# Justification and plots

Answer 1:

In this we have to find the top five locations in India that get maximum funding.
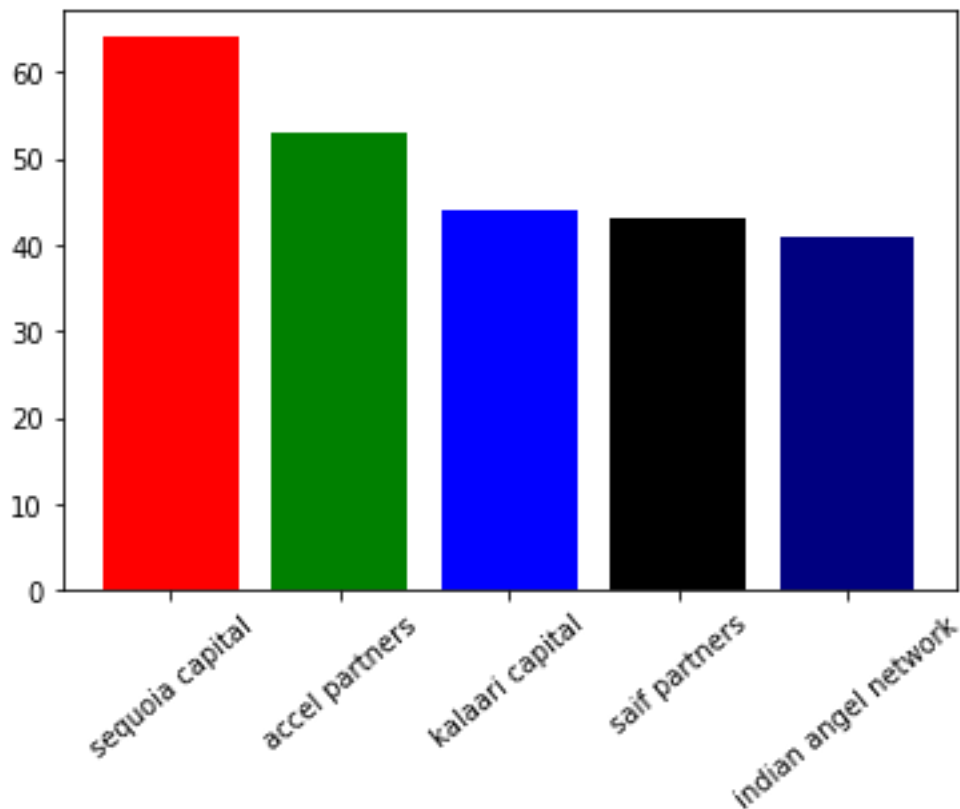
1. Read the concerned file and store it in a dataframe.
2. Then remove the entries that have nan values in the city location column.
3. Then write a function so as to split the cities that are mentioned with foreign cities and include them also with our count of cities. We applied the function for all the records in CityLocation column.
4. Then get the count of each city in the CityLocation column.
5. Add the count of "Delhi " and "New Delhi" taken as "New Delhi" and "bangalore" and "Bangalore" taken as "Bangalore".
6. Also take "New Delhi", "Gurgaon" and "Noida" as "Ncr".
7. Sort the result in descending order and then take the first three entries and display them as a bar graph.
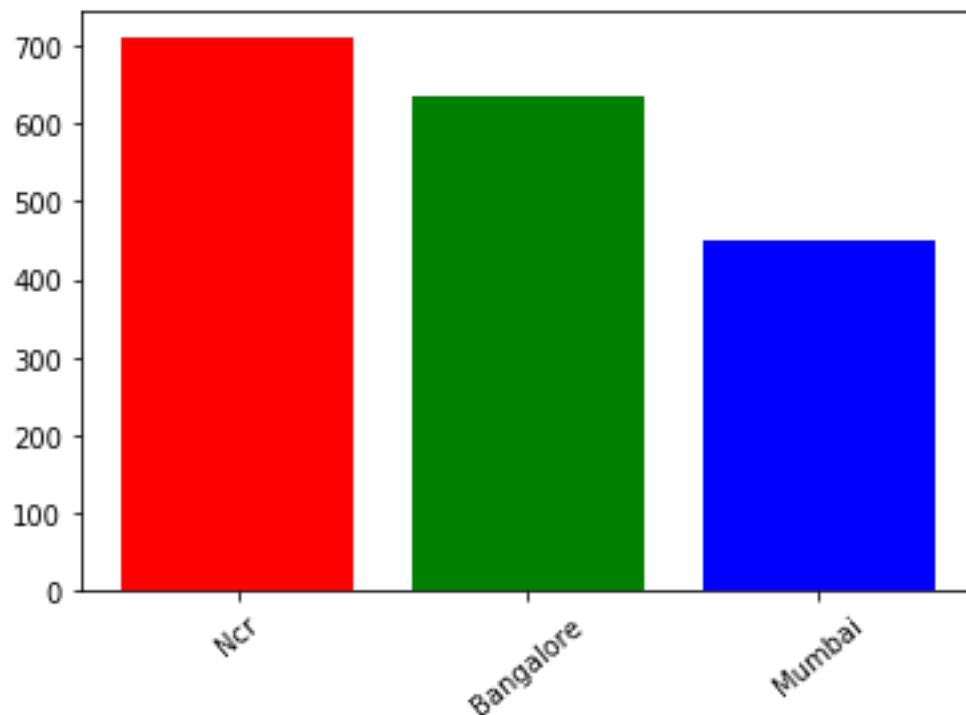
Answer 2

In this we have to find the top five investors who have invested maximum number of times (including the repeat investments in one company).Ignore the undisclosed investors.

1. Read the file "startup_funding.csv" and store the data in a dataframe.
2. Drop the entries that have nan values in the InvestorsName column.
3. Create a function named change() to create a list of investors for each record in InvestorsName column and remove the unnecessary space present.
4.   Change the case of all Investors Name to lower using function tolower function.
5. Create a dictionary and count the occurrence of Investors Name in all the lists present in InvestorsName column.
6. Sort the dictionary in descending order and take the first five entries and convert their name to camel case and display them.
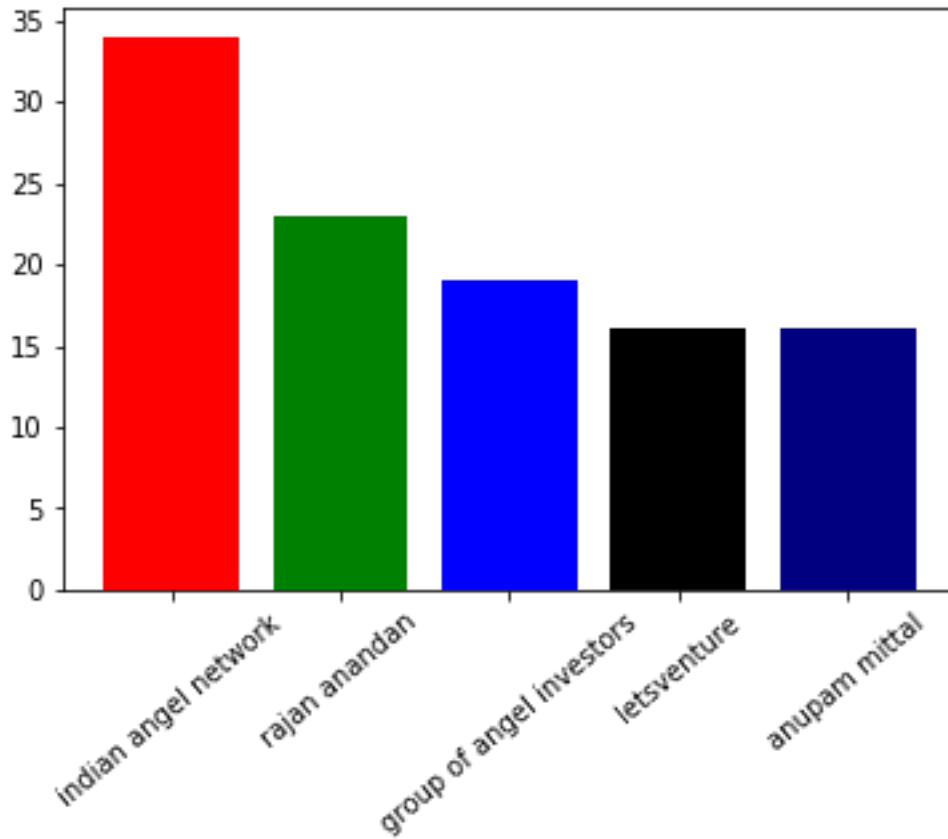
Answer 3

In this we have to display the top five investors that have invested in different startups.

