

Basic Analysis using Numpy and Pandas

import libraries

In [1]:

```
import pandas as pd
import numpy as np
import matplotlib as pp
```

import dataset

In [2]:

```
data=pd.read_csv(r"E:\154\5_Instagram data.csv")
```

In [3]:

```
display(data)
```

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
0	3920	2586	1028	619	56	98	9	5	162	35	2
1	5394	2727	1838	1174	78	194	7	14	224	48	10
2	4021	2085	1188	0	533	41	11	1	131	62	12
3	4528	2700	621	932	73	172	10	7	213	23	8
4	2518	1704	255	279	37	96	5	4	123	8	0
...
114	13700	5185	3041	5352	77	573	2	38	373	73	80

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
115	5731	1923	1368	2266	65	135	4	1	148	20	18
116	4139	1133	1538	1367	33	36	0	1	92	34	10
117	32695	11815	3147	17414	170	1095	2	75	549	148	214
118	36919	13473	4176	16444	2547	653	5	26	443	611	228

119 rows × 13 columns

To display top 10 rows

In [4]:

data.head()

Out[4]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
0	3920	2586	1028	619	56	98	9	5	162	35	2
1	5394	2727	1838	1174	78	194	7	14	224	48	10
2	4021	2085	1188	0	533	41	11	1	131	62	12

To display last 5 rows

In [5]:

```
data.tail()
```

Out[5]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
114	13700	5185	3041	5352	77	573	2	38	373	73	80
115	5731	1923	1368	2266	65	135	4	1	148	20	18
116	4139	1133	1538	1367	33	36	0	1	92	34	10
117	32695	11815	3147	17414	170	1095	2	75	549	148	214
118	36919	13473	4176	16444	2547	653	5	26	443	611	228

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

In [6]: `data.dtypes`

Out[6]:

Impressions	int64
From Home	int64
From Hashtags	int64
From Explore	int64
From Other	int64
Saves	int64
Comments	int64
Shares	int64
Likes	int64
Profile Visits	int64
Follows	int64
Caption	object
Hashtags	object
dtype:	object

To view statistical summary

In [7]: `data.describe()`

Out[7]:

	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	119.000000
mean	5703.991597	2475.789916	1887.512605	1078.100840	171.092437	153.310924	6.663866	6.663866
std	4843.780105	1489.386348	1884.361443	2613.026132	289.431031	156.317731	3.544576	3.544576
min	1941.000000	1133.000000	116.000000	0.000000	9.000000	22.000000	0.000000	0.000000
25%	3467.000000	1945.000000	726.000000	157.500000	38.000000	65.000000	4.000000	4.000000
50%	4289.000000	2207.000000	1278.000000	326.000000	74.000000	109.000000	6.000000	6.000000
75%	6138.000000	2602.500000	2363.500000	689.500000	196.000000	169.000000	8.000000	8.000000
max	36919.000000	13473.000000	11817.000000	17414.000000	2547.000000	1095.000000	19.000000	19.000000

To Print no of elements

In [8]: `data.size`

Out[8]: 1547

In [9]: `data.ndim`

Out[9]: 2

To print no of rows and columns

In [10]:

data.shape

Out[10]: (119, 13)

To find missing values

In [11]:

data.isna()

Out[11]:

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

To drop null values with constatns

In [12]:

data.fillna(5)

Out[12]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
0	3920	2586	1028	619	56	98	9	5	162	35	2

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
1	5394	2727	1838	1174	78	194	7	14	224	48	10
2	4021	2085	1188	0	533	41	11	1	131	62	12
3	4528	2700	621	932	73	172	10	7	213	23	8
4	2518	1704	255	279	37	96	5	4	123	8	0
...
114	13700	5185	3041	5352	77	573	2	38	373	73	80
115	5731	1923	1368	2266	65	135	4	1	148	20	18
116	4139	1133	1538	1367	33	36	0	1	92	34	10
117	32695	11815	3147	17414	170	1095	2	75	549	148	214
118	36919	13473	4176	16444	2547	653	5	26	443	611	228

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
0	3920	2586	1028	619	56	98	9	5	162	35	2
1	5394	2727	1838	1174	78	194	7	14	224	48	10
2	4021	2085	1188	0	533	41	11	1	131	62	12
3	4528	2700	621	932	73	172	10	7	213	23	8
4	2518	1704	255	279	37	96	5	4	123	8	0
...
114	13700	5185	3041	5352	77	573	2	38	373	73	80
115	5731	1923	1368	2266	65	135	4	1	148	20	18

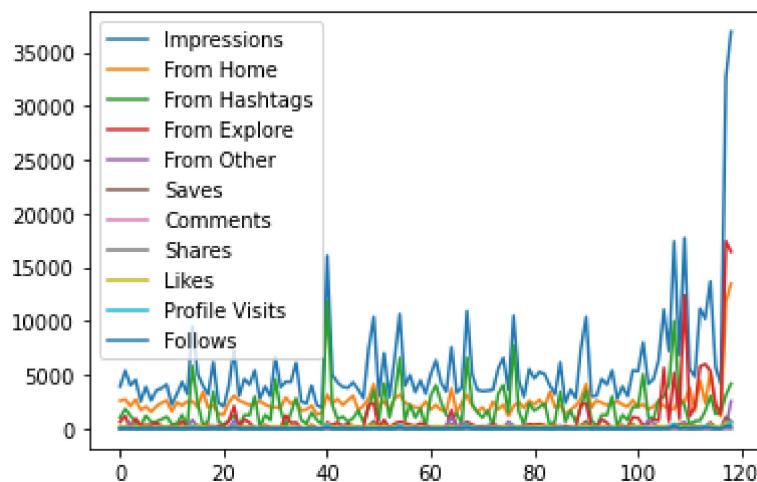
	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Follows
116	4139	1133	1538	1367	33	36	0	1	92	34	10
117	32695	11815	3147	17414	170	1095	2	75	549	148	214
118	36919	13473	4176	16444	2547	653	5	26	443	611	228

119 rows × 13 columns

Line Plot

In [14]: `data.plot.line()`

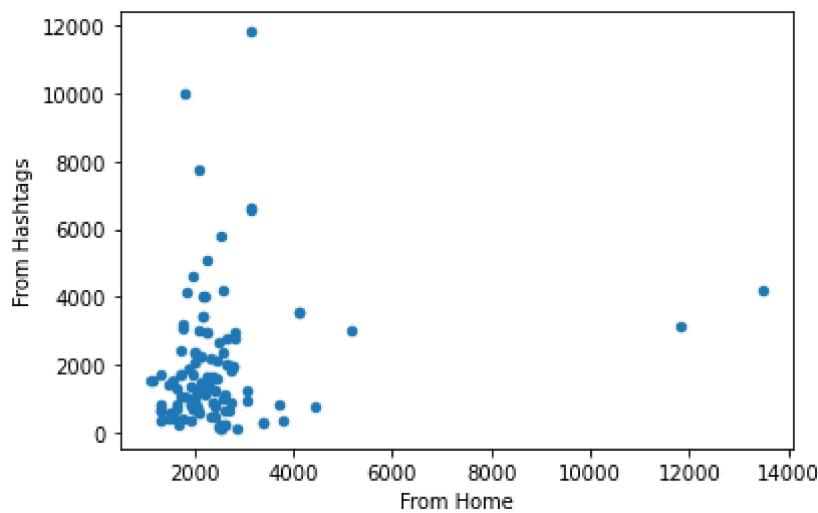
Out[14]: <AxesSubplot:>



Scatter Plot

In [16]: `data.plot.scatter(x='From Home', y='From Hashtags')`

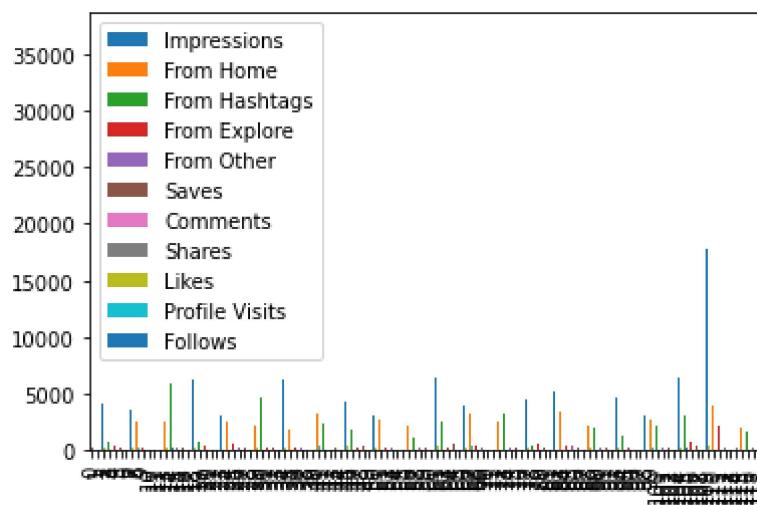
Out[16]: <AxesSubplot:xlabel='From Home', ylabel='From Hashtags'>



Bar Chart

In [17]: `data.plot.bar()`

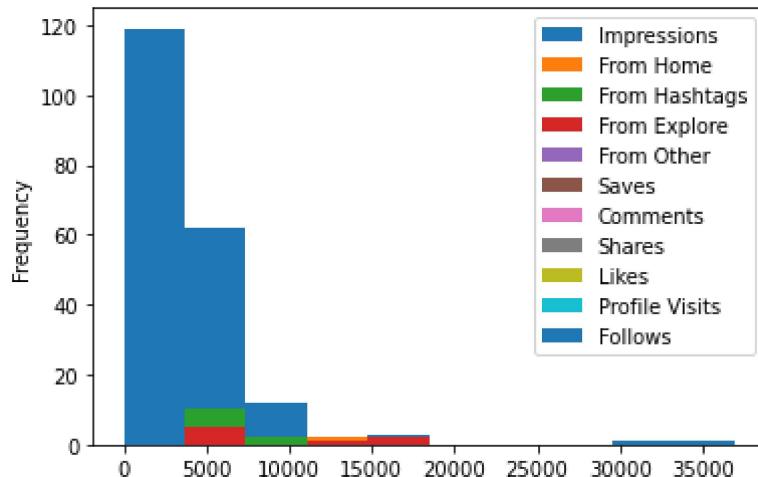
Out[17]: <AxesSubplot:>



Histogram

In [18]: `data.plot.hist()`

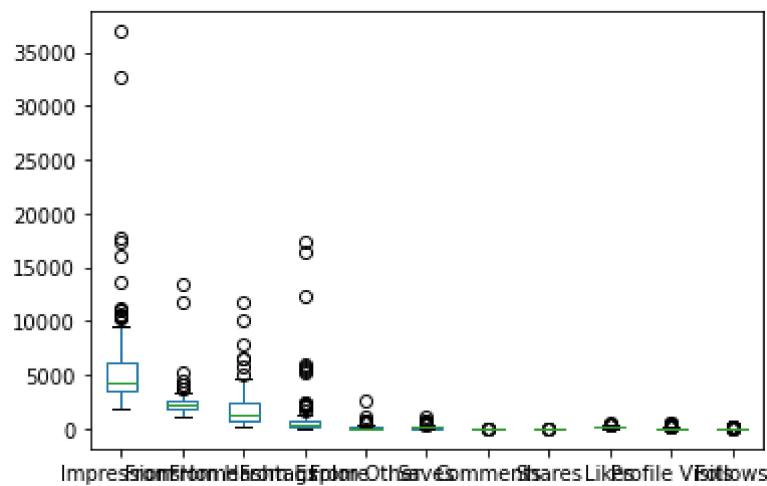
Out[18]: <AxesSubplot:ylabel='Frequency'>



Box Plot

```
In [19]: data.plot.box()
```

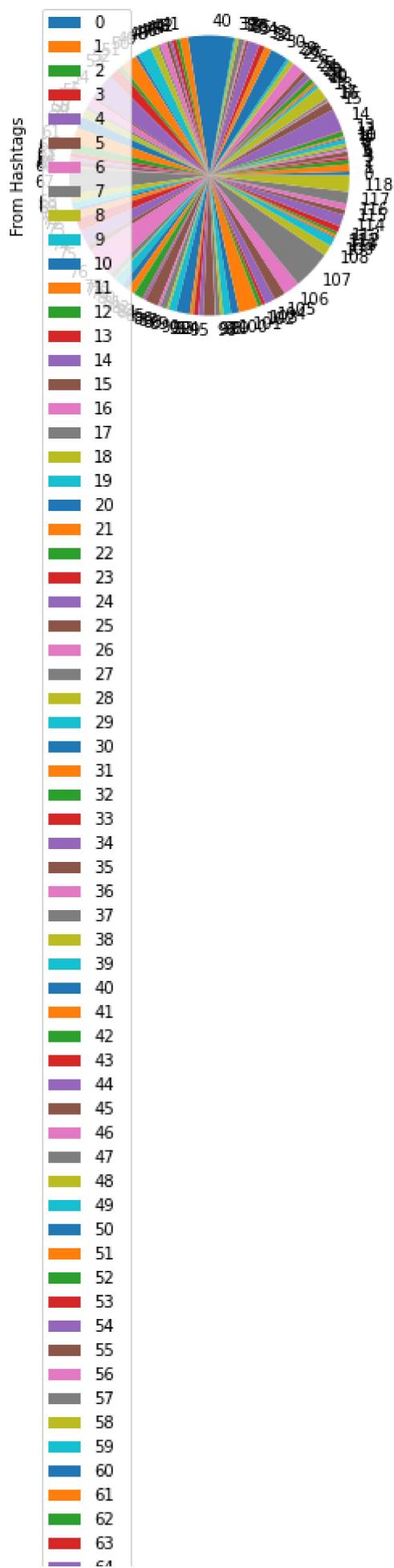
```
Out[19]: <AxesSubplot:>
```

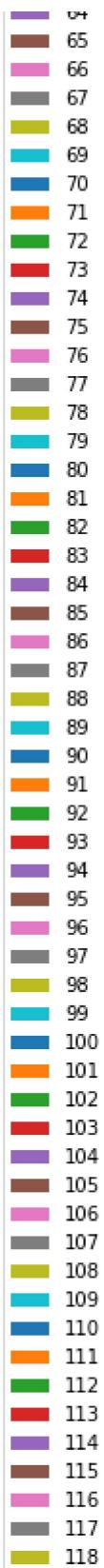


Pie Chart

```
In [20]: data.plot.pie(y="From Hashtags")
```

```
Out[20]: <AxesSubplot:ylabel='From Hashtags'>
```



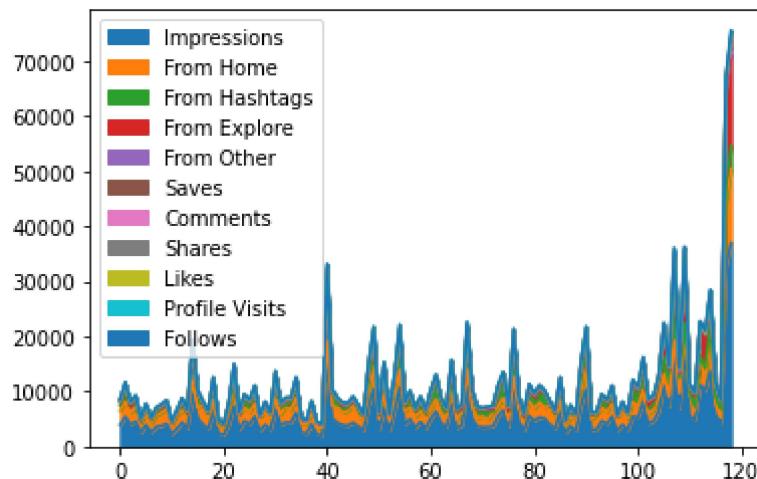


Area

In [21]:

```
data.plot.area()
```

Out[21]: <AxesSubplot:>



In []: