Analyza Startup_funding dataset and answer all the following questions:

Daniel Mehta

1 **How Does the Funding Ecosystem changes with respect to Time?**

2 **What is the General Amount that Startups get in India?**

3 **Which Kind of Industries are more preferred for Startups?**

4 **Does Location also play a role, In determining the Growth of a Startup?**

5 **Who plays the main role in Indian Startups Ecosystem?**

6 **What are the different Types of Funding for Startups?**

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
from collections import Counter
from pprint import pprint
```

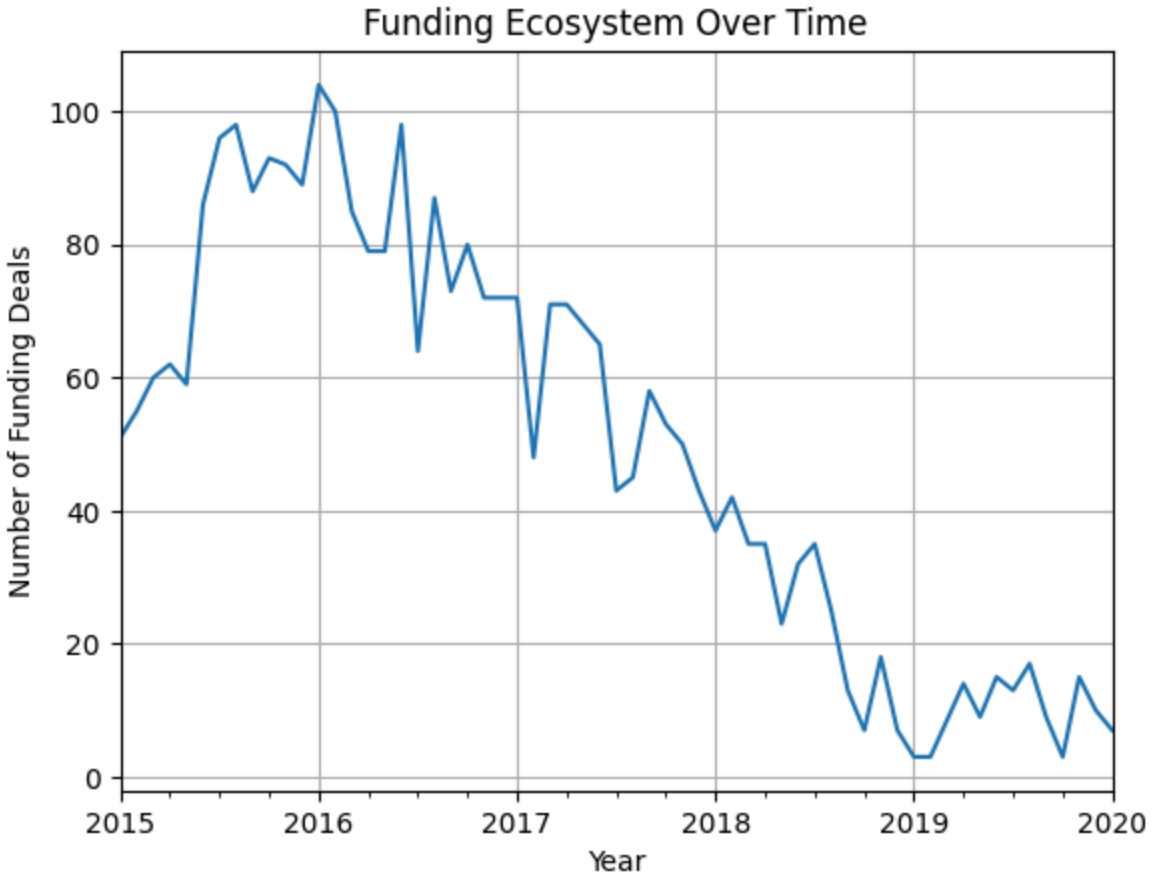
In [2]:
```
df = pd.read_csv('startup_funding.csv')
df.head()
```

Out[2]:

	Sr No	Date dd/mm/yyyy	Startup Name	Industry Vertical	SubVertical	City Location	Investors Name	InvestmentnType	Amount in USD	Re
0	1	09/01/2020	BYJU'S	E-Tech	E-learning	Bengaluru	Tiger Global Management	Private Equity Round	20,00,00,000	
1	2	13/01/2020	Shuttl	Transportation	App based shuttle service	Gurgaon	Susquehanna Growth Equity	Series C	80,48,394	
2	3	09/01/2020	Mamaearth	E-commerce	Retailer of baby and toddler products	Bengaluru	Sequoia Capital India	Series B	1,83,58,860	
3	4	02/01/2020	https://www.wealthbucket.in/	FinTech	Online Investment	New Delhi	Vinod Khatumal	Pre-series A	30,00,000	
4	5	02/01/2020	Fashor	Fashion and Apparel	Embroiled Clothes For Women	Mumbai	Sprout Venture Partners	Seed Round	18,00,000	

Q1: How Does the Funding Ecosystem changes with respect to Time?

In [3]:
```
df.rename(columns={'Date dd/mm/yyyy': 'Date'}, inplace=True)
df['Date'] = pd.to_datetime(df['Date'], errors='coerce', dayfirst=True)
df['YearMonth'] = df['Date'].dt.to_period('M')

funding_over_time = df.groupby('YearMonth').size()

funding_over_time.plot(kind='line')
plt.title('Funding Ecosystem Over Time')
plt.xlabel('Year')
plt.ylabel('Number of Funding Deals')
plt.grid(True)
plt.show()
```



Q2: What is the General Amount that Startups get in India?

In [4]:
```
df['Amount in USD'] = df['Amount in USD'].replace('[^0-9.]', '', regex=True)
df['Amount in USD'] = pd.to_numeric(df['Amount in USD'], errors='coerce')

amounts = df['Amount in USD'].dropna()

mean_amt = amounts.mean()
median_amt = amounts.median()
mode_amt = amounts.mode().iloc[0]


print(f"Mean Funding Amount: ${mean_amt:,.0f}")
print(f"Median Funding Amount: ${median_amt:,.0f}")
print(f"Most Common (Mode) Funding Amount: ${mode_amt:,.0f}")
```

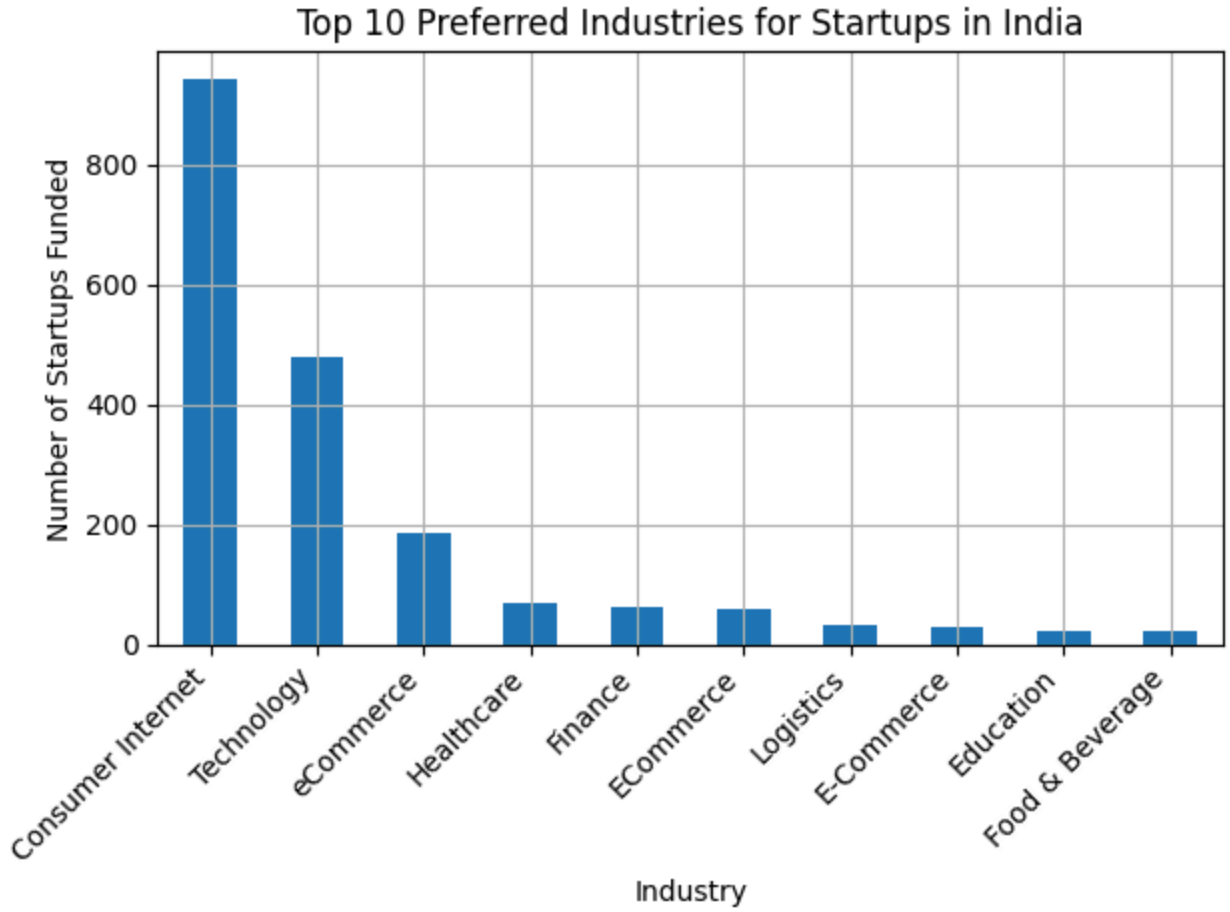
```
Mean Funding Amount: $22,332,169
Median Funding Amount: $1,750,000
Most Common (Mode) Funding Amount: $1,000,000
```

Q3: Which Kind of Industries are more preferred for Startups?

In [5]:
```python
df['Industry Vertical'] = df['Industry Vertical'].str.strip()

industry_counts = df['Industry Vertical'].value_counts().head(10)

print(industry_counts)
```

```
Industry Vertical
Consumer Internet    941
Technology           478
eCommerce            186
Healthcare            70
Finance               62
ECommerce             61
Logistics             32
E-Commerce            29
Education             24
Food & Beverage       23
Name: count, dtype: int64
```

In [6]:
```python
industry_counts.plot(kind='bar')
plt.title('Top 10 Preferred Industries for Startups in India')
plt.xlabel('Industry')
plt.ylabel('Number of Startups Funded')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.grid(True)
plt.show()
```



Therefore the indistries most preferred for startups are Consumer Internet, Technology, then eCommerce

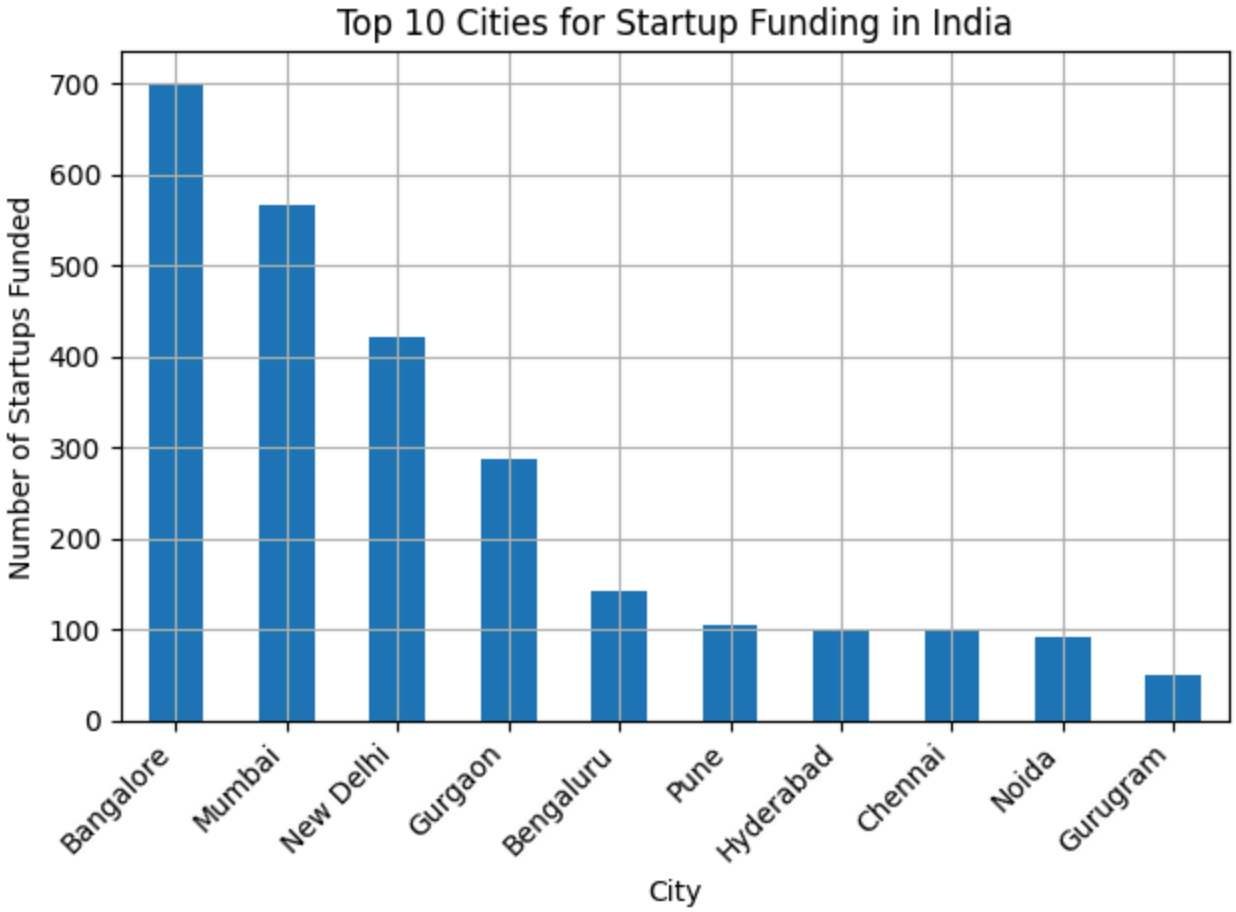
Q4: Does Location also play a role, In determining the Growth of a Startup?

In [7]:
```
df.rename(columns={'City  Location': 'City Location'}, inplace=True)
df['City Location'] = df['City Location'].str.strip()

city_counts = df['City Location'].value_counts().head(10)

print(city_counts)
```

```
City Location
Bangalore    700
Mumbai       567
New Delhi    421
Gurgaon      287
Bengaluru    141
Pune         105
Hyderabad     99
Chennai       97
Noida         92
Gurugram      50
Name: count, dtype: int64
```

In [8]:

```
city_counts.plot(kind='bar')
plt.title('Top 10 Cities for Startup Funding in India')
plt.xlabel('City')
plt.ylabel('Number of Startups Funded')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.grid(True)
plt.show()
```



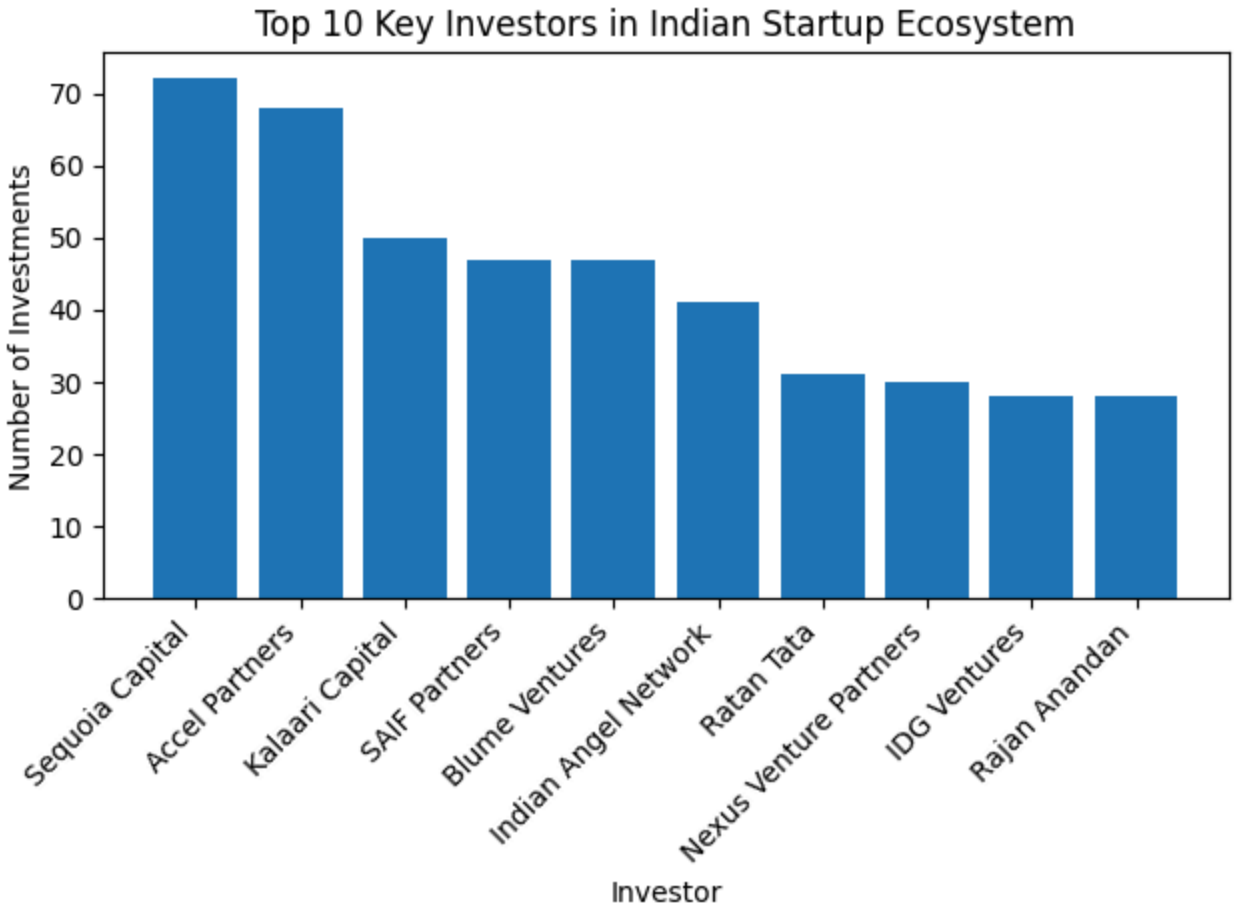
Therefore Location does play a role as Bangalore, Mumbai and New Delhi get majority of start up funding

Q5 Who plays the main role in Indian Startups Ecosystem?

In [9]:
```
investor_list = df['Investors Name'].dropna().str.split(',')
flat_list = [investor.strip() for sublist in investor_list for investor in sublist]

filtered_list = [i for i in flat_list if i not in ['', 'Undisclosed Investors', 'Undisclosed investors']]

investor_counts = Counter(filtered_list)
top_investors = dict(investor_counts.most_common(10))

print("Top 10 Investors (most deals):")
pprint(top_investors, sort_dicts=False)
```

```
Top 10 Investors (most deals):
{'Sequoia Capital': 72,
 'Accel Partners': 68,
 'Kalaari Capital': 50,
 'SAIF Partners': 47,
 'Blume Ventures': 47,
 'Indian Angel Network': 41,
 'Ratan Tata': 31,
 'Nexus Venture Partners': 30,
 'IDG Ventures': 28,
 'Rajan Anandan': 28}
```

In [10]:
```
plt.bar(top_investors.keys(), top_investors.values())
plt.title('Top 10 Key Investors in Indian Startup Ecosystem')
plt.xlabel('Investor')
plt.ylabel('Number of Investments')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```



Q6: What are the different Types of Funding for Startups?

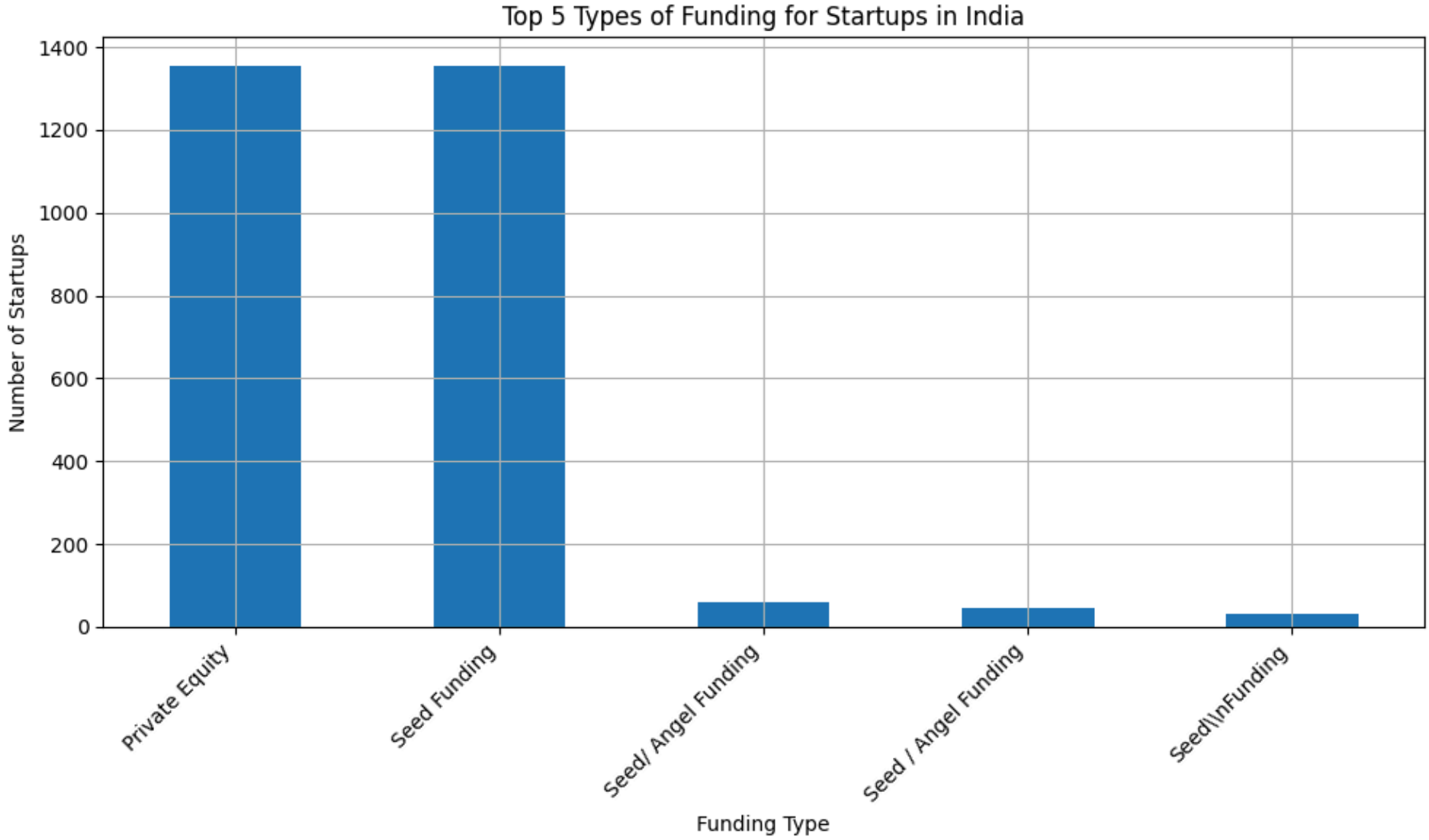
In [11]:
```
df.rename(columns={'InvestmentnType': 'Investment Type'}, inplace=True)

df['Investment Type'] = df['Investment Type'].str.strip()

funding_types = df['Investment Type'].value_counts()

print(funding_types)
```

Investment Type	
Private Equity	1356
Seed Funding	1355
Seed/ Angel Funding	60
Seed / Angel Funding	47
Seed\\nFunding	30
Debt Funding	25
Series A	24
Seed/Angel Funding	23
Series B	20
Series C	14
Series D	12
Angel / Seed Funding	8
Seed Round	7
Pre-Series A	4
Private Equity Round	4
Seed	4
Seed / Angle Funding	3
pre-Series A	2
Venture Round	2
Equity	2
Corporate Round	2
Series E	2
Series F	2
Private	1
Debt-Funding	1
Term Loan	1
Seed funding	1
PrivateEquity	1
Angel Funding	1
Private\\nEquity	1
Private Funding	1
Equity Based Funding	1
Crowd funding	1
Series B (Extension)	1
Mezzanine	1
Structured Debt	1
Venture - Series Unknown	1
Pre Series A	1
Debt	1
Pre-series A	1
Series G	1
Series H	1
Venture	1
Funding Round	1
Maiden Round	1
pre-series A	1
Seed Funding Round	1

```
Single Venture                  1
Angel                           1
Series J                        1
Angel Round                     1
Bridge Round                    1
Debt and Preference capital     1
Inhouse Funding                 1
Crowd Funding                   1
Name: count, dtype: int64
```

In [12]:

```python
top_funding_types = df['Investment Type'].value_counts().head()

plt.figure(figsize=(10,6))
top_funding_types.plot(kind='bar')
plt.title('Top 5 Types of Funding for Startups in India')
plt.xlabel('Funding Type')
plt.ylabel('Number of Startups')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.grid(True)
plt.show()
```


Top 5 Types of Funding for Startups in India
Number of Startups
0
200
400
600
800
1000
1200
1400
Private Equity
Seed Funding
Seed/ Angel Funding
Seed / Angel Funding
Seed\\nFunding
Funding Type

In [ ]: