In [50]:

```
# import the libraries
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

In [51]:

```
# importing the data
data=pd.read_csv("Billionaire.csv")
data.head()
```

Out[51]:

	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 [52]:

```
# Pre processing the data
print(data.isnull().sum())
data=data.dropna()
```

```
Name         0
NetWorth     0
Country      0
Source       0
Rank         0
Age         79
Industry     0
dtype: int64
```

In [53]:

```
print(data.isnull().sum())
```

```
Name        0
NetWorth    0
Country     0
Source      0
Rank        0
Age         0
Industry    0
dtype: int64
```

In [54]:

```
data["NetWorth"]=data["NetWorth"].str.strip("$")
data["NetWorth"]=data["NetWorth"].str.strip("B")
data["NetWorth"]=data["NetWorth"].astype(float)
```

In [55]:

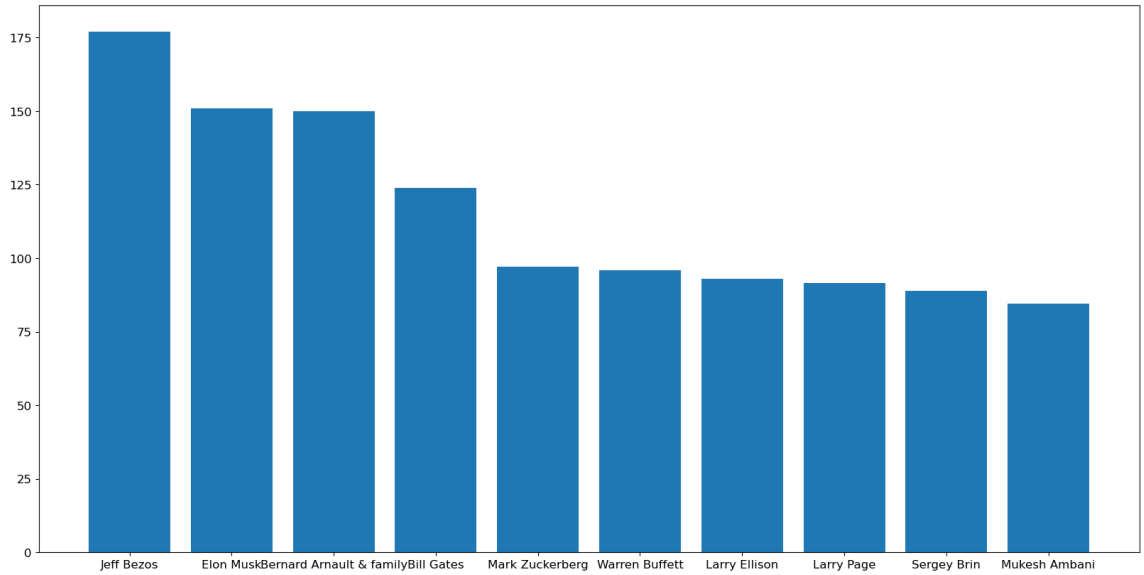
```
data.head()
```

Out[55]:

	Name	NetWorth	Country	Source	Rank	Age	Industry
0	Jeff Bezos	177.0	United States	Amazon	1	57.0	Technology
1	Elon Musk	151.0	United States	Tesla, SpaceX	2	49.0	Automotive
2	Bernard Arnault & family	150.0	France	LVMH	3	72.0	Fashion & Retail
3	Bill Gates	124.0	United States	Microsoft	4	65.0	Technology
4	Mark Zuckerberg	97.0	United States	Facebook	5	36.0	Technology

In [56]:

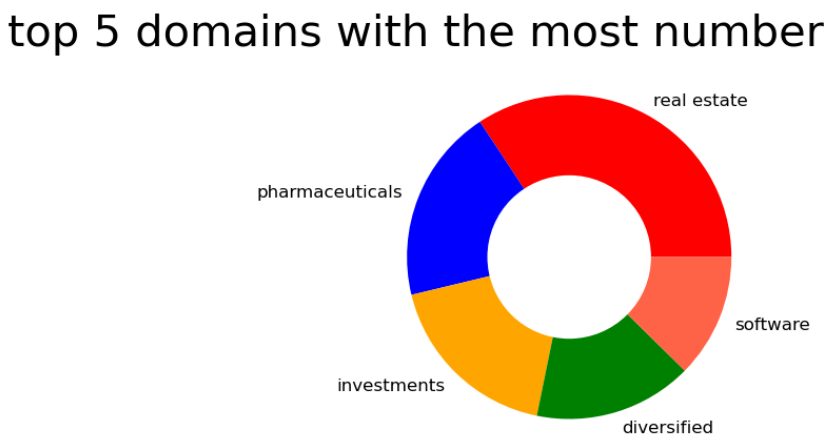
```
df=data.sort_values(by=["NetWorth"],ascending=False).head(10)
plt.figure(figsize=(20,10))
plt.bar(df['Name'],df['NetWorth'])
plt.show()
```



In [62]:

