

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [4]: data = pd.read_csv('Billionaire.csv')
data.head()
```

Out[4]:

	Name	NetWorth	Country	Source	Rank	Age	Industry
0	Jeff Bezos	\$177 B	United States	Amazon	1	57.0	Technology
1	Elon Musk	\$151 B	United States	Tesla, SpaceX	2	49.0	Automotive
2	Bernard Arnault & family	\$150 B	France	LVMH	3	72.0	Fashion & Retail
3	Bill Gates	\$124 B	United States	Microsoft	4	65.0	Technology
4	Mark Zuckerberg	\$97 B	United States	Facebook	5	36.0	Technology

```
In [5]: data.isnull().sum()
```

```
Out[5]: Name      0
NetWorth  0
Country   0
Source    0
Rank      0
Age       0
Industry  0
dtype: int64
```

```
In [6]: data = data.dropna()
```

```
In [7]: data.isnull().sum()
```

```
Out[7]: Name      0
NetWorth  0
Country   0
Source    0
Rank      0
Age       0
Industry  0
dtype: int64
```

```
In [8]: data.columns
```

```
Out[8]: Index(['Name', 'NetWorth', 'Country', 'Source', 'Rank', 'Age', 'Industry'], dtype='object')
```

```
In [10]: data['NetWorth'] = data['NetWorth'].str.strip('$')
data['NetWorth'] = data['NetWorth'].str.strip('B')
```

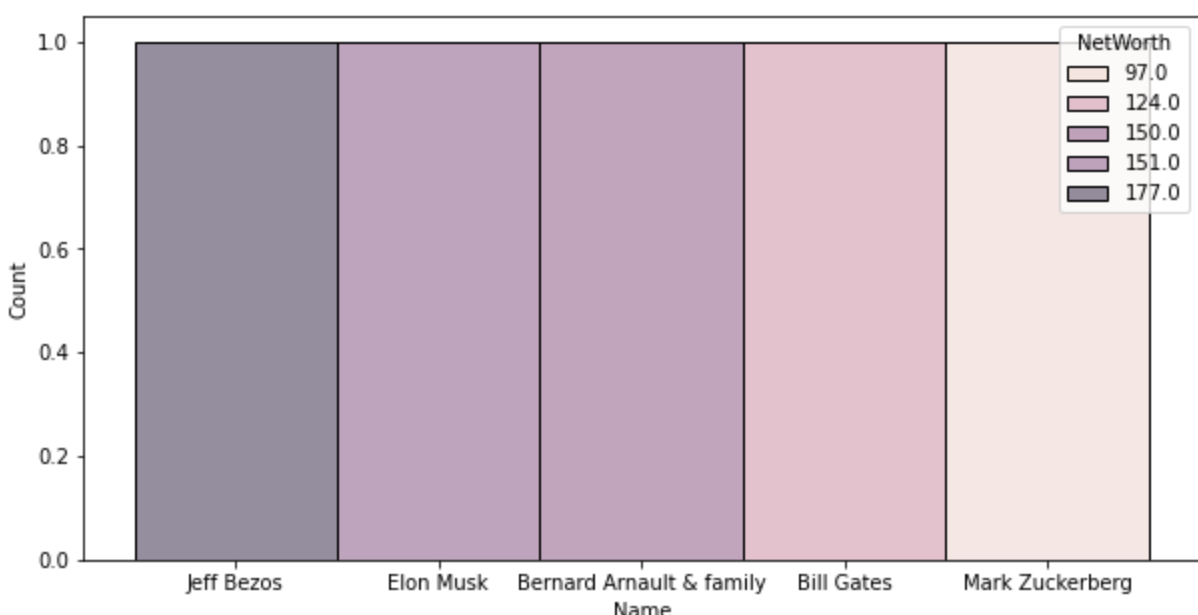
```
In [12]: data['NetWorth'] = data['NetWorth'].astype(float)
```

```
In [13]: data['NetWorth']
```

```
Out[13]: 0      177.0
1      151.0
2      150.0
3      124.0
4       97.0
...
2750     1.0
2751     1.0
2752     1.0
2753     1.0
2754     1.0
Name: NetWorth, Length: 2676, dtype: float64
Top 10 billionaires according to their NetWorth.
```

```
In [17]: df = data.sort_values(by='NetWorth', ascending=False).head()
```

```
In [20]: plt.figure(figsize=(10,5))
sns.histplot(x='Name', hue='NetWorth', data=df)
plt.show()
```



```
In [71]: data.columns
```

```
Out[71]: Index(['Name', 'NetWorth', 'Country', 'Source', 'Rank', 'Age', 'Industry'], dtype='object')
```

```
In [75]: a = data['Source'].value_counts().head()
a
```

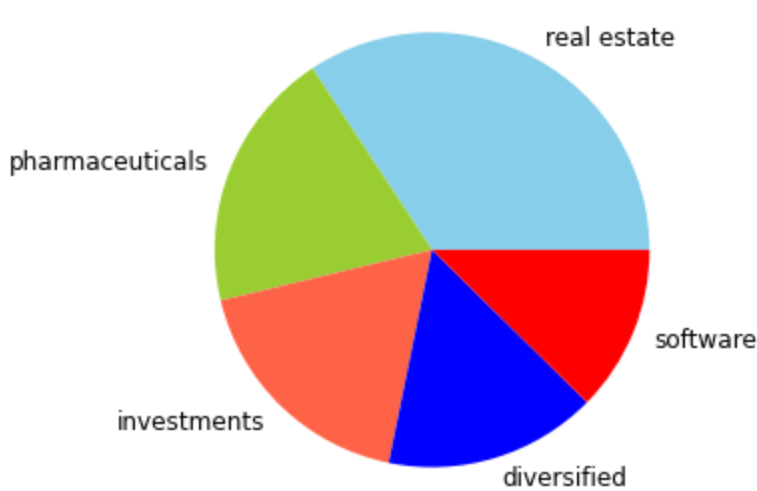
```
Out[75]: real estate      169
pharmaceuticals      96
investments          89
diversified          78
software             61
Name: Source, dtype: int64
```

```
In [76]: index = a.index
```

```
In [77]: Sources = a.values
```

```
In [87]: custom_colors = ['skyblue', 'yellowgreen', 'tomato', 'blue', 'red']
plt.figure(figsize=(5, 5))
plt.pie(Sources, labels=index, colors= custom_colors)
plt.title('Top 5 domains with the most number of billionaires')
plt.show()
```

Top 5 domains with the most number of billionaires



```
In [21]: data.columns
```

```
Out[21]: Index(['Name', 'NetWorth', 'Country', 'Source', 'Rank', 'Age', 'Industry'], dtype='object')
```

```
In [25]: a = data['Industry'].value_counts().head()
```

```
In [32]: a
```

```
Out[32]: Finance & Investments      365
Technology                      355
Manufacturing                   317
Fashion & Retail                 268
Healthcare                     213
Name: Industry, dtype: int64
```

```
In [34]: a.index
```

```
Out[34]: Index(['Finance & Investments', 'Technology', 'Manufacturing',
'Fashion & Retail', 'Healthcare'],
dtype='object')
```

```
In [30]: index = a.index
```

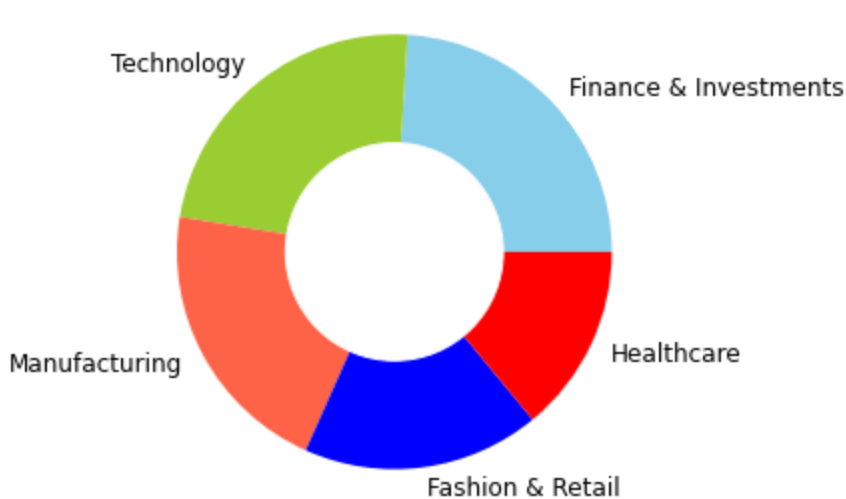
```
In [35]: a.values
```

```
Out[35]: array([365, 355, 317, 268, 213], dtype=int64)
```

```
In [31]: industries = a.values
```

```
In [50]: custom_colors = ['skyblue', 'yellowgreen', 'tomato', 'blue', 'red']
plt.figure(figsize=(5,5))
plt.pie(industries, labels = index, colors = custom_colors)
central_circle = plt.Circle((0,0), 0.5, color='white')
fig = plt.gcf()
fig.gca().add_artist(central_circle)
plt.rc('font', size=12)
plt.title('Top 5 Industries with Most Number of Billionaire')
plt.show()
```

Top 5 Industries with Most Number of Billionaire



```
In [63]: data.columns
```

```
Out[63]: Index(['Name', 'NetWorth', 'Country', 'Source', 'Rank', 'Age', 'Industry'], dtype='object')
```

```
In [64]: a = data['Country'].value_counts().head()
a
```

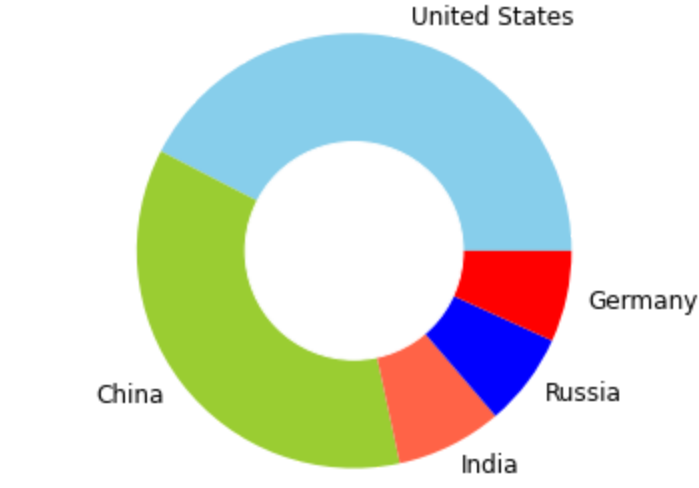
```
Out[64]: United States      720
China                    619
India                    134
Russia                   118
Germany                  115
Name: Country, dtype: int64
```

```
In [65]: index = a.index
```

```
In [68]: Countries = a.values
```

```
In [70]: custom_colors = ['skyblue', 'yellowgreen', 'tomato', 'blue', 'red']
plt.figure(figsize=(5, 5))
plt.pie(Countries, labels = index, colors = custom_colors)
central_circle = plt.Circle((0, 0), 0.5, color='white')
fig = plt.gcf()
fig.gca().add_artist(central_circle)
plt.rc('font', size=12)
plt.title('Top 5 Industries with Most Number of Billionaire')
plt.show()
```

Top 5 Industries with Most Number of Billionaire



```
In [ ]:
```