

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import csv

file_obj = open("startup_funding.csv")
file_data = csv.DictReader(file_obj, skipinitialspace = True)

d = {}

for row in file_data:
    if row["StartupName"] == "Ola Cabs" or row["StartupName"] == "Olacabs":
        row["StartupName"] = "Ola"
    if row["StartupName"] == "Flipkart.com":
        row["StartupName"] = "Flipkart"
    if row["StartupName"] == "Oyo Rooms" or row["StartupName"] == "OyoRooms" or row["StartupName"] == "Oyorooms" or row["StartupName"] == "OYO Rooms":
        row["StartupName"] = "Oyo"
    if row["StartupName"] == "Paytm Marketplace":
        row["StartupName"] = "Paytm"

    if row["InvestmentType"] == "PrivateEquity":
        row["InvestmentType"] = "Private Equity"

    value = row["InvestorsName"].split(",")
    for i in range(len(value)):
        value[i] = value[i].strip()

    for i in value:
        if (row["StartupName"] != None) and i != "" and i != "Undisclosed Investors" and i != "Undisclosed investors" and (row["InvestmentType"] == "Private Equity" ) :
            if i in d:
                d[i].add(row["StartupName"])
            else:
                d[i] = set()
                d[i].add(row["StartupName"])

for i in d.keys():
    d[i] = len(d[i])

d1 = sorted(d, key = d.get, reverse = True)

investors = d1[0:5]
num_of_companies = []

```

```
top_10_investors = d1[0:10]
number = []

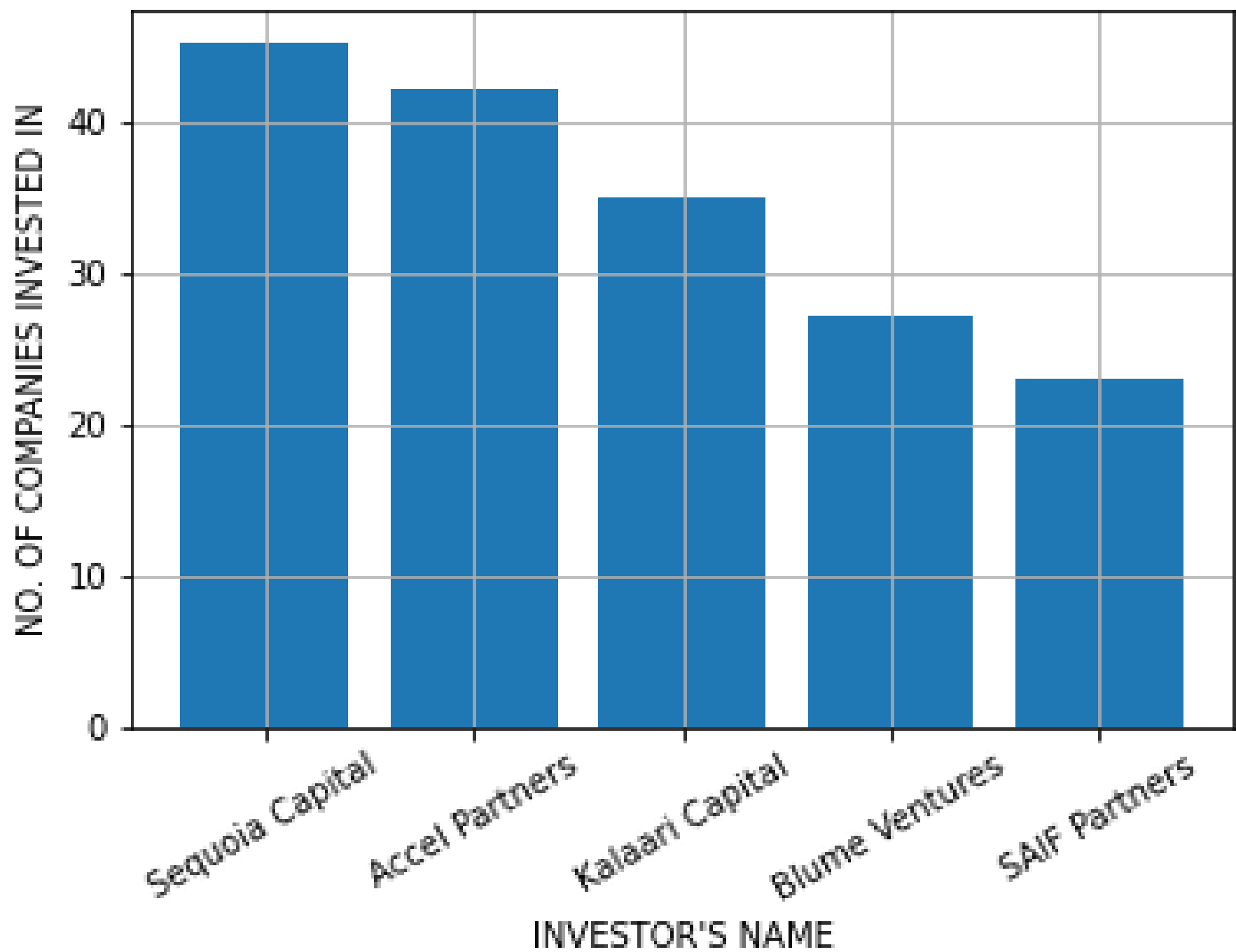
for i in top_10_investors:
    number.append(d[i])

for i in investors:
    num_of_companies.append(d[i])

plt.bar(investors, num_of_companies)
plt.title("Top 5 Investors in Distinct companies(Private Equity Funding)", fontsize = 15)
plt.xlabel("INVESTOR'S NAME",fontsize = 10)
plt.ylabel("NO. OF COMPANIES INVESTED IN",fontsize = 10)
plt.xticks(rotation = 30)
plt.grid()
plt.show()

for i in investors:
    print(i,d[i])
```

Top 5 Investors in Distinct companies(Private Equity Funding)



Sequoia Capital 45
Accel Partners 42
Kalaari Capital 35
Blume Ventures 27
SAIF Partners 23