Instagram Data set import labary

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

Import dataset

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

Out[2]:

	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

Print head first 20 rows

In [3]:
```
data.head(30)
```

14	9453	2525	5799	208	794	100	6	10	294	181
15	5055	2017	2351	298	108	101	7	11	159	17
16	4002	3401	278	128	73	111	17	18	205	16
17	3169	1979	707	341	32	106	8	1	121	21
18	6168	2177	3450	153	296	82	6	6	151	77

Print tail last 7 rows

In [4]:
```
data.tail(15)
```

Out[4]:

	Impressions	From Home	From Hashtags	From Explore	From Other	Saves	Comments	Shares	Likes	Profile Visits	F
104	6348	2517	2660	737	154	188	4	15	194	26	
105	11068	2099	2986	5634	122	214	7	8	250	39	
106	7231	1855	4156	703	309	73	8	3	171	74	
107	17396	1817	10008	5192	251	285	7	7	416	467	
108	6814	2816	2769	900	128	469	8	22	249	24	
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	

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

To print statistical data

```
In [5]: data.describe()
```

Out[5]:

	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

To print rows and coloum

```
In [6]: np.shape(data)
```

```
Out[6]: (119, 13)
```

To print no. of elements

In [7]:
```
np.size(data)
```

Out[7]:
```
1547
```

To print missing values

In [8]:
```
data.isna()
```

Out[8]:

	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

To drop the value in missing place

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

Out[9]:

	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 [10]:
```
dd=data[['Impressions','Likes']]
dd
```

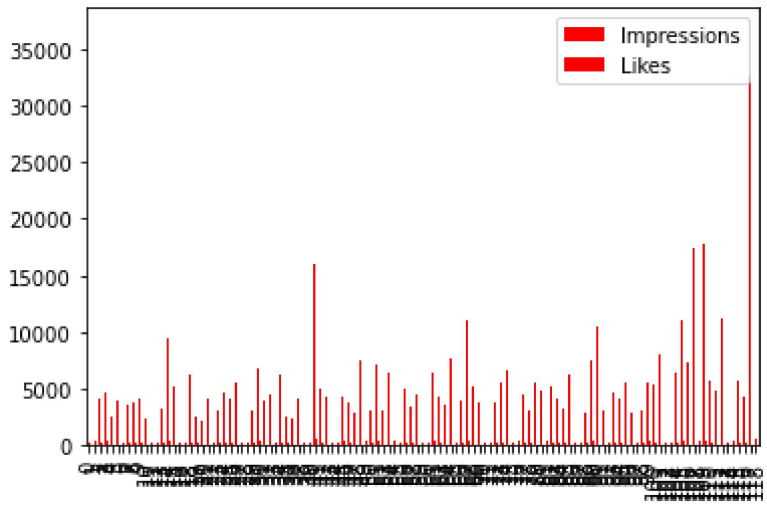
Out[10]:

	Impressions	Likes
0	3920	162
1	5394	224
2	4021	131
3	4528	213
4	2518	123
...	...	...
114	13700	373
115	5731	148
116	4139	92
117	32695	549
118	36919	443

119 rows × 2 columns

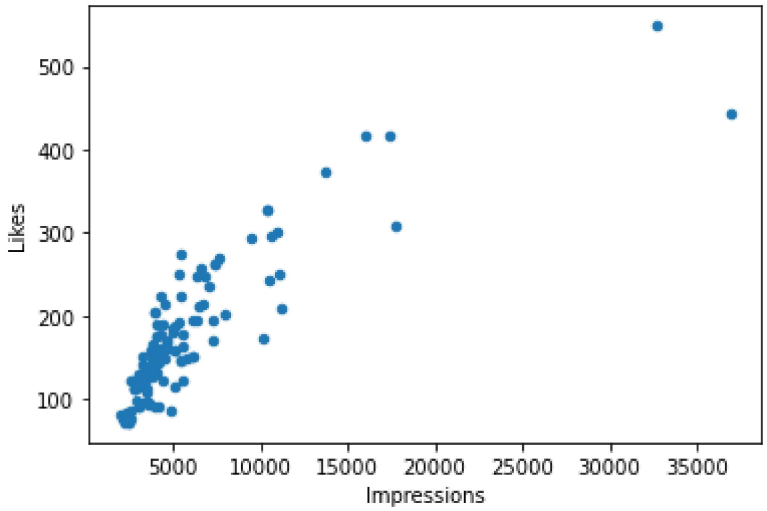
In [11]:
```
dd.plot.bar(color='r')
```

Out[11]: <AxesSubplot:>

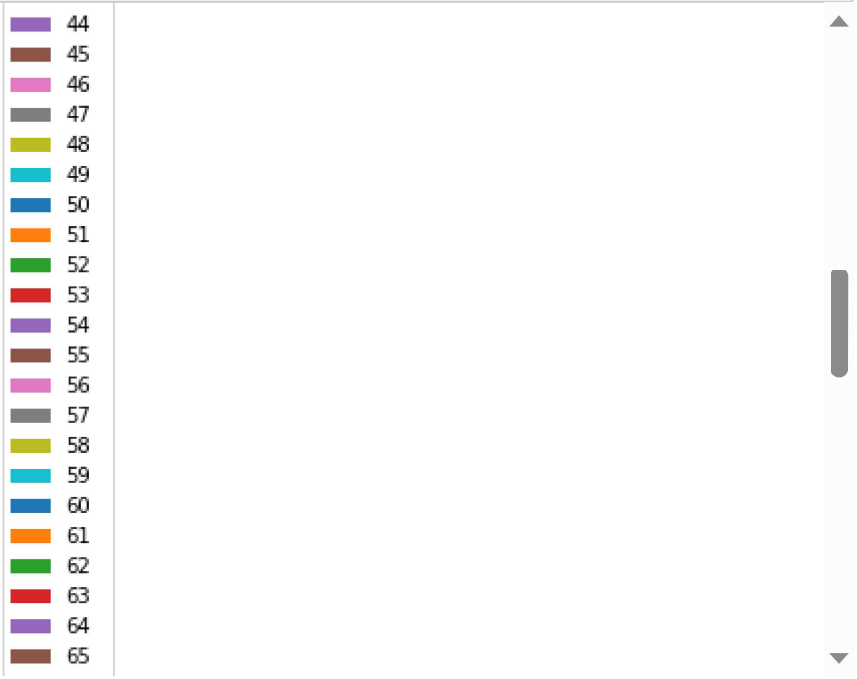


In [12]:
```
dd.plot.scatter(x='Impressions',y='Likes')
```

Out[12]:
```
<AxesSubplot:xlabel='Impressions', ylabel='Likes'>
```

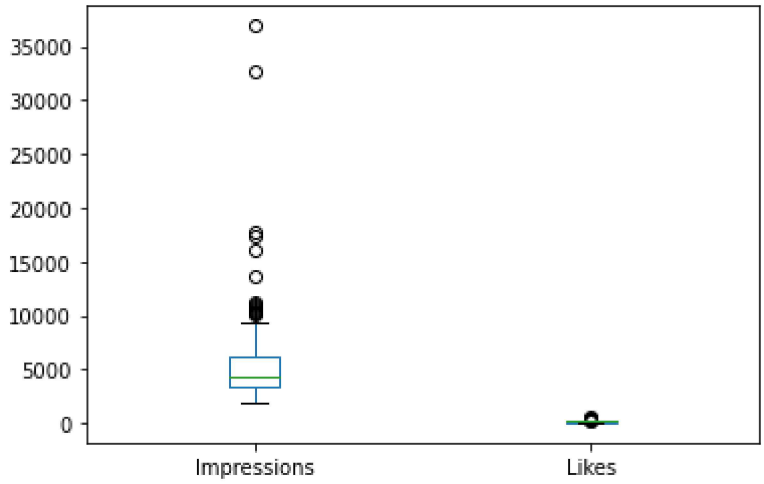


In [13]:
```
dd.plot.pie(y='Likes')
```



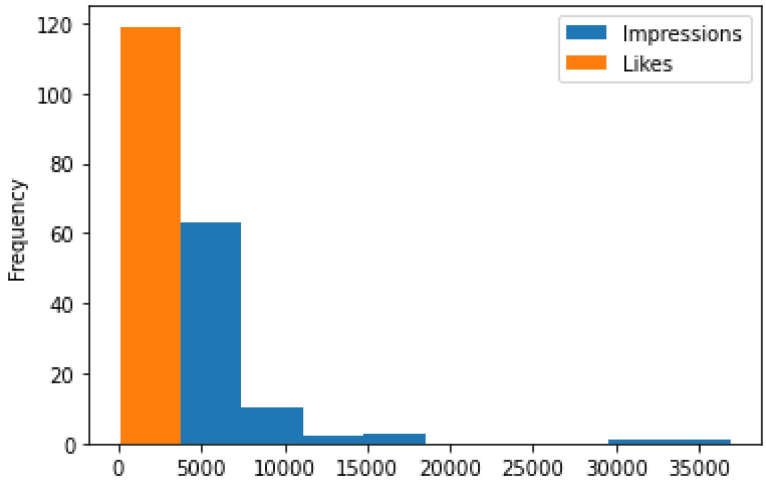
In [14]:
