# Project : Case Study (Part - II)

## 1. To Find Startup Location:

**Problem Statement:**

Help a friend to decide the startup location among the 5 cities.
Modify the city names in column
Count the number of funding each five cities have received.
Plot the bar graph for each city and number of fundings it has received.

**Explanation:**

Firstly, I have imported the libraries required for the above tasks. Then I have created the dataframe of the csv file. From the dataframe all the NaN entries from the CityLocation column have been removed.

From the insights that have been provided in the problem statement, all the city names have been modified and then I have defined the function to look upon the startups that have multiple locations and have at least one Indian location. I splitted those startup locations and if one of them is having the desired location then that has been counted for the output.
Inside this function only I have stored the preferred city and counted the corresponding number of funding it has received and stored inside the dictionary. Later, I applied this tailored function on the dataframe.

Then, all the cities and corresponding number of funding have been segregated in the separate list which later will be used to plot the graph.All the cities and number of funding it has received have been printed.

At the last stage of the code, the Bar graph has been plotted using the city on X axis and number of funding on Y axis.

**Result Statement:**

From the output of the code it is inferred that Bangalore has received the highest Number of funding(637) from the investors so Bangalore is the preferred location to establish the startup.

# Project : Case Study (Part - II)

## 2. To Find Top Five Investors:

**Problem Statements:**

Separate if multiple investors have invested in one startup.
Ignore the undisclosed Investors
Count the number of times each investor has invested.
Print the Top Five investor's name

**Explanation:**

First all the necessary libraries have been imported. Then the csv data is stored in the dataframe. All the NaN entries in the InvestorsName column have been dropped.

Then, I have iterated the loop over the InvestorsName column and converted the series in the string format and splitted the name if having multiple names in one startup. Later each separate name has been stored in the list.

Later, using the dictionary I counted the frequency of the name stored in the list. Then, the dictionary has been sorted using the lambda function in the descending order so that the name of the investor with the highest frequency comes first.

After that I iterated over the dictionary and stored each name in the investors list and number of investments in the no_of_investments list. Later I executed the for loop to print the first five names.

**Result Statement:**

Output of the code suggests that following are the top five investors:

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

So the startup founder should approach the above five investors directly to fill the fuel in his/her own startup.

# Project : Case Study (Part - II)

## 3. Top 5 Investors in Different Startups:

**Problem Statements:**

Find top 5 Investors.

Count only once if they have invested in one startup multiple times.

**Explanation :**

Firstly I have imported the necessary libraries. Then I loaded the csv dataset and converted it into a dataframe and dropped the NaN entries from the Startup_Name, Investors_Name column.

Then I checked the startup name errors and corrected all the errors and reflected in the dataframe. After correction I stored the startup name and investor name for that startup in a separate list.

Then I created the empty dictionary to store the investor name as a key and the number of companies they have invested in as the set to remove duplicate entries as a value in the dictionary by splitting the multiple names.After that I have stored the keys of the dictionary that is the name of investors inside numpy array and created another array to store the length of the values of dictionary i.e sets for startup names.

Then I sorted the array that holds the length of sets and later reversed the array to get the maximum number of invested companies at top. Based on this array I found the indices of max number and from that I got the name of investors having the highest number of investment counts. Then at last iterated over the names to get the top 5 Investors.

**Result Statement:**

Below are the top five investors who have maximum investment count and have invested in different startups. So founder should approach these investors.

Top 5 Investors are as below ::

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

# Project : Case Study (Part - II)

## 4. Top 5 Investors in Seed Funding & Crowd Funding:

**Problem Statement :**

Find top 5 investors who have invested as Seed Funding or Crowd Funding.

**Explanation :**

Firstly I have imported the necessary libraries. Then I loaded the csv dataset and converted it into a dataframe and dropped the NaN entries from the Startup_Name, Investors_Name and Investment_Type column.

Then I checked the startup name errors and investment type spelling errors, corrected all the errors and reflected in the dataframe. After correction I filtered the entries having investment type Seed Funding or Crowd Funding.

Then I created the empty dictionary to store the investor name as a key and the number of companies they have invested in as the set to remove duplicate entries as a value in the dictionary by splitting the multiple names.After that I have stored the keys of the dictionary that is the name of investors inside numpy array and created another array to store the length of the values of dictionary i.e sets for startup names.During this stage all InvestmentType column entries having empty value or Undisclosed Investors have been neglected.

Then I sorted the array that holds the length of sets and later reversed the array to get the maximum number of invested companies at top. Based on this array I found the indices of max number and from that I got the name of investors having the highest number of investment counts and investment type as Seed Funding or Crowd Funding. Then at last iterated over the names to get the top 5 Investors with Seed Funding or Crowd Funding.

**Result Statement:**

Below are the names of the Top 5 Investors who have invested in different startups as Seed Funding or Crowd Funding. So ideally Startup Founder should approach these founders to get the investment in the form of Seed Funding or Crowd Funding.

| | |
|---|---|
| Indian Angel Network | 33 |
| Rajan Anandan | 23 |
| LetsVenture | 16 |
| Anupam Mittal | 16 |
| Group of Angel Investors | 14 |

# Project : Case Study (Part - II)

## 5. Top 5 Investors with Private Equity:

**Problem Statements:**

Find the top 5 investors who have invested in multiple companies in Private Equity Form.

**Code Explanation:**

Firstly I have imported the necessary libraries. Then I loaded the csv dataset and converted it into a dataframe and dropped the NaN entries from the Startup_Name, Investors_Name and Investment_Type column.

Then I checked the startup name errors and investment type spelling errors, corrected all the errors and reflected in the dataframe. After correction I filtered the entries having investment type Private Equity.

Then I created the empty dictionary to store the investor name as a key and the number of companies they have invested in as the set to remove duplicate entries as a value in the dictionary by splitting the multiple names.After that I have stored the keys of the dictionary that is the name of investors inside numpy array and created another array to store the length of the values of dictionary i.e sets for startup names.During this stage all InvestmentType column entries having empty value or Undisclosed Investors have been neglected.

Then I sorted the array that holds the length of sets and later reversed the array to get the maximum number of invested companies at top. Based on this array I found the indices of max number and from that I got the name of investors having the highest number of investment counts and investment type as Private Equity. Then at last iterated over the names to get the top 5 Investors with Private Equity.

**Result Statement:**

Below are the top 5 Investors who have topped the table for the Private Equity investment category. As the startup is in the initial stage, the founder should go for Private Equity investment and therefore approach these investors directly for funding.

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