

Instagram Analysis Using Python

Python data analysis for Instagram is a potent tool for learning about user behavior, industry trends, and content performance. Utilize data scraping from Instagram profiles, posts, and comments to gain insight into the preferences, actions, and interests of your audience. Utilize metrics like follower numbers, engagement rate, and content relevancy to identify possible influencers. This is a list of the top 50 accounts on Instagram, a photo and video sharing social network.

Import Library

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

Uploading Csv file

```
In [ ]: df = pd.read_csv("/content/07_Instagram Analysis.csv")
```

Data Preprocessing

.head()

```
In [ ]: df.head()
```

	Rank	Username	Owner	Followers(millions)	Profession/Activity	Country
0	1	@instagram	Instagram	645.0	Social media platform	United States
1	2	@cristiano	Cristiano Ronaldo	594.0	Footballer	Portugal
2	3	@leomessi	Lionel Messi	476.0	Footballer	Argentina
3	4	@selenagomez	Selena Gomez	423.0	Musician, actress, and businesswoman	United States
4	5	@kyliejenner	Kylie Jenner	395.0	Television personality and businesswoman	United States

.tail()

```
In [ ]: df.tail()
```

	Rank	Username	Owner	Followers(millions)	Profession/Activity	Country
45	46	@snoopdogg	Snoop Dogg	80.0	Musician	United

						States
46	47	@davidbeckham	David Beckham	79.3	Former footballer, president of MLS club Inter...	United Kingdom
47	48	@jennierubyjane	Jennie	79.1	Musician	South Korea
48	49	@khaby00	Khaby Lame	79.1	Social media personality	Italy Senegal
49	50	@gigihadid	Gigi Hadid	78.5	Model	United States

.shape

In []: df.shape

Out[]: (50, 6)

.columns

In []: df.columns

Out[]: Index(['Rank', 'Username', 'Owner', 'Followers(millions)', 'Profession/Activity', 'Country'], dtype='object')

.dtypes

In []: df.dtypes

Out[]:

Rank	int64
Username	object
Owner	object
Followers(millions)	float64
Profession/Activity	object
Country	object
dtype: object	

.unique()

In []: df["Country"].unique()

Out[]: array(['United States', 'Portugal', 'Argentina', 'Canada', 'India', 'Trinidad and Tobago\\xa0United States', 'Brazil', 'Barbados', 'Spain', 'Europe', 'France', 'Israel', 'Thailand', 'United Kingdom\\xa0Albania', 'Colombia', 'United States\\xa0Canada', 'United Kingdom', 'South Korea', 'Italy\\xa0Senegal'], dtype=object)

.nunique()

In []: df.nunique()

Out[]:

Rank	50
Username	50
Owner	50

```
Followers(millions)    47  
Profession/Activity   30  
Country                19  
dtype: int64
```

.describe()

```
In [ ]: df.describe()
```

```
Out[ ]:
```

	Rank	Followers(millions)
count	50.000000	50.000000
mean	25.500000	208.976000
std	14.57738	136.262332
min	1.000000	78.500000
25%	13.250000	96.850000
50%	25.500000	161.000000
75%	37.750000	288.000000
max	50.000000	645.000000

.value_counts()

```
In [ ]: df["Country"].value_counts()
```

```
Out[ ]:
```

Country	count
United States	28
India	3
Canada	2
Spain	2
Israel	1
South Korea	1
United Kingdom	1
United States Canada	1
Colombia	1
United Kingdom Albania	1
Thailand	1
Europe	1
France	1
Portugal	1
Barbados	1
Brazil	1
Trinidad and Tobago United States	1
Argentina	1
Italy Senegal	1

Name: count, dtype: int64

.isnull

```
In [ ]: df.isnull()
```

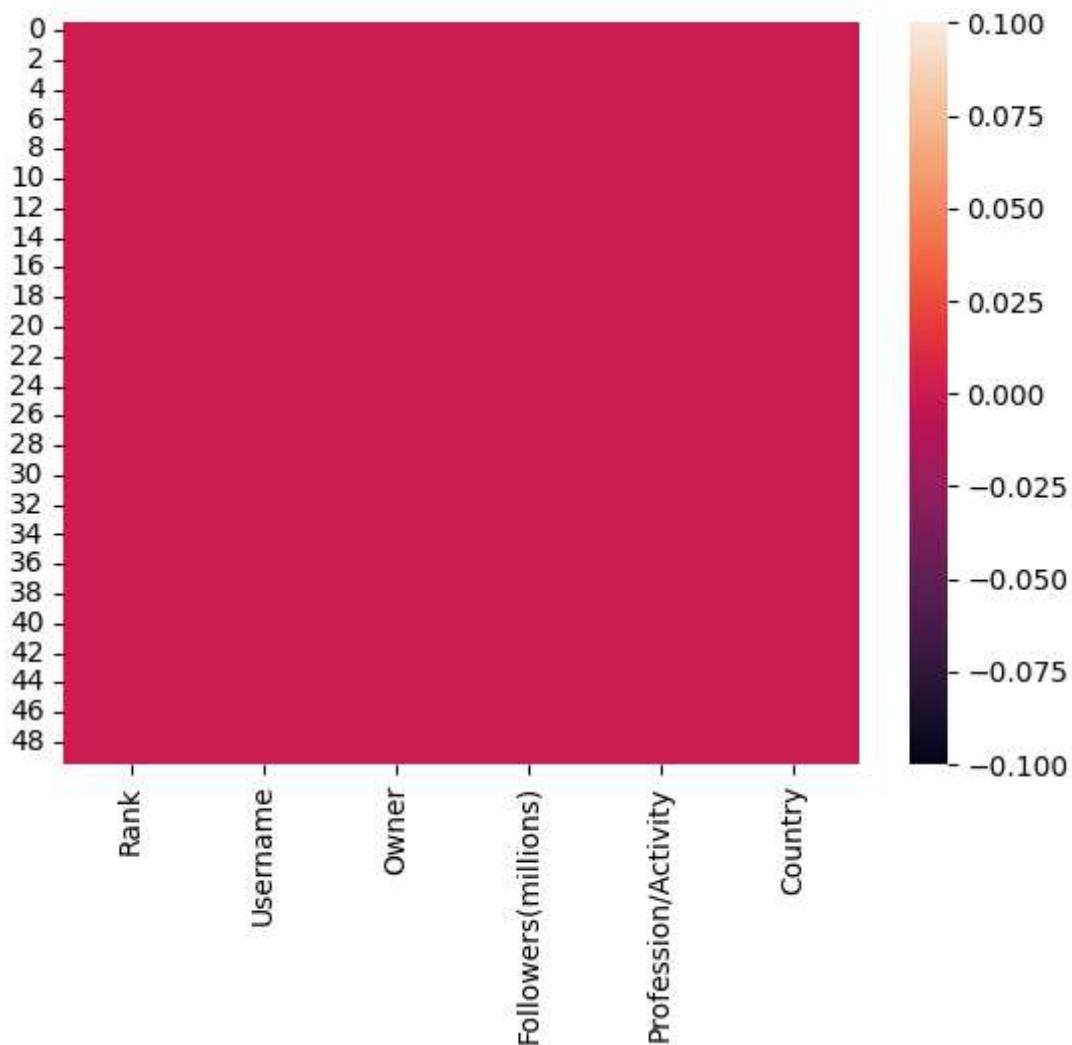
```
Out[ ]:
```

	Rank	Username	Owner	Followers(millions)	Profession/Activity	Country
0	False	False	False	False	False	False

42	False	False	False	False	False	False
43	False	False	False	False	False	False
44	False	False	False	False	False	False
45	False	False	False	False	False	False
46	False	False	False	False	False	False
47	False	False	False	False	False	False
48	False	False	False	False	False	False
49	False	False	False	False	False	False

In []: sns.heatmap(df.isnull())

Out[]: <Axes: >



In []: df["Followers(millions)"] = pd.to_numeric(df["Followers(millions)"], errors="coerce")
print(df)

	Rank	Username	Owner	Followers(millions)
0	1	@instagram	Instagram	645.0
1	2	@cristiano	Cristiano Ronaldo	594.0
2	3	@leomessi	Lionel Messi	476.0
3	4	@selenagomez	Selena Gomez	423.0
4	5	@kyliejenner	Kylie Jenner	395.0
5	6	@therock	Dwayne Johnson	385.0
6	7	@arianagrande	Ariana Grande	375.0
7	8	@kimkardashian	Kim Kardashian	360.0

8	9	@beyonce	Beyoncé	312.0
9	10	@khloekardashian	Khloé Kardashian	309.0
10	11	@nike	Nike	299.0
11	12	@justinbieber	Justin Bieber	292.0
12	13	@kendalljenner	Kendall Jenner	291.0
13	14	@natgeo	National Geographic	279.0
14	15	@taylorswift	Taylor Swift	264.0
15	16	@virat.kohli	Virat Kohli	252.0
16	17	@jlo	Jennifer Lopez	248.0
17	18	@kourtneykardash	Kourtney Kardashian	222.0
18	19	@nickiminaj	Nicki Minaj	222.0
19	20	@mileycyrus	Miley Cyrus	211.0
20	21	@neymarjr	Neymar	209.0
21	22	@katyperry	Katy Perry	201.0
22	23	@zendaya	Zendaya	182.0
23	24	@kevinhart4real	Kevin Hart	176.0
24	25	@iamcardib	Cardi B	166.0
25	26	@ddlovato	Demi Lovato	156.0
26	27	@kingjames	LeBron James	155.0
27	28	@badgalriri	Rihanna	150.0
28	29	@chrisbrownofficial	Chris Brown	141.0
29	30	@realmadrid	Real Madrid CF	139.0
30	31	@ellendegeneres	Ellen DeGeneres	137.0
31	32	@champagnepapi	Drake	139.0
32	33	@fcbarcelona	FC Barcelona	121.0
33	34	@billieeilish	Billie Eilish	109.0
34	35	@championsleague	UEFA Champions League	105.0
35	36	@k.mبابپے	Kylian Mbappé	104.0
36	37	@gal_gadot	Gal Gadot	103.0
37	38	@vindiesel	Vin Diesel	94.8
38	39	@lalalalisa_m	Lisa	94.1
39	40	@nasa	NASA	92.7
40	41	@dualipa	Dua Lipa	88.3
41	42	@priyankachopra	Priyanka Chopra	87.6
42	43	@shakira	Shakira	86.6
43	44	@nba	NBA	80.9
44	45	@shraddhakapoor	Shraddha Kapoor	80.8
45	46	@snoopdogg	Snoop Dogg	80.0
46	47	@davidbeckham	David Beckham	79.3
47	48	@jennierubyjane	Jennie	79.1
48	49	@khaby00	Khaby Lame	79.1
49	50	@gigihadid	Gigi Hadid	78.5

Profession/Activity \

0 Social media platform
1 Footballer
2 Footballer
3 Musician, actress, and businesswoman
4 Television personality and businesswoman
5 Actor and professional wrestler
6 Musician, actress and businesswoman
7 Television personality, model and businesswoman
8 Musician and businesswoman
9 Television personality and model
10 Sportswear multinational
11 Musician
12 Model and television personality
13 Magazine
14 Musician
15 Cricketer
16 Musician and actress
17 Television personality and model
18 Musician
19 Musician and actress
20 Footballer
21 Musician

22	Actress and singer
23	Comedian and actor
24	Musician and actress
25	Musician and actress
26	Basketball player
27	Musician and businesswoman
28	Musician
29	Football club
30	Comedian and television personality
31	Musician
32	Football club
33	Musician
34	Club football competition
35	Footballer at Paris Saint-Germain
36	Actress
37	Actor
38	Musician
39	Space agency
40	Musician
41	Actress and musician
42	Musician
43	Professional basketball league
44	Actress
45	Musician
46	Former footballer, president of MLS club Inter...
47	Musician
48	Social media personality
49	Model

	Country
0	United States
1	Portugal
2	Argentina
3	United States
4	United States
5	United States
6	United States
7	United States
8	United States
9	United States
10	United States
11	Canada
12	United States
13	United States
14	United States
15	India
16	United States
17	United States
18	Trinidad and Tobago
19	United States
20	Brazil
21	United States
22	United States
23	United States
24	United States
25	United States
26	United States
27	Barbados
28	United States
29	Spain
30	United States
31	Canada
32	Spain
33	United States
34	Europe
35	France

```

36           Israel
37           United States
38           Thailand
39           United States
40           United Kingdom Albania
41           India
42           Colombia
43           United States Canada
44           India
45           United States
46           United Kingdom
47           South Korea
48           Italy Senegal
49           United States

```

```
In [ ]: df_usa = df[df['Country'] == 'United States']
print(df)
```

	Rank	Username	Owner	Followers(millions)	\
0	1	@instagram	Instagram	645.0	
1	2	@cristiano	Cristiano Ronaldo	594.0	
2	3	@leomessi	Lionel Messi	476.0	
3	4	@selenagomez	Selena Gomez	423.0	
4	5	@kyliejenner	Kylie Jenner	395.0	
5	6	@therock	Dwayne Johnson	385.0	
6	7	@arianagrande	Ariana Grande	375.0	
7	8	@kimkardashian	Kim Kardashian	360.0	
8	9	@beyonce	Beyoncé	312.0	
9	10	@khloekardashian	Khloé Kardashian	309.0	
10	11	@nike	Nike	299.0	
11	12	@justinbieber	Justin Bieber	292.0	
12	13	@kendalljenner	Kendall Jenner	291.0	
13	14	@natgeo	National Geographic	279.0	
14	15	@taylorswift	Taylor Swift	264.0	
15	16	@virat.kohli	Virat Kohli	252.0	
16	17	@jlo	Jennifer Lopez	248.0	
17	18	@kourtneykardash	Kourtney Kardashian	222.0	
18	19	@nickiminaj	Nicki Minaj	222.0	
19	20	@mileycyrus	Miley Cyrus	211.0	
20	21	@neymarjr	Neymar	209.0	
21	22	@katyperry	Katy Perry	201.0	
22	23	@zendaya	Zendaya	182.0	
23	24	@kevinhart4real	Kevin Hart	176.0	
24	25	@iamcardib	Cardi B	166.0	
25	26	@ddlovato	Demi Lovato	156.0	
26	27	@kingjames	LeBron James	155.0	
27	28	@badgalriri	Rihanna	150.0	
28	29	@chrisbrownofficial	Chris Brown	141.0	
29	30	@realmadrid	Real Madrid CF	139.0	
30	31	@ellendegeneres	Ellen DeGeneres	137.0	
31	32	@champagnepapi	Drake	139.0	
32	33	@fcbarcelona	FC Barcelona	121.0	
33	34	@billieeilish	Billie Eilish	109.0	
34	35	@championsleague	UEFA Champions League	105.0	
35	36	@k.mbappe	Kylian Mbappé	104.0	
36	37	@gal_gadot	Gal Gadot	103.0	
37	38	@vindiesel	Vin Diesel	94.8	
38	39	@lalalalisa_m	Lisa	94.1	
39	40	@nasa	NASA	92.7	
40	41	@dualipa	Dua Lipa	88.3	
41	42	@priyankachopra	Priyanka Chopra	87.6	
42	43	@shakira	Shakira	86.6	
43	44	@nba	NBA	80.9	
44	45	@shraddhakapoor	Shraddha Kapoor	80.8	
45	46	@snoopdogg	Snoop Dogg	80.0	
46	47	@davidbeckham	David Beckham	79.3	

47	48	@jennierubyjane	Jennie	79.1
48	49	@khaby00	Khaby Lame	79.1
49	50	@gigihadid	Gigi Hadid	78.5

			Profession/Activity	\
0			Social media platform	
1			Footballer	
2			Footballer	
3			Musician, actress, and businesswoman	
4			Television personality and businesswoman	
5			Actor and professional wrestler	
6			Musician, actress and businesswoman	
7			Television personality, model and businesswoman	
8			Musician and businesswoman	
9			Television personality and model	
10			Sportswear multinational	
11			Musician	
12			Model and television personality	
13			Magazine	
14			Musician	
15			Cricketer	
16			Musician and actress	
17			Television personality and model	
18			Musician	
19			Musician and actress	
20			Footballer	
21			Musician	
22			Actress and singer	
23			Comedian and actor	
24			Musician and actress	
25			Musician and actress	
26			Basketball player	
27			Musician and businesswoman	
28			Musician	
29			Football club	
30			Comedian and television personality	
31			Musician	
32			Football club	
33			Musician	
34			Club football competition	
35			Footballer at Paris Saint-Germain	
36			Actress	
37			Actor	
38			Musician	
39			Space agency	
40			Musician	
41			Actress and musician	
42			Musician	
43			Professional basketball league	
44			Actress	
45			Musician	
46		Former footballer, president of MLS club	Inter...	
47			Musician	
48			Social media personality	
49			Model	

	Country
0	United States
1	Portugal
2	Argentina
3	United States
4	United States
5	United States
6	United States
7	United States
8	United States

```
9          United States
10         United States
11             Canada
12         United States
13         United States
14         United States
15             India
16         United States
17         United States
18 Trinidad and Tobago United States
19         United States
20             Brazil
21         United States
22         United States
23         United States
24         United States
25         United States
26         United States
27             Barbados
28         United States
29             Spain
30         United States
31             Canada
32             Spain
33         United States
34             Europe
35             France
36             Israel
37         United States
38             Thailand
39         United States
40         United Kingdom Albania
41             India
42             Colombia
43         United States Canada
44             India
45         United States
46         United Kingdom
47             South Korea
48             Italy Senegal
49         United States
```

