Name Nouman Ali

Roll no FA20-BST-009

Overview

Dataset statistics

Number of variables 7 Number of observations 50 Missing cells 0 Missing cells (%) 0.0% Duplicate rows 0 Duplicate rows (%) 0.0% Total size in memory 2.9 KiB Average record size in memory 58.6 B Variable types

Numeric 2 Categorical 4 Boolean 1

```
In [2]:  import pandas as pd
         import matplotlib.pyplot as plt
         import numpy as np
         import seaborn as sns

         data= pd.read_csv('most_followed_twitter.csv')
         data.head(5)
```

Out[2]:

	Rank	Account Name	Account Owner	Brand Account(Y/N)	Followers(in M)	Occupation	Country
0	1	@BarackObama	Barack Obama	N	133.4	44th President of the United States	United States
1	2	@elonmusk	Elon Musk	N	123.3	Business magnate, investor, and owner of Twitter	South Africa
2	3	@justinbieber	Justin Bieber	N	113.6	Musician	Canada
3	4	@katyperry	Katy Perry	N	108.7	Musician	United States
4	5	@rihanna	Rihanna	N	107.4	Musician	Barbados

```
In [3]:  data.tail()
```

Out[3]:

	Rank	Account Name	Account Owner	Brand Account(Y/N)	Followers(in M)	Occupation	Country
45	46	@NiallOfficial	Niall Horan	N	41.1	Musician	Ireland
46	47	@NBA	NBA	Y	41.1	Basketball league	United States
47	48	@KylieJenner	Kylie Jenner	N	40.3	Television personality and businesswoman	United States
48	49	@Drake	Drake	N	39.8	Rapper	Canada
49	50	@BBCWorld	BBC World News	Y	39.0	News channel	United Kingdom

In [4]:
```
data.shape
```

Out[4]:
```
(50, 7)
```

In [5]:
```
data.columns
```

Out[5]:
```
Index(['Rank', 'Account Name', 'Account Owner', 'Brand Account(Y/N)',
       'Followers(in M)', 'Occupation', 'Country'],
      dtype='object')
```

In [6]:
```
data.info
```

```
Out[6]: <bound method DataFrame.info of      Rank      Account Name       Account Owner Bra
        nd Account(Y/N)  \
0        1      @BarackObama         Barack Obama                       N
1        2         @elonmusk            Elon Musk                       N
2        3      @justinbieber       Justin Bieber                       N
3        4         @katyperry          Katy Perry                       N
4        5           @rihanna             Rihanna                       N
5        6         @Cristiano    Cristiano Ronaldo                       N
6        7      @taylorswift13        Taylor Swift                       N
7        8    @realDonaldTrump        Donald Trump                       N
8        9       @narendramodi       Narendra Modi                       N
9       10          @ladygaga           Lady Gaga                       N
10      11           @YouTube             YouTube                       Y
11      12      @TheEllenShow      Ellen DeGeneres                       N
12      13      @KimKardashian      Kim Kardashian                       N
13      14             @NASA                NASA                       Y
14      15      @selenagomez        Selena Gomez                       N
15      16          @Twitter             Twitter                       Y
16      17            @cnnbrk    CNN Breaking News                       Y
17      18       @jtimberlake    Justin Timberlake                       N
18      19          @BillGates          Bill Gates                       N
19      20              @CNN                 CNN                       Y
20      21         @neymarjr              Neymar                       N
21      22      @britneyspears       Britney Spears                       N
22      23          @nytimes    The New York Times                       Y
23      24          @ddlovato          Demi Lovato                       N
24      25          @shakira             Shakira                       N
25      26         @KingJames         LeBron James                       N
26      27          @imVkohli         Virat Kohli                       N
27      28         @PMOIndia            PMO India                       Y
28      29      @BBCBreaking      BBC Breaking News                       Y
29      30       @jimmyfallon        Jimmy Fallon                       N
30      31         @SrBachchan      Amitabh Bachchan                       N
31      32          @BTS_twt                 BTS                       N
32      33         @MileyCyrus          Miley Cyrus                       N
33      34         @realmadrid      Real Madrid CF                       Y
34      35      @ChampionsLeague  UEFA Champions League                   Y
35      36              @JLo        Jennifer Lopez                       N
36      37       @akshaykumar         Akshay Kumar                       N
37      38        @FCBarcelona         FC Barcelona                       Y
38      39    @BeingSalmanKhan          Salman Khan                       N
39      40             @espn                ESPN                       Y
40      41         @BrunoMars           Bruno Mars                       N
41      42            @Oprah        Oprah Winfrey                       N
42      43           @iamsrk        Shah Rukh Khan                       N
43      44        @bts_bighit                 BTS                       N
44      45      @SportsCenter        SportsCenter                       Y
45      46      @NiallOfficial         Niall Horan                       N
46      47              @NBA                 NBA                       Y
47      48       @KylieJenner        Kylie Jenner                       N
48      49            @Drake               Drake                       N
49      50         @BBCWorld      BBC World News                       Y

      Followers(in M)                                    Occupation  \
0              133.4             44th President of the United States
1              123.3   Business magnate, investor, and owner of Twitter
2              113.6                                       Musician
3              108.7                                       Musician
4              107.4                                       Musician
5              106.3                                 Football player
```

6	92.0	Musician
7	87.8	45th President of the United States
8	85.1	Prime Minister of India
9	84.9	Musician and actress
10	77.9	Online video sharing platform
11	76.9	Comedian and television host
12	74.3	Television personality and businesswoman
13	67.6	Space agency
14	66.4	Musician and actress
15	65.0	Social media platform
16	64.0	News channel
17	62.7	Musician and actor
18	61.2	Businessman and philanthropist
19	60.7	News channel
20	59.5	Football player
21	56.0	Musician
22	54.7	Newspaper
23	53.9	Musician and actress
24	53.2	Musician
25	52.4	Basketball player
26	52.2	Cricket player
27	51.6	Office of the Prime Minister of India
28	51.4	News channel
29	51.2	Comedian and television host
30	48.1	Actor
31	47.8	Musicians
32	46.7	Musician and actress
33	45.8	Football club
34	45.7	Football league
35	45.4	Musician and actress
36	45.4	Actor
37	45.2	Football club
38	44.6	Actor
39	43.6	Sports channel
40	43.1	Musician
41	43.0	Television personality
42	42.8	Actor
43	42.2	Musicians
44	41.3	Sports channel
45	41.1	Musician
46	41.1	Basketball league
47	40.3	Television personality and businesswoman
48	39.8	Rapper
49	39.0	News channel

	Country
0	United States
1	South Africa
2	Canada
3	United States
4	Barbados
5	Portugal
6	United States
7	United States
8	India
9	United States
10	United States
11	United States
12	United States
13	United States

```
14    United States
15    United States
16    United States
17    United States
18    United States
19    United States
20           Brazil
21    United States
22    United States
23    United States
24         Colombia
25    United States
26            India
27            India
28   United Kingdom
29    United States
30            India
31      South Korea
32    United States
33            Spain
34           Europe
35    United States
36            India
37            Spain
38            India
39    United States
40    United States
41    United States
42            India
43      South Korea
44    United States
45          Ireland
46    United States
47    United States
48           Canada
49   United Kingdom  >
```

In [8]:
```python
country = data['Country'].value_counts()
country
```

Out[8]:
```
United States     28
India              7
Canada             2
United Kingdom     2
South Korea        2
Spain              2
South Africa       1
Barbados           1
Portugal           1
Brazil             1
Colombia           1
Europe             1
Ireland            1
Name: Country, dtype: int64
```
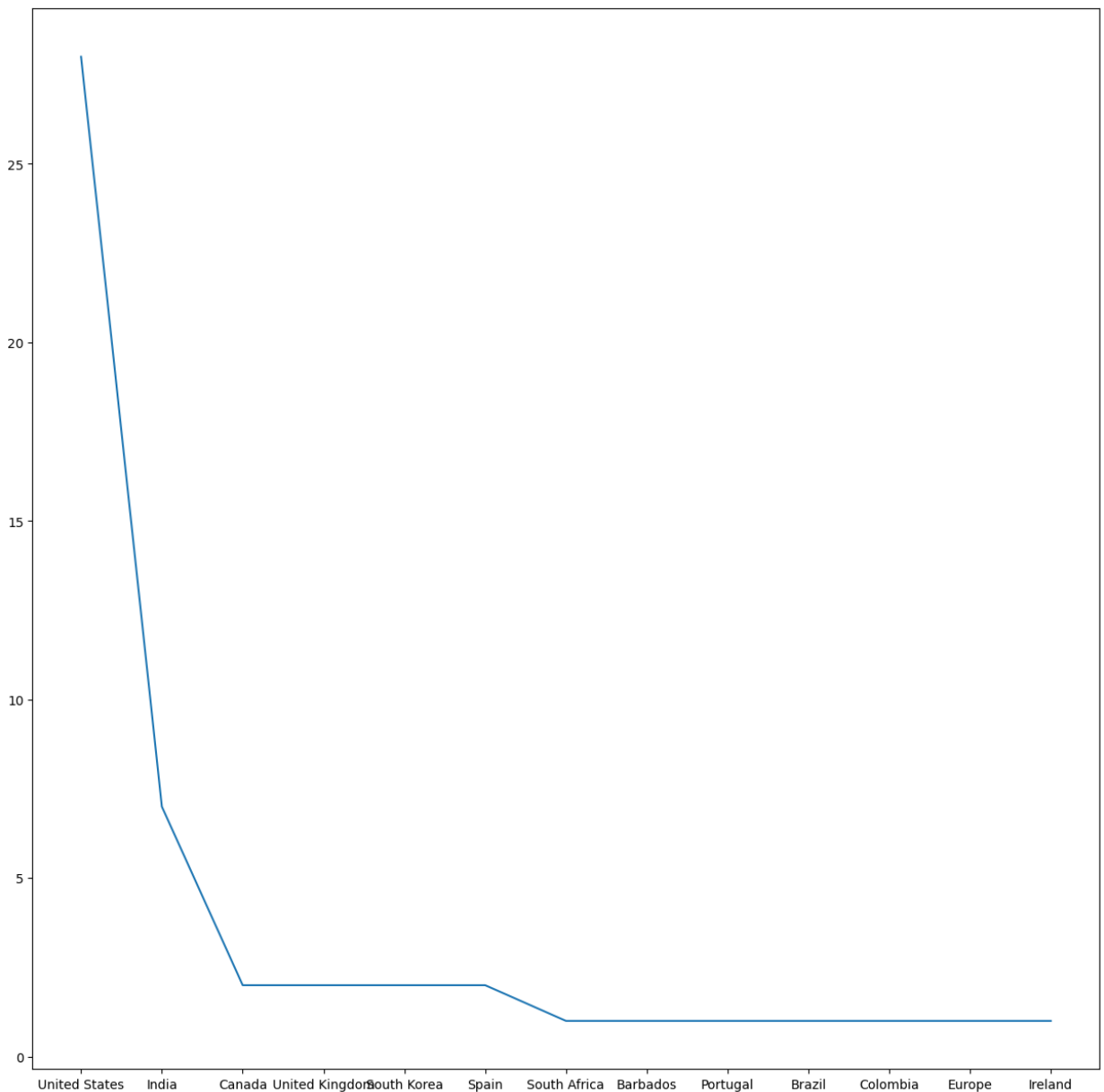
Line plot

In [9]:
```python
print('='*20, 'First Country has more influenceur',"="*20)
plt.figure(figsize=(15,15))
plt.plot(country)
```

==================== First Country has more influenceur ====================

Out[9]:
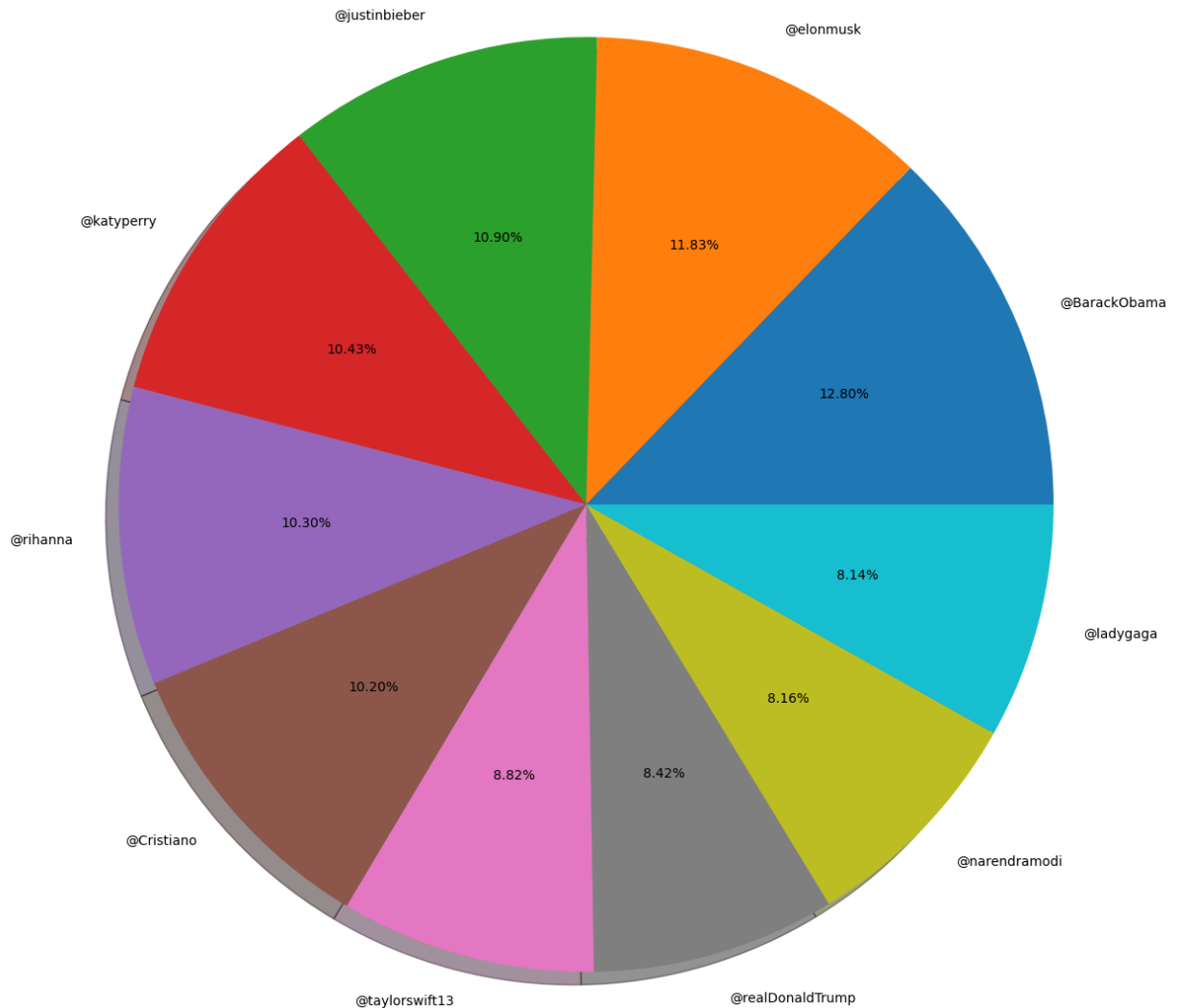[<matplotlib.lines.Line2D at 0x263ea1d8730>]

In [10]:
```python
top10_influenceur= data.drop([ 'Account Owner', 'Brand Account(Y/N)',
          'Occupation', 'Country'], axis=1)
top10=top10_influenceur.head(10)
top10
```

Out[10]:

	Rank	Account Name	Followers(in M)
0	1	@BarackObama	133.4
1	2	@elonmusk	123.3
2	3	@justinbieber	113.6
3	4	@katyperry	108.7
4	5	@rihanna	107.4
5	6	@Cristiano	106.3
6	7	@taylorswift13	92.0
7	8	@realDonaldTrump	87.8
8	9	@narendramodi	85.1
9	10	@ladygaga	84.9

Pi Chart

In [11]:
```python
plt.figure(figsize=(16,16))
plt.pie(top10['Followers(in M)'],shadow = True, labels=top10["Account Name"],autopct =
plt.show()
```
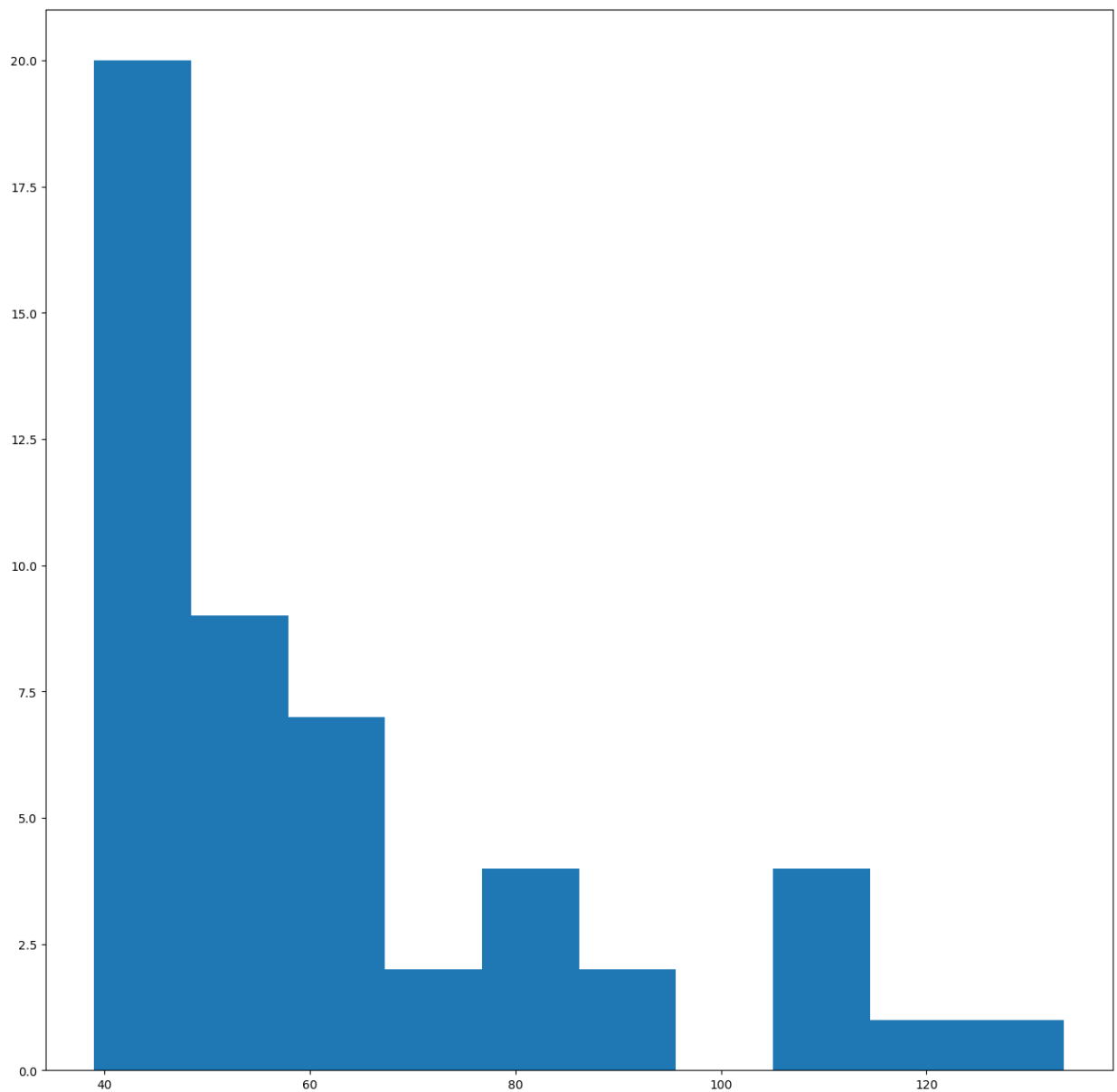
In [12]:
```python
followers_data= data['Followers(in M)']
followers_data.head(5)
```

Out[12]:
```
0    133.4
1    123.3
2    113.6
3    108.7
4    107.4
Name: Followers(in M), dtype: float64
```
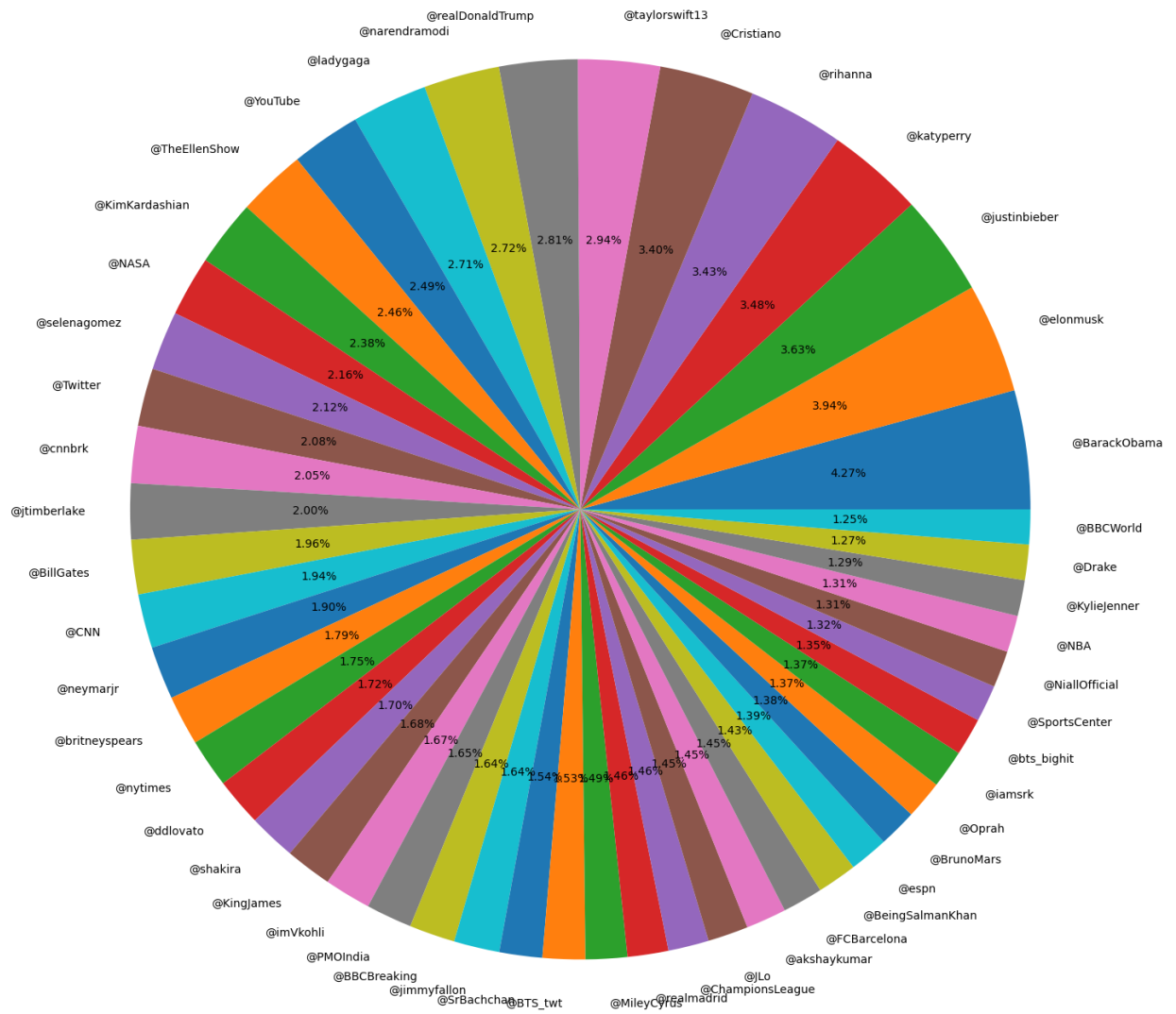
Bar chart

In [13]:
```python
plt.figure(figsize=(16,16))
plt.hist(followers_data)
```

Out[13]:
```
(array([20.,  9.,  7.,  2.,  4.,  2.,  0.,  4.,  1.,  1.]),
 array([ 39.  ,  48.44,  57.88,  67.32,  76.76,  86.2 ,  95.64, 105.08,
        114.52, 123.96, 133.4 ]),
 <BarContainer object of 10 artists>)
```
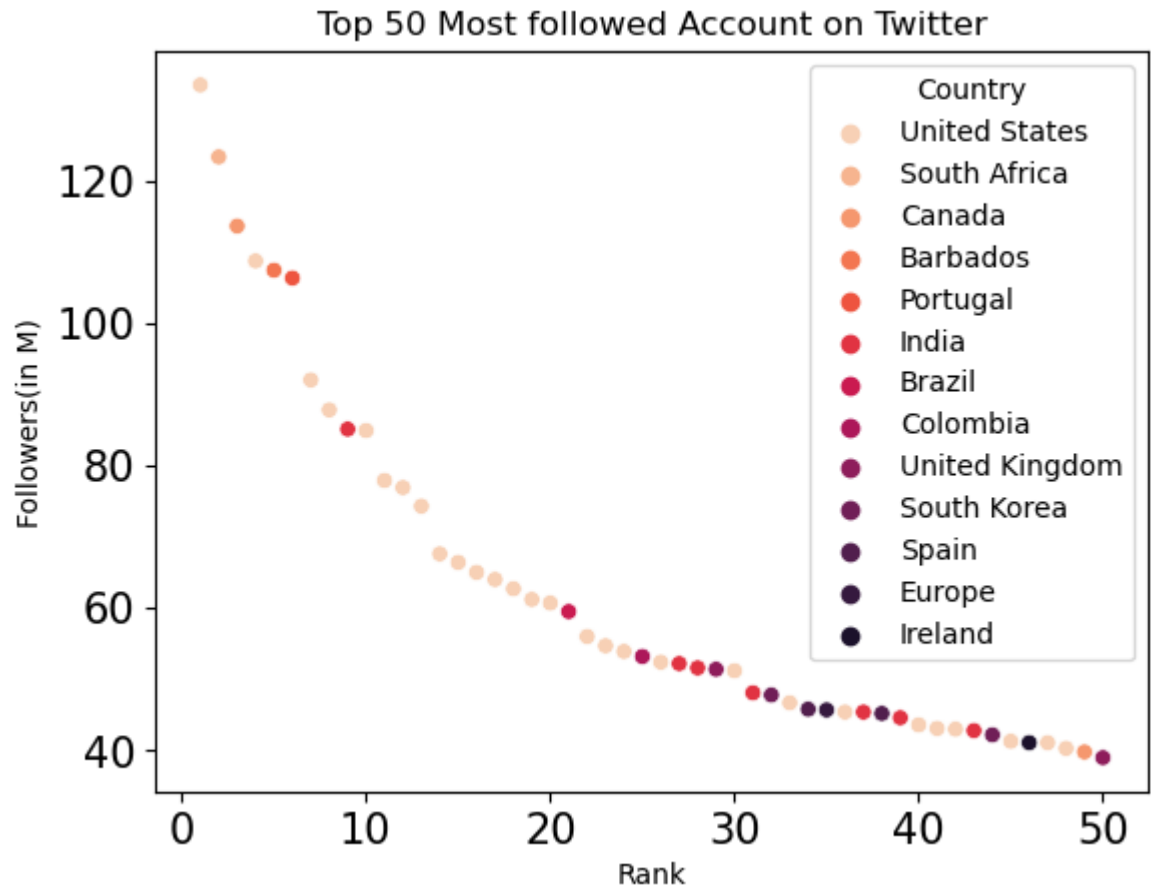
Extensive Pi Chart

```
In [14]:  Name_data= data['Account Name']
          plt.figure(figsize=(18,18))
          plt.pie(followers_data, autopct = "%0.2f%%",labels=Name_data)
          plt.show()
```

@realDonaldTrump @taylorswift13 @Cristiano @rihanna @katyperry @justinbieber @elonmusk @BarackObama @BBCWorld @Drake @KylieJenner @NBA @NiallOfficial @SportsCenter @bts_bighit @iamsrk @Oprah @BrunoMars @espn @BeingSalmanKhan @FCBarcelona @akshaykumar @JLo @ChampionsLeague @realmadrid @MileyCyrus @BTS_twt @SrBachchan @jimmyfallon @BBCBreaking @PMOIndia @imVkohli @KingJames @shakira @ddlovato @nytimes @britneyspears @neymarjr @CNN @BillGates @jtimberlake @cnnbrk @Twitter @selenagomez @NASA @KimKardashian @TheEllenShow @YouTube @ladygaga @narendramodi

2.94% 3.40% 3.43% 3.48% 3.63% 3.94% 4.27% 1.25% 1.27% 1.29% 1.31% 1.31% 1.32% 1.35% 1.37% 1.37% 1.38% 1.39% 1.43% 1.45% 1.45% 1.46% 1.46% 1.49% 1.53% 1.54% 1.64% 1.64% 1.65% 1.67% 1.68% 1.70% 1.72% 1.75% 1.79% 1.90% 1.94% 1.96% 2.00% 2.05% 2.08% 2.12% 2.16% 2.38% 2.46% 2.49% 2.71% 2.72% 2.81%

scatter plot
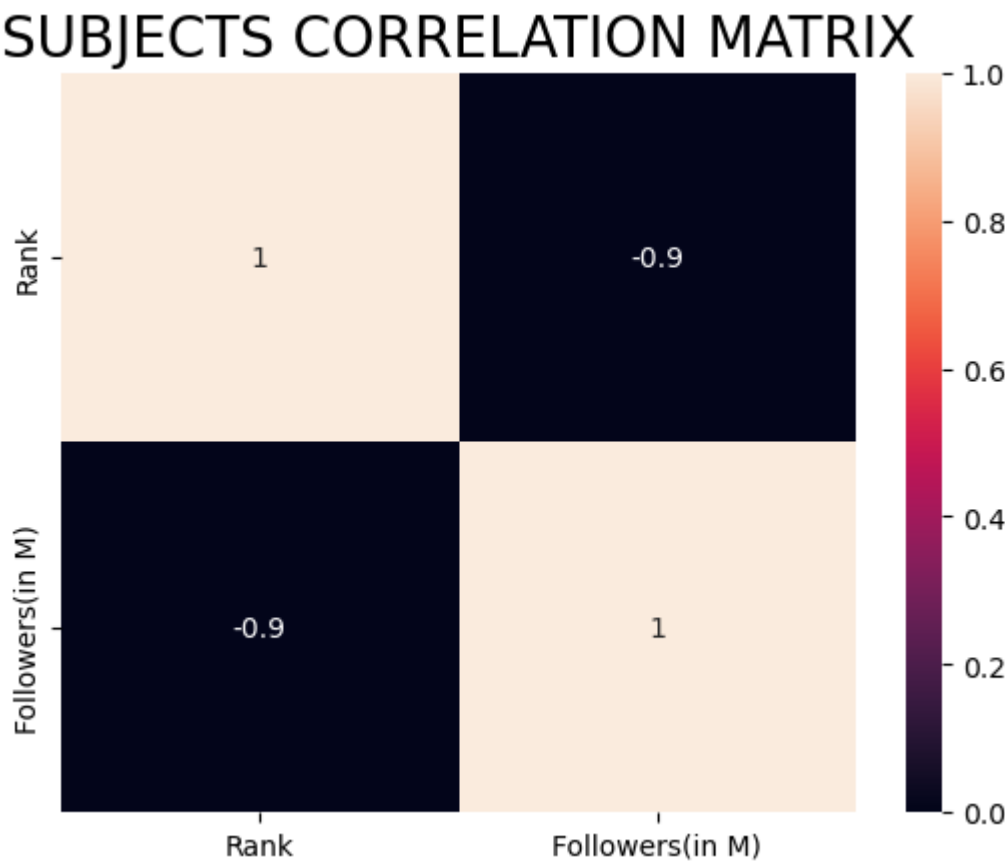
```
In [21]:  import pandas as pd
          import numpy as np
          import seaborn as sns
          import matplotlib.pyplot as plt
          x=pd.read_csv("most_followed_twitter.csv")
          sns.scatterplot(data=x, x = 'Rank' , y = 'Followers(in M)', hue = 'Country'  ,palette
          plt.title("Top 50 Most followed Account on Twitter")
          plt.ylabel('Followers(in M)')
          plt.xlabel('Rank')
          plt.yticks(size=15)
          plt.xticks(size=15)
          plt.show()
```

Top 50 Most followed Account on Twitter

Corelation Chart

```
In [22]:  sns.heatmap(x.corr().round(2),annot=True, vmin = 0)
          plt.title('SUBJECTS CORRELATION MATRIX' , size =20)
```

```
Out[22]:  Text(0.5, 1.0, 'SUBJECTS CORRELATION MATRIX')
```

SUBJECTS CORRELATION MATRIX

Conclusion

Hence we can see that there is a very strong corelation between follower and Ranks.And we can see that increasing the Ranks we have decreased Followers