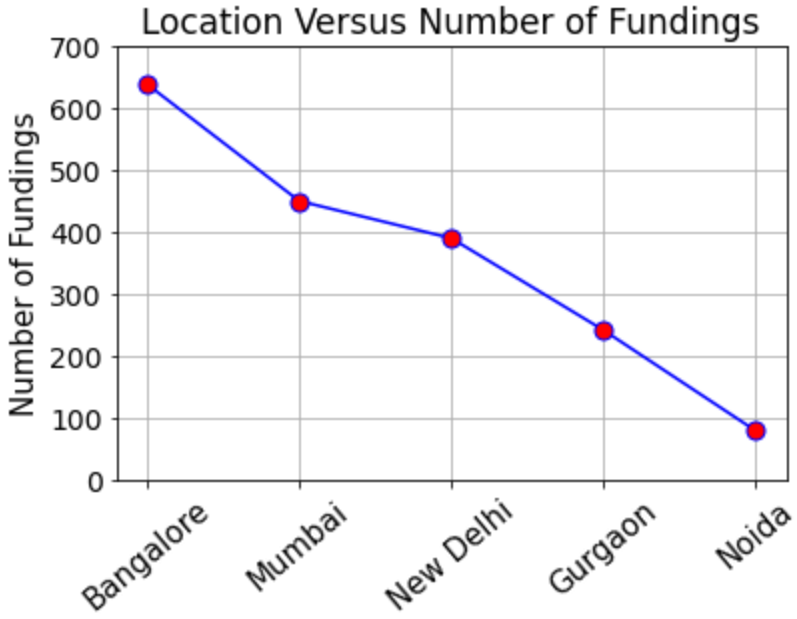
CASE STUDY 1

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df=pd.read_csv('startup_funding.csv', encoding='utf-8')
df['CityLocation'].dropna(inplace=True)
df['CityLocation'].astype(str)

def city_count(city):
    l = str(city).split('/')
    for c in l:
        c=c.strip()
        if c in locations:
            d[c]=d.get(c,0)+1
    return


df['CityLocation'].replace("Delhi","New Delhi",inplace=True)
df['CityLocation'].replace("bangalore","Bangalore",inplace=True)

locations=['Bangalore','Mumbai','Gurgaon','Noida','New Delhi']
d={}
df['CityLocation'].apply(city_count)

location = np.array(list(d.keys()))
number_of_fundings = np.array(list(d.values()))

print('List of locations where most number of times fundings for startups have been observed:')
for i in range(len(location)):
    print(location[i],"-",number_of_fundings[i])

ind=number_of_fundings.argmax()
```

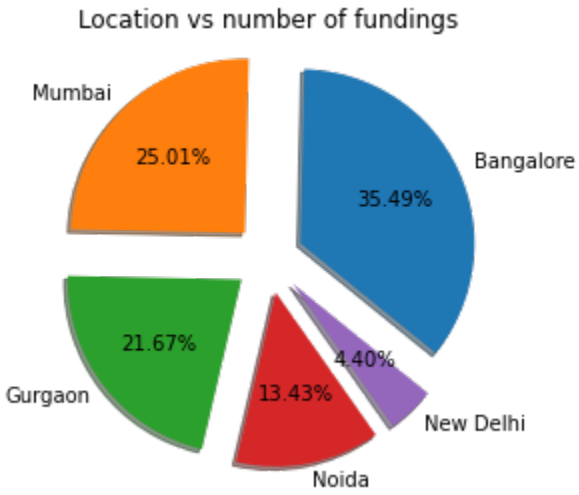
```
plt.plot(location,number_of_fundings, color = 'blue', marker = 'o', markerfacecolor = 'red', markersize = 9)
plt.ylabel('Number of Fundings',fontsize=15)
plt.title('Location Versus Number of Fundings',fontsize=17)
plt.yticks(fontsize=14)
plt.xticks(rotation=40,fontsize=15)
plt.ylim([0, 700])
plt.grid()
plt.show()
```

```
List of locations where most number of times fundings for startups have been observed:
Bangalore - 637
Mumbai - 449
New Delhi - 389
Gurgaon - 241
Noida - 79
Maximum Funding is done in Bangalore , 637 times.
```




```
d = np.true_divide(number_of_fundings, number_of_fundings.sum())*(100)
labels = locations
```

```
explode = [0.2, 0.2, 0.2, 0.2, 0.2]
plt.pie(d, labels = labels, radius = 1, autopct = '%.2f%%', explode = explode, shadow = True, startangle = -39, counterclc
plt.title('Location vs number of fundings')
plt.show()
```

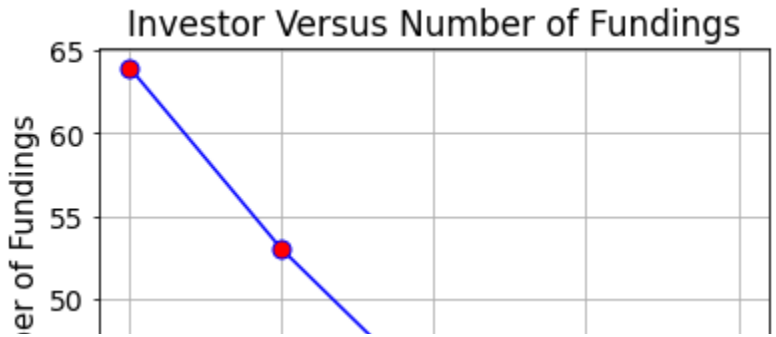


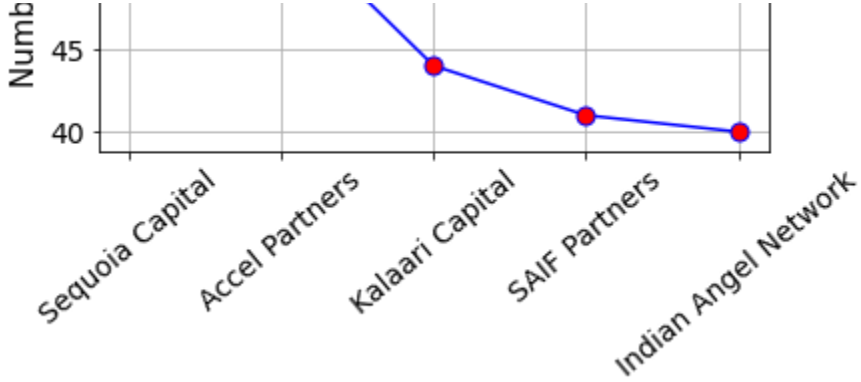
CASE STUDY 2

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df=pd.read_csv('startup_funding.csv', encoding='utf-8')
df['InvestorsName'].dropna(inplace=True)
d= {}

def investor_count(investor):
    l = str(investor).split(',')
    for i in l:
        i = i.strip()
        d[i]=d.get(i,0)+1
```

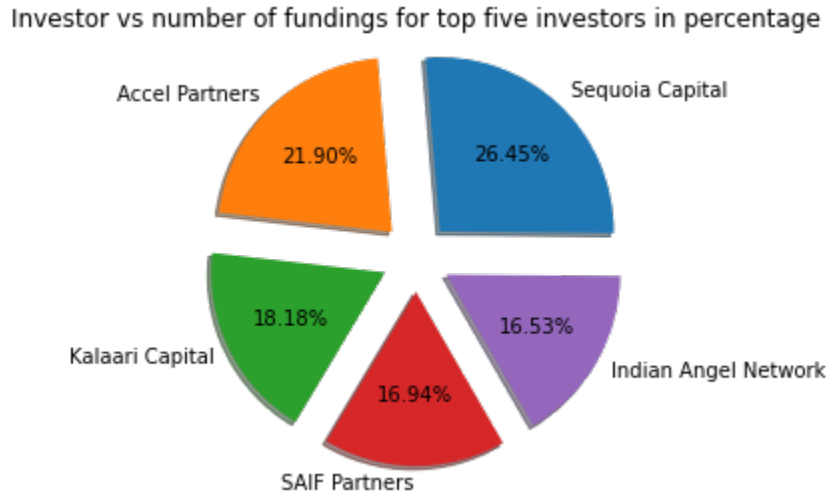
```
df['InvestorsName'].apply(investor_count)

investor_name=np.array(list(d.keys()))
number_of_fundings=np.array(list(d.values()))

ind=number_of_fundings.argsort()[::-1]
ind=ind[:5]

investor_name=investor_name[ind]
number_of_fundings=number_of_fundings[ind]

print('Top 5 Investors are:')
for i in range(len(investor_name)):
    print(investor_name[i],'-',number_of_fundings[i])

plt.plot(investor_name,number_of_fundings, color = 'blue', marker = 'o', markerfacecolor = 'red', markersize = 9)
plt.ylabel('Number of Fundings',fontsize=15)
plt.title('Investor Versus Number of Fundings',fontsize=17)
plt.yticks(fontsize=14)
plt.xticks(rotation=40,fontsize=14)
plt.grid()
plt.show()
```

```
Top 5 Investors are:
Sequoia Capital - 64
Accel Partners - 53
Kalaari Capital - 44
SAIF Partners - 41
Indian Angel Network - 40
```






```
d = np.true_divide(number_of_fundings, number_of_fundings.sum())*(100)
labels = investor_name
explode = [0.2, 0.2, 0.2, 0.2, 0.2]
plt.pie(d, labels = labels, radius = 1, autopct = '%.2f%%', explode = explode, shadow = True, startangle = -0.5, counterc]
plt.title('Investor vs number of fundings for top five investors in percentage')
plt.show()
```