Visualization

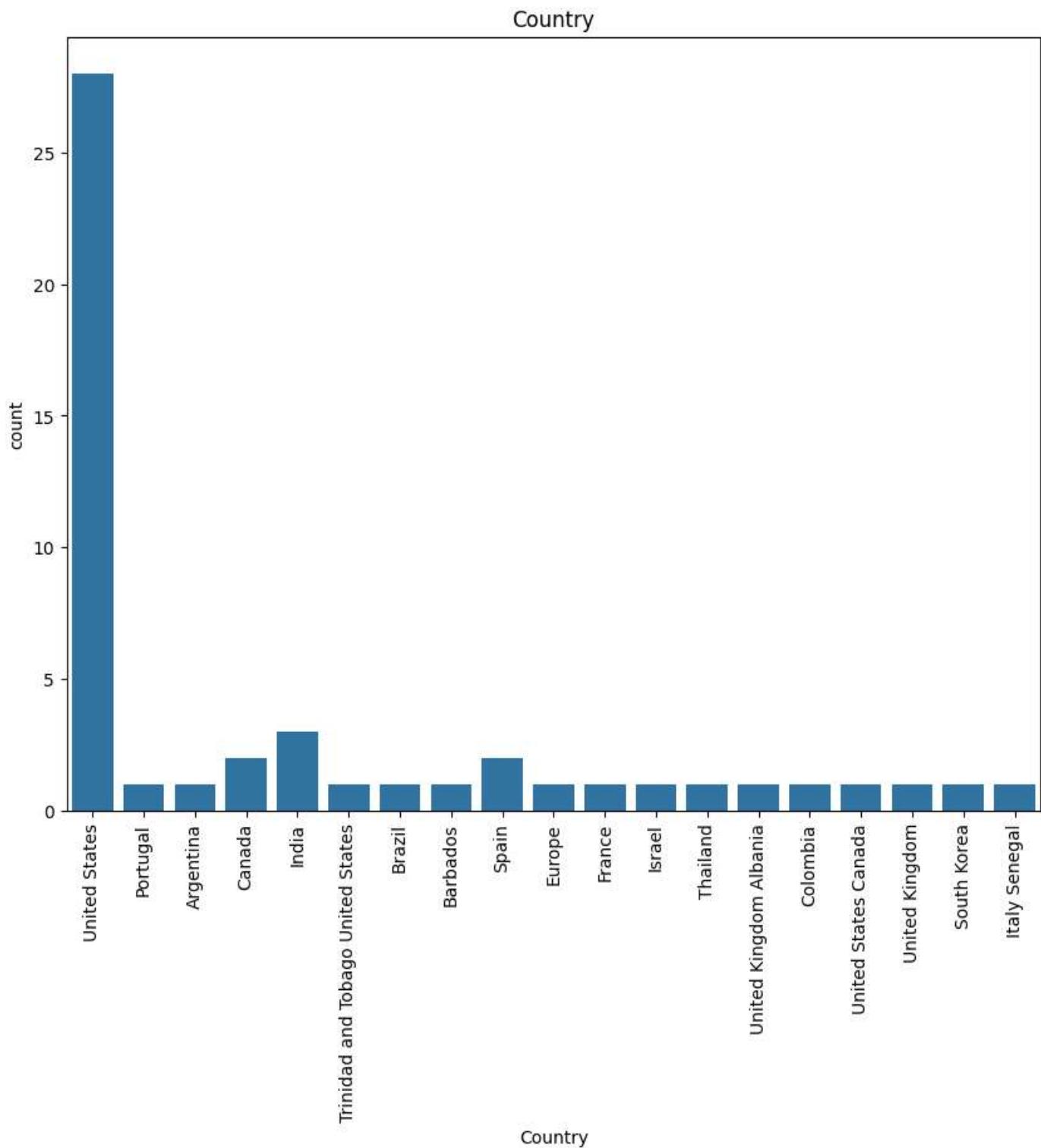
```
In [ ]: Country = df['Country'].value_counts()
```

```
In [ ]: Country
```

```
Out[ ]: Country
United States      28
India              3
Canada             2
Spain              2
Israel             1
South Korea        1
United Kingdom     1
United States Canada 1
Colombia           1
United Kingdom Albania 1
Thailand            1
Europe              1
France              1
Portugal            1
```

```
Barbados          1
Brazil            1
Trinidad and Tobago United States    1
Argentina         1
Italy             1
Senegal           1
Name: count, dtype: int64
```

```
In [ ]: plt.figure(figsize=(10,8))
plt.title("Country")
sns.countplot(x=df["Country"]);
plt.xticks(rotation=90);
```



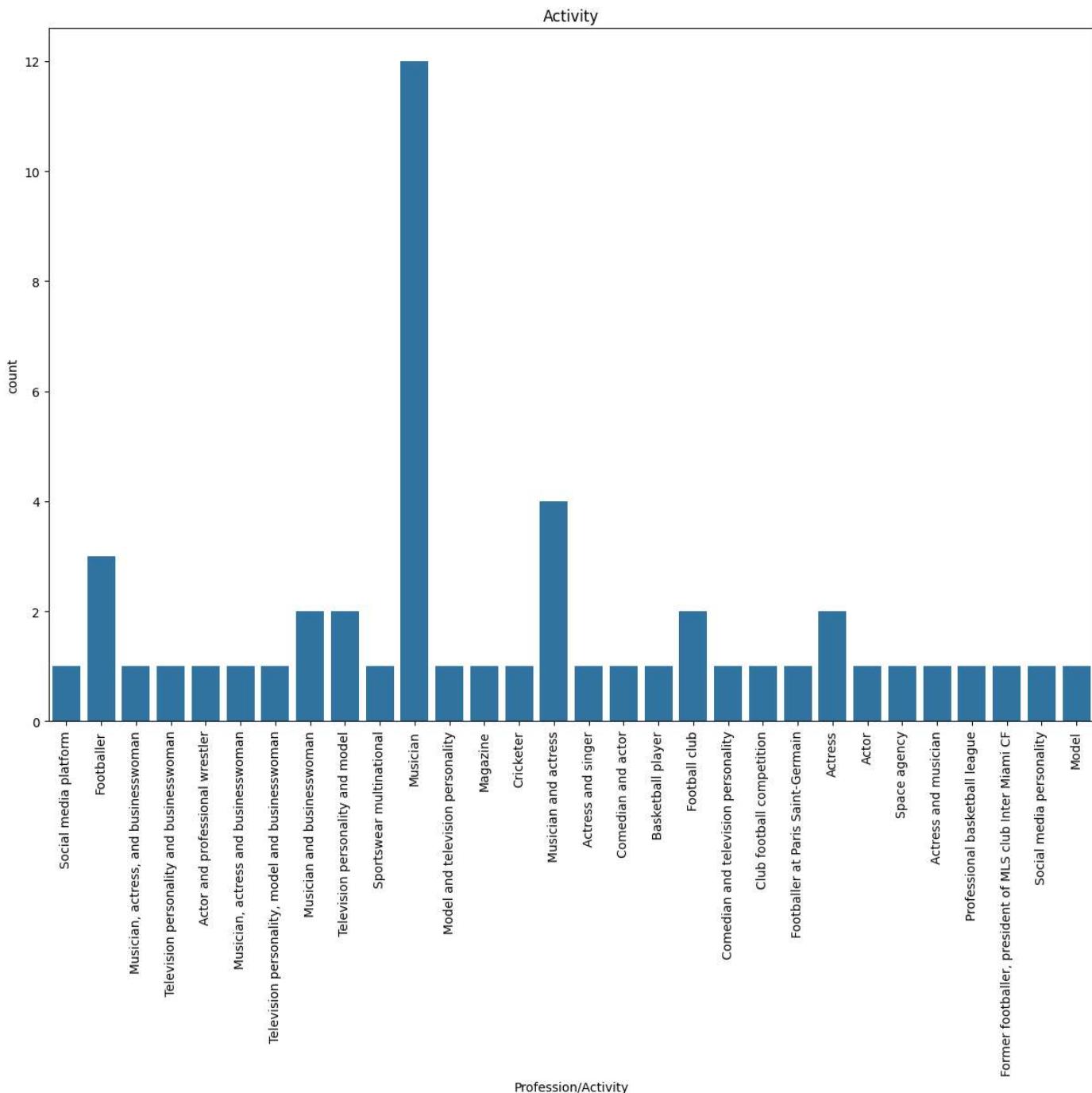
```
In [ ]: Profession = df['Profession/Activity'].value_counts()
Profession
```

```
Out[ ]: Profession/Activity
Musician
```

Musician and actress	4
Footballer	3
Television personality and model	2
Actress	2
Musician and businesswoman	2
Football club	2
Club football competition	1
Footballer at Paris Saint-Germain	1
Social media platform	1
Actor	1
Space agency	1
Actress and musician	1
Professional basketball league	1
Former footballer, president of MLS club Inter Miami CF	1
Social media personality	1
Comedian and television personality	1
Actress and singer	1
Basketball player	1
Comedian and actor	1
Cricketer	1
Magazine	1
Model and television personality	1
Sportswear multinational	1
Television personality, model and businesswoman	1
Musician, actress and businesswoman	1
Actor and professional wrestler	1
Television personality and businesswoman	1
Musician, actress, and businesswoman	1
Model	1

Name: count, dtype: int64

```
In [ ]: plt.subplots(figsize=(15,10))
plt.title("Activity")
sns.countplot(x=df["Profession/Activity"]);
plt.xticks(rotation=90);
```



```
In [ ]: top10_influenceur= df.drop(['Owner', 'Profession/Activity', 'Country'], axis=1)
top10=top10_influenceur.head(10)
```

```
In [ ]: top10
```

	Rank	Username	Followers(millions)
0	1	@instagram	645.0
1	2	@cristiano	594.0
2	3	@leomessi	476.0
3	4	@selenagomez	423.0
4	5	@kyliejenner	395.0
5	6	@therock	385.0
6	7	@arianagrande	375.0
7	8	@kimkardashian	360.0
8	9	@beyonce	312.0

```
In [ ]: plt.figure(figsize=(10,10))
plt.pie(top10['Followers(millions)'], shadow = True, labels=top10["Username"], autopct = "%1.2f%%")
plt.show()
```

