



```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
data = pd.read_csv("/content/archive (3).zip")
```

data.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

data.tail()



	Unnamed: 0	rank	name	networth	age	country	source	industry
2595	2595	2578	Jorge Gallardo Ballart	\$1 B	80	Spain	pharmaceuticals	Healthcare
2596	2596	2578	Nari Genomal	\$1 B	82	Philippines	apparel	Fashion & Retail
2597	2597	2578	Ramesh Genomal	\$1 B	71	Philippines	apparel	Fashion & Retail

data.describe()



	Unnamed: 0	rank	age
count	2600.000000	2600.000000	2600.000000
mean	1299.500000	1269.570769	64.271923
std	750.699674	728.146364	13.220607
min	0.000000	1.000000	19.000000
25%	649.750000	637.000000	55.000000
50%	1299.500000	1292.000000	64.000000
75%	1949.250000	1929.000000	74.000000
max	2599.000000	2578.000000	100.000000

data.isnull().sum()

```

⇒
0
-----
Unnamed: 0    0
   rank      0
   name      0
  networth   0
   age      0
  country    0
   source    0
  industry   0

dtype: int64

```

```
print("Data Shape: ", data.shape)
```

```
⇒ Data Shape: (2600, 8)
```

```
data.info()
```

```

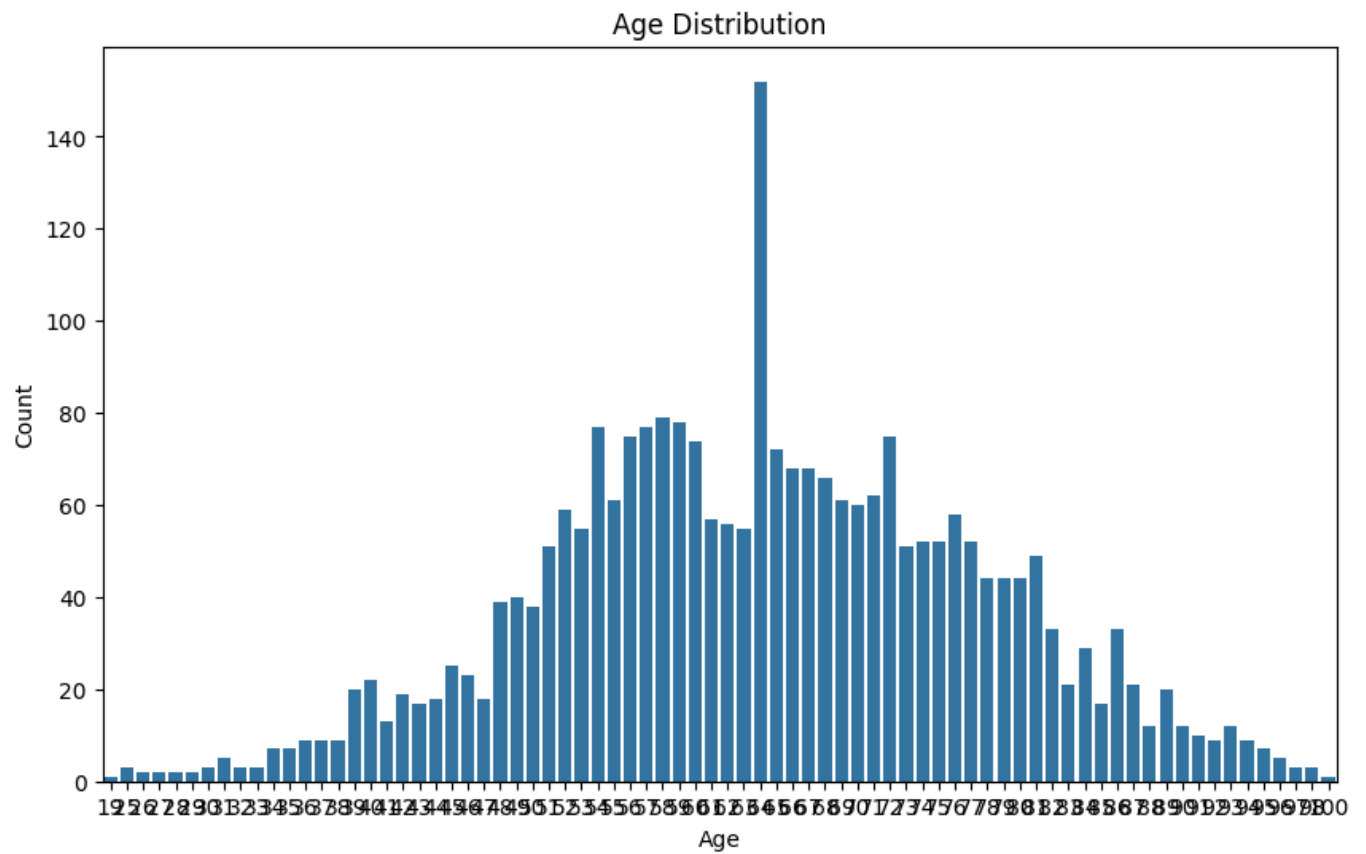
⇒ <class 'pandas.core.frame.DataFrame'>
RangeIndex: 2600 entries, 0 to 2599
Data columns (total 8 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Unnamed: 0   2600 non-null   int64
1   rank         2600 non-null   int64
2   name         2600 non-null   object
3   networth     2600 non-null   object
4   age          2600 non-null   int64
5   country      2600 non-null   object
6   source       2600 non-null   object
7   industry     2600 non-null   object
dtypes: int64(3), object(5)
memory usage: 162.6+ KB

```

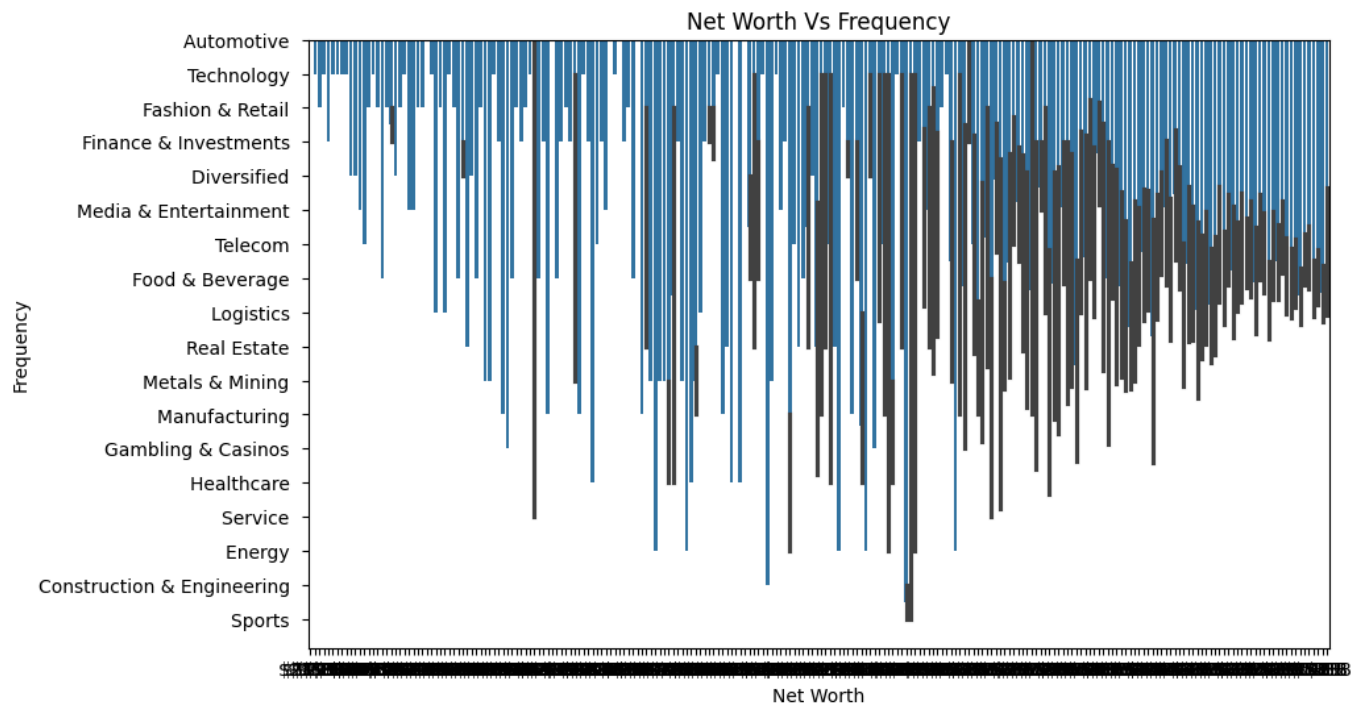
```

plt.figure(figsize=(10, 6))
sns.countplot(x='age', data=data)
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Count')
plt.show()

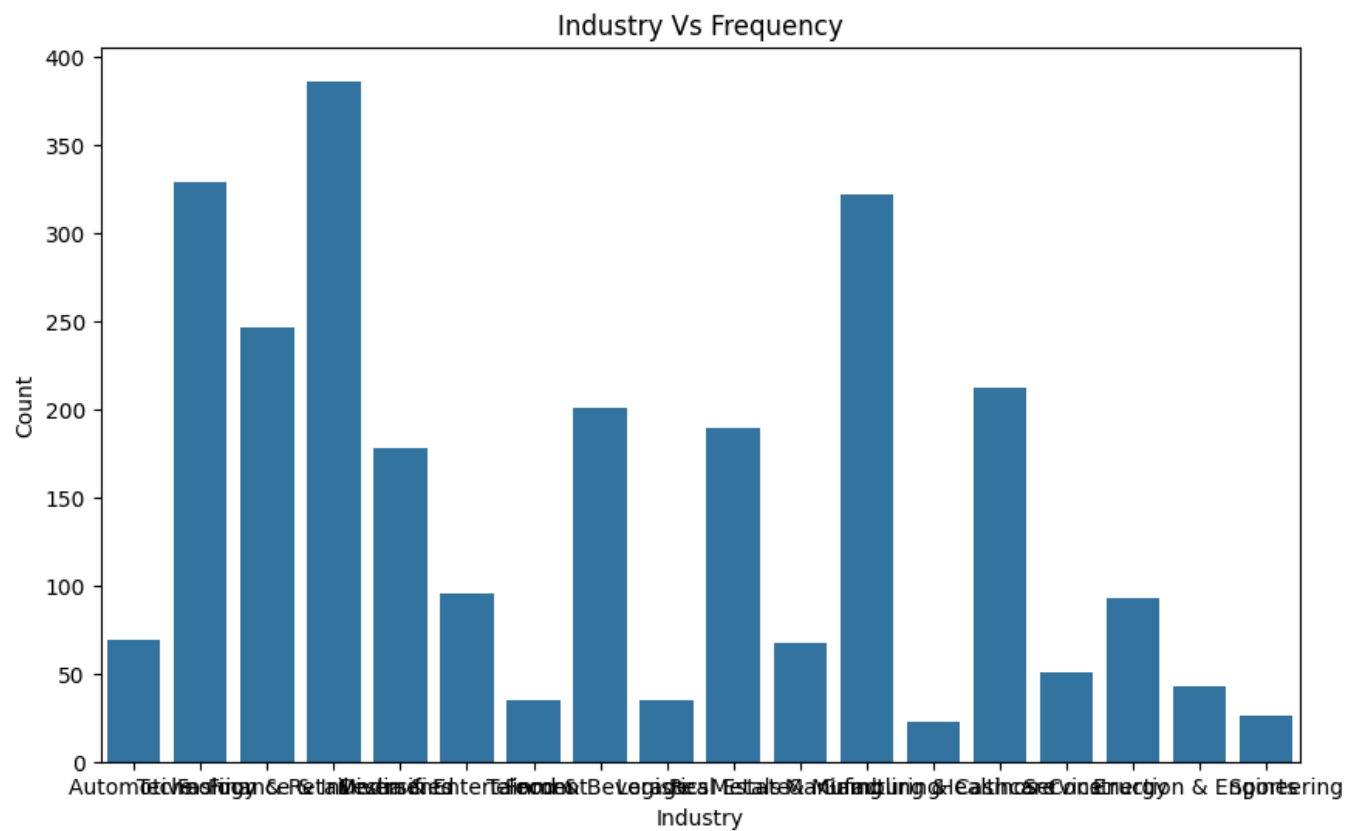
```



```
plt.figure(figsize=(10, 6))
sns.barplot(x='networth', y='industry', data=data)
plt.title('Net Worth Vs Frequency')
plt.xlabel('Net Worth')
plt.ylabel('Frequency')
plt.show()
```

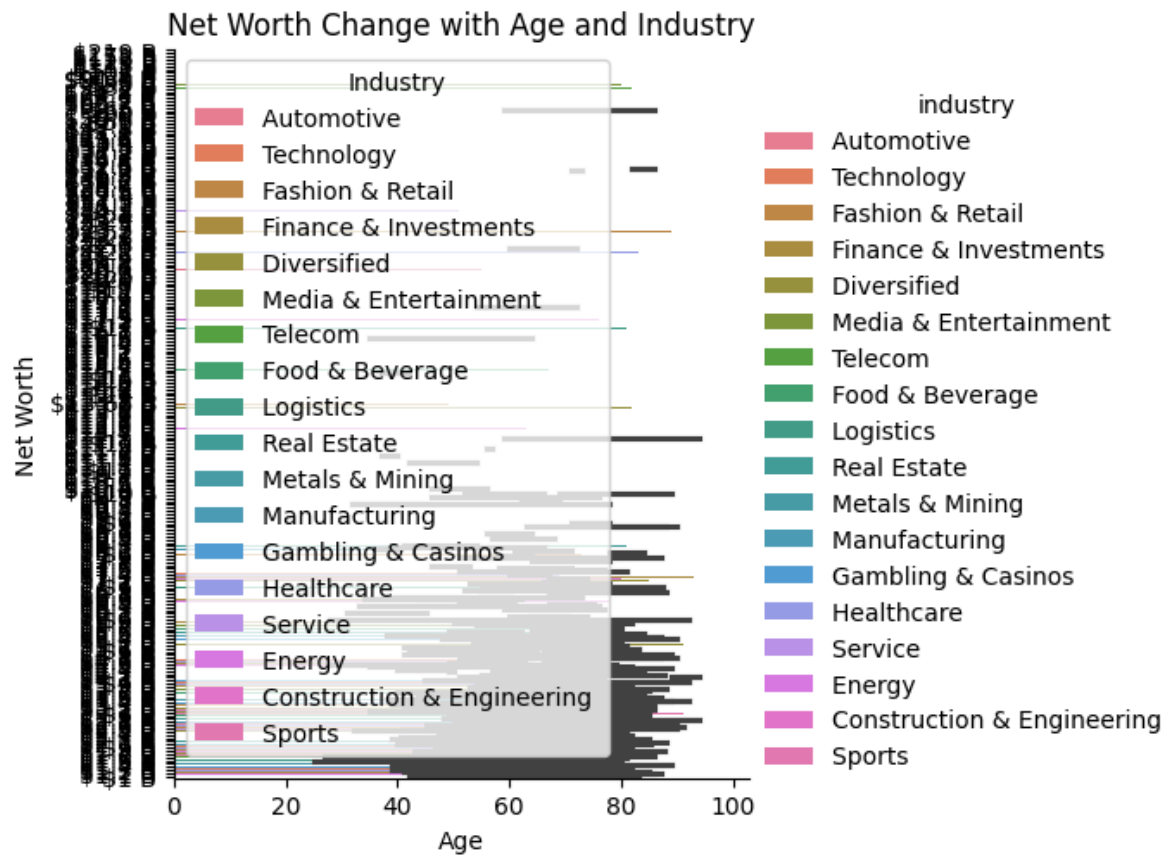


```
plt.figure(figsize=(10, 6))
sns.countplot(x='industry', data=data)
plt.title('Industry Vs Frequency')
plt.xlabel('Industry')
plt.ylabel('Count')
plt.show()
```



```
plt.figure(figsize=(10, 8))
sns.catplot(x='age', y='networth', hue='industry', kind='bar', data=data)
plt.title('Net Worth Change with Age and Industry')
plt.xlabel('Age')
plt.ylabel('Net Worth')
plt.legend(title='Industry')
plt.show()
```

<Figure size 1000x800 with 0 Axes>



```
top_10_rich = data.nlargest(10, 'rank')
plt.figure(figsize=(10, 6))
sns.barplot(x='name', y='rank', data=top_10_rich)
plt.title('Top 10 Richest People')
plt.xlabel('Name')
plt.ylabel('Net Worth ($B)')
plt.xticks(rotation=90)
```

```

[0, 1, 2, 3, 4, 5, 6, 7, 8, 9],
[Text(0, 0, 'Krishna Kumar Bangur '),
Text(1, 0, "Wilbur 'Ed' Bosarge, Jr. "),
Text(2, 0, 'Johanna Braun '),
Text(3, 0, 'Karl Friedrich Braun '),
Text(4, 0, 'Jean-Pierre Cayard '),
Text(5, 0, 'Tony Chen '),
Text(6, 0, 'Vivien Chen '),
Text(7, 0, 'Chey Ki-won '),
Text(8, 0, 'Pollyanna Chu '),
Text(9, 0, 'Roy Chi Ping Chung ')]])

```



```

plt.figure(figsize=(10, 6))
sns.barplot(x='name', y='networth', data=data)
plt.title('Richest People from India')
plt.xlabel('Name')
plt.ylabel('Net Worth')
plt.show()

```

