

In [18]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

In [19]:

```
data = pd.read_csv("Indian_Startup.csv")
```

In [21]:

```
data.head()
```

Out[21]:

	Company/Brand	Founded	Headquarters	Sector	What it does	Founder/s	Investor/s
0	CollegeDekho	2015.0	Gurgaon	E-learning	Collegedekho.com is Student's Partner, Friend ...	Ruchir Arora	Disrupt ADQ, QIC
1	BOX8	2012.0	Mumbai	Food & Beverages	India's Largest Desi Meals Brand	Anshul Gupta, Amit Raj	Tiger Global
2	Simpl	2015.0	Bangalore	Consumer Services	Simpl empowers merchants to build trusted rela...	Nitya, Chaitra Chidanand	Valar Ventures, IA Ventures
3	8i Ventures	2018.0	Mumbai	Venture Capital & Private Equity	8i is a Mumbai & Bangalore based early stage f...	Vikram Chachra	NaN
4	PayGlocal	2021.0	Bangalore	Financial Services	PayGlocal is a FinTech solving for global paym...	Prachi Dharani, Rohit Sukhija, Yogesh Lokhande	Sequoia Capital India

In [22]:

data.tail()

Out[22]:

	Company/Brand	Founded	Headquarters	Sector	What it does	Founder/s	Investor/s
1204	HobSpace	2019.0	Mumbai	EdTech	HobSpace is a platform for extra-curricular ac...	Priya Goel Sheth, Harsh Jain	Siddharth Bhaskar Shah, Upsparks
1205	DcodeAI	2020.0	New Delhi	EdTech	EMPOWER LEARNERS TO BE FUTURE READY USING AI	Kartik Sharma, Ashish Aggarwal	S Chand
1206	Biocon Biologics	1978.0	Bangalore	BioTechnology	Biocon is a fully integrated pure play biosimi...	Kiran Mazumdar-Shaw	Goldman Sachs, Tata Capital
1207	KiranaKart	2020.0	Mumbai	E-commerce	A better model for instant grocery delivery in...	Kaivalya Vohra, Aadit Palicha	Global Founders Capital
1208	Meddo	2018.0	Gurgaon	Healthcare	Meddo is an end-to-end medical services provider	Saurabh Kochhar	growX ventures, Venture Gurukool

In [23]:

data.shape

Out[23]:

(1209, 10)

In [24]:

```
data.columns
```

Out[24]:

```
Index(['Company/Brand', 'Founded', 'Headquarters', 'Sector', 'What it doe  
s',  
      'Founder/s', 'Investor/s', 'Amount', 'Stage', 'Date '],  
      dtype='object')
```

In [25]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1209 entries, 0 to 1208  
Data columns (total 10 columns):  
#   Column                Non-Null Count  Dtype  
---  ---  
0   Company/Brand         1209 non-null  object  
1   Founded               1208 non-null  float64  
2   Headquarters          1208 non-null  object  
3   Sector                1209 non-null  object  
4   What it does          1209 non-null  object  
5   Founder/s            1205 non-null  object  
6   Investor/s           1147 non-null  object  
7   Amount               1206 non-null  object  
8   Stage                 781 non-null   object  
9   Date                 1209 non-null  object  
dtypes: float64(1), object(9)  
memory usage: 94.6+ KB
```

In [26]:

```
data.describe()
```

Out[26]:

	Founded
count	1208.000000
mean	2016.655629
std	4.517364
min	1963.000000
25%	2015.000000
50%	2018.000000
75%	2020.000000
max	2021.000000

In [27]:

```
data.isnull().sum()
```

Out[27]:

```
Company/Brand      0
Founded            1
Headquarters       1
Sector             0
What it does       0
Founder/s          4
Investor/s         62
Amount             3
Stage             428
Date              0
dtype: int64
```

In [28]:

```
data.drop('Stage', axis = 1, inplace = True)
```

In [30]:

```
data.head()
```

Out[30]:

	Founded	Headquarters	Sector	What it does	Founder/s	Investor/s	Amount	Date
0	2015.0	Gurgaon	E-learning	Collegedekho.com is Student's Partner, Friend ...	Ruchir Arora	Disrupt ADQ, QIC	\$35,000,000	Dec-21
3	2012.0	Mumbai	Food & Beverages	India's Largest Desi Meals Brand	Anshul Gupta, Amit Raj	Tiger Global	\$40,000,000	Dec-21
1	2015.0	Bangalore	Consumer Services	Simpl empowers merchants to build trusted rela...	Nitya, Chaitra Chidanand	Valar Ventures, IA Ventures	\$40,000,000	Dec-21
5	2018.0	Mumbai	Venture Capital & Private Equity	8i is a Mumbai & Bangalore based early stage f...	Vikram Chachra	NaN	\$50,000,000	Dec-21
1	2021.0	Bangalore	Financial Services	PayGlocal is a FinTech solving for global paym...	Prachi Dharani, Rohit Sukhija, Yogesh Lokhande	Sequoia Capital India	\$4,900,000	Dec-21

In [31]:



```
data['Headquarters'].value_counts()
```

Out[31]:

Bangalore	426
Mumbai	203
Gurugram	155
New Delhi	144
Pune	47
...	
The Nilgiris	1
Rajsamand	1
West Bengal	1
Guwahati	1
Powai	1

Name: Headquarters, Length: 70, dtype: int64

In [32]:



```
data['Headquarters'].value_counts().head()
```

Out[32]:

Bangalore	426
Mumbai	203
Gurugram	155
New Delhi	144
Pune	47

Name: Headquarters, dtype: int64

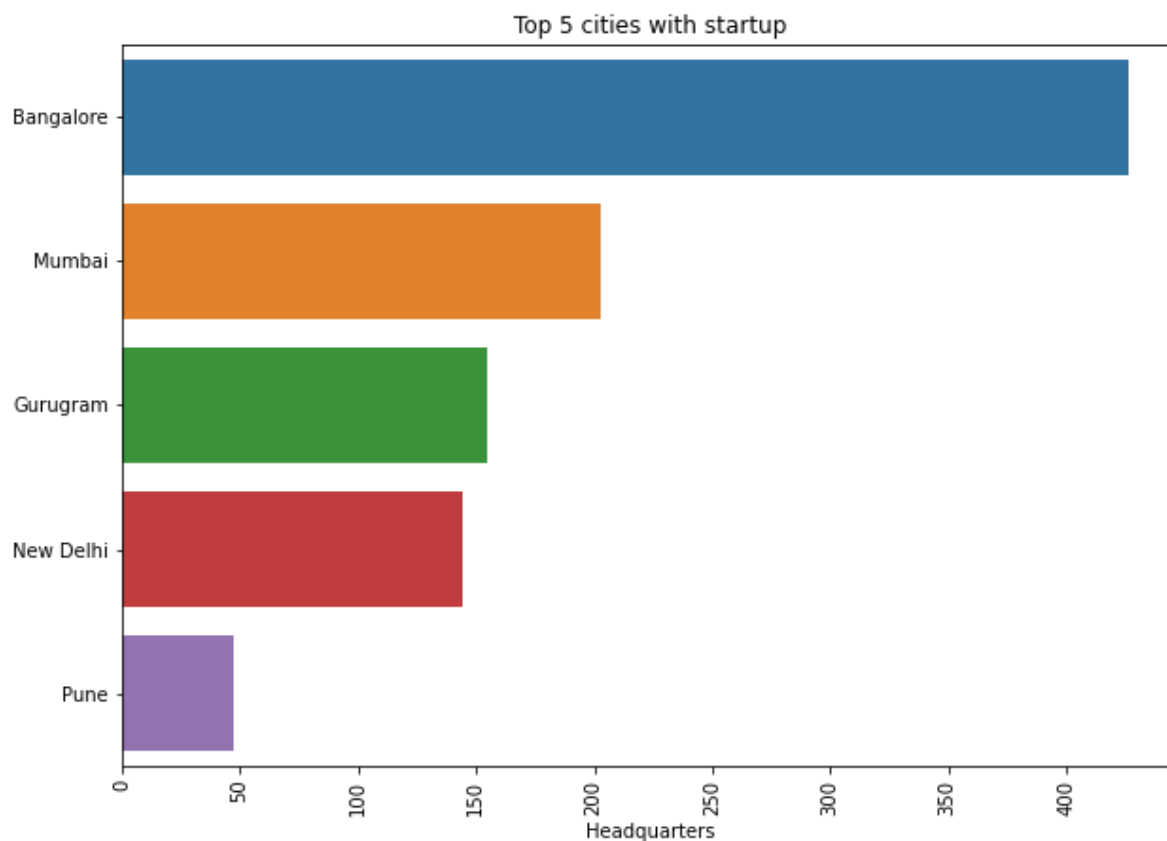
In [34]:



```
top_5_cities = pd.DataFrame(data['Headquarters'].value_counts().head())
```

In [36]:

```
plt.figure(figsize = (10,7))
sns.barplot(data = data, x = top_5_cities['Headquarters'],y=top_5_cities.index)
plt.xticks(rotation=90)
plt.title("Top 5 cities with startup")
plt.show()
```



In [37]:



```
data['Sector'].value_counts()
```

Out[37]:

FinTech	121
EdTech	102
Financial Services	61
E-commerce	37
Food & Beverages	37
...	
FemTech	1
Community	1
Mechanical & Industrial Engineering	1
Nutrition	1
Manufacturing startup	1

Name: Sector, Length: 256, dtype: int64

In [38]:



```
data['Sector'].value_counts().head()
```

Out[38]:

FinTech	121
EdTech	102
Financial Services	61
E-commerce	37
Food & Beverages	37

Name: Sector, dtype: int64

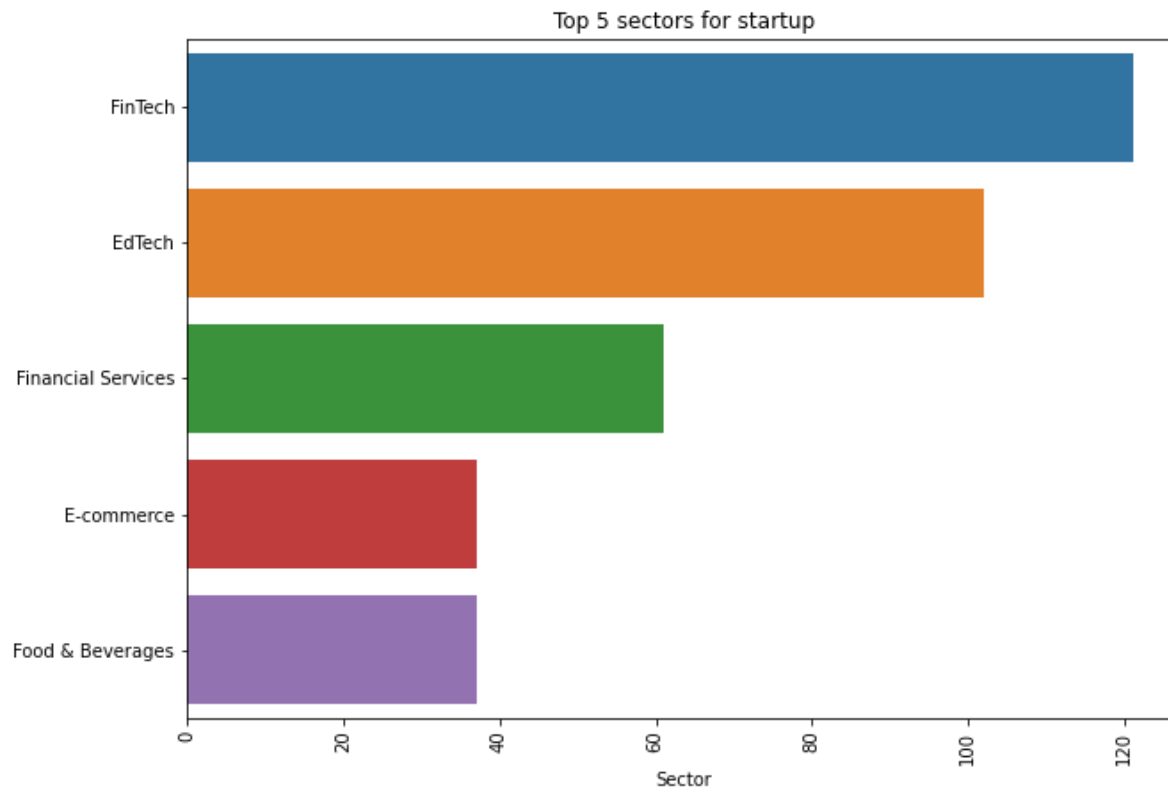
In [39]:



```
top_5_sectors = pd.DataFrame(data['Sector'].value_counts().head())
```

In [40]:

```
plt.figure(figsize = (10,7))  
sns.barplot(data = data, x = top_5_sectors['Sector'],y=top_5_sectors.index)  
plt.xticks(rotation=90)  
plt.title("Top 5 sectors for startup")  
plt.show()
```



In [41]:



```
data['Investor/s'].value_counts()
```

Out[41]:

Inflection Point Ventures	24
Venture Catalysts	11
Tiger Global	11
Elevation Capital	7
Titan Capital	6
..	
Kunal Shah, Snehal Fulzele	1
LetsVenture, Paradigm Shift VC, Binny Bansal	1
Appit Simple Infotek	1
Titan Capital, Archana Priyadarshini, Gaurav Ahuja	1
Sanjay Baldwa, Ajay Jain	1

Name: Investor/s, Length: 938, dtype: int64

In [42]:



```
data['Investor/s'].value_counts().head()
```

Out[42]:

Inflection Point Ventures	24
Venture Catalysts	11
Tiger Global	11
Elevation Capital	7
Titan Capital	6

Name: Investor/s, dtype: int64

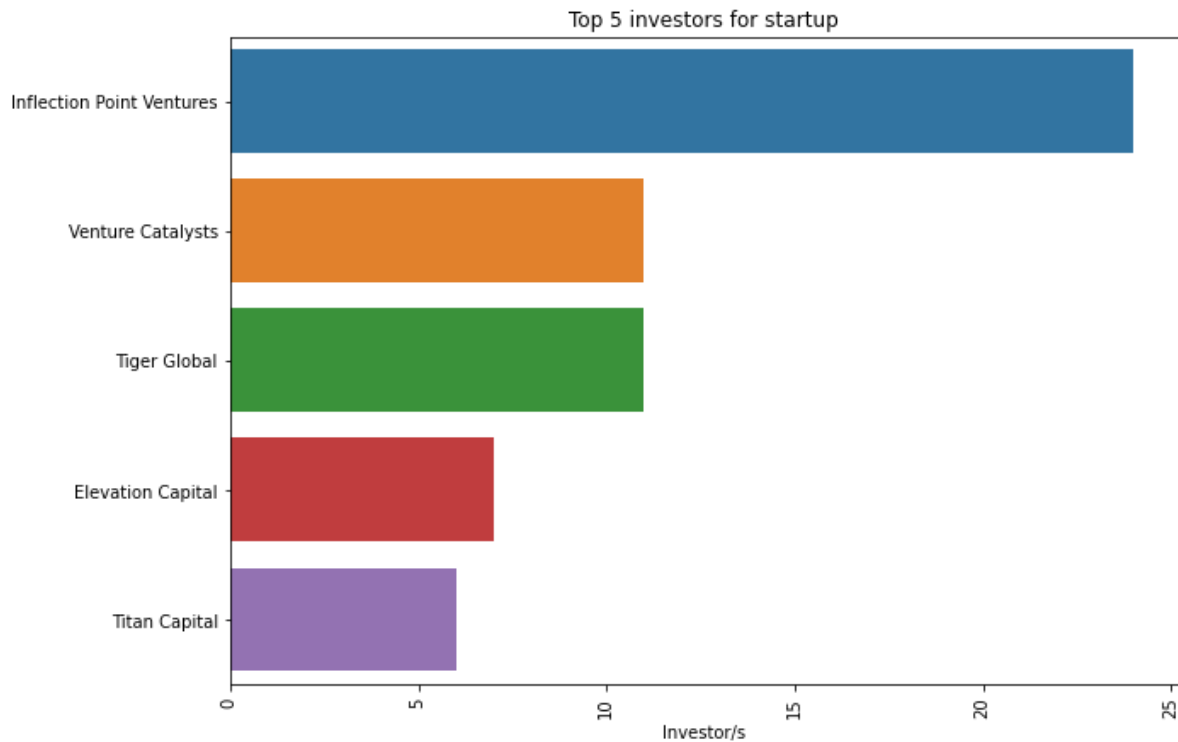
In [44]:



```
top_5_investors = pd.DataFrame(data['Investor/s'].value_counts().head())
```

In [45]:

```
plt.figure(figsize = (10,7))  
sns.barplot(data = data, x = top_5_investors['Investor/s'],y=top_5_investors.index)  
plt.xticks(rotation=90)  
plt.title("Top 5 investors for startup")  
plt.show()
```



In [49]:

```
data['Founded'] = data['Founded'].fillna(0)
```

In [50]:

```
data['Founded'] = data['Founded'].astype(int)
```

In [51]:



```
data['Founded'].value_counts()
```

Out[51]:

2020	237
2019	177
2018	143
2015	129
2016	118
2017	112
2021	76
2014	57
2011	34
2012	31
2013	22
2008	15
2010	14
2009	9
2000	6
2007	6
2006	3
2004	2
2002	2
1999	2
1998	2
1994	2
1989	2
2003	1
2005	1
1993	1
1991	1
1984	1
1978	1
1963	1
0	1

Name: Founded, dtype: int64

In [64]:



```
data['Founded'].value_counts().head()
```

Out[64]:

2020	237
2019	177
2018	143
2015	129
2016	118

Name: Founded, dtype: int64

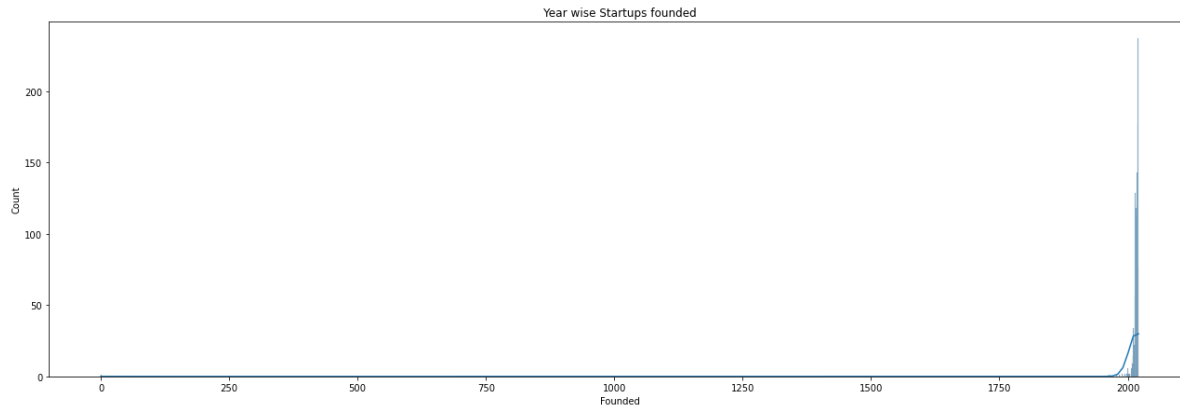
In [65]:



```
top_5_years = pd.DataFrame(data['Founded'].value_counts().head())
```

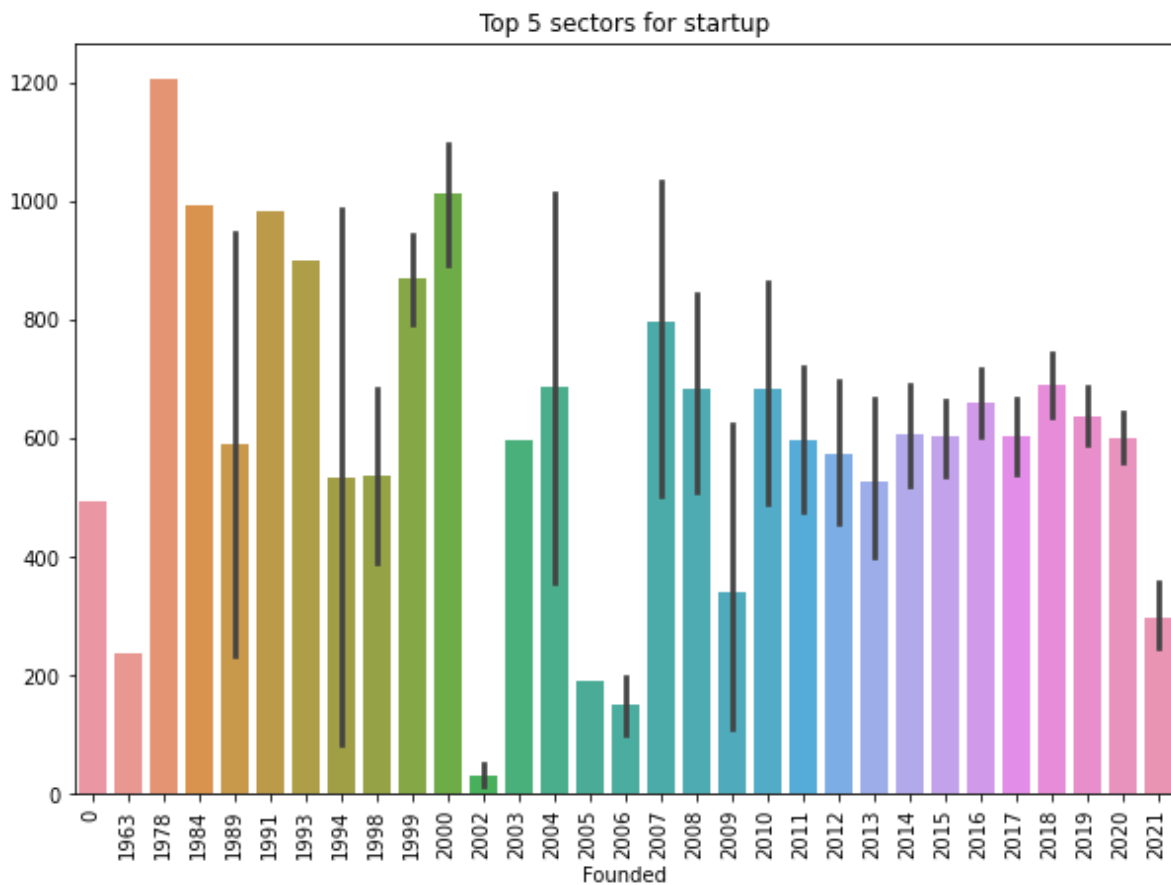
In [113]:

```
plt.figure(figsize = (22,7))  
sns.histplot(data['Founded'], kde = True)  
plt.title('Year wise Startups founded')  
plt.show()
```



In [73]:

```
plt.figure(figsize = (10,7))  
sns.barplot(x = data['Founded'],y=data.index)  
plt.xticks(rotation=90)  
plt.title("Top 5 sectors for startup")  
plt.show()
```



In [74]:

```
data['Amount_New'] = pd.DataFrame(data['Amount'].str.replace('$', '', regex = True))
data['Amount_New'] = data['Amount_New'].str.replace(',', '', regex = True)
data['Amount_New'] = pd.to_numeric(data['Amount_New'], errors = 'coerce')
data.head()
```

Out[74]:

Headquarters	Sector	What it does	Founder/s	Investor/s	Amount	Date	Amount_New
Gurgaon	E-learning	Collegedekho.com is Student's Partner, Friend ...	Ruchir Arora	Disrupt ADQ, QIC	\$35,000,000	Dec-21	35000000.0
Mumbai	Food & Beverages	India's Largest Desi Meals Brand	Anshul Gupta, Amit Raj	Tiger Global	\$40,000,000	Dec-21	40000000.0
Bangalore	Consumer Services	Simpl empowers merchants to build trusted rela...	Nitya, Chaitra Chidanand	Valar Ventures, IA Ventures	\$40,000,000	Dec-21	40000000.0
Mumbai	Venture Capital & Private Equity	8i is a Mumbai & Bangalore based early stage f...	Vikram Chachra	NaN	\$50,000,000	Dec-21	50000000.0
Bangalore	Financial Services	PayGlocal is a FinTech solving for global paym...	Prachi Dharani, Rohit Sukhija, Yogesh Lokhande	Sequoia Capital India	\$4,900,000	Dec-21	4900000.0

In [75]:

```
data_group = data.groupby(by = 'Headquarters').sum().drop('Founded', axis=1).reset_index()
```

In [76]:



```
data_group
```

Out[76]:

	Headquarters	Amount_New
0	Ahmadabad	4.300000e+06
1	Ahmedabad	1.981550e+08
2	Ambernath	2.000000e+05
3	Andheri	1.000000e+06
4	Bangalore	1.574084e+10
...
65	The Nilgiris	2.000000e+05
66	Thiruvananthapuram	0.000000e+00
67	Trivandrum	3.000000e+05
68	Vadodara	1.000000e+07
69	West Bengal	1.500000e+06

70 rows × 2 columns

In [77]:



```
data_group.head()
```

Out[77]:

	Headquarters	Amount_New
0	Ahmadabad	4.300000e+06
1	Ahmedabad	1.981550e+08
2	Ambernath	2.000000e+05
3	Andheri	1.000000e+06
4	Bangalore	1.574084e+10

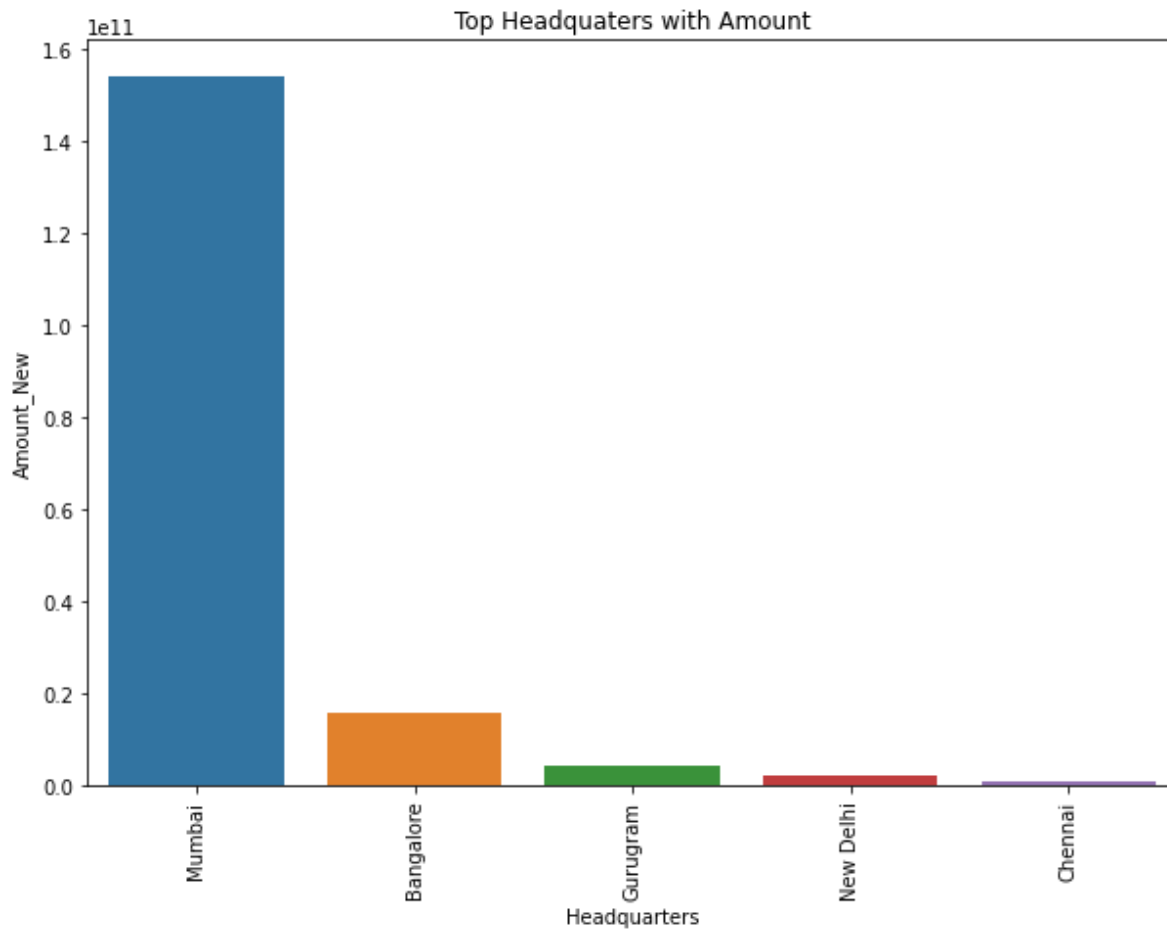
In [88]:



```
data_top_group = data_group.sort_values(by = 'Amount_New',ascending = False).head()
```

In [89]:

```
plt.figure(figsize = (10,7))
sns.barplot(x = data_top_group['Headquarters'],y=data_top_group['Amount_New'])
plt.xticks(rotation=90)
plt.title("Top Headquarters with Amount")
plt.show()
```



In [81]:

```
data_group_invest = data.groupby(by = 'Investor/s').sum().drop('Founded',axis=1).reset_
```

In [82]:

```
data_group_invest
```

Out[82]:

	Investor/s	Amount_New
0	\$1,000,000	0.0
1	\$22,000,000	0.0
2	\$5,000,000	0.0
3	\$Undisclosed	0.0
4	021 Capital, Sparrow Capital	1800000.0
...
933	http://100x.vc/	2500000.0
934	iSeed, Titan Capital	500000.0
935	pi Ventures	1800000.0
936	responsAbility Investments, Blue Orchard	10000000.0
937	velocity.in	200000.0

938 rows × 2 columns

In [83]:

```
data_group_invest.head()
```

Out[83]:

	Investor/s	Amount_New
0	\$1,000,000	0.0
1	\$22,000,000	0.0
2	\$5,000,000	0.0
3	\$Undisclosed	0.0
4	021 Capital, Sparrow Capital	1800000.0

In [86]:

```
data_group_invest_top = data_group_invest.sort_values(by = 'Amount_New', ascending = False)
```


In [87]:

```
plt.figure(figsize = (10,7))  
sns.barplot(x = data_group_invest_top['Investor/s'],y=data_group_invest_top['Amount_New  
plt.xticks(rotation=90)  
plt.title("Top Investors with Amount")  
plt.show()
```

