04/08/2023 (BOOK-13)

```
In [110]: 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 [111]: a=pd.read_csv(r"C:\Users\user\Downloads\Book13.csv")
          a
```

Out[111]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	2622	MADHYA MAHARASHTRA	1901	18.8	0.6	7.7	36.6	30.4	107.7	215.9	194.1	83.7	68.7
1	2623	MADHYA MAHARASHTRA	1902	7.8	0.0	0.1	5.0	9.8	102.6	210.9	114.5	169.5	60.4
2	2624	MADHYA MAHARASHTRA	1903	7.6	0.0	0.0	3.2	77.2	86.3	281.8	155.5	142.3	74.2
3	2625	MADHYA MAHARASHTRA	1904	0.4	4.7	1.7	3.0	18.7	114.6	126.5	59.5	183.0	91.1
4	2626	MADHYA MAHARASHTRA	1905	0.0	1.2	0.0	2.3	23.6	65.0	252.8	79.0	52.6	52.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...
110	2732	MADHYA MAHARASHTRA	2011	0.0	0.3	0.3	5.0	2.9	133.3	261.4	238.1	148.4	62.8
111	2733	MADHYA MAHARASHTRA	2012	0.0	0.0	0.0	3.0	1.4	67.9	203.0	187.8	129.5	95.2
112	2734	MADHYA MAHARASHTRA	2013	0.1	5.3	0.8	5.7	6.0	212.4	311.8	147.0	210.3	57.8
113	2735	MADHYA MAHARASHTRA	2014	3.1	6.2	24.4	7.5	29.8	44.0	277.9	240.3	120.4	38.5
114	2736	MADHYA MAHARASHTRA	2015	1.4	0.8	41.2	9.6	24.4	177.0	111.7	67.2	146.6	48.3

115 rows × 20 columns

```
In [112]: 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 [113]: b=a.fillna(method='ffill')
          b
```
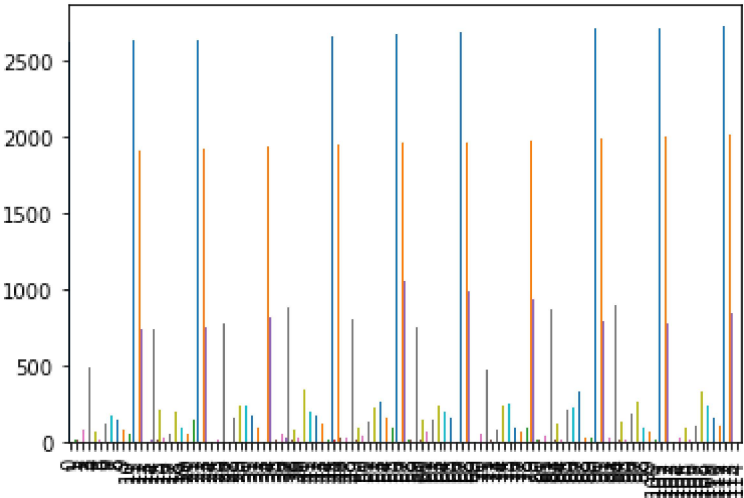
Out[113]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	2622	MADHYA MAHARASHTRA	1901	18.8	0.6	7.7	36.6	30.4	107.7	215.9	194.1	83.7	68.7
1	2623	MADHYA MAHARASHTRA	1902	7.8	0.0	0.1	5.0	9.8	102.6	210.9	114.5	169.5	60.4
2	2624	MADHYA MAHARASHTRA	1903	7.6	0.0	0.0	3.2	77.2	86.3	281.8	155.5	142.3	74.2
3	2625	MADHYA MAHARASHTRA	1904	0.4	4.7	1.7	3.0	18.7	114.6	126.5	59.5	183.0	91.7
4	2626	MADHYA MAHARASHTRA	1905	0.0	1.2	0.0	2.3	23.6	65.0	252.8	79.0	52.6	52.9
...	...	...	...	...	...	...	...	...	...	...	...	...	..
110	2732	MADHYA MAHARASHTRA	2011	0.0	0.3	0.3	5.0	2.9	133.3	261.4	238.1	148.4	62.8
111	2733	MADHYA MAHARASHTRA	2012	0.0	0.0	0.0	3.0	1.4	67.9	203.0	187.8	129.5	95.2
112	2734	MADHYA MAHARASHTRA	2013	0.1	5.3	0.8	5.7	6.0	212.4	311.8	147.0	210.3	57.8
113	2735	MADHYA MAHARASHTRA	2014	3.1	6.2	24.4	7.5	29.8	44.0	277.9	240.3	120.4	38.5
114	2736	MADHYA MAHARASHTRA	2015	1.4	0.8	41.2	9.6	24.4	177.0	111.7	67.2	146.6	48.3

115 rows × 20 columns
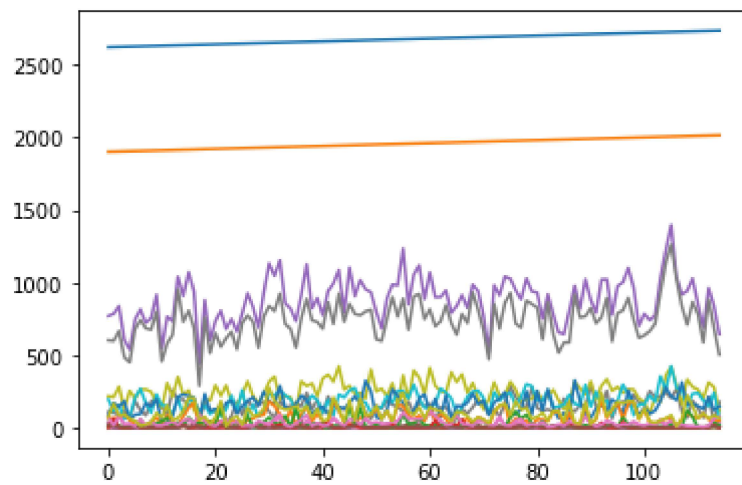
```
In [114]: b.plot.bar(legend=None)
```

Out[114]: <AxesSubplot:>
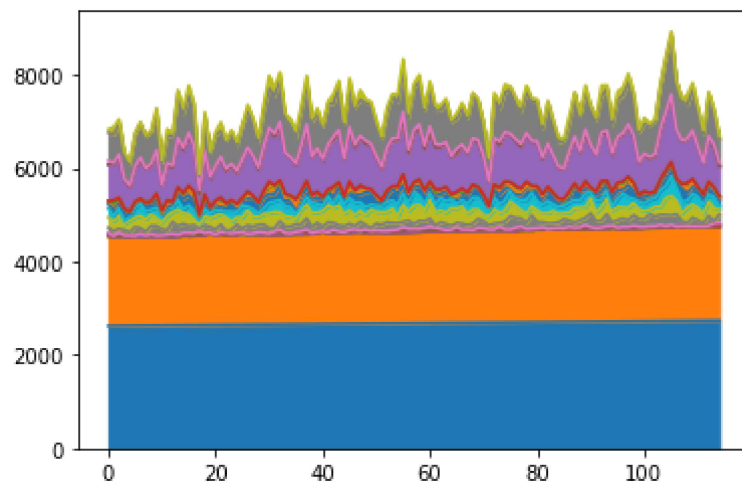
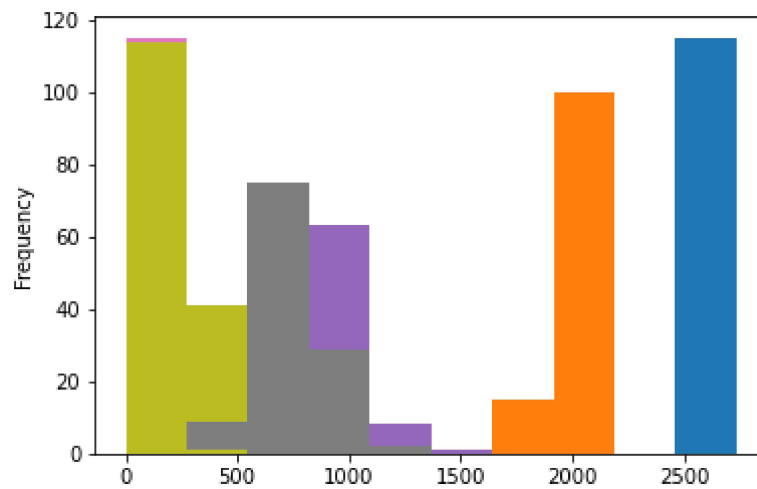In [115]: `b.plot.line(legend=None)`

Out[115]: `<AxesSubplot:>`

In [116]: `b.plot.area(legend=None)`
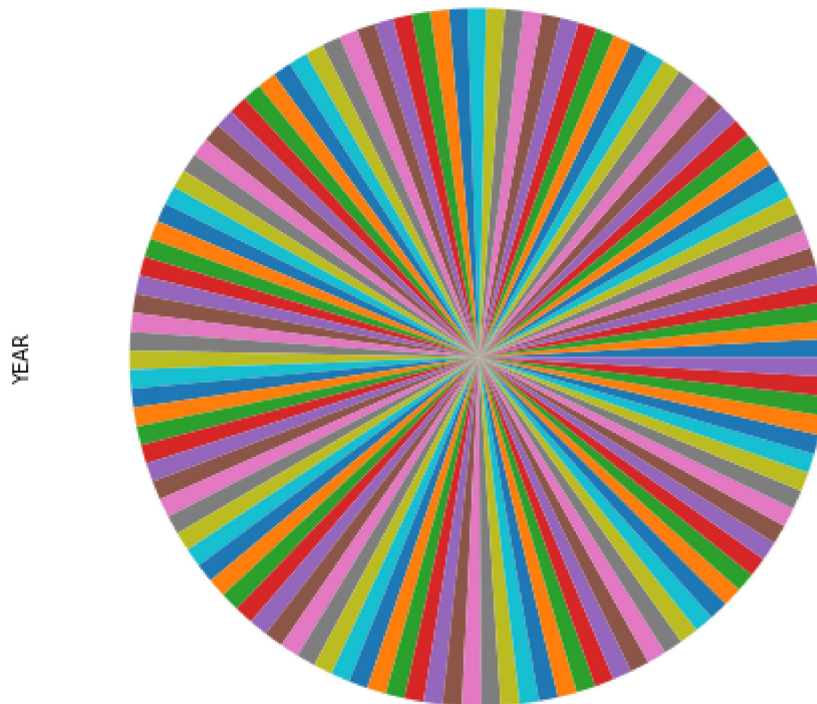
Out[116]: `<AxesSubplot:>`

In [117]: `b.plot.hist(legend=None)`

Out[117]: `<AxesSubplot:ylabel='Frequency'>`

In [118]: `b.plot.pie(y='YEAR',figsize=(8,8),labels=None,legend=None)`

Out[118]: `<AxesSubplot:ylabel='YEAR'>`

In [ ]: