In [1]:
```
import pandas as pd
```

In [2]:
```
dataset= pd.read_csv('D:\\PYTHON ASSIGNMENT AND DATASET\\Unicorn_Companies.csv')
```

In [3]:
```
dataset.head()
```

Out[3]:

	Company	Valuation	Date Joined	Industry	City	Country	Continent	Year Founded	Funding	Select Investors
0	Bytedance	$180B	2017-04-07	Artificial intelligence	Beijing	China	Asia	2012	$8B	Sequoia Capital China, SIG Asia Investments, S...
1	SpaceX	$100B	2012-12-01	Other	Hawthorne	United States	North America	2002	$7B	Founders Fund, Draper Fisher Jurvetson, Rothen...
2	SHEIN	$100B	2018-07-03	E-commerce & direct-to-consumer	Shenzhen	China	Asia	2008	$2B	Tiger Global Management, Sequoia Capital China...
3	Stripe	$95B	2014-01-23	Fintech	San Francisco	United States	North America	2010	$2B	Khosla Ventures, LowercaseCapital, capitalG
4	Klarna	$46B	2011-12-12	Fintech	Stockholm	Sweden	Europe	2005	$4B	Institutional Venture Partners, Sequoia Capita...

In [ ]:
```
##Which unicorn companies have had the biggest return on investment?

```

In [21]:
```
sorted_dataset = dataset.sort_values(by='Valuation', ascending=True)
```

In [25]:
```
print(sorted_dataset.head(2)[['Company', 'Valuation']])
```

```
  Company Valuation
1  SpaceX     $100B
2   SHEIN     $100B
```

In [ ]:
```
###How long does it usually take for a company to become a unicorn? Has it always been this way
```

In [26]:
```
98,12010, 2005, 2012, 2012, 2015, 1991, 2013, 2015, 2018, 2002, 2008, 2017, 2015, 2016, 2011, 2017, 2016, 2013, 2011, 2017, 2016]
```

In [27]:
```
avg_time_to_unicorn = (max(founding_years) - min(founding_years)) / len(founding_years)
```

In [28]:
```
print("On average, it takes around", round(avg_time_to_unicorn), "years for a company to become a unicorn.")
```

```
On average, it takes around 1 years for a company to become a unicorn.
```

In [29]:
```
###Which countries have the most unicorns? Are there any cities that appear to be industry hubs
```

In [30]:
```
unicorn_count = dataset['Country'].value_counts()
```

In [31]:
```
top_countries = unicorn_count.head()
```

In [32]:
```
print("Countries with the most unicorns:")
print(top_countries)
```

```
Countries with the most unicorns:
Country
United States     562
China             173
India              65
United Kingdom     43
Germany            26
Name: count, dtype: int64
```

In [34]:
```
# Identify industry hubs (cities with multiple unicorns)
city_hubs = dataset['City'].value_counts()
```

In [35]:
```
1
2 print("\nCities that appear to be industry hubs:")
3 print(city_hubs[city_hubs > 1])
```

```
Cities that appear to be industry hubs:
City
San Francisco    152
New York         103
Beijing           63
Shanghai          44
London            34
                 ...
Suzhou             2
Montreal           2
Chennai            2
Washington         2
Bogota             2
Name: count, Length: 100, dtype: int64
```

In [36]:
```
1 ##Which investors have funded the most unicorns?
```

In [37]:
```
1 dataset['Investors'] = dataset['Select Investors'].str.split(', ')
```

In [38]:
```
1 dataset = dataset.explode('Investors')
```

In [40]:
```
1 investor_count = dataset['Investors'].value_counts()
```

In [41]:
```
1 top_investors = investor_count.head()
```

In [42]:
```
1 print("Investors that have funded the most unicorns:")
2 print(top_investors)
```

```
Investors that have funded the most unicorns:
Investors
Accel                     60
Tiger Global Management   53
Andreessen Horowitz       53
Sequoia Capital China     48
Insight Partners          47
Name: count, dtype: int64
```

In [43]:
```
1
```

In [ ]:
```
1
```

In [ ]:
```
1
```

In [ ]:
```
1
```

In [ ]:
```
1
```