```
dd.plot.box()
```

Out[14]: <AxesSubplot:>



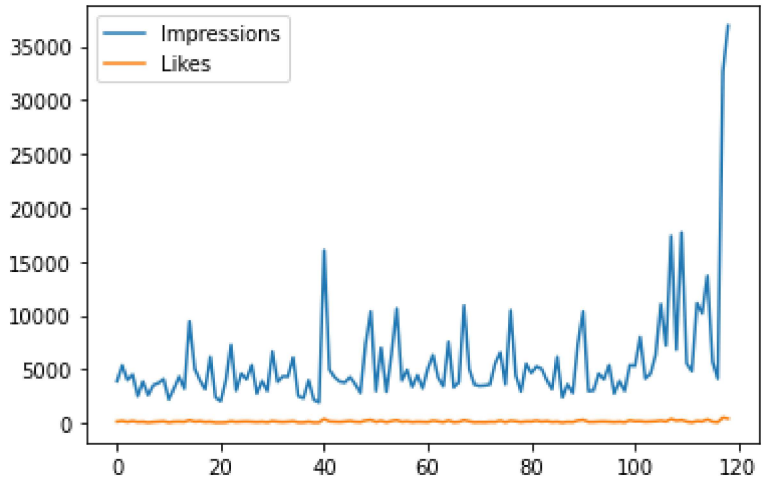
In [15]:
```
dd.plot.hist()
```

Out[15]: <AxesSubplot:ylabel='Frequency'>



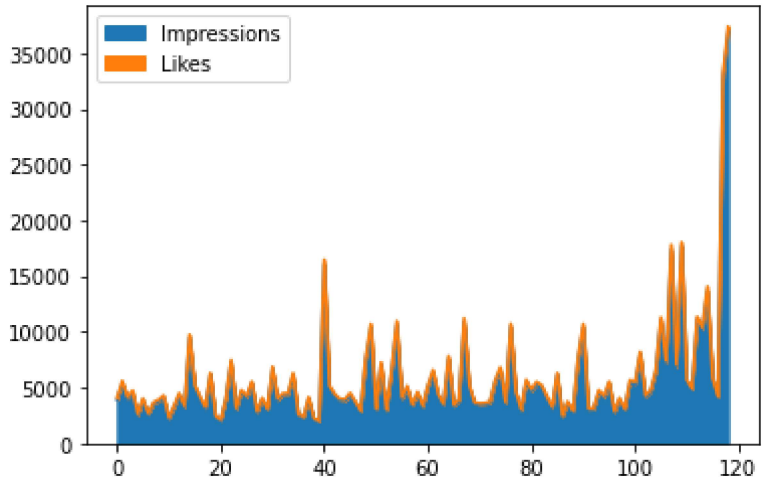
In [16]:
```
dd.plot.line()
```

Out[16]: <AxesSubplot:>



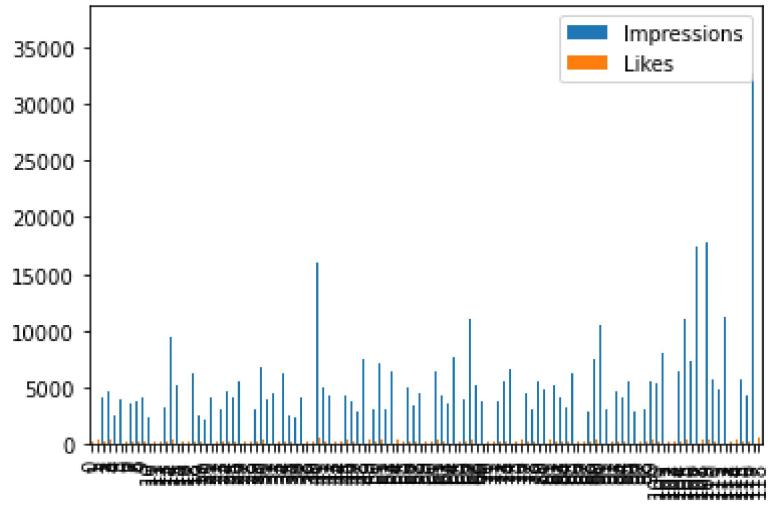
In [17]:
```
dd.plot.area()
```

Out[17]: <AxesSubplot:>



In [18]:
```
dd.plot.bar()
```

Out[18]: <AxesSubplot:>



In [ ]: