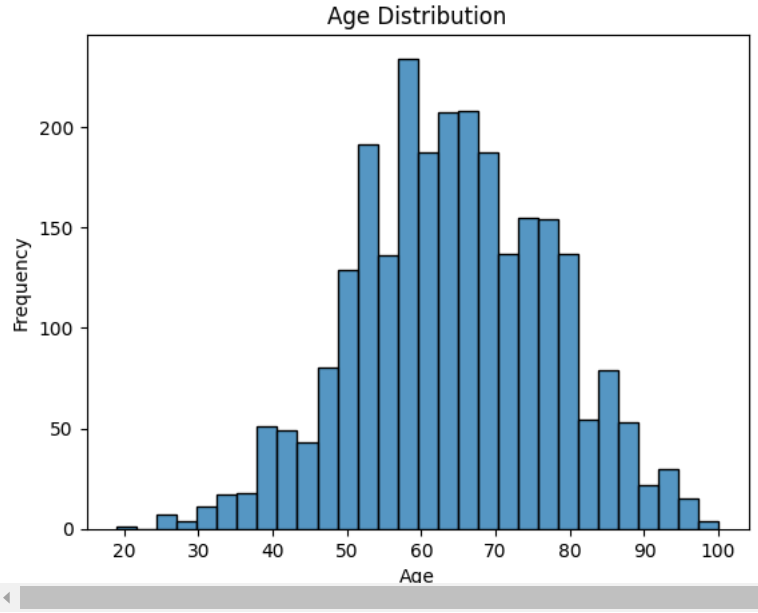
```
# Project-3
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

#df=pd.read_csv("/content/archive.zip")
df.head()
```

	Unnamed: 0	rank	name	networth	age	country	source	industry
0	0	1	Elon Musk	$219 B	50	United States	Tesla, SpaceX	Automotive
1	1	2	Jeff Bezos	$171 B	58	United States	Amazon	Technology
2	2	3	Bernard Arnault & family	$158 B	73	France	LVMH	Fashion & Retail
3	3	4	Bill Gates	$129 B	66	United States	Microsoft	Technology
4	4	5	Warren Buffett	$118 B	91	United States	Berkshire Hathaway	Finance & Investments

```
#Visualization
# 1. Bar plot for 'Age'
sns.histplot (df['age'],bins=30,kde=False)
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show
```

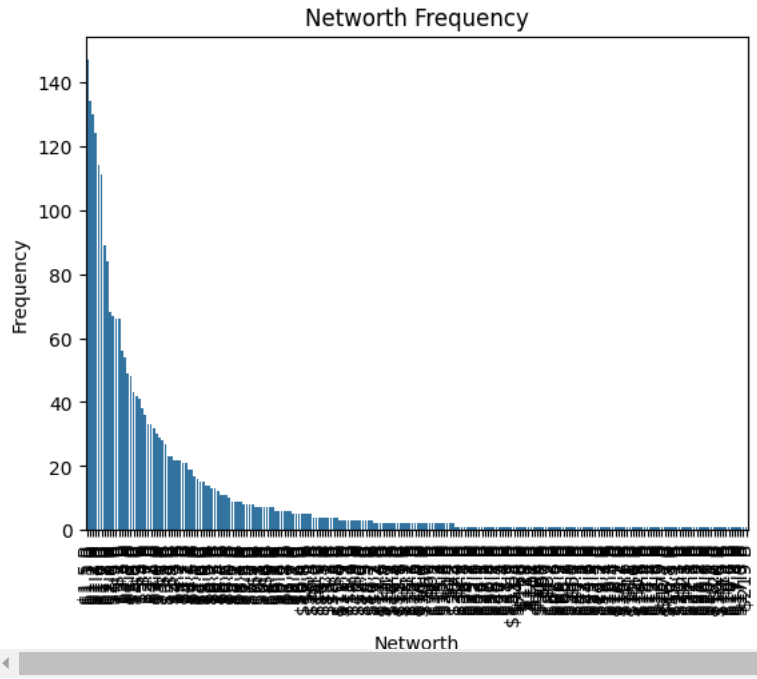
```
matplotlib.pyplot.show
def show(*args, **kwargs)

Display all open figures.

Parameters
----------
block : bool, optional
    Whether to wait for all figures to be closed before returning.
```



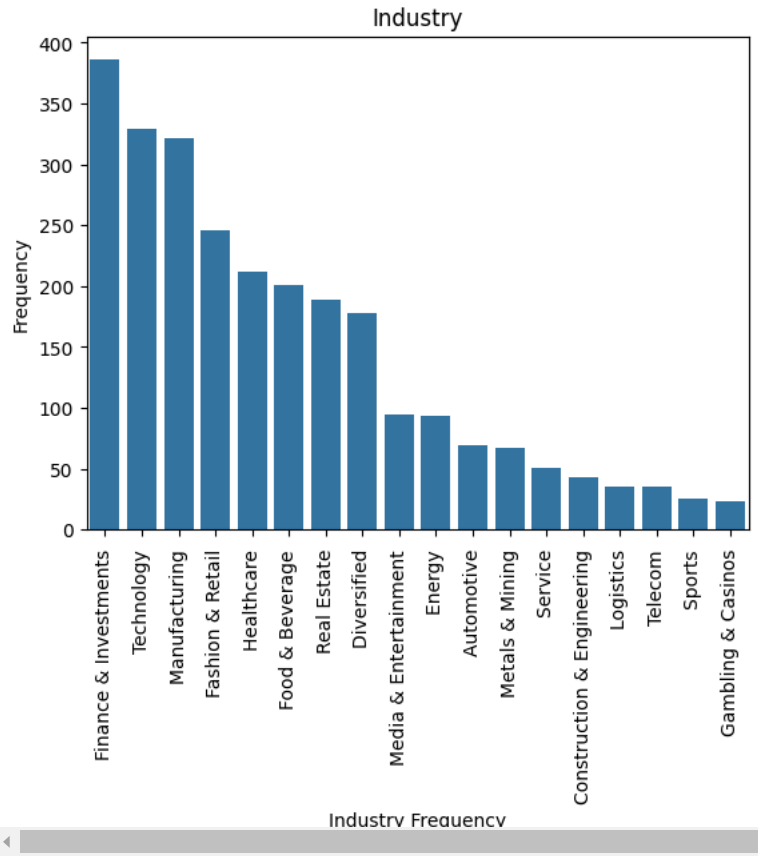

```
# 2. Bar plot for Networth vs Frequency
data = df.dropna(subset=['networth'])

sns.countplot(data=df, x='networth', order = data['networth'].value_counts().index)
plt.xticks(rotation=90)
plt.title('Networth Frequency')
plt.xlabel('Networth')
plt.ylabel('Frequency')
plt.xticks(rotation=90)
plt.show
```

```
matplotlib.pyplot.show
def show(*args, **kwargs)
```

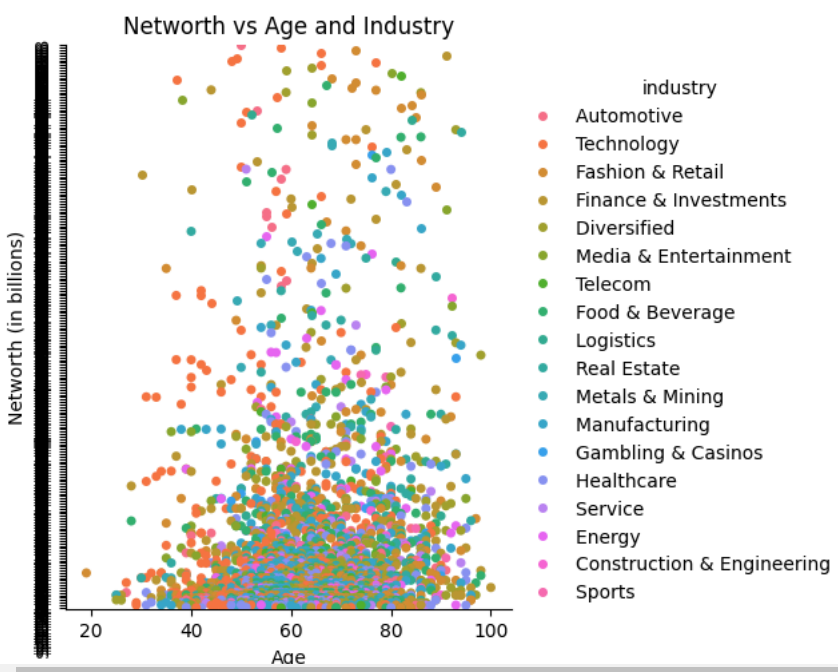
```
Display all open figures.

Parameters
----------
block : bool, optional
    Whether to wait for all figures to be closed before returning.
```



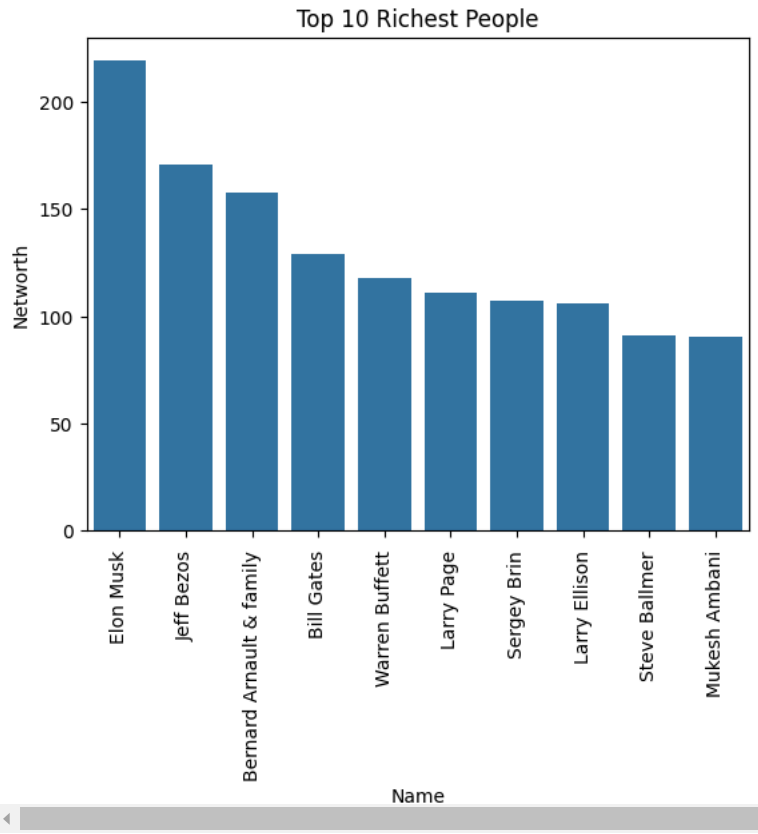

```
# 3. Industry frequency

sns.countplot(data=df, x='industry',order=data['industry'].value_counts().index)
plt.xticks(rotation=90)
plt.title('Industry')
plt.xlabel('Industry Frequency')
plt.ylabel('Frequency')
plt.show()
```



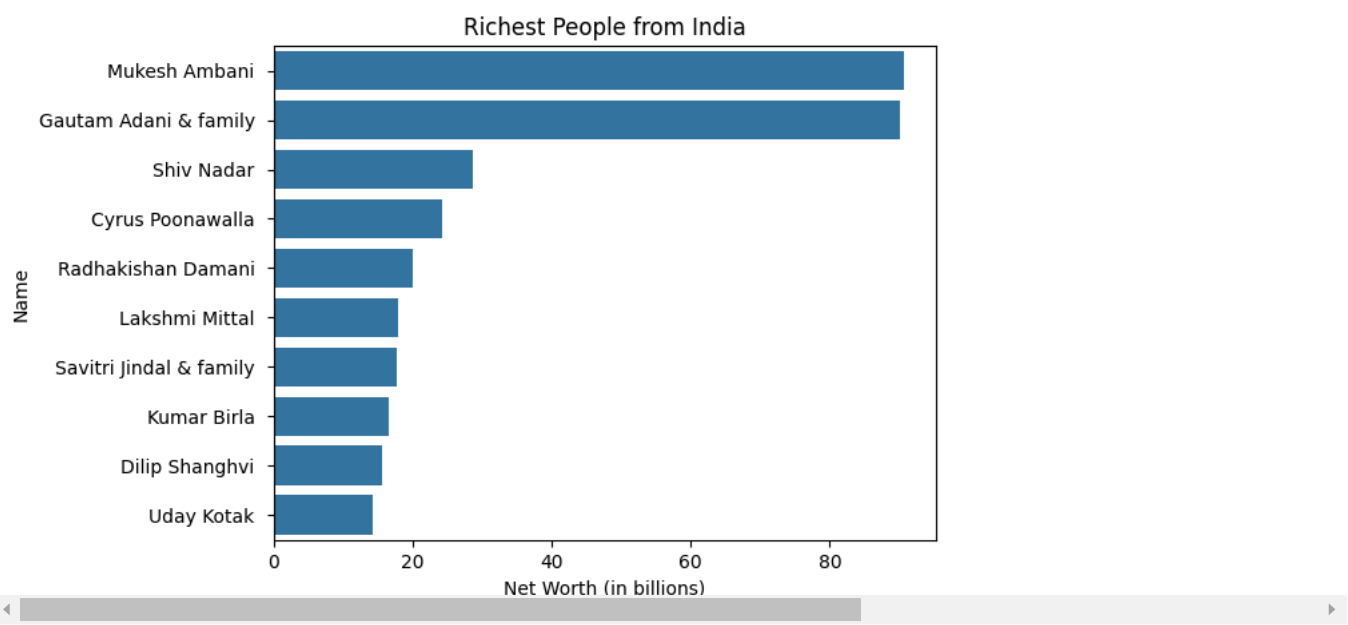
```
# 4. Networth vs Age and Industry

sns.catplot(x='age', y='networth',hue='industry',data=df)
plt.title('Networth vs Age and Industry')
plt.xlabel('Age')
plt.ylabel('Networth (in billions)')
plt.yticks(rotation=90)
plt.show()
```




```
# 5. Top 10 richest people
top=df.copy()
top['networth']=top['networth'].str.replace('$','').str.replace('B','').astype(float)
top = top.nlargest(10,'networth')
sns.barplot(x='name', y='networth', data=top)
plt.title('Top 10 Richest People')
plt.xlabel('Name')
plt.ylabel( 'Networth')
plt.xticks(rotation=90)
plt.show()
```



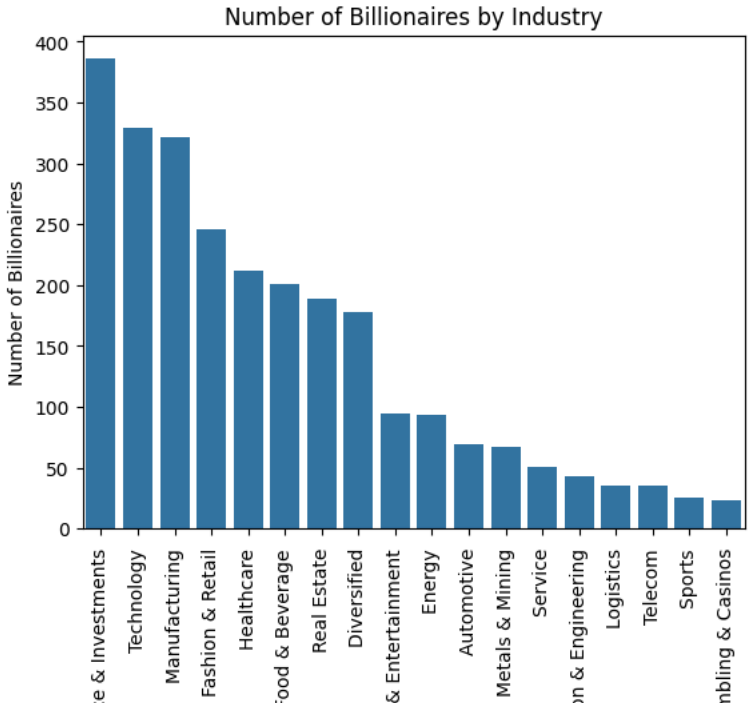
```
# 6.Richest people from India
india_richest = df[df['country'] == 'India'].copy()
india_richest['networth'] = india_richest['networth'].str.replace('\n', '').str.replace('B', '').str.replace('$', '').astype(float)
india_richest = india_richest.nlargest(10, 'networth')
sns.barplot(x='networth', y='name', data=india_richest)
plt.title('Richest People from India')
plt.xlabel('Net Worth (in billions)')
plt.ylabel( 'Name')
plt.show()
```




```
# 7.Minimum age billionaire <=50
young_billionaires = df[df['age'] <= 50]
young_billionaires = young_billionaires[['name', 'age', 'industry']].sort_values(by='age')
print(young_billionaires)
```

```
                     name  age                industry
1311  Kevin David Lehmann    19       Fashion & Retail
2190   Alexandra Andresen    25            Diversified
1975     Pedro Franceschi    25  Finance & Investments
2062          Wang Zelong    25        Metals & Mining
2191   Katharina Andresen    26            Diversified
...                   ...   ...                    ...
2395         Park Kwan-ho    50  Media & Entertainment
1572        Evan Williams    50             Technology
575           Daniel Ziff    50  Finance & Investments
1485         David Mindus    50            Real Estate
0               Elon Musk    50             Automotive

[359 rows x 3 columns]
```

```
# 8.Industry with the most billionaires
industry_counts = data['industry'].value_counts()
sns.barplot(x=industry_counts.index, y=industry_counts.values)
plt.title('Number of Billionaires by Industry')
plt.xlabel('Industry')
plt.ylabel( 'Number of Billionaires')
plt.xticks(rotation=90)
plt.show()
```


Number of Billionaires by Industry
Number of Billionaires
400
350
300
250
200
150
100
50
0
Technology
Manufacturing
Fashion & Retail
Healthcare
Real Estate
Diversified
Energy
Automotive
Metals & Mining
Service
Logistics
Telecom
Sports