

## Answer 1:

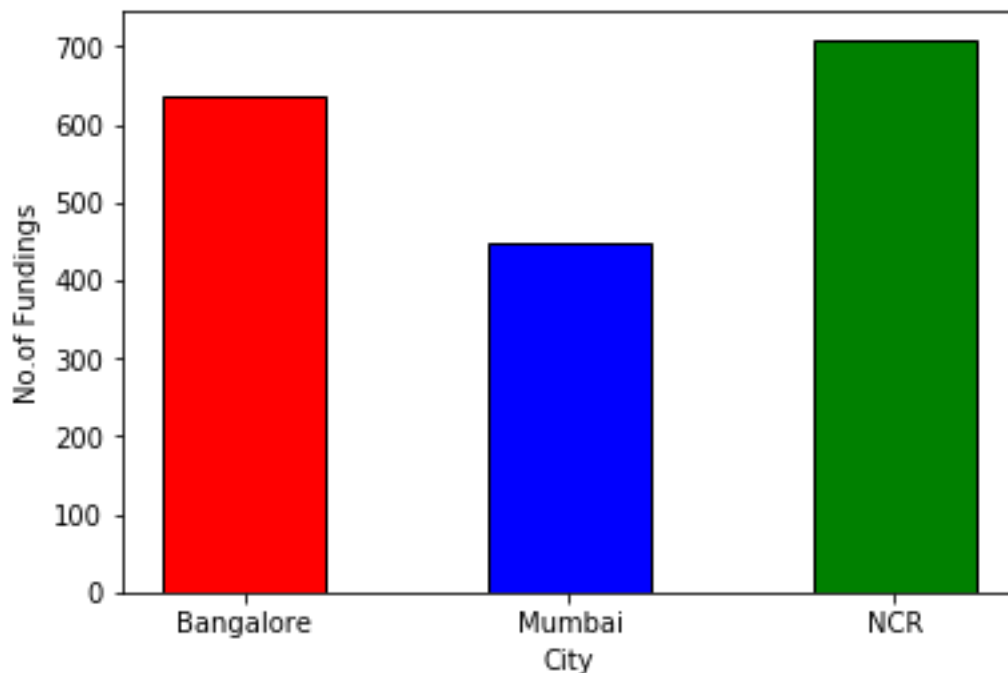
Bangalore

City

Bangalore 637

Mumbai 449

NCR 709



## Justification:

As we AIM to find the location in cities including Bangalore, Mumbai, and NCR(Gurgaon, Noida, New Delhi) where startups has received funding maximum number of times.

So, according to the result **NCR** received funding maximum number of times with count of **709**.

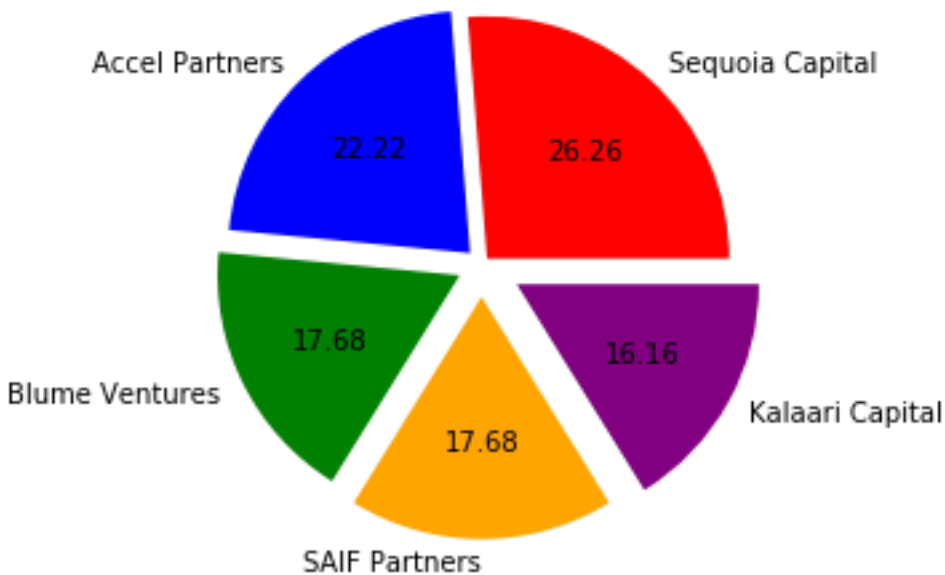
As count of funding received by given cities is as followed:

**Bangalore 637    Mumbai 449    NCR 709**

- In the code, firstly csv library is used to read the content line by line so that a **new list of city** is created after '/' has been removed by collecting the city locations(keeping in mind the condition i.e. if any city in either side of '/' is among Mumbai, Bangalore, Gurgaon, Noida and New Delhi,It should be considered).
- Then using pandas library, data frame is created of given csv file and the city list is now created as **Column City** in data frame.
- All the **data is cleaned** in the start only as mentioned in the questions for startup Name, City Name, Investment type,Date and Industry vertical.
- Now, **df\_loc data frame** is created which is data frame of locations including Mumbai, Bangaloreand NCR (Gurgaon, Noida and New Delhi).
- In df\_loc, the **cities is counted** by groupby() to know the start ups count in 3 selected locations .
- The maximum count out of 3 locations is found and the **id of maximum count(city name) is printed**.
- At the end, **printed and plotted bar graph** for cities vs count.

## Answer 2:

Sequoia Capital 52  
Accel Partners 44  
Blume Ventures 35  
SAIF Partners 35  
Kalaari Capital 32



## Justification:

As we AIM to find the top 5 investors who have invested maximum number of times at location in cities including Bangalore, Mumbai, and NCR(Gurgaon, Noida, NewDelhi).

So, according to the result **Sequoia Capital** invested maximum number of times with count of **52** and percentage of **26.26** among top 5 investors.

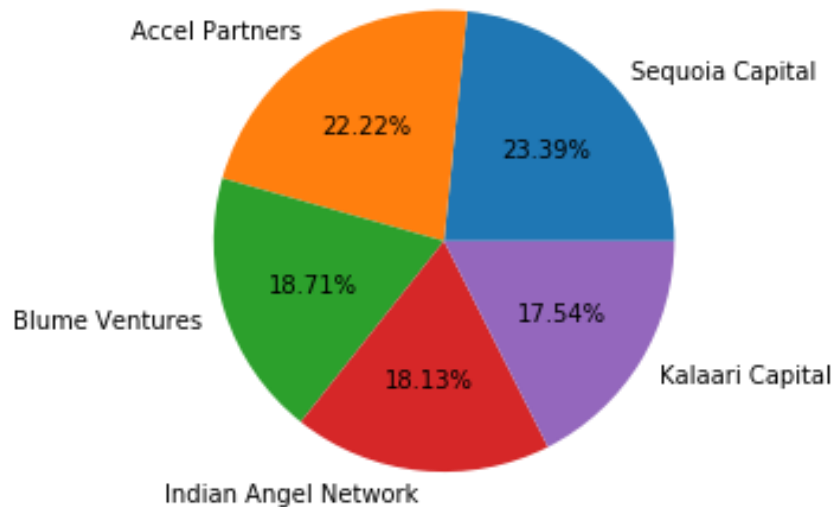
Also, Top 5 investors are as followed:

**Sequoia Capital 52   Accel Partners 44   Blume Ventures 35**  
**SAIF Partners 35   Kalaari Capital 32**

- In this question, **df\_loc data frame** is used and filled nan in investors with 'N/A'.
- In the df\_loc of investors, all the **investors are splitted** by ',' for each sub list in it and got a **list 'a'** for each sub list.
- With that list 'a', a **dictionary dict\_inv** of all the investors in df\_loc data frame is created with key as investor name and value as their count of occurances.
- **Extracted top 5 investors** from the dictionary dict\_inv using heapq nlargest() function.
- At the end, **printed and plotted pie chart** for the same.

### Answer 3:

Sequoia Capital 40  
Accel Partners 38  
Blume Ventures 32  
Indian Angel Network 31  
Kalaari Capital 30



### Justification:

As we AIM to find the top 5 investors who have invested maximum number of times in different companies at location in cities including Bangalore, Mumbai, and NCR(Gurgaon, Noida, New Delhi).

So, according to the result **Sequoia Capital** invested maximum number of times in different companies with count of **40** and percentage of **23.39** among top 5 investors.

Also, Top 5 investors are as followed:

**Sequoia Capital 40   Accel Partners 38   Blume Ventures 32**  
**Indian Angel Network 31   Kalaari Capital 30**

- In this question, **df\_loc data frame** is used and filled nan in investors with 'N/A'.
- Now, a **dictionary dict\_st** with Startup Name as key and list of investors as value is made.
- In that, using **iloc the data is extracted** (row by row).
- For value(investors name), firstly investors name got **split** by ',' and the **names are appended in dictionary dict\_st** one by one into a list.
- After that, **dictionary of investors dict\_investors** using dict\_st is created.
- In this, firstly, a loop for **each key in dictionary dict\_st** and then a nested loop the **set of investors** is created.
- The **dictionary dict\_investors** which have list of investors as key and value of unique startups to which they invested is created.
- **Sorted the dictionary dict\_investors** using sorted() function and **extracted the top 5 investors**.
- At the end, printed **and plotted pie chart** for the same.

## Answer 4:

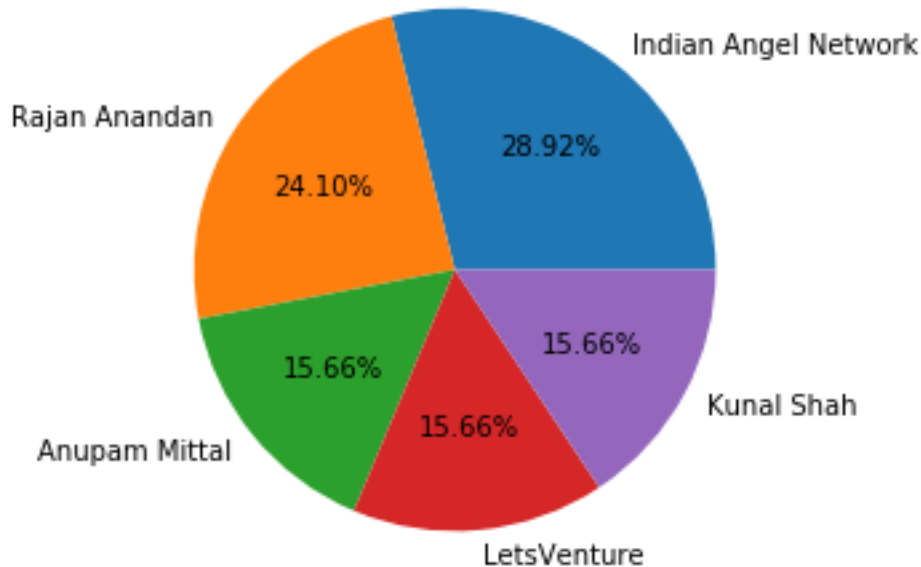
Indian Angel Network 24

Rajan Anandan 20

Kunal Shah 13

Anupam Mittal 13

LetsVenture 13



## Justification:

As we AIM to find the top 5 investors who have invested in a different number of startups and their investment type is Crowdfunding or Seed Funding at location in cities including Bangalore, Mumbai, and NCR(Gurgaon, Noida, New Delhi). So, according to the result **Indian Angel Network** invested maximum number of times in different companies having investment type Crowdfunding or Seed Funding with count of **24** and percentage of **28.92** among top 5 investors.

Also, Top 5 investors are as followed:

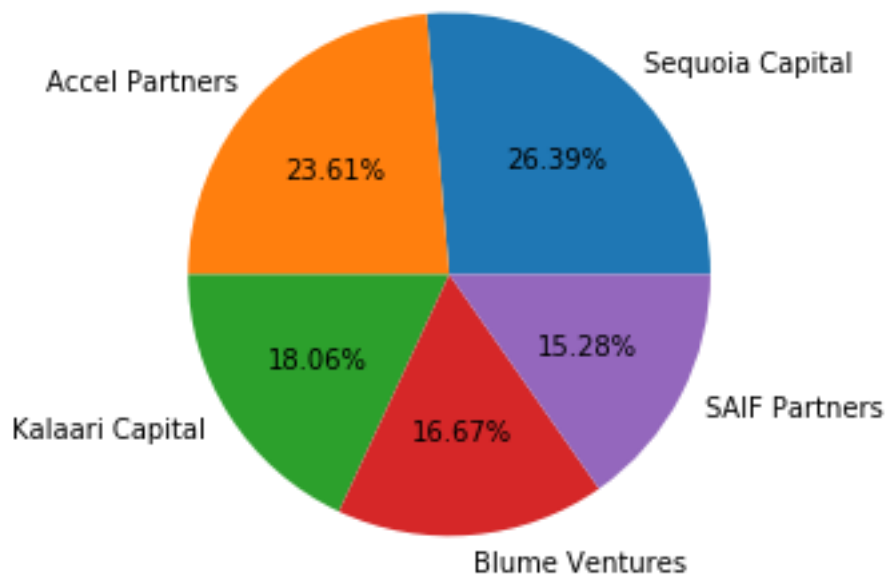
**Indian Angel Network 24      Rajan Anandan 20      Kunal Shah 13**  
**Anupam Mittal 13      LetsVenture 13**

- In this question, df\_loc data frame is used with restricted Investment type to Seed Funding and Crowd Funding and stored it in a new **data frame CSFunding**.
- In this question, **df\_loc data frame** is used and filled nan in investors with 'N/A'.
- Now, a **dictionary dict\_st** with Startup Name as key and list of investors as value is made.
- In that, using **iloc the data is extracted** (row by row).
- For value(investors name), firstly investors name got **splitted** by ',' and the **names are appended in dictionary dict\_st** one by one into a list.
- After that, **dictionary of investors dict\_investors** using dict\_st is created.
- In this, firstly, a loop for **each key in dictionary dict\_st** and then a nested loop the **set of investors** is created.
- The **dictionary dict\_investors** which have list of investors as key and value of unique startups to which they invested is created.
- **Sorted the dictionary dict\_investors** using sorted() function and **extracted the top 5 investors**.
- At the end, printed **and plotted pie chart** for the same.



## Answer 5:

Sequoia Capital 38  
Accel Partners 34  
Kalaari Capital 26  
Blume Ventures 24  
SAIF Partners 22



## Justification:

As we AIM to find the top 5 investors who have invested in a different number of startups and their investment type is Private Equity at location in cities including Bangalore, Mumbai, and NCR(Gurgaon, Noida, New Delhi).

So, according to the result **Sequoia Capital** invested maximum number of times in different companies having investment type Private Equity with count of **38** and percentage of **26.39** among top 5 investors.

Also, Top 5 investors are as followed:

**Sequoia Capital 38 Accel Partners 34 Kalaari Capital 26**  
**Blume Ventures 24 SAIF Partners 22**

- In this question, df\_loc data frame is used with restricted Investment type to Private Equity and stored it in a new **data frame PEFunding**.
- In this question, **df\_loc data frame** is used and filled nan in investors with 'N/A'.
- Now, a **dictionary dict\_st** with Startup Name as key and list of investors as value is made.
- In that, using **iloc the data is extracted** (row by row).
- For value(investors name), firstly investors name got **splitted** by ',' and the **names are appended in dictionary dict\_st** one by one into a list.
- After that, **dictionary of investors dict\_investors** using dict\_st is created.
- In this, firstly, a loop for **each key in dictionary dict\_st** and then a nested loop the **set of investors** is created.
- The **dictionary dict\_investors** which have list of investors as key and value of unique startups to which they invested is created.
- **Sorted the dictionary dict\_investors** using sorted() function and **extracted the top 5 investors**.
- At the end, printed **and plotted pie chart** for the same.