```
a=data["Source"].value_counts().head(5)
index=a.index
print(index)
sources=a.values
print(sources)
col=["red","blue","orange","green","tomato"]
plt.figure(figsize=(5,5))
plt.pie(sources,labels=index,colors=col)
center=plt.Circle((0,0),0.5,color='white')
fig=plt.gcf()
fig.gca().add_artist(center)
plt.rc('font',size=12)
plt.title("top 5 domains with the most number of billionaires",fontsize=30)
plt.show()
```

```
Index(['real estate', 'pharmaceuticals', 'investments', 'diversified',
       'software'],
      dtype='object')
[169  96  89  78  61]
```

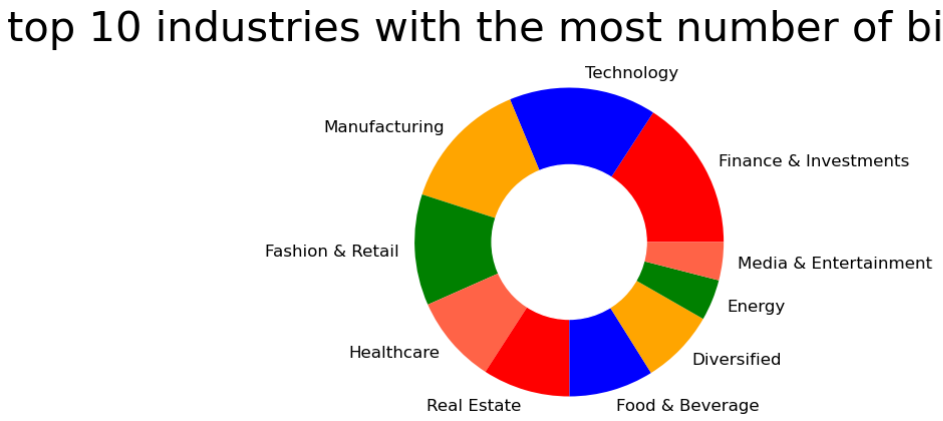


In [61]:

```
# top 10 industries with the most number of billionaries

a=data["Industry"].value_counts().head(10)
index=a.index
print(index)
sources=a.values
print(sources)
col=["red","blue","orange","green","tomato"]
plt.figure(figsize=(5,5))
plt.pie(sources,labels=index,colors=col)
center=plt.Circle((0,0),0.5,color='white')
fig=plt.gcf()
fig.gca().add_artist(center)
plt.rc('font',size=12)
plt.title("top 10 industries with the most number of billionaires",fontsize=30)
plt.show()
```

```
Index(['Finance & Investments', 'Technology', 'Manufacturing',
       'Fashion & Retail', 'Healthcare', 'Real Estate', 'Food & Beverage',
       'Diversified', 'Energy', 'Media & Entertainment'],
      dtype='object')
[365 355 317 268 213 211 204 180  99  92]
```

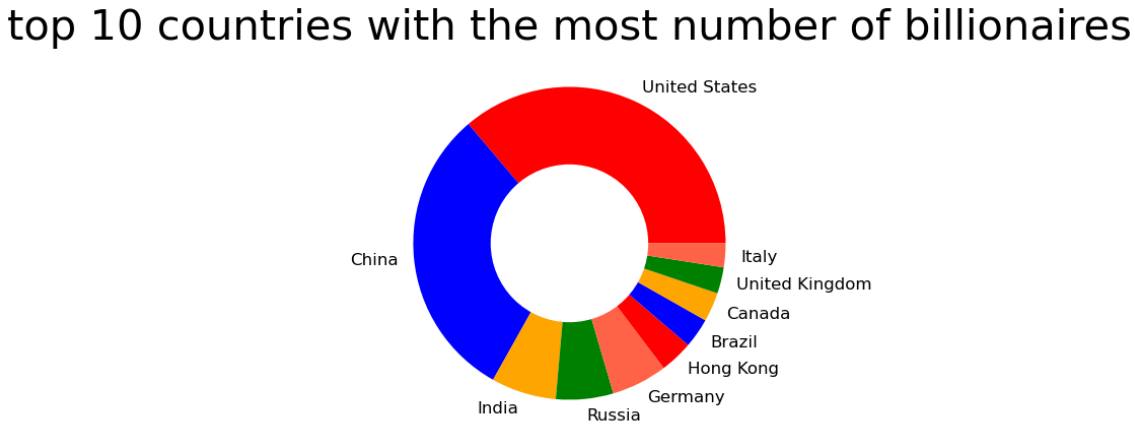


In [60]:

```
# top 10 countries with the most number of billionaires

a=data["Country"].value_counts().head(10)
index=a.index
print(index)
sources=a.values
print(sources)
col=["red","blue","orange","green","tomato"]
plt.figure(figsize=(5,5))
plt.pie(sources,labels=index,colors=col)
center=plt.Circle((0,0),0.5,color='white')
fig=plt.gcf()
fig.gca().add_artist(center)
plt.rc('font',size=12)
plt.title("top 10 countries with the most number of billionaires",fontsize=30)
plt.show()
```

```
Index(['United States', 'China', 'India', 'Russia', 'Germany', 'Hong Kon
g',
       'Brazil', 'Canada', 'United Kingdom', 'Italy'],
      dtype='object')
[720 610 134 118 115  68  61  60  54  49]
```



In [ ]: