Importing necessary libraries, we need.

- Pandas
- NumPy
- Matplotlib

Opening the file using the pd_csv function and then creating a copy ‘df’ of the data ,So that the original data doesn’t get affected.

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 decide the location. NCR includes Gurgaon, Noida and New Delhi. Find the location where the most number of funding is done. That means, find the location where startups have received funding maximum number of times. Plot the bar graph between location and number of funding. Take the 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 the city name as "Bangalore". For a few startups multiple locations are given, one Indian and one Foreign. Consider the startup if any one of the cities lies in given locations.

- Firstly, I got rid off the nan values by giving the command

  ```
  #df= df[df['CityLocation'].notna()]
  ```

- And then resetting the Index’s
- Then I replaced the Cities where Wrong name were given and also Cities with lower Case-Sensitiveness’ i.e :"Delhi" as "New Delhi” & "bangalore" as "Bangalore”

- Then I run a for loop for range of Len of City Location to in order to transverse through all Locations and values which Contained cities with one foreign Location were replaced with Indian city mentioned in the Question.
- Next, only those locations were taken which were mentioned in the question using the `df.loc` function.
- Using the values_count() function we got the cities which got funded maximum times.

Interpretation: The result shows that Banglore has the maximum number of fundings for a new startup . So it would be better to set up your startup in Banglore to increase the chances of getting funding.

---

2. Even after trying 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.

- Deleted nan values in Investors Name Column using

  `#df = df[df['Investors Name'].notna()]` & then resetting the Index.

- Some values contained multiple values in it separated by ', '.
- So, In order to take them into consideration I created a function which creates a dictionary in which I split the values which contained ',' and added them to the dictionary, where key is the Investor Name and value is the number of times he invested.
- Then I converted the dictionary into a data frame by using `pd.Dataframe`
- then sorted the values and saved them in descending order in df2
- Slicing them to get the first five values & reshaping Y variable and printing first 5 values.

Interpretation:The result shows that Sequoia Capital has invested maximum times followed by Accel Partners, Kalaari Capital, SAIF Partners ,Indian Angel Network.So the chances of getting funded are more if my friend goes to these companies showing them his startup plan,

---

3. After re-analyzing the dataset you found out that some investors have invested in the same startup at different number of funding rounds. So before finalizing the previous list, you want to improvise it by finding the top 5 investors who have invested in a different number of startups. This list will be more helpful than your previous list in finding the investment for your friend's startup. Find the top 5 investors who have invested a 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.

- dropping rows having Investors and startup names as nan..
- Then, replacing the wrong word with the correct word.. ex: OYO as Oyo etc..
- Firstly created a dictionary for each investor names and maintained a set to counter duplicity meaning that each key (investor's name) having a value set (names of startup’s in which they invested)
- set is taken as a value to avoid count of multiple investment in a single startup by an investor
- In the set , there are startup names in which investor's had already invested . In case there are multiple investors for a single startup so here I used split function to split that and traversed through each name separately as asked in the question.
- Then created a dictionary where key is investor's name and value is count of startup's in which they had invested
- Then sorting the keys according to the values in descending Order and printing the top 5 Investors name.

Interpretation:The result shows that Sequoia Capital has the maximum unique investment followed by Accel Partners, Kalaari Capital,Indian Angel Network,Blume Ventures .So the chances of getting funded are more if my friend goes to these companies showing them his startup plan/Idea.

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's 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 mistakes. 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.

- removing row's having nan's in Investors name column
- Checking and correcting spelling checks of investment types
- replacing the undisclosed investors name by null values
- keeping only those rows having investment type seed funding and crowdfunding by using
  `df = df[(df["InvestmentType"] == "Seed Funding") | (df["InvestmentType"] == "Crowd Funding")]`

- Firstly created a dictionary for each investor names and maintained a set to counter duplicity meaning that each key (investor's name) having a value set (names of startups in which they invested)
- set is taken as a value to avoid count of multiple investment in a single startup by an investor. In the set , there are startup names in which investor's had already invested, in case there are multiple investors for a single startup so here used split function to

split that and traversed through each name separately as asked in the question.

- created a dictionary where key is investor's name and value is count of startup's in which they had invested
- sorting the keys according to there values in descending order..and taking the top 5 investors from the group

Interpretation:Since my friend's startup is in its initial phase So its better to seek funding in Investment type Seed Funding and Crowdfunding.So, the result shows investors who have invested maximum times in different startups where investment type is either Seed funding or Crowd funding.

5.Due to your immense help, your friend startup successfully got seed funding and it is in 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 investors. Before moving forward you remember your investor friend's advice to find the investors by analysing the investment type. Since your friend's startup is not in the early phase it is in a 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 mistakes. 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.

- removing row's having nan's in Investors name column
- Checking and correcting spelling checks of investment types
- replacing the undisclosed investors name by null values
- keeping only those rows having investment type seed funding and crowdfunding by using

```
df = df[(df["InvestmentType"] == "Private Equity")]
```

- Firstly created a dictionary for each investor names and maintained a set to counter duplicity meaning that each key (investor's name) having a value set (names of startups in which they invested)
- set is taken as a value to avoid count of multiple investment in a single startup by an investor. In the set , there are startup names in which investor's had already invested, in case there are multiple investors for a single startup so here used split function to split that and traversed through each name separately as asked in the question.
- created a dictionary where key is investor's name and value is count of startup's in which they had invested
- sorting the keys according to there values in descending order..and taking the top 5 investors from the group

Interpretation:Since my friend's startup is not in the early phase it is in growth stage so the best-suited investment type is Private Equity for the growth of the start up .So, the result shows investors who have invested maximum times in different startups where Investment type is Private Equity.