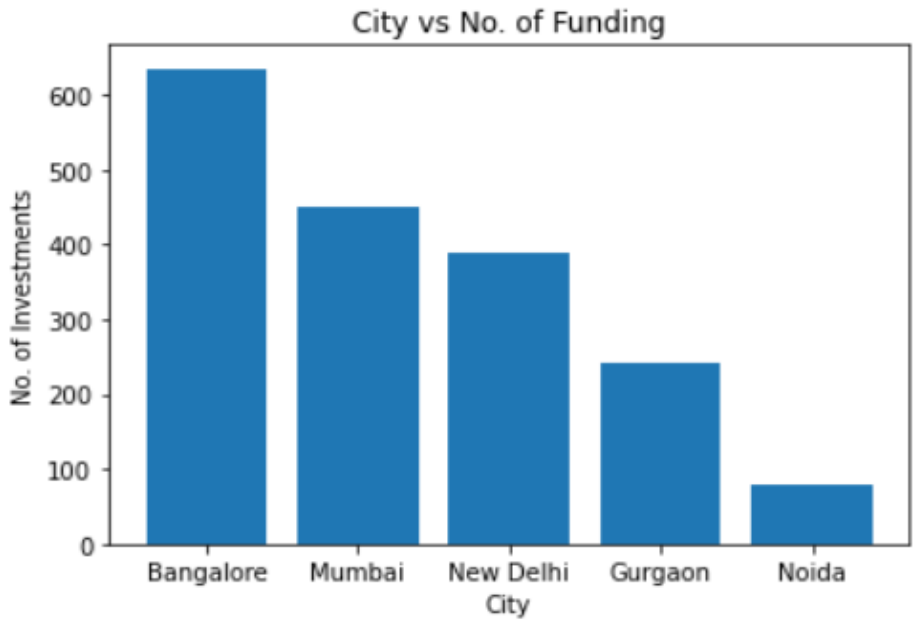
CASE STUDY – 2

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.

ANSWER: -

```
Bangalore 635
Mumbai 449
New Delhi 389
Gurgaon 241
Noida 79
Max funding is done in Bangalore , 635 times.
```



EXPLANATION: -

The code cleans the data by dropping rows with missing city locations, standardizes city names, and filters the Data Frame to include only specific cities. It then counts the occurrences of each city and prints the city name and its count. Finally, it identifies the city with the highest count, representing the location where the maximum funding has been done.

The output shows the count of funding occurrences in different cities. Here's the interpretation:

- Bangalore: 635 times
- Mumbai: 449 times
- New Delhi: 389 times
- Gurgaon: 241 times
- Noida: 79 times

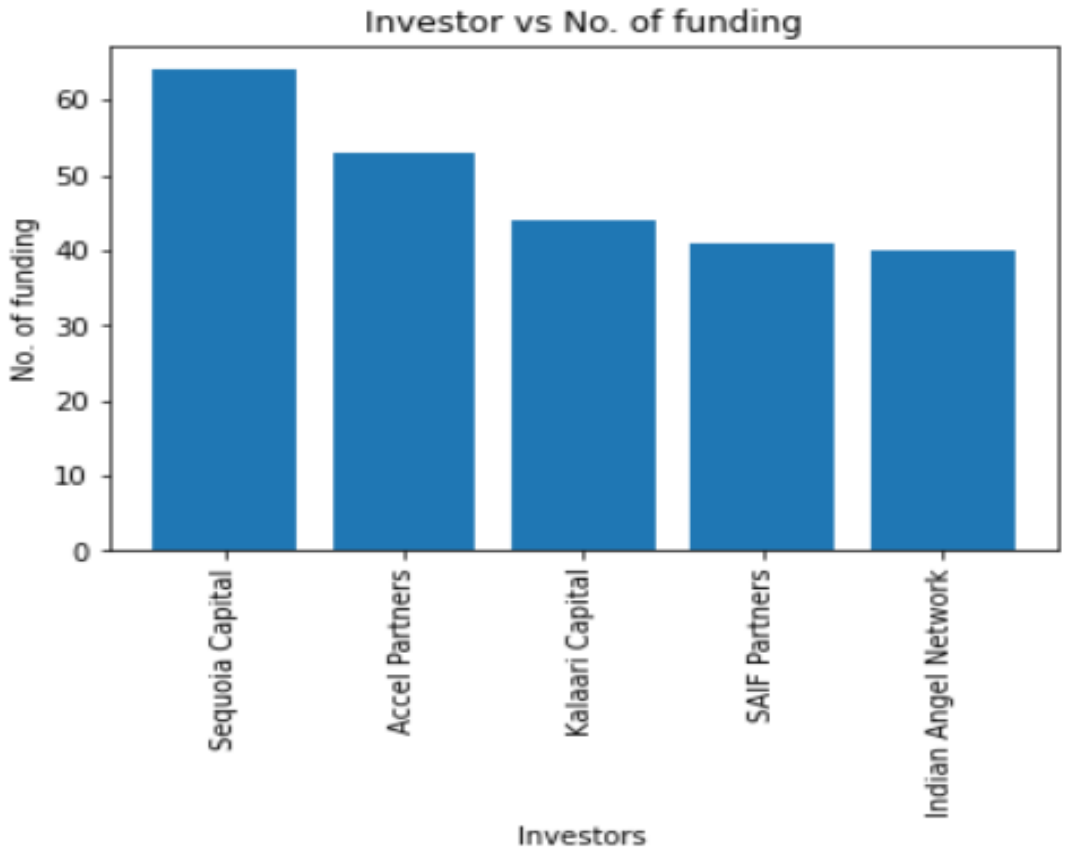
The city with the highest count is Bangalore, which had funding occur 635 times. This means that Bangalore had the highest number of funding events among the cities in the dataset.

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.

ANSWER: -

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



EXPLANATION: -

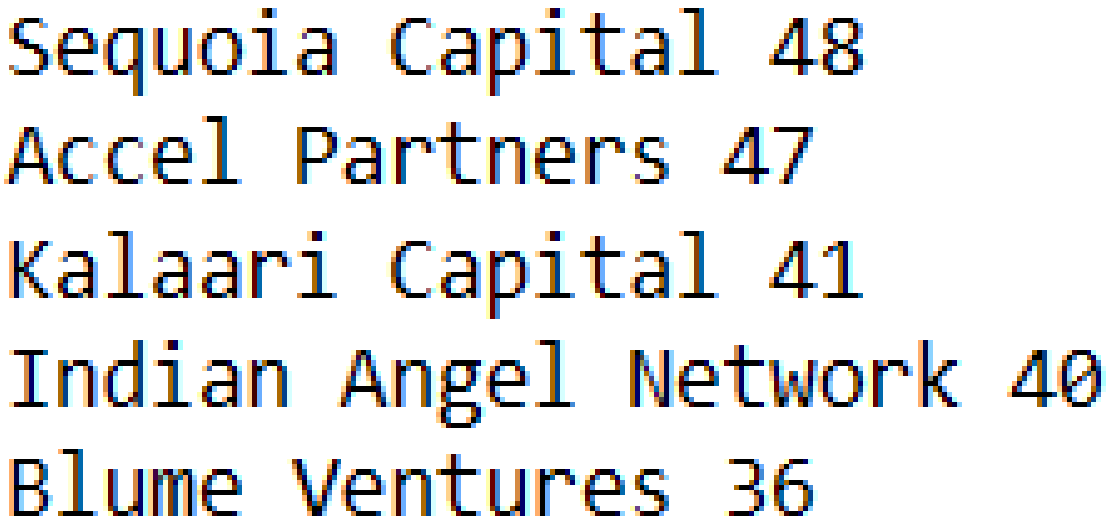
The code performs the following tasks:

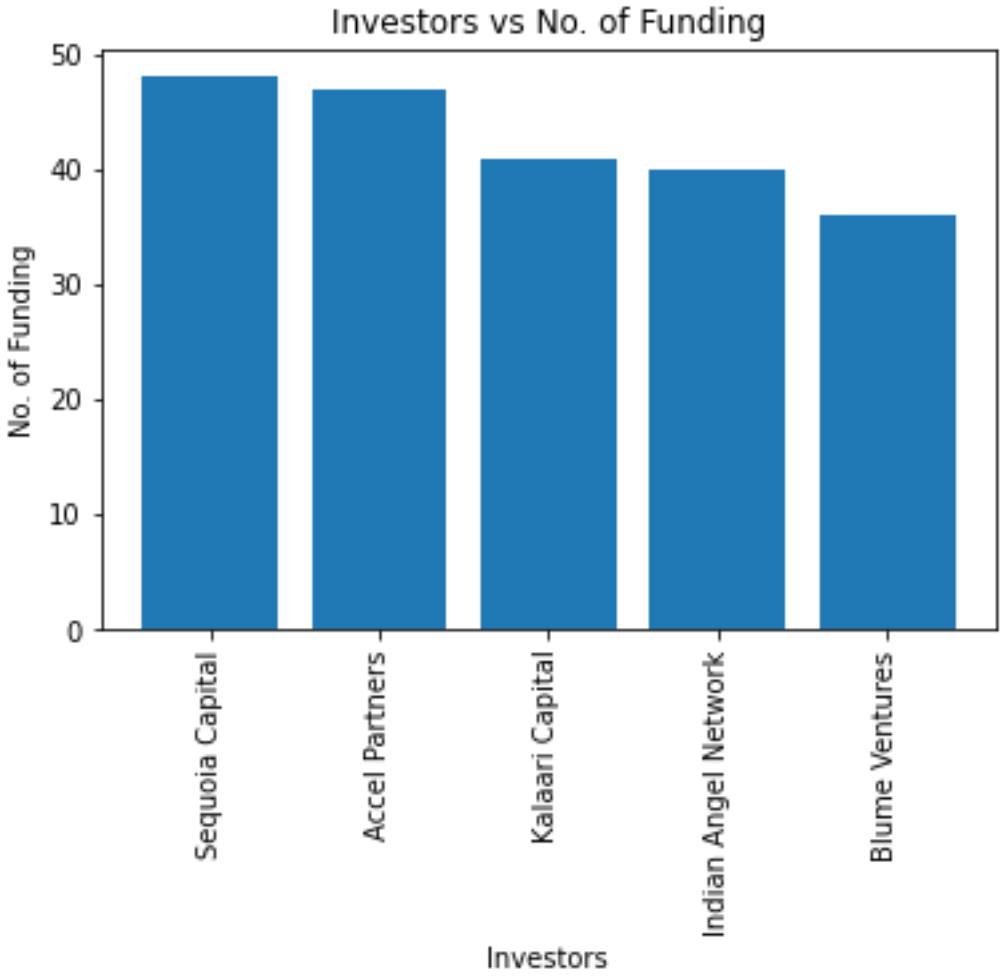
1. Drops rows with missing values in the 'InvestorsName' column.
2. Initializes an empty dictionary to store investor names and their counts.
3. Defines a function that splits a comma-separated string of investor names, strips whitespace, and updates the count in the dictionary.
4. Applies the function to each element in the 'InvestorsName' column, populating the dictionary.
5. Converts the dictionary keys (investor names) and values (counts) into NumPy arrays.
6. Sorts the count values in descending order and retrieves the corresponding indices.
7. Selects the top 5 investor names and their counts based on the sorted indices.
8. Prints the names and counts of the top 5 investors.

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.

ANSWER: -





EXPLANATION: -

The code performs the following tasks:

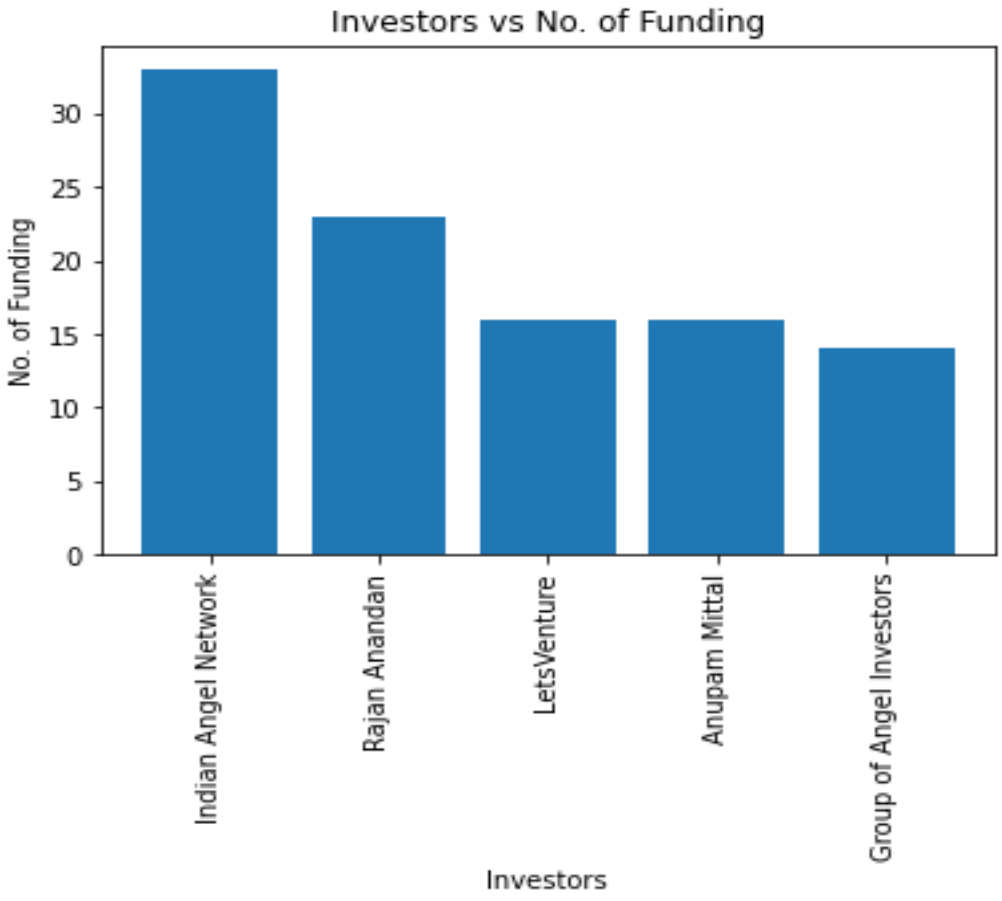
1. Drops rows with missing values in the 'StartupName' and 'InvestorsName' columns.
2. Replaces variations of startup names with standardized versions using the replace() method.
3. Initializes an empty dictionary d to store investors and their associated startups.
4. Iterates over each investor and startup pair from the respective lists.
5. Splits the investor names by comma and strips any leading or trailing whitespace.
6. Checks if the investor name is not empty or "Undisclosed Investors".
7. Updates the dictionary d with the investor and the associated startup.
8. Calculates the number of startups associated with each investor in the dictionary.
9. Converts the dictionary keys (investors) and values (counts) into NumPy arrays.
10. Sorts the count values in descending order and retrieves the corresponding indices.
11. Selects the top 5 investors and their counts based on the sorted indices.
12. Prints the names of the top 5 investors along with the corresponding counts.

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.

ANSWER: -

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



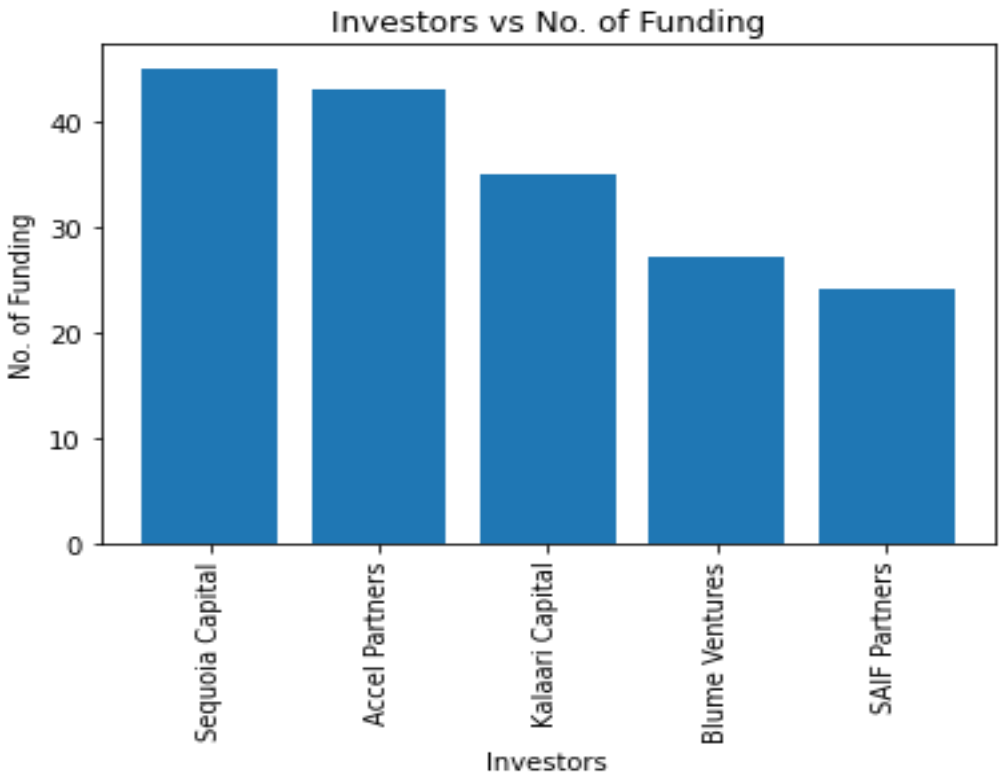
EXPLANATION: -

1. Drops rows with missing values in the 'InvestorsName', 'InvestmentType', and 'StartupName' columns.
2. Replaces variations of investment types with standardized versions using the replace() method.
3. Replaces variations of startup names with standardized versions using the replace() method.
4. Filters the DataFrame to include only rows where the 'InvestmentType' is either 'Seed Funding' or 'Crowd Funding'.
5. Initializes an empty dictionary d to store investors and their associated startups.
6. Iterates over each investor and startup pair from the respective lists.
7. Splits the investor names by comma and strips any leading or trailing whitespace.
8. Checks if the investor name is not empty or "Undisclosed Investors".
9. Updates the dictionary d with the investor and the associated startup.
10. Calculates the number of startups associated with each investor in the dictionary.
11. Converts the dictionary keys (investors) and values (counts) into NumPy arrays.
12. Sorts the count values in descending order and retrieves the corresponding indices.
13. Selects the top 5 investors and their counts based on the sorted indices.
14. Prints the names of the top 5 investors along with the corresponding counts.

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.

ANSWER: -



EXPLANATION: -

The code performs the following tasks:

1. Replaces variations of investment types with standardized versions using the replace() method.
2. Replaces variations of startup names with standardized versions using the replace() method.
3. Drops rows with missing values in the 'InvestorsName', 'StartupName', and 'InvestmentType' columns.
4. Filters the DataFrame to include only rows where the 'InvestmentType' is 'Private Equity'.
5. Initializes an empty dictionary d to store investors and their associated startups.
6. Iterates over each investor and startup pair from the respective lists.
7. Splits the investor names by comma and strips any leading or trailing whitespace.
8. Checks if the investor name is not empty and not equal to "Undisclosed Investors".
9. Updates the dictionary d with the investor and the associated startup.
10. Calculates the number of startups associated with each investor in the dictionary.
11. Converts the dictionary keys (investors) and values (counts) into NumPy arrays.
12. Sorts the count values in descending order and retrieves the corresponding indices.
13. Selects the top 5 investors and their counts based on the sorted indices.
14. Prints the names of the top 5 investors along with the corresponding counts.