1. Read the file "startup_funding.csv" and store it in a data frame.
2. Drop the records having nan values for StartupName and the entries having nan values in InvestorsName column.
3. Change the case of all startup names to lower and correct the startup names of Ola, Oyo, Paytm and Flipkart.
4. Create a function that returns a list of investors by splitting it at "," and removing the extra space and apply it to all the records in the InvestorsName column.
5. Change the case of InvestorsName column to lower to avoid case sensitivity.
6. Group the values of InvestorsName column according to the StartupName column.
7. After that the grouped result in merged into one list and unique function is applied to take only one occurrence of investor name for each startup.
8. Create a dictionary of InvestorsNames and count the occurrence of each investor name. Ignore the undisclosed investor and empty values.
9. Sort the dictionary in descending order and take the first five entries.
10. Display the result by converting the investor names in camel case.
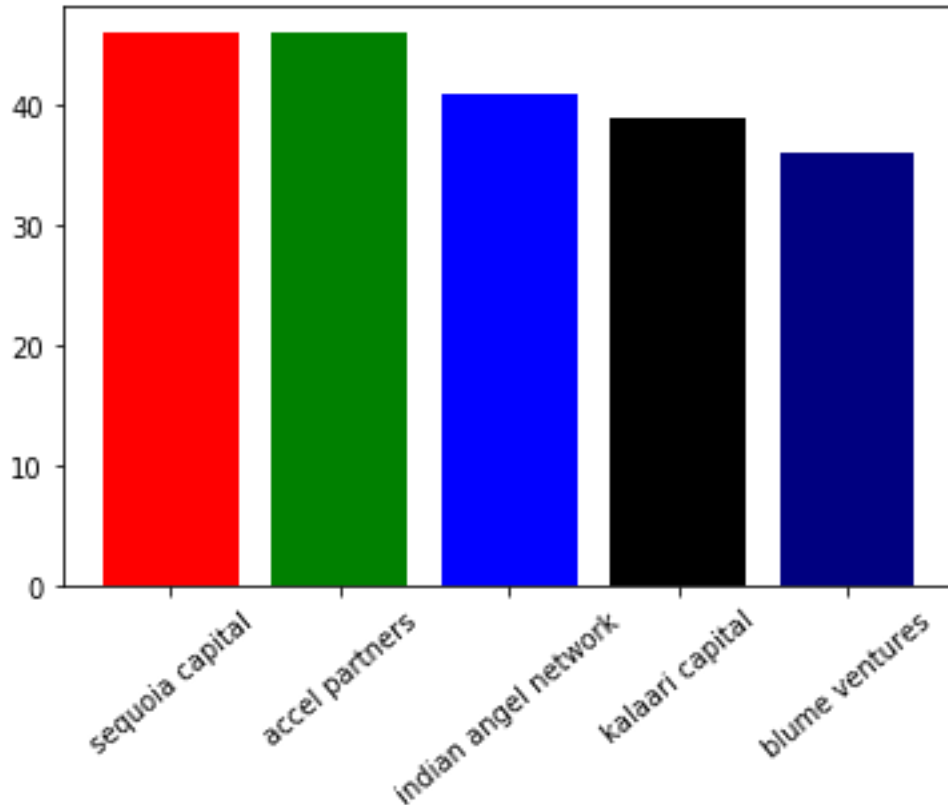
Answer 4

1. Repeat steps 1 to 2 in question 3.
2. Correct the spelling of values in InvestmentType column.
3. Filter the data frame by checking if the investment type is "Seed Funding" or "Crowd Funding".
4. Repeat steps 3 to 10.

Answer 5

1. Repeat steps 1 to 2 in question 3.
2. Correct the spelling of values in InvestmentType column.
3. Filter the data frame by checking if the investment type is "Private Equity".
4. Repeat steps 3 to 10.