

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

```
In [135]: data=pd.read_csv(r"C:\Users\user\Desktop\Vicky\5 Instagram data.csv")
```

```
In [137]: data.head()
```

Out[137]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Fol
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	



In [138]: `data.tail()`

Out[138]:

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



In [139]: `data.isna()`

Out[139]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	F
0	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False
...
114	False	False	False	False	False	False	False	False	False	False	False
115	False	False	False	False	False	False	False	False	False	False	False
116	False	False	False	False	False	False	False	False	False	False	False
117	False	False	False	False	False	False	False	False	False	False	False
118	False	False	False	False	False	False	False	False	False	False	False

119 rows × 13 columns



In [140]: `data.fillna(value=5)`

Out[140]:

	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 [141]: `data.shape`

Out[141]: (119, 13)

In [142]: `data.size`

Out[142]: 1547

In [143]: `pd.isna(data)`

Out[143]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	F
0	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False
...
114	False	False	False	False	False	False	False	False	False	False	False
115	False	False	False	False	False	False	False	False	False	False	False
116	False	False	False	False	False	False	False	False	False	False	False
117	False	False	False	False	False	False	False	False	False	False	False
118	False	False	False	False	False	False	False	False	False	False	False

119 rows × 13 columns



In [144]: `data.dropna()`

Out[144]:

	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

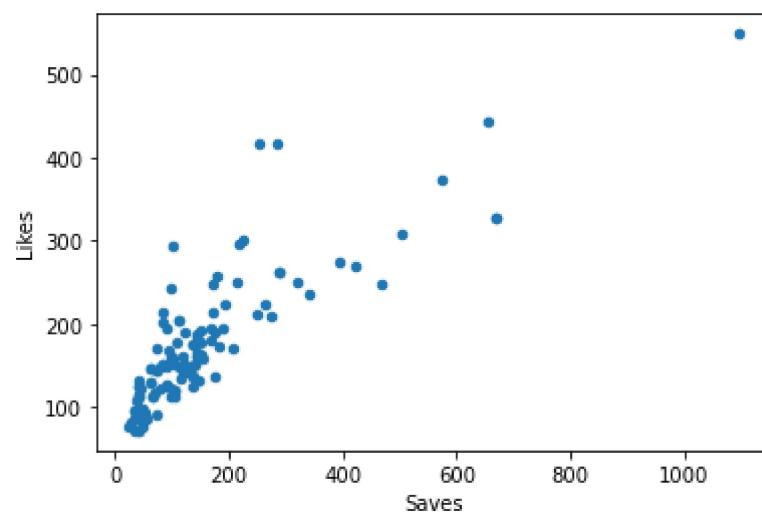
In [145]: `data.describe()`

Out[145]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments
count	119.000000	119.000000	119.000000	119.000000	119.000000	119.000000	119.000000
mean	5703.991597	2475.789916	1887.512605	1078.100840	171.092437	153.310924	6.666667
std	4843.780105	1489.386348	1884.361443	2613.026132	289.431031	156.317731	3.544671
min	1941.000000	1133.000000	116.000000	0.000000	9.000000	22.000000	0.000000
25%	3467.000000	1945.000000	726.000000	157.500000	38.000000	65.000000	4.000000
50%	4289.000000	2207.000000	1278.000000	326.000000	74.000000	109.000000	6.000000
75%	6138.000000	2602.500000	2363.500000	689.500000	196.000000	169.000000	8.000000
max	36919.000000	13473.000000	11817.000000	17414.000000	2547.000000	1095.000000	19.000000

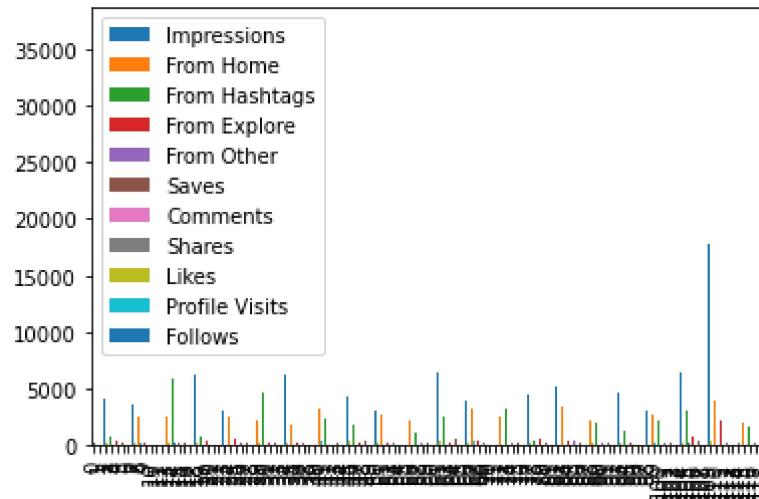
In [147]: `data.plot.scatter("Saves", "Likes")`

Out[147]: <AxesSubplot:xlabel='Saves', ylabel='Likes'>



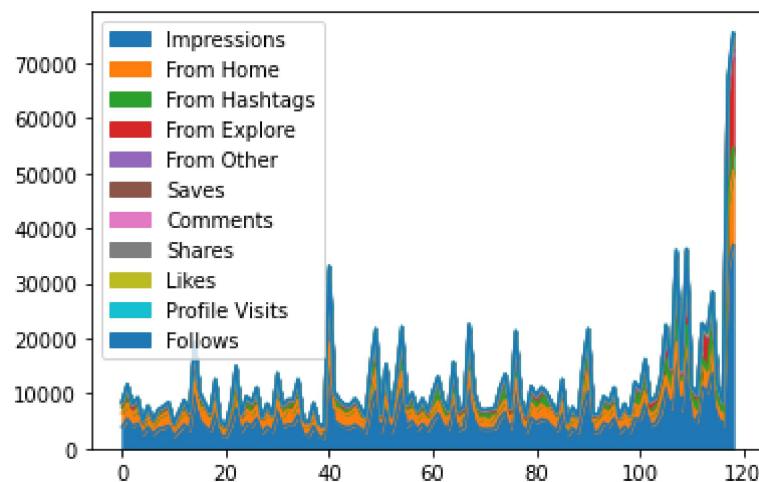
In [148]: `data.plot.bar()`

Out[148]: <AxesSubplot:>



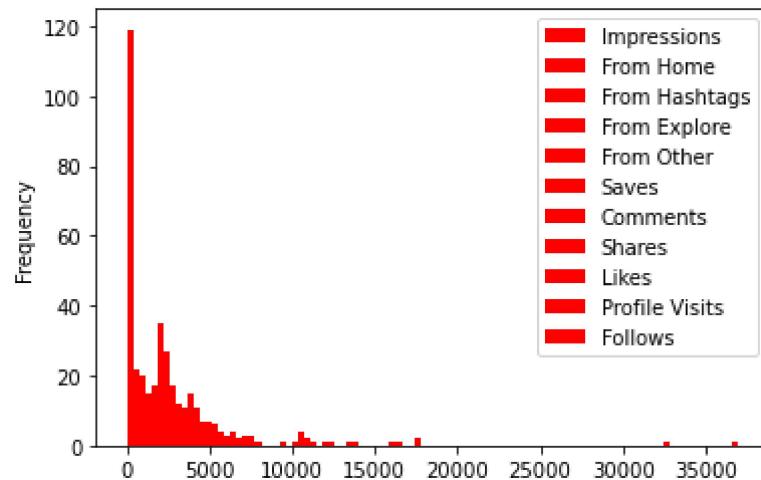
In [149]: `data.plot.area()`

Out[149]: <AxesSubplot:>



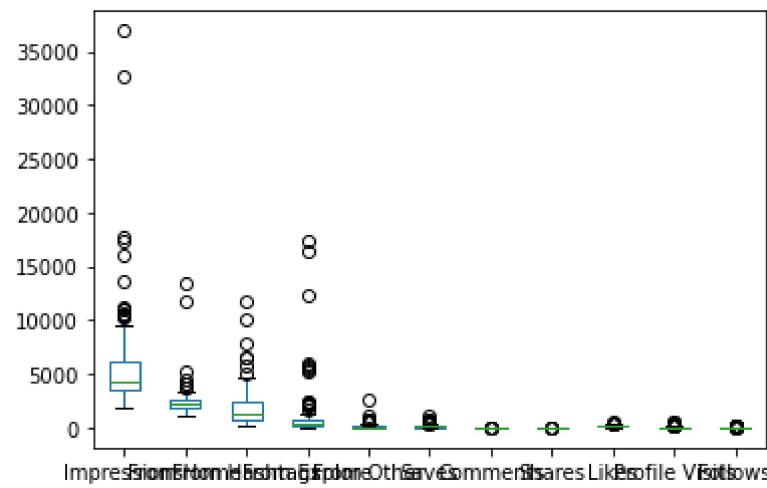
```
In [150]: data.plot.hist(bins=100,color="r")
```

```
Out[150]: <AxesSubplot:ylabel='Frequency'>
```

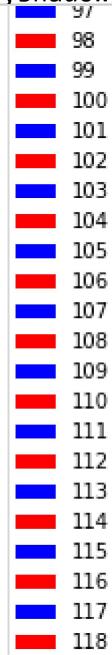


```
In [151]: data.plot.box()
```

```
Out[151]: <AxesSubplot:>
```



```
In [152]: cols=["r","b"]
data.plot.pie(y="Saves", shadow=True, startangle=90, colors=cols)
```

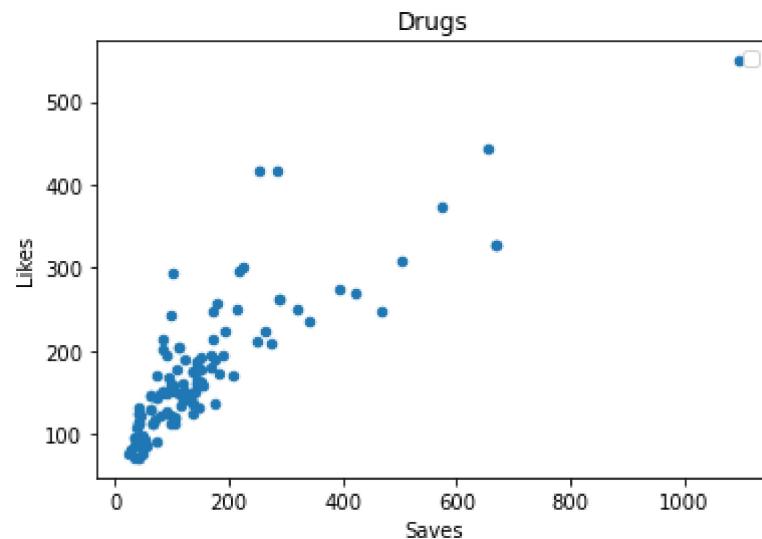


```
In [153]: data.plot.scatter(x="Saves",y="Likes")
```

```
plt.title("Drugs")
plt.legend()
```

No handles with labels found to put in legend.

```
Out[153]: <matplotlib.legend.Legend at 0x244005f1400>
```



```
In [154]: from numpy import linalg as la
```

In [155]: `data.mean()`

Out[155]:

Impressions	5703.991597
From Home	2475.789916
From Hashtags	1887.512605
From Explore	1078.100840
From Other	171.092437
Saves	153.310924
Comments	6.663866
Shares	9.361345
Likes	173.781513
Profile Visits	50.621849
Follows	20.756303

dtype: float64

In [156]: `data.median()`

Out[156]:

Impressions	4289.0
From Home	2207.0
From Hashtags	1278.0
From Explore	326.0
From Other	74.0
Saves	109.0
Comments	6.0
Shares	6.0
Likes	151.0
Profile Visits	23.0
Follows	8.0

dtype: float64

In [157]: `data.mode()`

Out[157]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	F
0	5394.0	1975.0		116	45.0	34.0	40.0		6.0	3.0	114.0
1	NaN	NaN		201	84.0	NaN	135.0		NaN	NaN	151.0
2	NaN	NaN		278	NaN	NaN	144.0		NaN	NaN	NaN
3	NaN	NaN		362	NaN	NaN	NaN		NaN	NaN	NaN
4	NaN	NaN		411	NaN	NaN	NaN		NaN	NaN	NaN
5	NaN	NaN		583	NaN	NaN	NaN		NaN	NaN	NaN
6	NaN	NaN		655	NaN	NaN	NaN		NaN	NaN	NaN
7	NaN	NaN		707	NaN	NaN	NaN		NaN	NaN	NaN
8	NaN	NaN		771	NaN	NaN	NaN		NaN	NaN	NaN
9	NaN	NaN		794	NaN	NaN	NaN		NaN	NaN	NaN
10	NaN	NaN		1248	NaN	NaN	NaN		NaN	NaN	NaN
11	NaN	NaN		1260	NaN	NaN	NaN		NaN	NaN	NaN
12	NaN	NaN		1278	NaN	NaN	NaN		NaN	NaN	NaN
13	NaN	NaN		1693	NaN	NaN	NaN		NaN	NaN	NaN
14	NaN	NaN		1938	NaN	NaN	NaN		NaN	NaN	NaN
15	NaN	NaN		2351	NaN	NaN	NaN		NaN	NaN	NaN
16	NaN	NaN		2975	NaN	NaN	NaN		NaN	NaN	NaN
17	NaN	NaN		3450	NaN	NaN	NaN		NaN	NaN	NaN
18	NaN	NaN		3551	NaN	NaN	NaN		NaN	NaN	NaN



In [158]: `data.std()`

Out[158]:

Impressions	4843.780105
From Home	1489.386348
From Hashtags	1884.361443
From Explore	2613.026132
From Other	289.431031
Saves	156.317731
Comments	3.544576
Shares	10.089205
Likes	82.378947
Profile Visits	87.088402
Follows	40.921580

dtype: float64

In [159]: `data.var()`

Out[159]:

Impressions	2.346221e+07
From Home	2.218272e+06
From Hashtags	3.550818e+06
From Explore	6.827906e+06
From Other	8.377032e+04
Saves	2.443523e+04
Comments	1.256402e+01
Shares	1.017921e+02
Likes	6.786291e+03
Profile Visits	7.584390e+03
Follows	1.674576e+03

dtype: float64

In [160]: `data.max()`

Out[160]:

Impressions	36919
From Home	13473
From Hashtags	11817
From Explore	17414
From Other	2547
Saves	1095
Comments	19
Shares	75
Likes	549
Profile Visits	611
Follows	260
Caption	You must have seen the news divided into categ...
Hashtags	#timeseries◆#time◆#statistics◆#datascience◆#bi...

dtype: object

In [161]: `data.min()`

Out[161]:

Impressions	1941
From Home	1133
From Hashtags	116
From Explore	0
From Other	9
Saves	22
Comments	0
Shares	0
Likes	72
Profile Visits	4
Follows	0
Caption	170 Python Projects with Source Code solved an...
Hashtags	#career♦#job♦#jobs♦#jobsearch♦#education♦#busi...

`dtype: object`

In [170]: `data1=data[["Saves","Likes"]]`
`data1`

Out[170]:

	Saves	Likes
0	98	162
1	194	224
2	41	131
3	172	213
4	96	123
...
114	573	373
115	135	148
116	36	92
117	1095	549
118	653	443

119 rows × 2 columns

In [163]: `data1.sum()`

Out[163]:

Sum of Jan	5.62%	4.21%	9.83%	2.81%	25.28%	8.15%	18.54%	25.56%	100...
Sum of Feb	7.73%	17.27%	11.60%	21.91%	10.57%	16.24%	8.76%	5.93%	1...

`dtype: object`

In [164]: `data.cumsum()`

Out[164]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits
0	3920	2586	1028	619	56	98	9	5	162	35
1	9314	5313	2866	1793	134	292	16	19	386	83
2	13335	7398	4054	1793	667	333	27	20	517	145
3	17863	10098	4675	2725	740	505	37	27	730	168
4	20381	11802	4930	3004	777	601	42	31	853	176
...
114	599291	266275	214385	90803	17545	16325	782	1011	19448	5211
115	605022	268198	215753	93069	17610	16460	786	1012	19596	5231
116	609161	269331	217291	94436	17643	16496	786	1013	19688	5265
117	641856	281146	220438	111850	17813	17591	788	1088	20237	5413

Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	
118	678775	294619	224614	128294	20360	18244	793	1114	20680	6024

119 rows × 13 columns

In [165]: `data1.cumsum()`

Out[165]:

Sum of Jan

0	5.62%
1	5.62%4.21%
2	5.62%4.21%9.83%
3	5.62%4.21%9.83%2.81%
4	5.62%4.21%9.83%2.81%25.28%
5	5.62%4.21%9.83%2.81%25.28%8.15%
6	5.62%4.21%9.83%2.81%25.28%8.15%18.54%
7	5.62%4.21%9.83%2.81%25.28%8.15%18.54%25.56%
8	5.62%4.21%9.83%2.81%25.28%8.15%18.54%25.56%100...

In [168]: `from scipy.stats import spearmanr
from scipy.stats import pearsonr`

```
In [171]: print(spearmanr(data,data1))
```



```

1.54251120e-01, 8.94486953e-06, 2.91561315e-14],
[2.11639002e-13, 1.73105178e-07, 9.89702195e-03, 0.00000000e+00,
 4.45956535e-03, 1.77560737e-15, 8.71459370e-01, 1.78110529e-06,
 4.99691207e-10, 3.06196498e-04, 2.76348912e-08, 1.07773566e-02,
 7.31524496e-01, 1.77560737e-15, 4.99691207e-10],
[5.69101183e-08, 5.36505126e-02, 2.42059331e-05, 4.45956535e-03,
 0.00000000e+00, 9.07782657e-04, 1.45847886e-01, 1.80090003e-03,
 1.63641548e-05, 3.13480037e-13, 8.28733146e-14, 1.68450863e-01,
 6.51827905e-01, 9.07782657e-04, 1.63641548e-05],
[5.57179233e-18, 3.42866380e-19, 8.94486953e-06, 1.77560737e-15,
 9.07782657e-04, 0.00000000e+00, 4.64953934e-02, 7.05458788e-14,
 2.59014227e-34, 4.58086204e-03, 7.52530656e-07, 4.28584449e-04,
 1.73387188e-01, 0.00000000e+00, 2.59014227e-34],
[1.40602879e-02, 2.33805392e-04, 9.13997462e-02, 8.71459370e-01,
 1.45847886e-01, 4.64953934e-02, 0.00000000e+00, 1.46417080e-01,
 7.21460549e-04, 7.56880163e-01, 3.96005685e-01, 1.85711169e-01,
 6.52547062e-01, 4.64953934e-02, 7.21460549e-04],
[9.87370888e-08, 7.40589087e-12, 5.87294543e-03, 1.78110529e-06,
 1.80090003e-03, 7.05458788e-14, 1.46417080e-01, 0.00000000e+00,
 1.42478205e-11, 2.21394985e-01, 1.27208465e-02, 8.51689519e-04,
 3.94533941e-03, 7.05458788e-14, 1.42478205e-11],
[5.99576983e-35, 3.52449378e-19, 2.91561315e-14, 4.99691207e-10,
 1.63641548e-05, 2.59014227e-34, 7.21460549e-04, 1.42478205e-11,
 0.00000000e+00, 2.25503976e-08, 1.62421738e-11, 7.46317514e-03,
 4.67894001e-01, 2.59014227e-34, 0.00000000e+00],
[7.44719262e-16, 7.31369934e-03, 3.26524361e-12, 3.06196498e-04,
 3.13480037e-13, 4.58086204e-03, 7.56880163e-01, 2.21394985e-01,
 2.25503976e-08, 0.00000000e+00, 2.01952335e-23, 9.67042936e-03,
 4.72067533e-01, 4.58086204e-03, 2.25503976e-08],
[7.20467362e-24, 9.49535634e-05, 2.92493122e-11, 2.76348912e-08,
 8.28733146e-14, 7.52530656e-07, 3.96005685e-01, 1.27208465e-02,
 1.62421738e-11, 2.01952335e-23, 0.00000000e+00, 8.03766594e-01,
 6.96127300e-01, 7.52530656e-07, 1.62421738e-11],
[1.96498538e-01, 1.11578948e-02, 5.55718716e-01, 1.07773566e-02,
 1.68450863e-01, 4.28584449e-04, 1.85711169e-01, 8.51689519e-04,
 7.46317514e-03, 9.67042936e-03, 8.03766594e-01, 0.00000000e+00,
 5.93637244e-01, 4.28584449e-04, 7.46317514e-03],
[3.97031859e-01, 6.23169718e-01, 1.54251120e-01, 7.31524496e-01,
 6.51827905e-01, 1.73387188e-01, 6.52547062e-01, 3.94533941e-03,
 4.67894001e-01, 4.72067533e-01, 6.96127300e-01, 5.93637244e-01,
 0.00000000e+00, 1.73387188e-01, 4.67894001e-01],
[5.57179233e-18, 3.42866380e-19, 8.94486953e-06, 1.77560737e-15,
 9.07782657e-04, 0.00000000e+00, 4.64953934e-02, 7.05458788e-14,
 2.59014227e-34, 4.58086204e-03, 7.52530656e-07, 4.28584449e-04,
 1.73387188e-01, 0.00000000e+00, 2.59014227e-34],
[5.99576983e-35, 3.52449378e-19, 2.91561315e-14, 4.99691207e-10,
 1.63641548e-05, 2.59014227e-34, 7.21460549e-04, 1.42478205e-11,
 0.00000000e+00, 2.25503976e-08, 1.62421738e-11, 7.46317514e-03,
 4.67894001e-01, 2.59014227e-34, 0.00000000e+00]])))

```

In [75]:

```

from numpy import mean, std, cov
from numpy.random import randn, seed
from matplotlib import pyplot

```

In []:

In [173]: `data.fillna(value=5)`

Out[173]:

	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 [174]: `data.count()`

Out[174]:

Impressions	119
From Home	119
From Hashtags	119
From Explore	119
From Other	119
Saves	119
Comments	119
Shares	119
Likes	119
Profile Visits	119
Follows	119
Caption	119
Hashtags	119

`dtype: int64`

In [175]: `data.count`

Out[175]:

	From Explore	From Other	Saves	Impressions	From Home	From Hashtags
0	3920	2586		1028	619	56
1	5394	2727		1838	1174	78
2	4021	2085		1188	0	533
3	4528	2700		621	932	73
4	2518	1704		255	279	37
..
114	13700	5185		3041	5352	77
115	5731	1923		1368	2266	65
116	4139	1133		1538	1367	33
117	32695	11815		3147	17414	170
118	36919	13473		4176	16444	2547

	Comments	Shares	Likes	Profile Visits	Follows	\
0	9	5	162	35	2	
1	7	14	224	48	10	
2	11	1	131	62	12	
3	10	7	213	23	8	
4	5	4	123	8	0	
..	
114	2	38	373	73	80	
115	4	1	148	20	18	
116	0	1	92	34	10	
117	2	75	549	148	214	
118	5	26	443	611	228	

	Caption	\
0	Here are some of the most important data visua...	
1	Here are some of the best data science project...	
2	Learn how to train a machine learning model an...	
3	Here's how you can write a Python program to d...	
4	Plotting annotations while visualizing your da...	
..	...	
114	Here are some of the best data science certifi...	
115	Clustering is a machine learning technique use...	
116	Clustering music genres is a task of grouping ...	
117	Here are some of the best data science certifi...	
118	175 Python Projects with Source Code solved an...	

	Hashtags
0	#finance◆#money◆#business◆#investing◆#investme...
1	#healthcare◆#health◆#covid◆#data◆#datascience◆...
2	#data◆#datascience◆#dataanalysis◆#dataanalytic...
3	#python◆#pythonprogramming◆#pythonprojects◆#py...
4	#datavisualization◆#datascience◆#data◆#dataana...
..	...
114	#datascience◆#datasciencejobs◆#datascientrai...
115	#machinelearning◆#machinelearningalgorithms◆#d...
116	#machinelearning◆#machinelearningalgorithms◆#d...
117	#datascience◆#datasciencejobs◆#datascientrai...
118	#python◆#pythonprogramming◆#pythonprojects◆#py...

[119 rows x 13 columns]>

```
In [176]: data.columns
```

```
Out[176]: Index(['Impressions', 'From Home', 'From Hashtags', 'From Explore',
       'From Other', 'Saves', 'Comments', 'Shares', 'Likes', 'Profile Visit
s',
       'Follows', 'Caption', 'Hashtags'],
      dtype='object')
```

```
In [132]: data.iloc[3]
```

```
Out[132]: Row Labels          D
Sum of Jan        2.81%
Sum of Feb        21.91%
Sum of Mar        7.88%
Sum of Total Sales    127
Name: 3, dtype: object
```

In [177]: `data.dropna(axis=1, how="any")`

Out[177]:

	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 [179]: `data.index`

Out[179]: `RangeIndex(start=0, stop=119, step=1)`

In []: