida	13	29	7	29
Ahmedabad	6	12	6	11
Kolkata	4	8	1	6
Singapore	2	1	0	2
Jaipur	2	12	3	8
Kozhikode	1	0	0	0
Missourie	1	0	0	0
Kochi	1	1	0	0

```
In [76]: tech=list(final_observed['Technology'])
         cons_int=list(final_observed['consumer internet'])
```

```
In [77]: spearmanr(final_observed['Technology'],final_observed['consumer internet'])
```

```
Out[77]: SpearmanrResult(correlation=0.9450979550283518, pvalue=1.1045589280798846e-07)
```

RESULT:

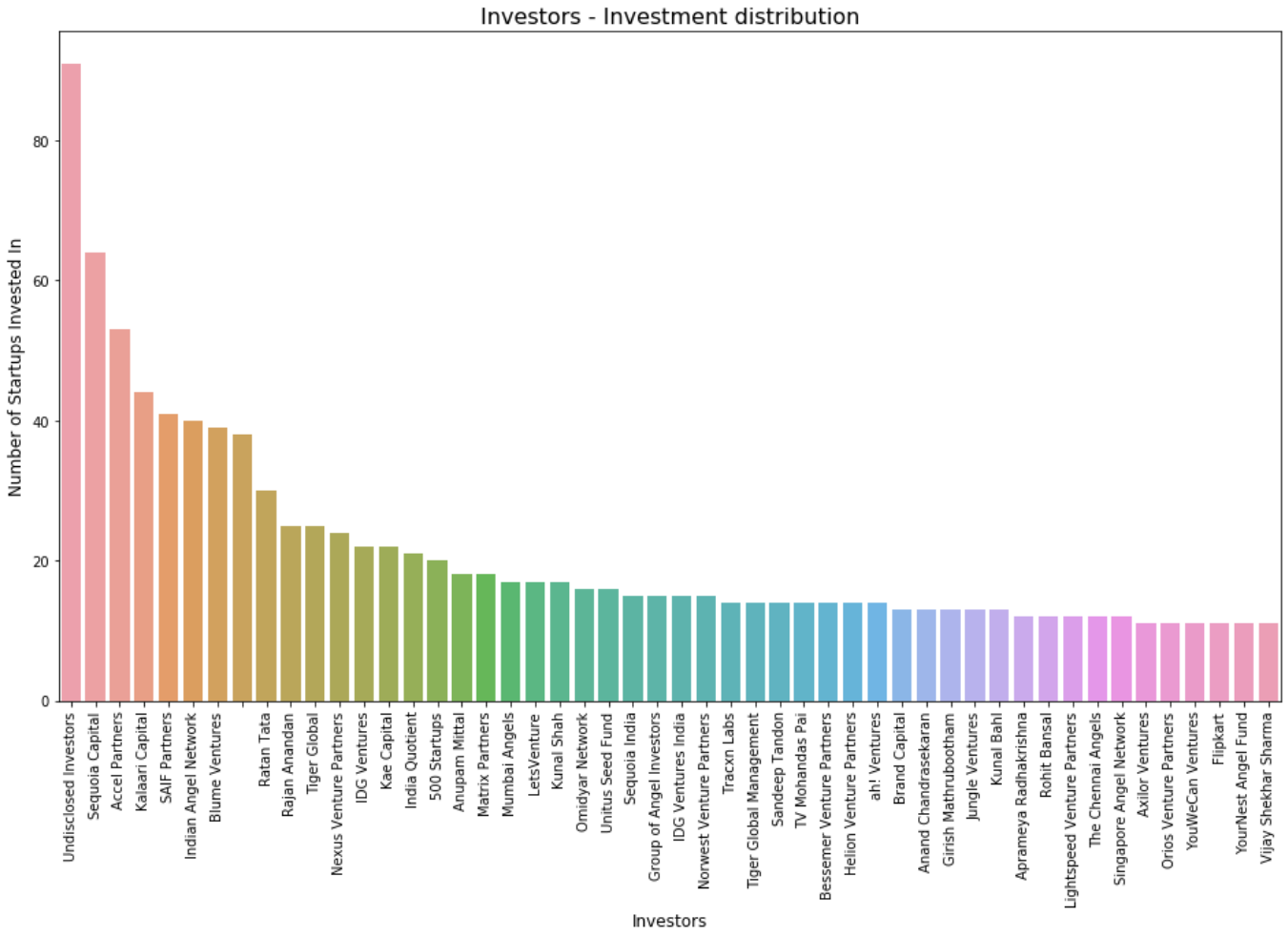
The city ranking based on technology and consumer internet sector is highly correlated.
This can be understood as the rate of growing up of city in technology sector is like that of consumer internet sector.

CONCLUSION:

The following points can be concluded:

1. The city advancing in technology sector is almost similar in that of consumer internet sector.
2. Entrepreneurs can take this as an opportunity of collaboration with others to expand or diversify startups.
3. This will surely lead to employment opportunities.

4. The diversification of numerous startups will lead to innovation and industry growth at a rapid rate.



FUTURE SCOPE

The following points can be included in the study in future to diversify it:

1. Revenue generation by the startups can be worked upon, to compare the funding startups get and the revenue they generate.

2. The collaborations made by startups can be studied to know in detail the industry growth in terms of innovation and employment factor.

REFERNCES

1 Upadhyay1, C. S., & P. R. (2017, October). Start Ups; Let's Start Them Up - An Inside View in the Indian Start Up Scenario. Retrieved from https://www.ijraset.com/fileserve.php?FID=10609

2 Goal 8: Decent work and economic growth. (n.d.). Retrieved from https://www.undp.org/content/undp/en/home/sustainable-development-goals/goal-8-decent-work-and-economic-growth.html

3 Goal 8: Decent work and economic growth. (n.d.). Retrieved from https://www.undp.org/content/undp/en/home/sustainable-development-goals/goal-8-decent-work-and-economic-growth.html

4 M., & F. (n.d.). Inferences concerning proportions. In *Probability and Statistics*(p. 308).

5 Spearman Rank Correlation (Spearman's Rho): Definition and How to Calculate it. (2018, September 02). Retrieved from https://www.statisticshowto.datasciencecentral.com/spearman-rank-correlation-definition-calculate/

Government Scheme. (n.d.). Retrieved from https://www.startupindia.gov.in/content/sih/en/reources/government-schemes.htmlhttp://in.one.un.org/wp-content/uploads/2018/10/goal-9_32479504515_0.png

Spearman's Rank Correlation part 3. (2017, November 09). Retrieved from https://youtu.be/Fim8B0SAVqw

APPENDIX

DATA:

```
In [2]: data = pd.read_csv("startup_funding.csv")
        data
```

Out[2]:

	SNo	Date	StartupName	IndustryVertical	SubVertical	CityLocation	InvestorsName	InvestmentType	AmountInUSD	Remarks
0	0	01/08/2017	TouchKin	Technology	Predictive Care Platform	Bangalore	Kae Capital	Private Equity	1,300,000	NaN
1	1	02/08/2017	Ethinos	Technology	Digital Marketing Agency	Mumbai	Triton Investment Advisors	Private Equity	NaN	NaN
2	2	02/08/2017	Leverage Edu	Consumer Internet	Online platform for Higher Education Services	New Delhi	Kashyap Deorah, Anand Sankeshwar, Deepak Jain,...	Seed Funding	NaN	NaN
3	3	02/08/2017	Zepo	Consumer Internet	DIY Ecommerce platform	Mumbai	Kunal Shah, LetsVenture, Anupam Mittal, Hetal ...	Seed Funding	500,000	NaN
4	4	02/08/2017	Click2Clinic	Consumer Internet	healthcare service aggregator	Hyderabad	Narottam Thudi, Shireesh Palle	Seed Funding	850,000	NaN
5	5	01/07/2017	Billion Loans	Consumer Internet	Peer to Peer Lending platform	Bangalore	Reliance Corporate Advisory Services Ltd	Seed Funding	1,000,000	NaN
6	6	03/07/2017	Ecolibriumenergy	Technology	Energy management solutions provider	Ahmedabad	Infuse Ventures, JLL	Private Equity	2,600,000	NaN

The size of the data is 2372 rows × 10 columns.