04/08/2023 (BOOK-2)

In [11]:
```
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 [12]:
```
a=pd.read_csv(r"C:\Users\user\Downloads\Book2.csv")
a
```

Out[12]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	O
0	3887	KERALA	1901	28.7	44.7	51.6	160.0	174.7	824.6	743.0	357.5	197.7	26
1	3888	KERALA	1902	6.7	2.6	57.3	83.9	134.5	390.9	1205.0	315.8	491.6	35
2	3889	KERALA	1903	3.2	18.6	3.1	83.6	249.7	558.6	1022.5	420.2	341.8	35
3	3890	KERALA	1904	23.7	3.0	32.2	71.5	235.7	1098.2	725.5	351.8	222.7	32
4	3891	KERALA	1905	1.2	22.3	9.4	105.9	263.3	850.2	520.5	293.6	217.2	38
...	...	...	...	...	...	...	...	...	...	...	...	...	...
110	3997	KERALA	2011	20.5	45.7	24.1	165.2	124.2	788.5	536.8	492.7	391.2	22
111	3998	KERALA	2012	7.4	11.0	21.0	171.1	95.3	430.3	362.6	501.6	241.1	18
112	3999	KERALA	2013	3.9	40.1	49.9	49.3	119.3	1042.7	830.2	369.7	318.6	25
113	4000	KERALA	2014	4.6	10.3	17.9	95.7	251.0	454.4	677.8	733.9	298.8	35
114	4001	KERALA	2015	3.1	5.8	50.1	214.1	201.8	563.6	406.0	252.2	292.9	30

115 rows × 20 columns

In [13]:
```
a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 115 entries, 0 to 114
Data columns (total 20 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   index        115 non-null    int64  
 1   SUBDIVISION  115 non-null    object 
 2   YEAR         115 non-null    int64  
 3   JAN          115 non-null    float64
 4   FEB          115 non-null    float64
 5   MAR          115 non-null    float64
 6   APR          115 non-null    float64
 7   MAY          115 non-null    float64
 8   JUN          115 non-null    float64
 9   JUL          115 non-null    float64
 10  AUG          115 non-null    float64
 11  SEP          115 non-null    float64
 12  OCT          115 non-null    float64
 13  NOV          115 non-null    float64
 14  DEC          115 non-null    float64
 15  ANNUAL       115 non-null    float64
 16  Jan-Feb      115 non-null    float64
 17  Mar-May      115 non-null    float64
 18  Jun-Sep      115 non-null    float64
 19  Oct-Dec      115 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 18.1+ KB
```

In [14]:
```
b=a.fillna(method='ffill')
b
```

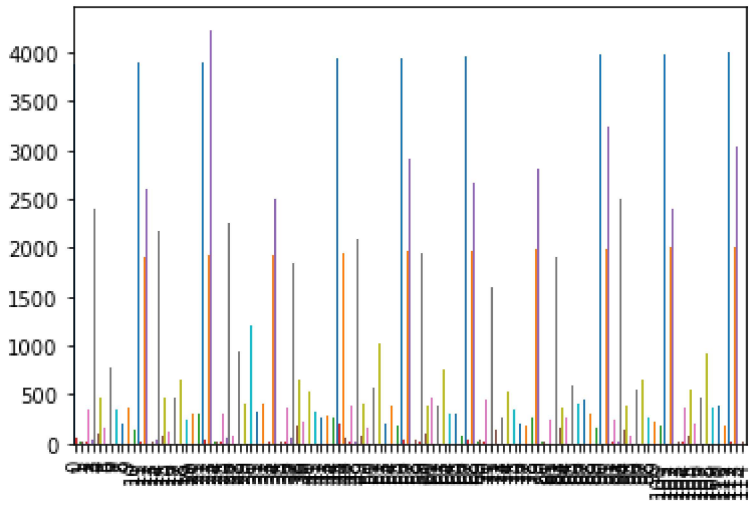
Out[14]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	C
0	3887	KERALA	1901	28.7	44.7	51.6	160.0	174.7	824.6	743.0	357.5	197.7	26
1	3888	KERALA	1902	6.7	2.6	57.3	83.9	134.5	390.9	1205.0	315.8	491.6	35
2	3889	KERALA	1903	3.2	18.6	3.1	83.6	249.7	558.6	1022.5	420.2	341.8	35
3	3890	KERALA	1904	23.7	3.0	32.2	71.5	235.7	1098.2	725.5	351.8	222.7	32
4	3891	KERALA	1905	1.2	22.3	9.4	105.9	263.3	850.2	520.5	293.6	217.2	38
...	...	...	...	...	...	...	...	...	...	...	...	...	
110	3997	KERALA	2011	20.5	45.7	24.1	165.2	124.2	788.5	536.8	492.7	391.2	22
111	3998	KERALA	2012	7.4	11.0	21.0	171.1	95.3	430.3	362.6	501.6	241.1	18
112	3999	KERALA	2013	3.9	40.1	49.9	49.3	119.3	1042.7	830.2	369.7	318.6	25
113	4000	KERALA	2014	4.6	10.3	17.9	95.7	251.0	454.4	677.8	733.9	298.8	35
114	4001	KERALA	2015	3.1	5.8	50.1	214.1	201.8	563.6	406.0	252.2	292.9	30

115 rows × 20 columns

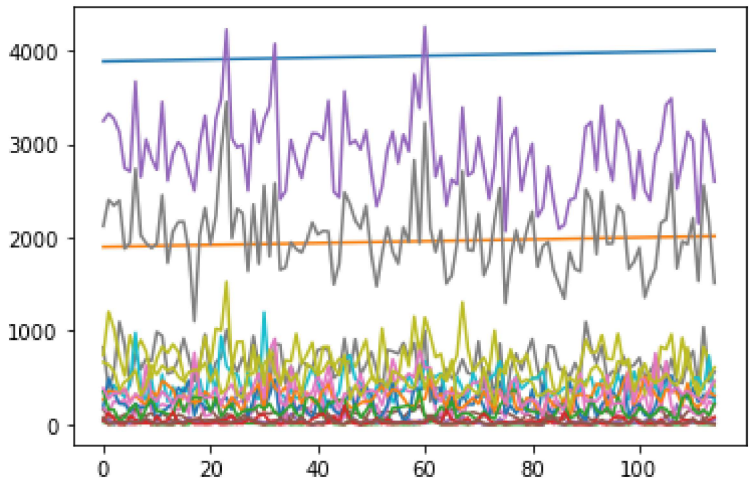
In [15]:
```
b.plot.bar(legend=None)
```

Out[15]: <AxesSubplot:>



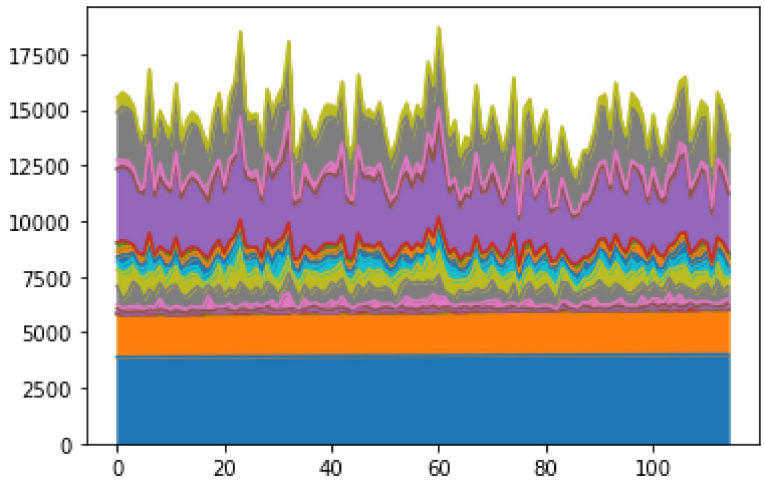
In [16]:
```
b.plot.line(legend=None)
```

Out[16]: <AxesSubplot:>



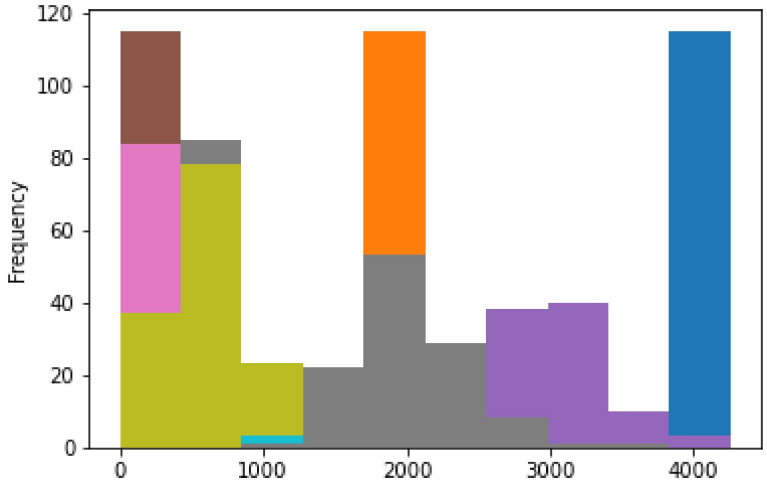
In [17]:
```
b.plot.area(legend=None)
```

Out[17]: <AxesSubplot:>



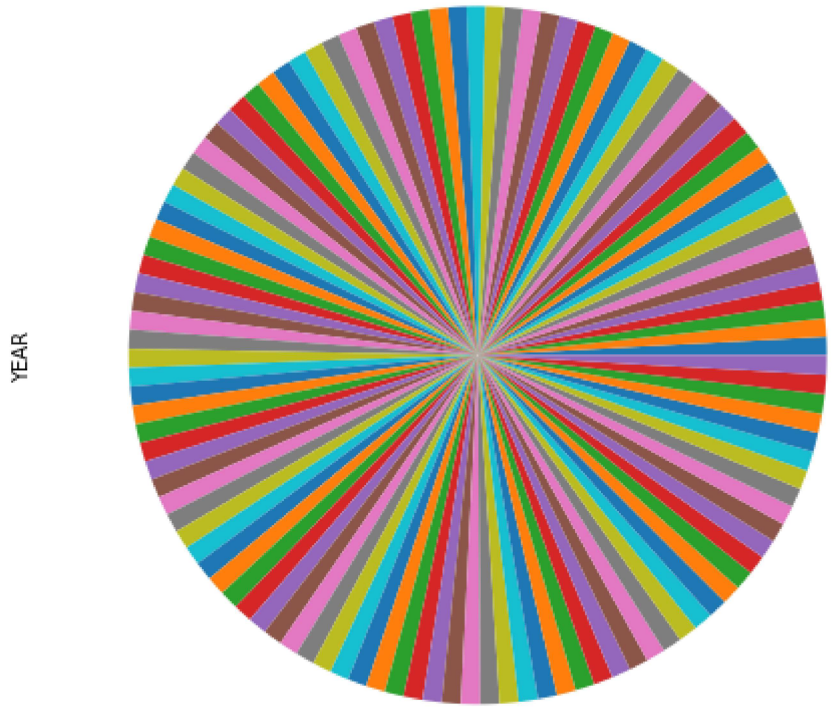
In [18]:
```
b.plot.hist(legend=None)
```

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



In [19]:
```
b.plot.pie(y='YEAR',figsize=(8,8),labels=None,legend=None)
```

Out[19]:
```
<AxesSubplot:ylabel='YEAR'>
```



In [ ]:
```

```