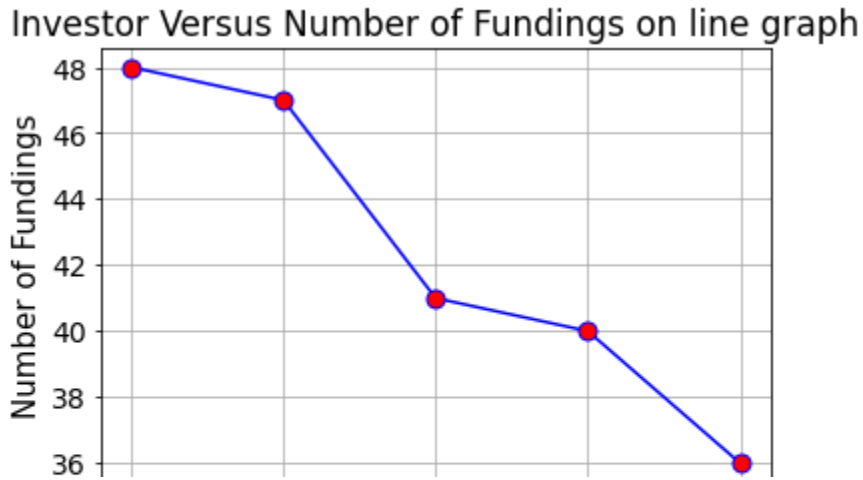


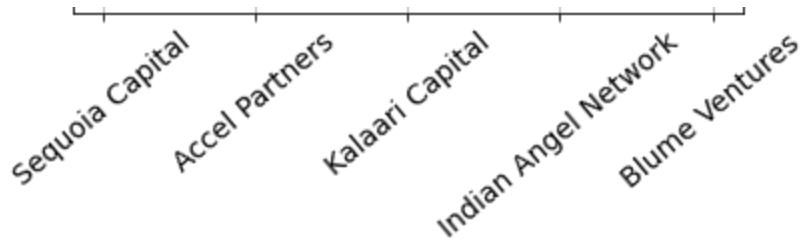
CASE STUDY 3

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df=pd.read_csv('startup_funding.csv', encoding='utf-8')
df.dropna(subset=['InvestorsName','StartupName'],inplace=True)
df['StartupName'].replace('Flipkart.com','Flipkart',inplace=True)
df['StartupName'].replace('Ola Cabs','Ola',inplace=True)
df['StartupName'].replace('Olacabs','Ola',inplace=True)
df['StartupName'].replace('Ola Cabs','Ola',inplace=True)
df['StartupName'].replace('Olacabs','Ola',inplace=True)
df['StartupName'].replace('Oyo Rooms','Oyo',inplace=True)
df['StartupName'].replace('Oyorooms','Oyo',inplace=True)
df['StartupName'].replace('OyoRooms','Oyo',inplace=True)
df['StartupName'].replace('OYO Rooms','Oyo',inplace=True)
df['StartupName'].replace('Paytm Marketplace','Paytm',inplace=True)

startup=list(df.StartupName)
investor_list=list(df.InvestorsName)

d={}
for i in range(len(investor_list)):
    investor = investor_list[i].split(',')
    for invest in investor:
        invest=invest.strip()
        if invest != "":
            if invest in d:
                s=d[invest]
                s.add(startup[i])
                d[invest]=s
            else:
                d[invest]={startup[i]}

for key in d:
    d[key]=len(d[key])

investor_name=np.array(list(d.keys()))
```

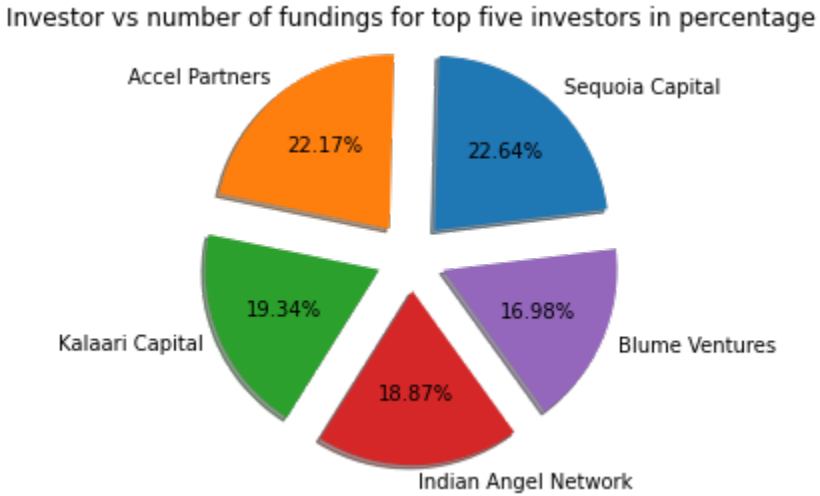
```
number_of_fundings=np.array(list(d.values()))

ind=number_of_fundings.argsort()[::-1]
ind=ind[:5]

investor_name=investor_name[ind]
number_of_fundings=number_of_fundings[ind]

print('Top 5 Investors are:')
for i in range(len(investor_name)):
    print(investor_name[i],'-',number_of_fundings[i])

plt.plot(investor_name,number_of_fundings, color = 'blue', marker = 'o', markerfacecolor = 'red', markersize = 9)
plt.ylabel('Number of Fundings',fontsize=15)
plt.title('Investor Versus Number of Fundings on line graph',fontsize=17)
plt.yticks(fontsize=14)
plt.xticks(rotation=40,fontsize=14)
plt.grid()
plt.show()
```

```
Top 5 Investors are:
Sequoia Capital - 48
Accel Partners - 47
Kalaari Capital - 41
Indian Angel Network - 40
Blume Ventures - 36
```






```
d = np.true_divide(number_of_fundings, number_of_fundings.sum())*(100)
labels = investor_name
explode = [0.2, 0.2, 0.2, 0.2, 0.2]
plt.pie(d, labels = labels, radius = 1, autopct = '%.2f%%', explode = explode, shadow = True, startangle = 7, counterclock
plt.title('Investor vs number of fundings for top five investors in percentage')
plt.show()
```



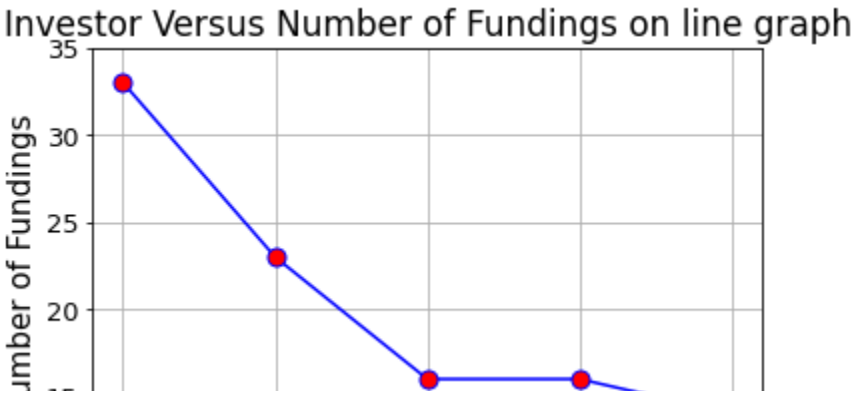
CASE STUDY 4

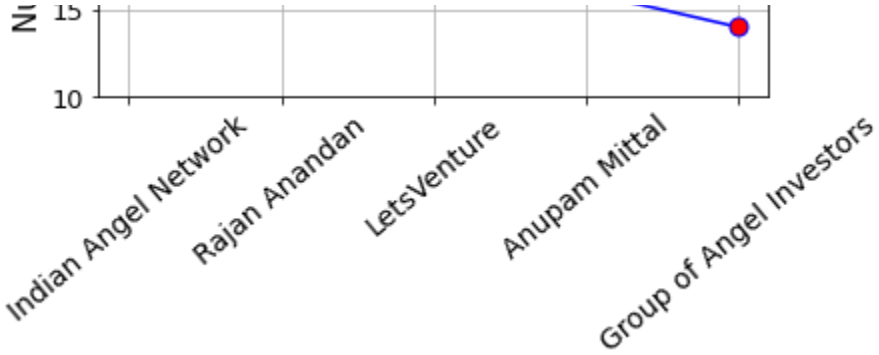
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
df=pd.read_csv('startup_funding.csv', encoding='utf-8')
df.dropna(subset=['InvestorsName','StartupName','InvestmentType'],inplace=True)
df['InvestmentType'].replace('PrivateEquity','Private Equity',inplace=True)
df['InvestmentType'].replace('SeedFunding','Seed Funding',inplace=True)
df['InvestmentType'].replace('Crowd funding','Crowd Funding',inplace=True)
df['InvestorsName'].replace('Undisclosed investors','Undisclosed Investors',inplace=True)
df['StartupName'].replace('Flipkart.com','Flipkart',inplace=True)
df['StartupName'].replace('Ola Cabs','Ola',inplace=True)
df['StartupName'].replace('Olacabs','Ola',inplace=True)
df['StartupName'].replace('Ola Cabs','Ola',inplace=True)
df['StartupName'].replace('Olacabs','Ola',inplace=True)
df['StartupName'].replace('Oyo Rooms','Oyo',inplace=True)
df['StartupName'].replace('Oyorooms','Oyo',inplace=True)
df['StartupName'].replace('OyoRooms','Oyo',inplace=True)
df['StartupName'].replace('OYO Rooms','Oyo',inplace=True)
df['StartupName'].replace('Paytm Marketplace','Paytm',inplace=True)

df=df[(df.InvestmentType=='Seed Funding') | (df.InvestmentType=='Crowd Funding')]

startup=list(df.StartupName)
investor_list=list(df.InvestorsName)

d={}
for i in range(len(investor_list)):
    investor = investor_list[i].split(',')
    for invest in investor:
        invest=invest.strip()
        if (invest != "") and (invest != "Undisclosed Investors"):
            if invest in d:
                s=d[invest]
                s.add(startup[i])
                d[invest]=s
            else:
                d[invest] = {startup[i]}
for key in d:
    d[key]=len(d[key])
```

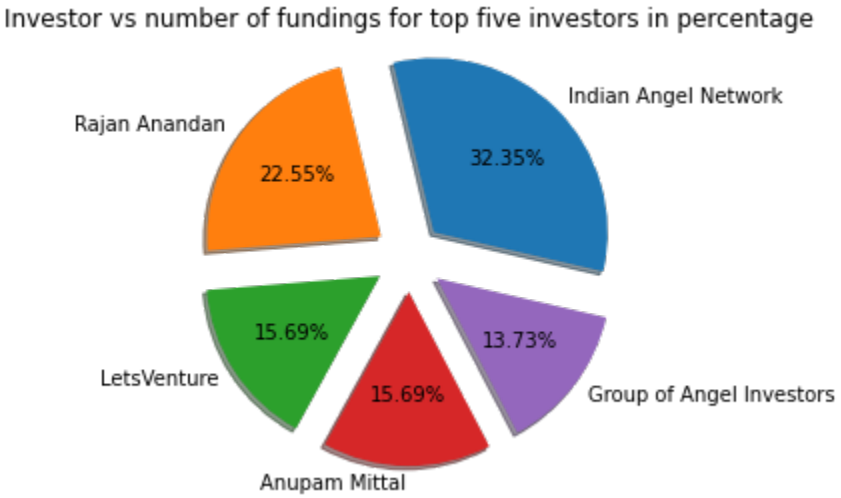
```
investor_name=np.array(list(d.keys()))
number_of_fundings=np.array(list(d.values()))

ind=number_of_fundings.argsort()[::-1]
ind=ind[:5]

investor_name=investor_name[ind]
number_of_fundings=number_of_fundings[ind]

print('Top 5 Investors for Investment type- Seed Funding and Crowd Funding are:')
for i in range(len(investor_name)):
    print(investor_name[i],'-',number_of_fundings[i])

plt.plot(investor_name,number_of_fundings, color = 'blue', marker = 'o', markerfacecolor = 'red', markersize = 9)
plt.ylabel('Number of Fundings',fontsize=15)
plt.title('Investor Versus Number of Fundings on line graph',fontsize=17)
plt.yticks(fontsize=13)
plt.xticks(rotation=40,fontsize=14)
plt.ylim([10, 35])
plt.grid()
plt.show()
```

```
Top 5 Investors for Investment type- Seed Funding and Crowd Funding are:
Indian Angel Network - 33
Rajan Anandan - 23
LetsVenture - 16
Anupam Mittal - 16
Group of Angel Investors - 14
```






```
d = np.true_divide(number_of_fundings, number_of_fundings.sum())*(100)
labels = investor_name
explode = [0.2, 0.2, 0.2, 0.2, 0.2]
plt.pie(d, labels = labels, radius = 1, autopct = '%.2f%%', explode = explode, shadow = True, startangle = -13, counterclo
plt.title('Investor vs number of fundings for top five investors in percentage')
plt.grid()
plt.show()
```

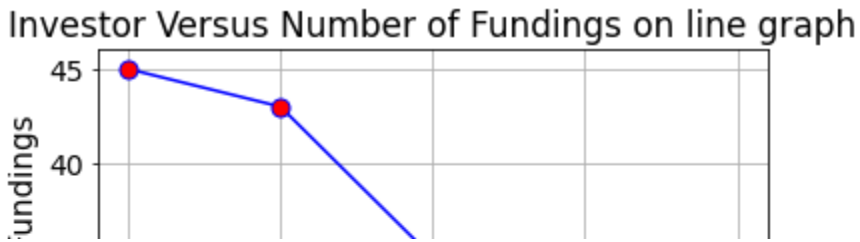


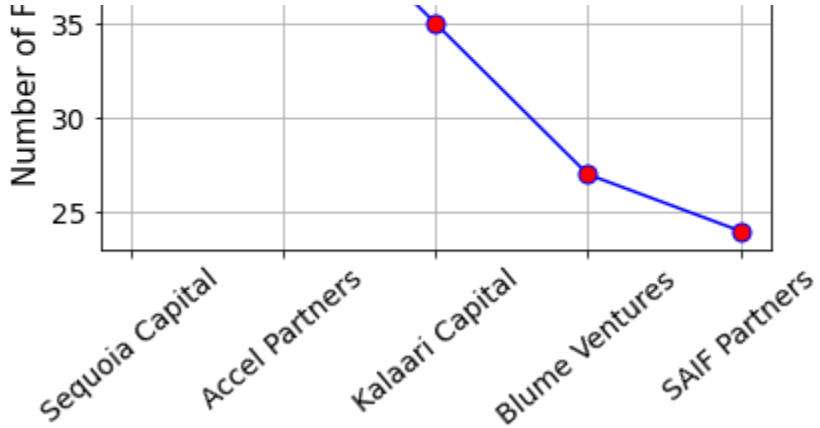
CASE STUDY 5

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df=pd.read_csv('startup_funding.csv', encoding='utf-8')
df.dropna(subset=['InvestorsName','StartupName','InvestmentType'],inplace=True)
df['InvestmentType'].replace('PrivateEquity','Private Equity',inplace=True)
df['InvestmentType'].replace('SeedFunding','Seed Funding',inplace=True)
df['InvestmentType'].replace('Crowd funding','Crowd Funding',inplace=True)
df['InvestorsName'].replace('Undisclosed investors','Undisclosed Investors',inplace=True)
df['StartupName'].replace('Flipkart.com','Flipkart',inplace=True)
df['StartupName'].replace('Ola Cabs','Ola',inplace=True)
df['StartupName'].replace('Olacabs','Ola',inplace=True)
df['StartupName'].replace('Ola Cabs','Ola',inplace=True)
df['StartupName'].replace('Olacabs','Ola',inplace=True)
df['StartupName'].replace('Oyo Rooms','Oyo',inplace=True)
df['StartupName'].replace('Oyorooms','Oyo',inplace=True)
df['StartupName'].replace('OyoRooms','Oyo',inplace=True)
df['StartupName'].replace('OYO Rooms','Oyo',inplace=True)
df['StartupName'].replace('Paytm Marketplace','Paytm',inplace=True)

df=df[df.InvestmentType=='Private Equity']

startup=list(df.StartupName)
investor_list=list(df.InvestorsName)

d={}
for i in range(len(investor_list)):
    investor = investor_list[i].split(',')
    for invest in investor:
        invest=invest.strip()
        if (invest != "") and (invest != "Undisclosed Investors"):
            if invest in d:
                s=d[invest]
                s.add(startup[i])
                d[invest]=s
            else:
```

```
                d[invest]={startup[i]}

for key in d:
    d[key]=len(d[key])

investor_name=np.array(list(d.keys()))
number_of_fundings=np.array(list(d.values()))

ind=number_of_fundings.argsort()[::-1]
ind=ind[:5]

investor_name=investor_name[ind]
number_of_fundings=number_of_fundings[ind]

print('Top 5 Investors for Investment type- Private Equity are:')
for i in range(len(investor_name)):
    print(investor_name[i],'-',number_of_fundings[i])

plt.plot(investor_name,number_of_fundings, color = 'blue', marker = 'o', markerfacecolor = 'red', markersize = 9)
plt.ylabel('Number of Fundings',fontsize=15)
plt.title('Investor Versus Number of Fundings on line graph',fontsize=17)
plt.yticks(fontsize=14)
plt.xticks(rotation=40,fontsize=14)
plt.grid()
plt.show()
```

```
Top 5 Investors for Investment type- Private Equity are:
Sequoia Capital - 45
Accel Partners - 43
Kalaari Capital - 35
Blume Ventures - 27
SAIF Partners - 24
```






```
d = np.true_divide(number_of_fundings, number_of_fundings.sum())*(100)
labels = investor_name
explode = [0.2, 0.2, 0.2, 0.2, 0.2]
plt.pie(d, labels = labels, radius = 1, autopct = '%.2f%%', explode = explode, shadow = True, startangle = -3, counterclo
plt.title('Investor vs number of fundings for top five investors in percentage')
plt.show()
```