04/08/2023 (BOOK-27)

In [237]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler
import re
from sklearn.datasets import load_digits
from sklearn.model_selection import train_test_split
```

In [238]:
```
a=pd.read_csv(r"C:\Users\user\Downloads\Book27.csv")
a
```

Out[238]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	C
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	N
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	NaN	269.6	107.9	823.8	909.1	628.4	411.5	19
2	112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	12
3	113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	94
4	114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	10
...	...	...	...	...	...	...	...	...	...	...	...	...	
92	202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.4	277.6	286.7	5
93	203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.9	316.0	724.9	24
94	204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.6	230.2	316.1	16
95	205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.4	599.6	343.0	3
96	206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.3	595.5	374.2	6

97 rows × 20 columns

In [239]:
```
a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 97 entries, 0 to 96
Data columns (total 20 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   index        97 non-null     int64  
 1   SUBDIVISION  97 non-null     object 
 2   YEAR         97 non-null     int64  
 3   JAN          96 non-null     float64
 4   FEB          96 non-null     float64
 5   MAR          95 non-null     float64
 6   APR          97 non-null     float64
 7   MAY          97 non-null     float64
 8   JUN          96 non-null     float64
 9   JUL          96 non-null     float64
 10  AUG          97 non-null     float64
 11  SEP          97 non-null     float64
 12  OCT          95 non-null     float64
 13  NOV          95 non-null     float64
 14  DEC          95 non-null     float64
 15  ANNUAL       91 non-null     float64
 16  Jan-Feb      96 non-null     float64
 17  Mar-May      95 non-null     float64
 18  Jun-Sep      95 non-null     float64
 19  Oct-Dec      94 non-null     float64
dtypes: float64(17), int64(2), object(1)
memory usage: 15.3+ KB
```

In [240]:
```
b=a.fillna(method='ffill')
b
```

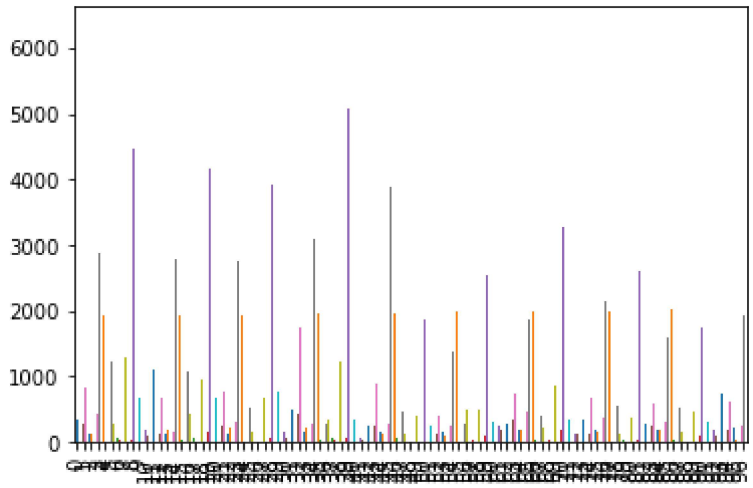
Out[240]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	C
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	N
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	71.1	269.6	107.9	823.8	909.1	628.4	411.5	19
2	112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	12
3	113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	94
4	114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	10
...	...	...	...	...	...	...	...	...	...	...	...	...	...
92	202	ARUNACHAL PRADESH	2011	40.0	51.3	174.5	240.8	219.6	288.4	531.4	277.6	286.7	5
93	203	ARUNACHAL PRADESH	2012	57.8	35.8	134.2	403.4	187.4	645.8	638.9	316.0	724.9	24
94	204	ARUNACHAL PRADESH	2013	18.5	40.5	115.1	175.1	335.8	290.0	329.6	230.2	316.1	16
95	205	ARUNACHAL PRADESH	2014	19.0	101.9	80.3	86.7	299.0	415.8	392.4	599.6	343.0	3
96	206	ARUNACHAL PRADESH	2015	30.8	47.5	97.5	287.1	238.9	637.9	329.3	595.5	374.2	6

97 rows × 20 columns

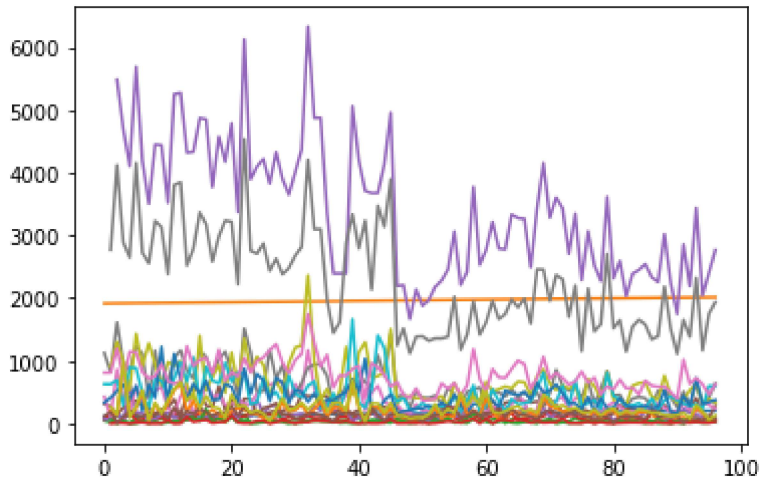
In [241]:
```
b.plot.bar(legend=None)
```

Out[241]: <AxesSubplot:>



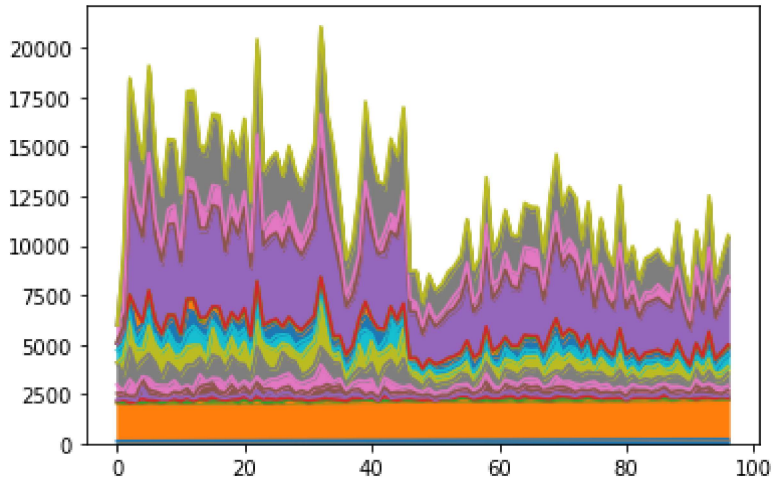
In [242]:
```
b.plot.line(legend=None)
```

Out[242]: <AxesSubplot:>



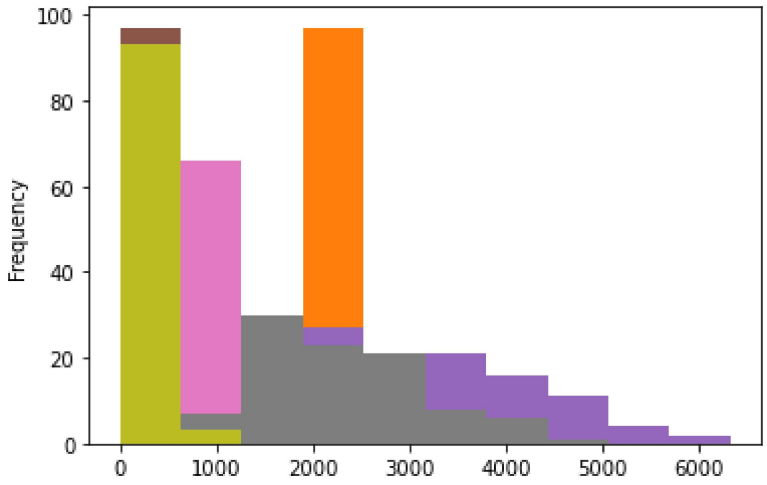
In [243]:
```
b.plot.area(legend=None)
```

Out[243]: <AxesSubplot:>



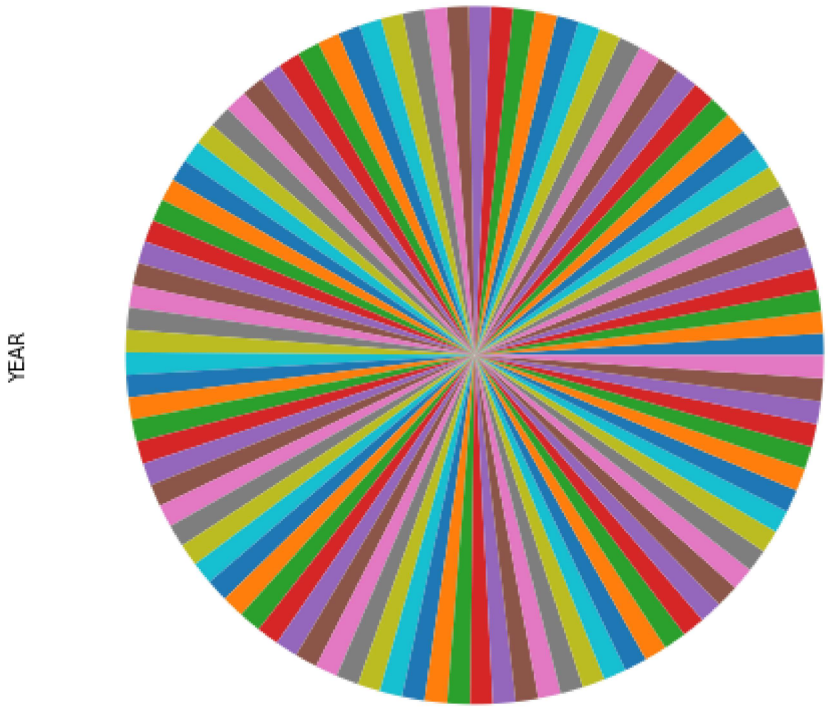
In [244]:
```
b.plot.hist(legend=None)
```

Out[244]: <AxesSubplot:ylabel='Frequency'>



In [245]:
```
b.plot.pie(y='YEAR',figsize=(8,8),labels=None,legend=None)
```

Out[245]: <AxesSubplot:ylabel='YEAR'>



In [ ]: