

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

```
In [2]: df=pd.read_csv(r"C:\Users\user\Downloads\5_Instagram data.csv")
df.fillna(0,inplace=True)
df
```

Out[2]:

pressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
3920	2586	1028	619	56	98	9	5	162	35	2
5394	2727	1838	1174	78	194	7	14	224	48	10
4021	2085	1188	0	533	41	11	1	131	62	12
4528	2700	621	932	73	172	10	7	213	23	8
2518	1704	255	279	37	96	5	4	123	8	0
...	...	...	...	...	...	...	...	...	...	...
13700	5185	3041	5352	77	573	2	38	373	73	80
5731	1923	1368	2266	65	135	4	1	148	20	18
4139	1133	1538	1367	33	36	0	1	92	34	10
32695	11815	3147	17414	170	1095	2	75	549	148	214

Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
36919	13473	4176	16444	2547	653	5	26	443	611	228

5

15 × 13 columns

In [3]: df.head()

Out[3]:

Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
3920	2586	1028	619	56	98	9	5	162	35	2
5394	2727	1838	1174	78	194	7	14	224	48	10
4021	2085	1188	0	533	41	11	1	131	62	12
4528	2700	621	932	73	172	10	7	213	23	8
2518	1704	255	279	37	96	5	4	123	8	0

6



In [4]: 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 [7]: import seaborn as sns

In [8]: df.describe()

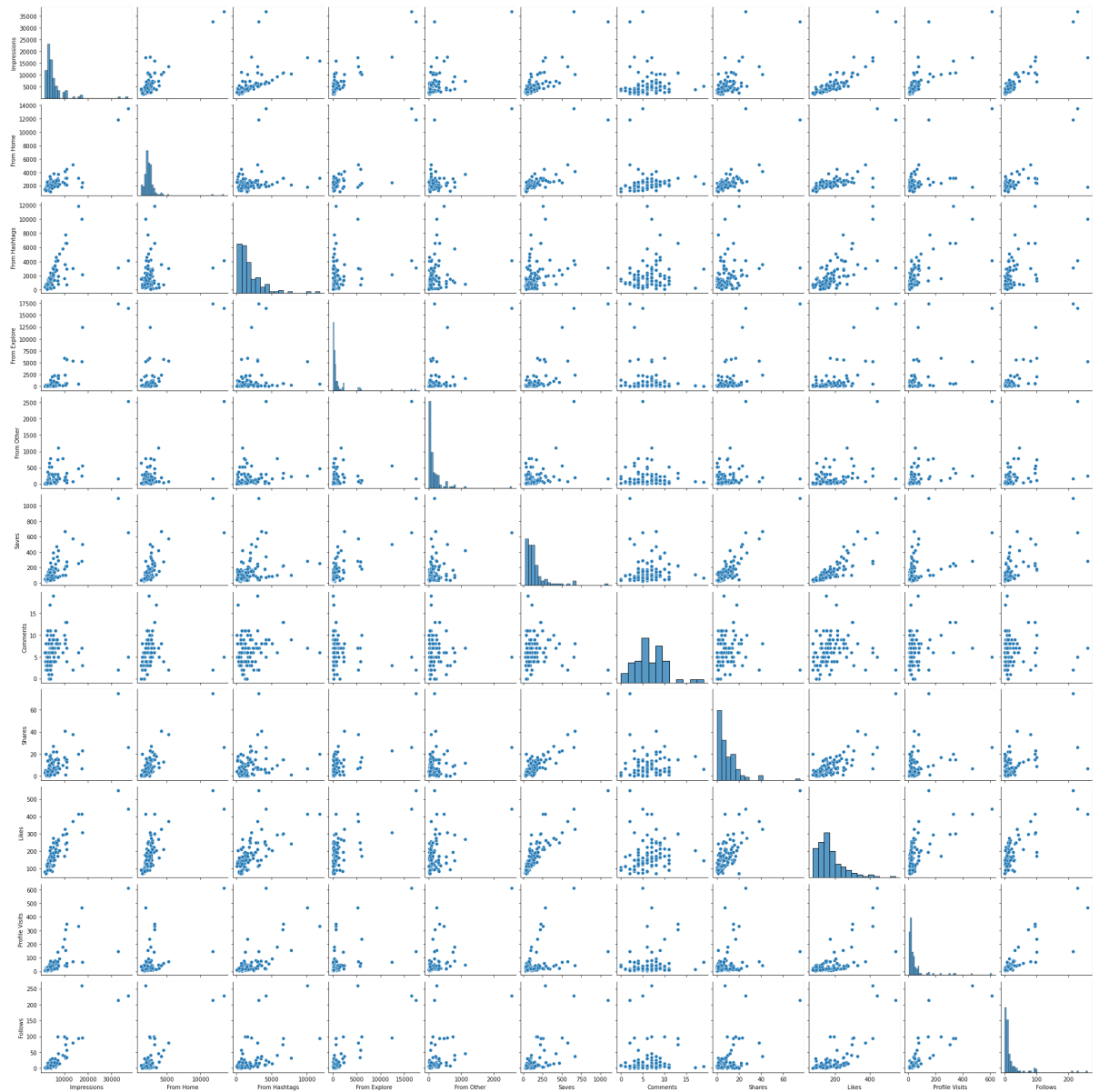
Out[8]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comm
count	119.000000	119.000000	119.000000	119.000000	119.000000	119.000000	119.00
mean	5703.991597	2475.789916	1887.512605	1078.100840	171.092437	153.310924	6.66
std	4843.780105	1489.386348	1884.361443	2613.026132	289.431031	156.317731	3.54
min	1941.000000	1133.000000	116.000000	0.000000	9.000000	22.000000	0.00
25%	3467.000000	1945.000000	726.000000	157.500000	38.000000	65.000000	4.00
50%	4289.000000	2207.000000	1278.000000	326.000000	74.000000	109.000000	6.00
75%	6138.000000	2602.500000	2363.500000	689.500000	196.000000	169.000000	8.00
max	36919.000000	13473.000000	11817.000000	17414.000000	2547.000000	1095.000000	19.00

In [10]: df=pd.read\_csv("5\_Instagram data.csv")

```
In [11]: sns.pairplot(df)
```

```
Out[11]: <seaborn.axisgrid.PairGrid at 0x1c34f7d7f40>
```

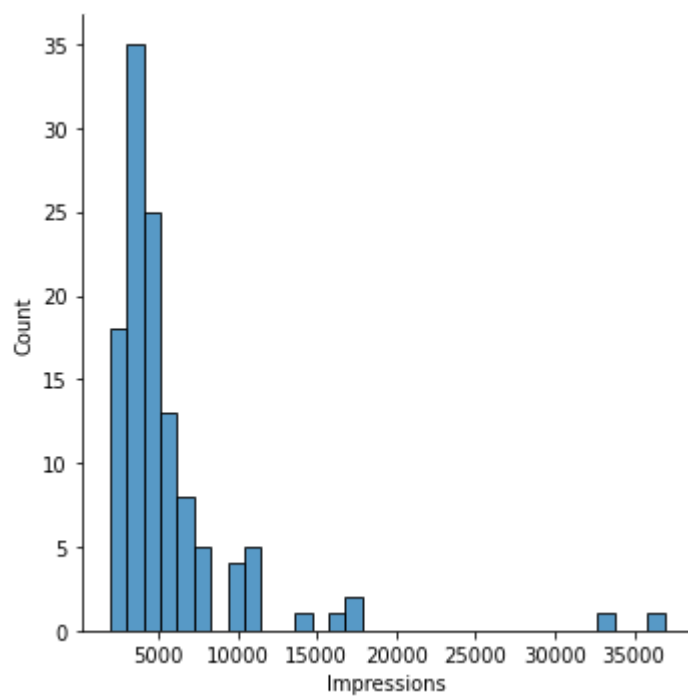


```
In [14]: df1=df.drop(['Comments'],axis=1)
df1
df1=df1.drop(df1.index[1537:])
df1.isna().sum()
```

```
Out[14]: Impressions      0
From Home      0
From Hashtags  0
From Explore   0
From Other     0
Saves          0
Shares         0
Likes          0
Profile Visits 0
Follows        0
Caption        0
Hashtags       0
dtype: int64
```

