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

In [28]:
```
df=pd.read_csv(r"C:\Users\Admin\Downloads\5_Instagram data - 5_Instagram data.c
df
```

Out[28]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	F
0	3920	2586	1028	619	56	98	9	5	162	35	
1	5394	2727	1838	1174	78	194	7	14	224	48	
2	4021	2085	1188	0	533	41	11	1	131	62	
3	4528	2700	621	932	73	172	10	7	213	23	
4	2518	1704	255	279	37	96	5	4	123	8	
...	...	...	...	...	...	...	...	...	...	...	
114	13700	5185	3041	5352	77	573	2	38	373	73	
115	5731	1923	1368	2266	65	135	4	1	148	20	
116	4139	1133	1538	1367	33	36	0	1	92	34	
117	32695	11815	3147	17414	170	1095	2	75	549	148	

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	F
118	36919	13473	4176	16444	2547	653	5	26	443	611	

119 rows × 13 columns

In [29]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 119 entries, 0 to 118
Data columns (total 13 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Impressions     119 non-null    int64 
 1   From Home       119 non-null    int64 
 2   From Hashtags   119 non-null    int64 
 3   From Explore    119 non-null    int64 
 4   From Other      119 non-null    int64 
 5   Saves           119 non-null    int64 
 6   Comments        119 non-null    int64 
 7   Shares          119 non-null    int64 
 8   Likes           119 non-null    int64 
 9   Profile Visits  119 non-null    int64 
 10  Follows         119 non-null    int64 
 11  Caption         119 non-null    object
 12  Hashtags        119 non-null    object
dtypes: int64(11), object(2)
memory usage: 12.2+ KB
```

In [30]:
```
df.columns
```

Out[30]:
```
Index(['Impressions', 'From Home', 'From Hashtags', 'From Explore',
       'From Other', 'Saves', 'Comments', 'Shares', 'Likes', 'Profile Visit
s',
       'Follows', 'Caption', 'Hashtags'],
      dtype='object')
```

In [31]:
```
df1=df.head(100)
df1
```

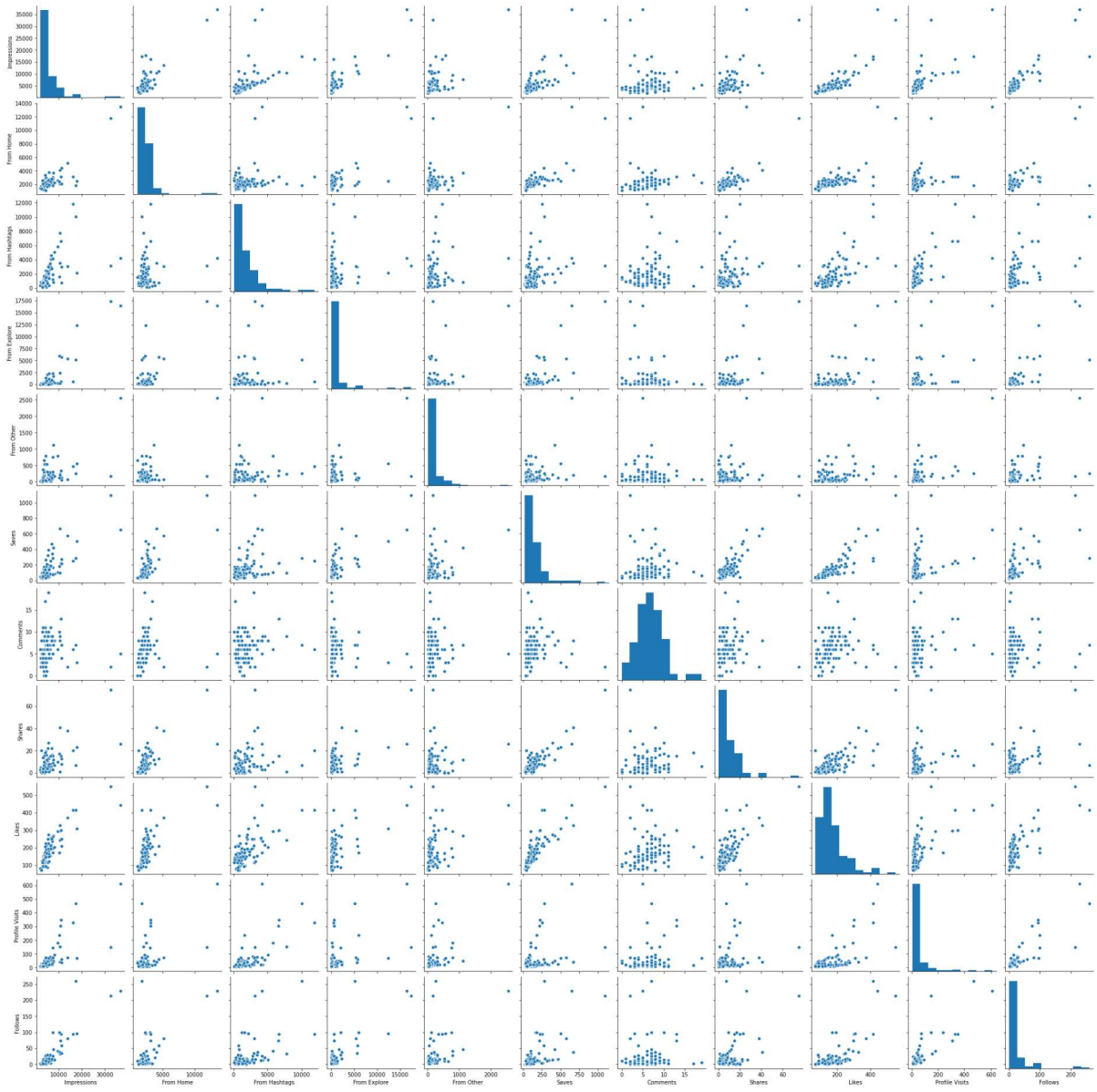
Out[31]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Fo
0	3920	2586	1028	619	56	98	9	5	162	35	
1	5394	2727	1838	1174	78	194	7	14	224	48	
2	4021	2085	1188	0	533	41	11	1	131	62	
3	4528	2700	621	932	73	172	10	7	213	23	
4	2518	1704	255	279	37	96	5	4	123	8	
...	...	...	...	...	...	...	...	...	...	...	
95	5394	2275	2975	45	65	61	19	6	147	69	
96	2766	2541	116	51	9	40	10	4	114	11	
97	3924	2244	1278	326	34	139	11	3	151	19	
98	3015	2034	771	115	41	52	11	4	92	9	
99	5409	2643	2006	1068	230	393	10	27	275	38	

100 rows × 13 columns

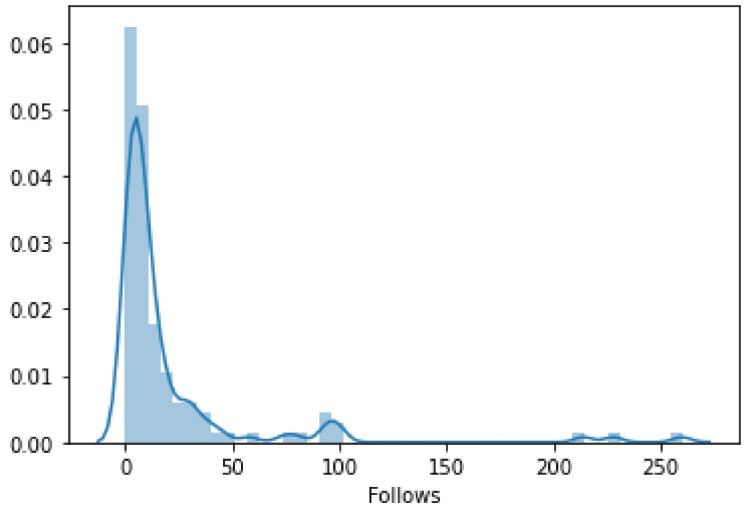
In [32]:
```
sns.pairplot(df)
```

Out[32]:
```
<seaborn.axisgrid.PairGrid at 0x1a8e60c2490>
```



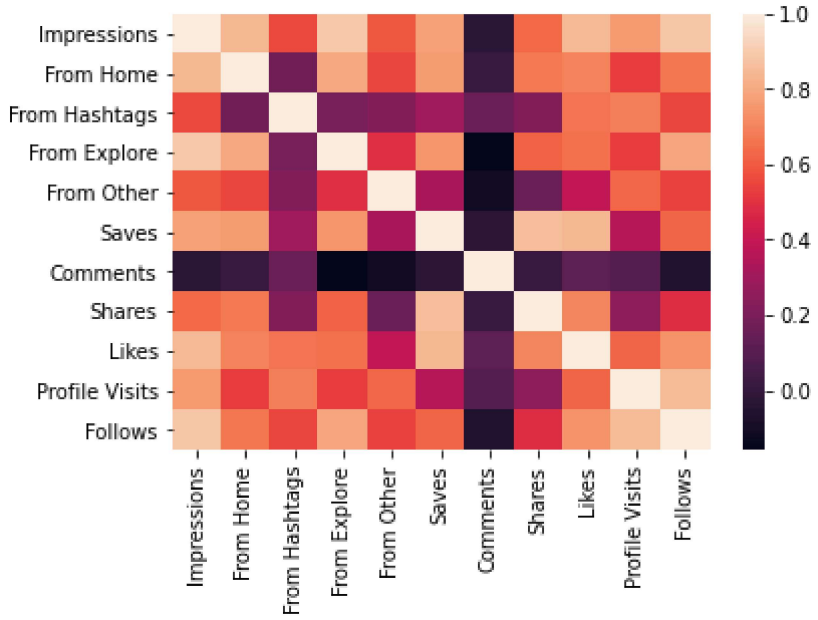
In [35]:
```
sns.distplot(df['Follows'])
```

Out[35]: <matplotlib.axes._subplots.AxesSubplot at 0x1a8e9c4f4c0>



In [36]:
```
sns.heatmap(df.corr())
```

Out[36]: <matplotlib.axes._subplots.AxesSubplot at 0x1a8ec53d7f0>



In [38]:
```
x=df[['Impressions', 'From Home', 'From Hashtags', 'From Explore',
       'From Other', 'Saves', 'Comments', 'Shares', 'Likes', 'Profile Visits']]
y=df['Follows']
```

In [39]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3)
```

In [40]:
```
from sklearn.linear_model import LinearRegression
lr= LinearRegression()
lr.fit(x_train,y_train)
```

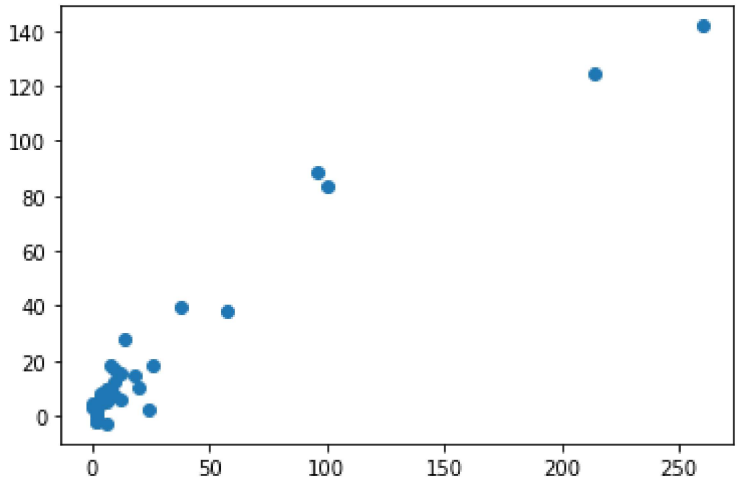
Out[40]:
```
LinearRegression()
```

In [41]:
```
print(lr.intercept_)
```

```
0.6714692062463641
```

In [42]:
```
prediction= lr.predict(x_test)
plt.scatter(y_test,prediction)
```

Out[42]:
```
<matplotlib.collections.PathCollection at 0x1a8d5720520>
```



In [43]:
```
print(lr.score(x_test,y_test))
```

```
0.7874406216325516
```

In [ ]: