

CASE STUDY PART-2 PROJECT

Question-1)

Your Friend has developed the Product and he wants to establish the product startup and he is searching for a perfect location where getting the investment has a high chance. But due to its financial restriction, he can choose only between three locations - Bangalore, Mumbai, and NCR. As a friend, you want to help your friend deciding the location. NCR include Gurgaon, Noida and New Delhi. Find the location where the most number of funding is done. That means, find the location where startups has received funding maximum number of times. Plot the bar graph between location and number of funding. Take city name "Delhi" as "New Delhi". Check the case-sensitiveness of cities also. That means, at some place instead of "Bangalore", "bangalore" is given. Take city name as "Bangalore". For few startups multiple locations are given, one Indian and one Foreign. Consider the startup if any one of the city lies in given locations.

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
In [2]: #importing libraries
import pandas as pd
import matplotlib.pyplot as plt
startup = pd.read_csv("startup_funding.csv")
df = startup.copy()
#dropping nan values
df.dropna(subset = ["CityLocation"],inplace = True)
df["CityLocation"].replace("bangalore","Bangalore",inplace = True)
df["CityLocation"].replace("Delhi","New Delhi",inplace = True)
d1 = {}
for i in df["CityLocation"]:
    if "Bangalore" in i:
        d1["Bangalore"] = d1.get("Bangalore",0) + 1
    elif "Mumbai" in i:
        d1["Mumbai"] = d1.get("Mumbai",0) + 1
    elif "New Delhi" in i:
        d1["New Delhi"] = d1.get("New Delhi",0) + 1
    elif "Noida" in i:
        d1["Noida"] = d1.get("Noida",0) + 1
    elif "Gurgaon" in i:
        d1["Gurgaon"] = d1.get("Gurgaon",0) + 1
d1['NCR'] = d1['New Delhi'] + d1['Gurgaon'] + d1['Noida']
del d1['New Delhi']
del d1['Gurgaon']
del d1['Noida']
# sorting the Dictionary in descending order according to it's values
cities = sorted(d1, key=d1.get , reverse=True)
fundings = []

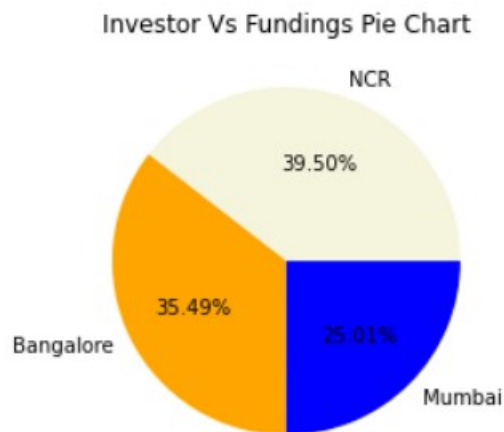
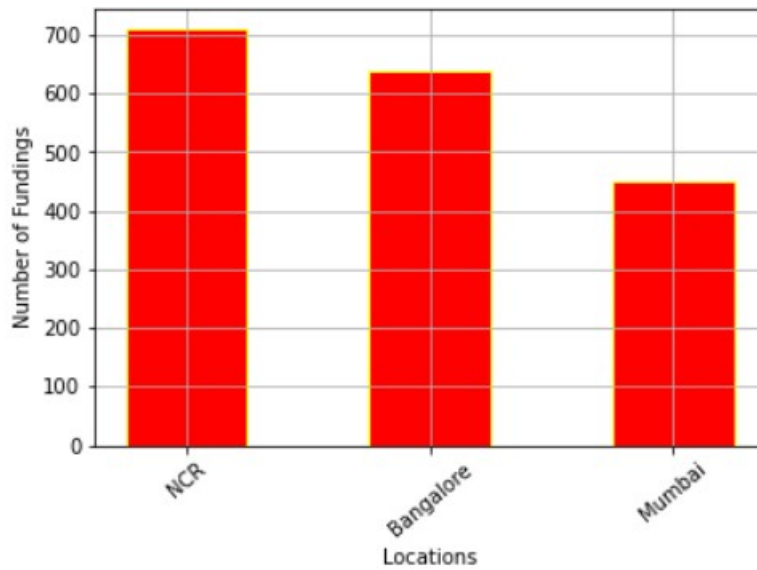
for i in cities:
    fundings.append(d1[i])
y = fundings
x = cities
# plotting the bar graph for 3 cities
plt.bar(cities,fundings,width = 0.5,color = "red" , edgecolor = "yellow")
plt.xlabel("Locations")
plt.ylabel("Number of Fundings")
plt.xticks(rotation = 40)
plt.grid()
plt.show()
```

```

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
plt.grid()
plt.show()
# plotting the pie chart b/w top 5 investors and number of fundings.
color = ["beige", "orange", "blue", "pink", "grey"]
plt.pie(y, labels = x, colors = color, autopct = "%.2f%%")
plt.title("Investor Vs Fundings Pie Chart")
plt.show()

```

GRAPHS:-



RESULT AND JUSTIFICATION:-

As the result of the code shows and according to graphs also,
The Top location is NCR which has maximum number of funding's .
So, I would recommend my friend to establish his start-up in **NCR LOCATION** because
that's where maximum chances of getting investment.

QUESTION 2

Even after trying for so many times, your friend's startup could not find the investment. So you decided to take this matter in your hand and try to find the list of investors who probably can invest in your friend's startup. Your list will increase the chance of your friend startup getting some initial investment by contacting these investors. Find the top 5 investors who have invested maximum number of times (consider repeat investments in one company also). In a startup, multiple investors might have invested. So consider each investor for that startup. Ignore undisclosed investors.

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
In [3]: import pandas as pd
import matplotlib.pyplot as plt
startup = pd.read_csv("startup_funding.csv")
df = startup.copy()
df.dropna(subset = ["InvestorsName"],inplace = True)
d = {}
for i in df["InvestorsName"].values:
    if "," in i:
        for j in i.strip().split(','):
            d[j.strip()] = d.get(j.strip(),0) + 1
    else:
        d[i.strip()] = d.get(i.strip(),0) + 1
d1 = sorted(d, key=d.get , reverse=True)[0:5]
Investor_name=[]
values=[]
for i in d1: #printing the top 5 investors funded maximum number of times
    print(i,d[i])
    Investor_name.append(i)
    values.append(d[i])

y = values
x = Investor_name

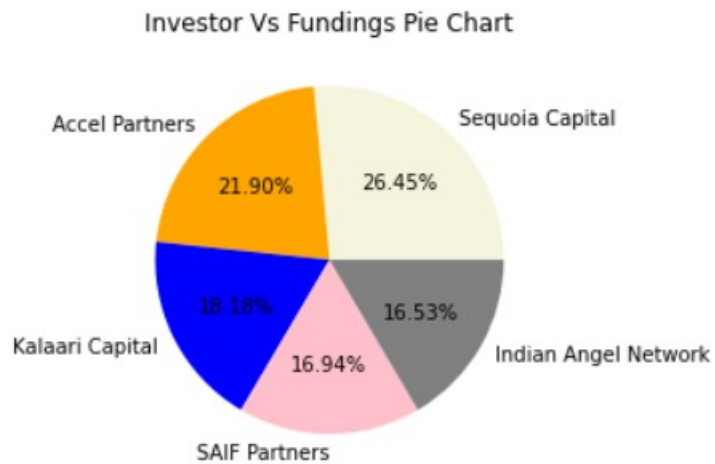
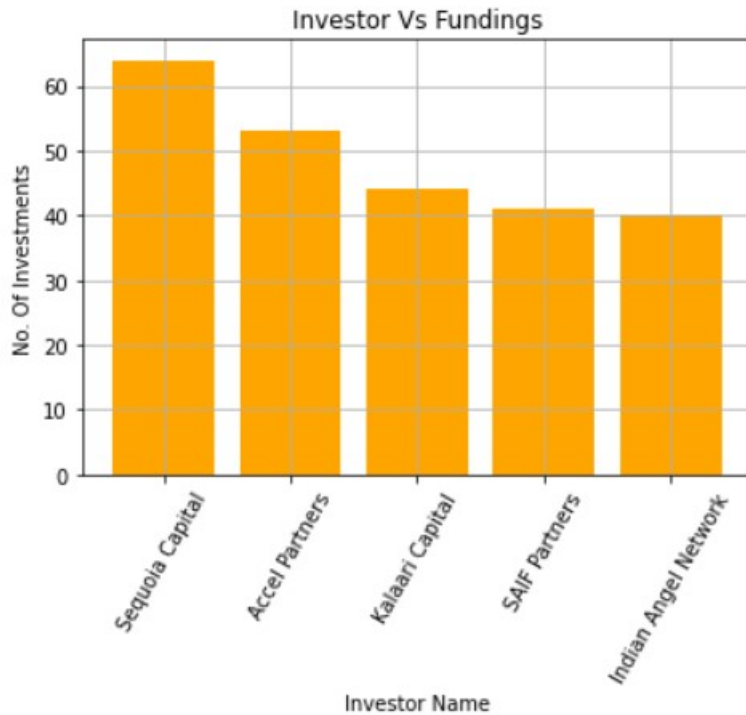
# plotting the bar graph for top 5 investors
plt.bar(x,y, color = 'orange')
plt.xlabel("Investor Name")
plt.ylabel("No. Of Investments")
plt.title("Investor Vs Fundings")
plt.grid()
plt.xticks(rotation = 60)
plt.show()

# plotting the pie chart b/w top 5 investors and number of fundings.
color = ["beige", "orange", "blue", "pink", "grey"]
plt.pie(y, labels = x, colors = color, autopct = "%.2f%%")
plt.title("Investor Vs Fundings Pie Chart")
plt.show()

plt.show()

Sequoia Capital 64
Accel Partners 53
Kalaari Capital 44
SAIF Partners 41
Indian Angel Network 40
```

GRAPHS:-



RESULT AND JUSTIFICATION:-

As the result of the code , we get the TOP 5 Investors who had invested maximum number of times in a start – up.

So, I would suggest my Friend to contact these 5 INVESTORS who can probably invest in his start-up and thereby increasing the chances of getting an initial investment.

QUESTION 3

After re-analysing the dataset you found out that some investors have invested in the

same startup at different number of funding rounds. So before finalising the previous list, you want to improvise it by finding the top 5 investors who have invested in different number of startups. This list will be more helpful than your previous list in finding the investment for your friend startup. Find the top 5 investors who have invested maximum number of times in different companies. That means, if one investor has invested multiple times in one startup, count one for that company. There are many errors in startup names. Ignore correcting all, just handle the important ones - Ola, Flipkart, Oyo and Paytm.

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
In [20]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

file = pd.read_csv("startup_funding.csv")
df = file.copy()

# dropping all the missing elements.
df.dropna(subset = ['InvestorsName'], inplace = True)

# Handling the errors in the names of some important start-ups
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('Oyo Rooms', 'Oyo', inplace = True)
df['StartupName'].replace('Paytm Marketplace', 'Paytm', 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.reset_index(drop = True, inplace = True)

# Making dictionary of all the investors and startup combination
def createDict(df):
    d1 = {}
    for i in range(len(df)):
        a = df['InvestorsName'][i].split(',')
        for j in a:
            j = j.strip()
            d1[(df['StartupName'][i], j)] = d1.get((df['StartupName'][i], j), 0) + 1

    return d1

# calling dictionary
dictionary = createDict(df)
```

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
+ %< %> Run Code

# calling dictionary
dictionary = createDict(df)

# function to create dictionary of non - repetitive investors
def unique_investors(dic):
    x = [] # List of investors
    for i in dictionary:
        x.append(i[1])
    d2 = {} #for finding investors without repetitive invesment done by the investor
    for i in x:
        d2[i] = d2.get(i,0) + 1
    return d2

final_dic = unique_investors(dictionary)
final_dic[''] = 0

# sorting the Dictionary in descending order according to the number of investments
sorted_d = sorted(final_dic.items(), key=lambda kv: kv[1],reverse = True)

investor_count = []
investor_name = []

# adding top 5 investors to the lists.
for i in range(5):
    print(sorted_d[i][0],sorted_d[i][1])
    investor_count.append(sorted_d[i][1])
    investor_name.append(sorted_d[i][0])

y = investor_count
x = investor_name

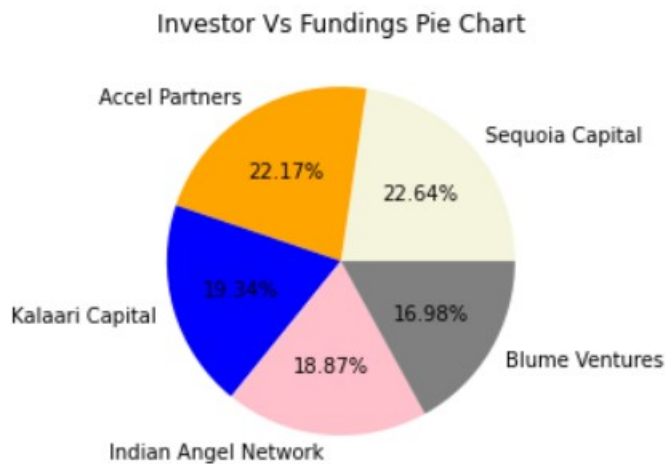
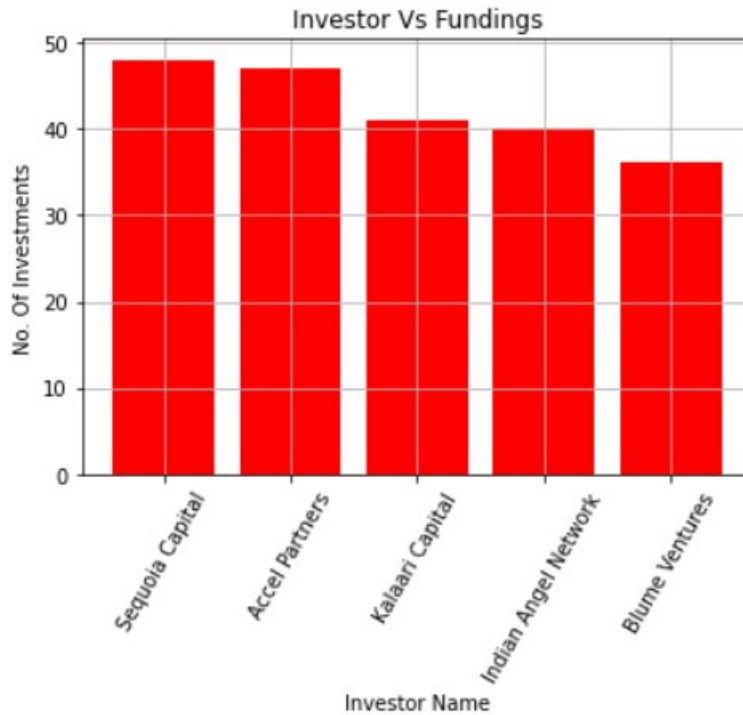
# plotting the bar graph for top 5 investors
plt.bar(x,y, color = 'red')
plt.xlabel("Investor Name")
plt.ylabel("No. Of Investments")
plt.title("Investor Vs Fundings")
plt.grid()
plt.xticks(rotation = 60)
plt.show()

plt.show()

# plotting the pie chart b/w top 5 investors and number of fundings.
color = ["beige", "orange", "blue", "pink", "grey"]
plt.pie(y, labels = x, colors = color, autopct = "%.2f%%")
plt.title("Investor Vs Fundings Pie Chart")
plt.show()

Sequoia Capital 48
Accel Partners 47
Kalaari Capital 41
Indian Angel Network 40
Blume Ventures 36
```

GRAPHS:-



RESULT AND JUSTIFICATION:-

As the result of the code , we get the improved list of TOP 5 Investors who had invested maximum number of times in different start – up's.

So, I would suggest my Friend to contact these 5 INVESTORS who can probably invest in his start-up and thereby increasing the chances of getting an initial investment.

QUESTION 4

Even after putting so much effort in finding the probable investors, it didn't turn out to be helpful for your friend. So you went to your investor friend to understand the situation better and your investor friend explained to you about the different Investment Types and their features. This new information will be helpful in finding the right investor. Since your friend startup is at an early stage startup, the best-suited investment type would be - Seed Funding and Crowdfunding. Find the top 5 investors who have invested in a different number of startups and their investment type is Crowdfunding or Seed Funding. Correct spelling of investment types are - "Private Equity", "Seed Funding", "Debt Funding", and "Crowd Funding". Keep an eye for any spelling mistake. You can find this by printing unique values from this column. There are many errors in startup names. Ignore correcting all, just handle the important ones - Ola, Flipkart, Oyo and Paytm.

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
In [24]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
startup = pd.read_csv("startup_funding.csv")
df = startup.copy()
df.dropna(subset = ["InvestorsName"],inplace = True)
df["InvestorsName"].replace("Undisclosed Investors","",inplace = True)
df["InvestorsName"].replace("Undisclosed investors","",inplace = True)
df = df[(df["InvestmentType"] == "Seed Funding") | (df["InvestmentType"] == "Crowd Funding")]

d = {}
for i in df["InvestorsName"].values:
    if "," in i:
        for j in i.strip().split(','):
            d[j.strip()] = d.get(j.strip(),0) + 1
    else:
        d[i.strip()] = d.get(i.strip(),0) + 1
del d[""]
investor_count = []
investor_name = []
d1 = sorted(d, key=d.get , reverse=True)[0:5]
for i in d1:
    print(i,d[i])
    investor_name.append(i)
    investor_count.append(d[i])
y = investor_count
x = investor_name

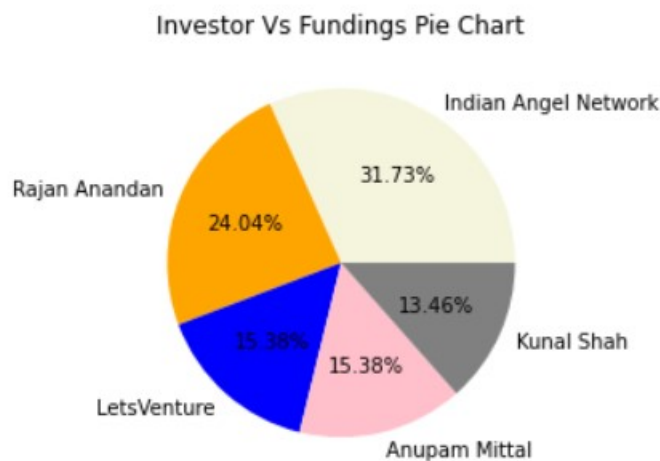
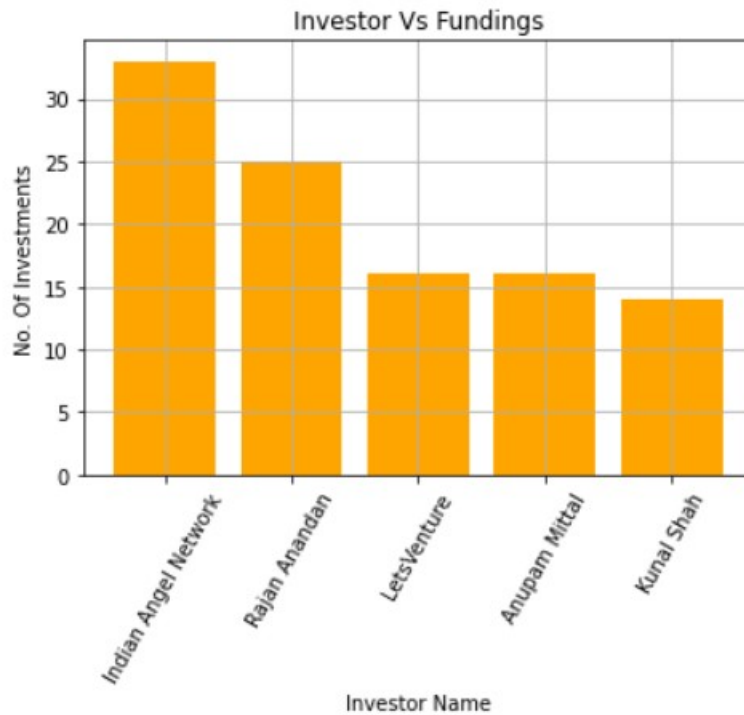
# plotting the bar graph for top 5 investors
plt.bar(x,y, color = 'orange')
plt.xlabel("Investor Name")
plt.ylabel("No. Of Investments")
plt.title("Investor Vs Fundings")
plt.grid()
plt.xticks(rotation = 60)
plt.show()
```

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
plt.xticks(rotation = 60)
plt.show()

# plotting the pie chart b/w top 5 investors and number of fundings.
color = ["beige", "orange", "blue", "pink", "grey"]
plt.pie(y, labels = x, colors = color, autopct = "%.2f%%")
plt.title("Investor Vs Fundings Pie Chart")
plt.show()

Indian Angel Network 33
Rajan Anandan 25
LetsVenture 16
Anupam Mittal 16
Kunal Shah 14
```

GRAPHS:-



RESULT AND JUSTIFICATION:-

As the result of the code , we get the improved list of TOP 5 Investors who had invested maximum number of times in different start – up's and invested in Seed and Crowd type of Funding's. So, I would suggest my Friend to contact these 5 INVESTORS who can probably invest in his start-up and thereby increasing the chances of getting an initial investment.

QUESTION 5

Due to your immense help, your friend startup successfully got seed funding and it is on the operational mode. Now your friend wants to expand his startup and he is looking for new investors for his startup. Now you again come as a saviour to help your friend and want to create a list of probable new new investors. Before moving forward you remember your investor friend advice that finding the investors by analysing the investment type. Since your friend startup is not in early phase it is in growth stage so the best-suited investment type is Private Equity. Find the top 5 investors who have invested in a different number of startups and their investment type is Private Equity. Correct spelling of investment types are - "Private Equity", "Seed Funding", "Debt Funding", and "Crowd Funding". Keep an eye for any spelling mistake. You can find this by printing unique values from this column. There are many errors in startup names. Ignore correcting all, just handle the important ones - Ola, Flipkart, Oyo and Paytm.

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
In [7]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

file = pd.read_csv("startup_funding.csv")
df = file.copy()

# dropping all the missing elements.
df.dropna(subset = ['InvestorsName'], inplace = True)
df['InvestmentType'].dropna(inplace = True)

# Handling the errors in the names of some important start-ups
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('Oyo Rooms', 'Oyo', inplace = True)
df['StartupName'].replace('Paytm Marketplace', 'Paytm', 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['InvestmentType'].replace('PrivateEquity', 'Private Equity', inplace = True)

df.reset_index(drop = True, inplace = True)

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

df.reset_index(drop = True, inplace = True)
# Making dictionary of all the investors and startup combination
def createDict(df):
    d1 = {}
    for i in range(len(df)):
        a = df['InvestorsName'][i].split(',')
        for j in a:
            j = j.strip()
            d1[(df['StartupName'][i], j)] = d1.get((df['StartupName'][i], j), 0) + 1
    return d1
```

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 O
+ < > Run Code
return d1

# calling dictionary
dictionary = createDict(df)

# function to create dictionary of non - repetitive investors
def unique_investors(dic):
    x = [] # list of investors
    for i in dictionary:
        x.append(i[1])

    d2 = {} #for finding investors without repetitive invesment done by the investor
    for i in x:
        d2[i] = d2.get(i,0) + 1
    return d2

final_dic = unique_investors(dictionary)
final_dic[''] = 0
final_dic['Undisclosed Investors'] = 0
final_dic['Undisclosed investors'] = 0

# sorting the Dictionary in descending order according to the number of investments
sorted_d = sorted(final_dic.items(), key=lambda kv: kv[1],reverse = True)

investor_count = []
investor_name = []

# adding top 5 investors to the Lists.
for i in range(5):
    print(sorted_d[i][0],sorted_d[i][1])
    investor_count.append(sorted_d[i][1])
    investor_name.append(sorted_d[i][0])

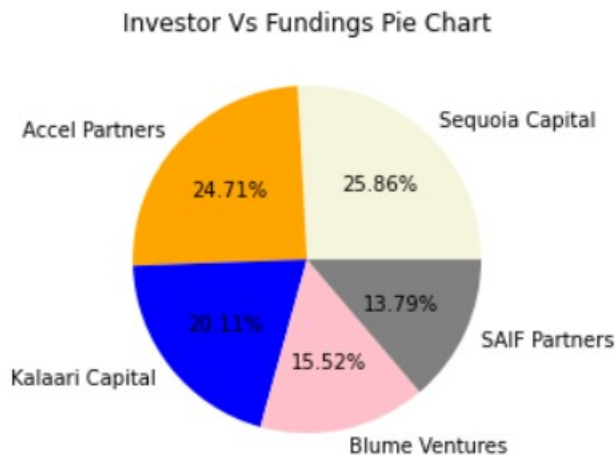
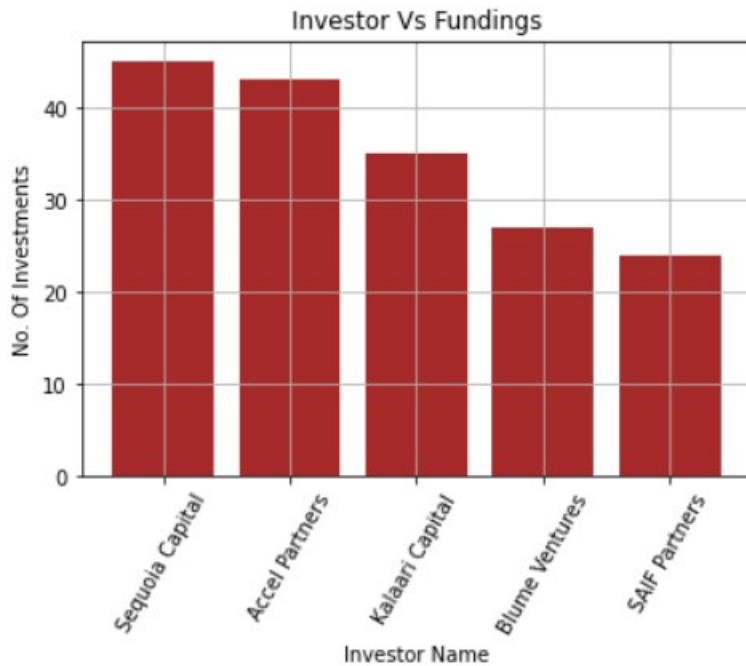
y = investor_count
x = investor_name
```

```
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 C
+ < > Run Code
# plotting the bar graph for top 5 investors
plt.bar(x,y, color = 'brown')
plt.xlabel("Investor Name")
plt.ylabel("No. Of Investments")
plt.title("Investor Vs Fundings")
plt.grid()
plt.xticks(rotation = 60)
plt.show()

# plotting the pie chart b/w top 5 investors and number of fundings.
color = ["beige", "orange", "blue", "pink", "grey"]
plt.pie(y, labels = x, colors = color, autopct = "%.2f%%")
plt.title("Investor Vs Fundings Pie Chart")
plt.show()

Sequoia Capital 45
Accel Partners 43
Kalaari Capital 35
Blume Ventures 27
SAIF Partners 24
```

GRAPHS:-



RESULT AND JUSTIFICATION:-

As the result of the code , we get the improved list of TOP 5 Investors who had invested maximum number of times in different start – up's and invested in Private Equity.
So, I would suggest my Friend to contact these 5 INVESTORS who can probably invest in

his start-up and thereby
increasing the chances of getting investments to expand his start-up and hence get funding
in Growth stage.