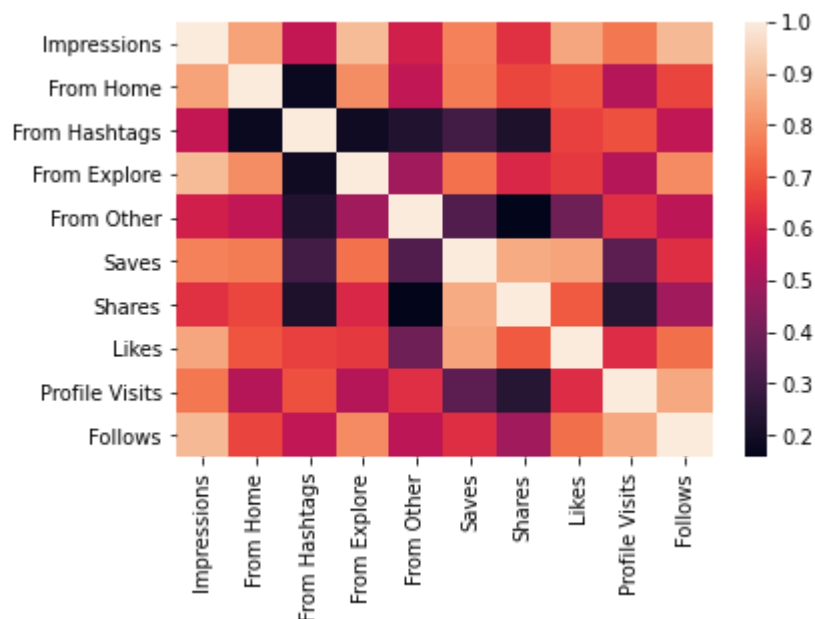
```
In [15]: sns.displot(df['Impressions'])
```

```
Out[15]: <seaborn.axisgrid.FacetGrid at 0x1c352fee100>
```



```
In [16]: sns.heatmap(df1.corr())
```

```
Out[16]: <AxesSubplot:>
```



```
In [17]: from sklearn.model_selection import train_test_split  
from sklearn.linear_model import LinearRegression
```

```
In [18]: df1.isna().sum()
```

```
Out[18]: Impressions      0  
From Home      0  
From Hashtags  0  
From Explore   0  
From Other     0  
Saves          0  
Shares         0  
Likes          0  
Profile Visits 0  
Follows        0  
Caption        0  
Hashtags       0  
dtype: int64
```



```
In [20]: y=df1['Likes']
x=df1.drop(['Caption','Hashtags'],axis=1)
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3)
print(x_train)
```

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves
\						
46	3630	1747	1693	72	86	137
33	4343	2278	1460	521	27	105
20	2064	1304	362	249	37	49
27	2766	2541	116	51	9	40
73	5542	1782	3212	213	168	46
..	...	...	...	...	...	...
117	32695	11815	3147	17414	170	1095
61	6339	2190	4036	48	27	171
24	4628	2406	1260	861	26	144
87	3630	1747	1693	72	86	137
40	16062	3144	11817	564	468	252

	Shares	Likes	Profile Visits	Follows
46	10	137	14	4
33	3	152	24	8
20	5	76	9	0
27	4	114	11	6
73	6	122	52	16
..	...	...	...	...
117	75	549	148	214
61	5	248	21	10
24	3	160	10	4
87	10	137	14	4
40	20	416	330	94

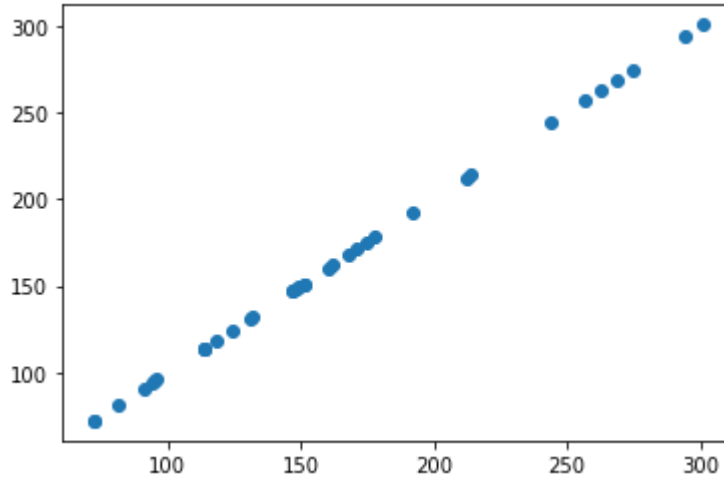
[83 rows x 10 columns]

```
In [22]: model=LinearRegression()
model.fit(x_train,y_train)
model.intercept_
```

Out[22]: -9.663381206337363e-13

```
In [23]: prediction=model.predict(x_test)
plt.scatter(y_test,prediction)
```

```
Out[23]: <matplotlib.collections.PathCollection at 0x1c358311fa0>
```



```
In [24]: model.score(x_test,y_test)
```

```
Out[24]: 1.0
```

```
In [25]: from sklearn.linear_model import Ridge,Lasso
```

```
In [27]: rr=Ridge(alpha=10)
rr.fit(x_train,y_train)
```

```
Out[27]: Ridge(alpha=10)
```

```
In [28]: rr.score(x_test,y_test)
```

```
Out[28]: 0.9999999926258979
```

```
In [29]: la =Lasso(alpha=10)
la.fit(x_train,y_train)
```

```
Out[29]: Lasso(alpha=10)
```

```
In [30]: la.score(x_test,y_test)
```

```
Out[30]: 0.9999697539644843
```