Data visualization

(a)Import library

In [1]:
```
import numpy as np
import pandas as pd
```

b) Import dataset

In [2]:
```
data=pd.read_csv(r"C:\Users\user\Downloads\5_Instagram data.csv")
```

In [3]:
```
data
```

Out[3]:

npressions	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

npressions	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

ws × 13 columns

c)head

In [4]:
```
data.head(10)
```

Out[4]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Foll
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	
5	3884	2046	1214	329	43	74	7	10	144	9	
6	2621	1543	599	333	25	22	5	1	76	26	
7	3541	2071	628	500	60	135	4	9	124	12	
8	3749	2384	857	248	49	155	6	8	159	36	

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	Foll
9	4115	2609	1104	178	46	122	6	3	191	31	

d) tail

In [5]:
```
data.tail(10)
```

Out[5]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	F
109	17713	2449	2141	12389	561	504	3	23	308	70	
110	5563	3813	362	1135	76	149	5	8	163	22	
111	4842	1658	694	2036	310	55	6	4	86	46	
112	11149	4439	747	5762	53	273	4	13	210	61	
113	10206	2371	1624	6000	117	182	10	17	172	237	
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	

e) describe

In [6]:
```
data.describe()
```

Out[6]:

	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

f) shape

In [7]:
```
data.shape
```

Out[7]:
```
(119, 13)
```

g) size

In [8]:
```
data.size
```

Out[8]:
```
1547
```

h) find missing values

In [9]:
```
data.isna()
```

Out[9]:

	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	
1	False	False	False	False	False	False	False	False	False	False	
2	False	False	False	False	False	False	False	False	False	False	
3	False	False	False	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	False	False	False	
...	...	...	...	...	...	...	...	...	...	...	
114	False	False	False	False	False	False	False	False	False	False	
115	False	False	False	False	False	False	False	False	False	False	
116	False	False	False	False	False	False	False	False	False	False	
117	False	False	False	False	False	False	False	False	False	False	
118	False	False	False	False	False	False	False	False	False	False	

119 rows × 13 columns

In [10]:
```
data.isnull().sum()
```

Out[10]:
```
Impressions       0
From Home         0
From Hashtags     0
From Explore      0
From Other        0
Saves             0
Comments          0
Shares            0
Likes             0
Profile Visits    0
Follows           0
Caption           0
Hashtags          0
dtype: int64
```

In [11]:
```
data.dropna(axis=1,how='any')
```

Out[11]:

	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 [13]:
```
data1 = data[['From Home','Saves']]
data1
```

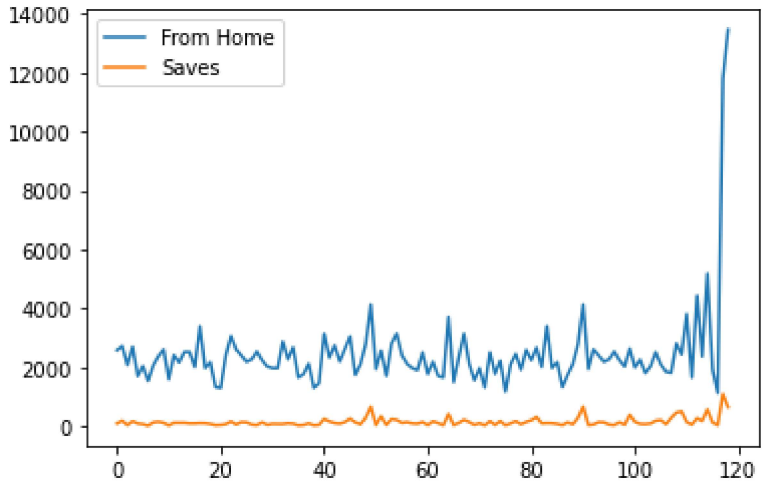
Out[13]:

	From Home	Saves
0	2586	98
1	2727	194
2	2085	41
3	2700	172
4	1704	96
...	...	...
114	5185	573
115	1923	135
116	1133	36
117	11815	1095
118	13473	653

119 rows × 2 columns

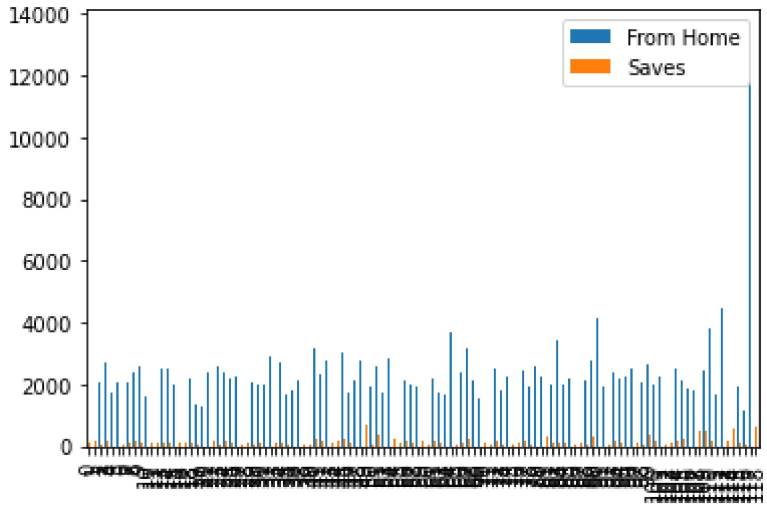
In [14]:
```
data1.plot.line()
```

Out[14]: <AxesSubplot:>



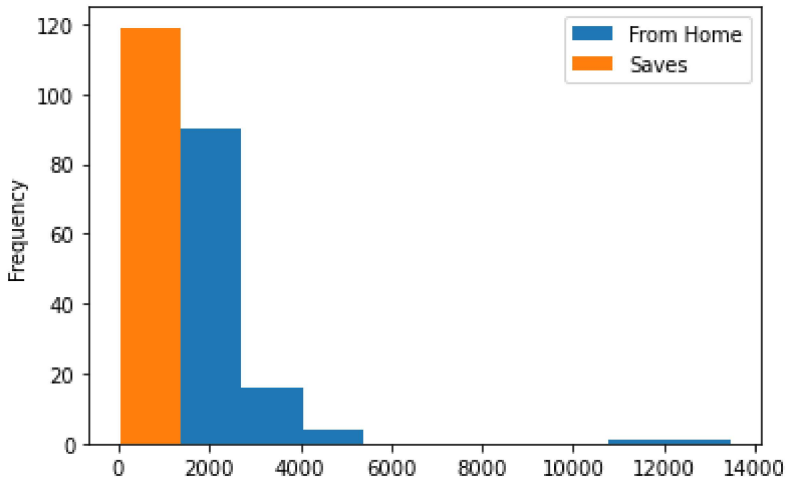
In [15]:
```
data1.plot.bar()
```

Out[15]: <AxesSubplot:>



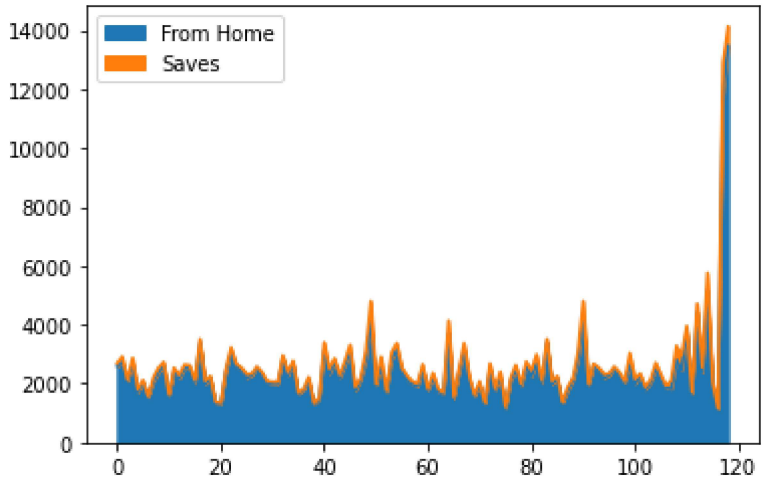
In [16]:
```
data1.plot.hist()
```

Out[16]: <AxesSubplot:ylabel='Frequency'>



In [17]:
```
data1.plot.area()
```

Out[17]: <AxesSubplot:>

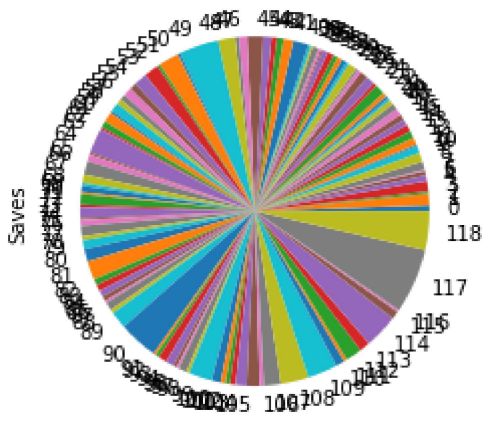


In [19]:
```
data2 = data1['Saves']
data2
```

Out[19]:
```
0        98
1       194
2        41
3       172
4        96
       ...
114     573
115     135
116      36
117    1095
118     653
Name: Saves, Length: 119, dtype: int64
```

In [20]:
```
data2.plot.pie()
```

Out[20]: <AxesSubplot:ylabel='Saves'>



In [ ]: