

mk 04/08/2023

```
In [1]: 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 [4]: a=pd.read_csv(r"C:\Users\user\Downloads\Book1.csv")
a
```

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	C	
	0	4002	LAKSHADWEEP	1901	22.6	86.4	114.8	263.8	37.3	459.0	0.0	0.0	46.7	18
	1	4003	LAKSHADWEEP	1902	99.3	9.6	32.6	40.4	179.1	374.2	413.3	170.0	214.3	38
	2	4004	LAKSHADWEEP	1903	63.5	95.0	0.0	29.5	144.1	212.4	261.8	202.0	292.1	7
	3	4005	LAKSHADWEEP	1904	0.0	0.0	13.5	13.2	143.3	261.3	256.0	38.9	219.9	18
	4	4006	LAKSHADWEEP	1905	62.4	0.0	0.0	0.0	166.7	400.7	68.7	377.5	107.5	28
	...	...	...	...	...	...	...	...	...	...	...	...	...	...
	109	4111	LAKSHADWEEP	2011	5.1	2.8	3.1	85.9	107.2	153.6	350.2	254.0	255.2	18
	110	4112	LAKSHADWEEP	2012	19.2	0.1	1.6	76.8	21.2	327.0	231.5	381.2	179.8	14
	111	4113	LAKSHADWEEP	2013	26.2	34.4	37.5	5.3	88.3	426.2	296.4	154.4	180.0	7
	112	4114	LAKSHADWEEP	2014	53.2	16.1	4.4	14.9	57.4	244.1	116.1	466.1	132.2	16
	113	4115	LAKSHADWEEP	2015	2.2	0.5	3.7	87.1	133.1	296.6	257.5	146.4	160.4	16

114 rows × 20 columns



```
In [5]: a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 0 to 113
Data columns (total 20 columns):
#   Column                Non-Null Count  Dtype
---  -
0   index                  114 non-null    int64
1   SUBDIVISION            114 non-null    object
2   YEAR                   114 non-null    int64
3   JAN                    112 non-null    float64
4   FEB                    113 non-null    float64
5   MAR                    112 non-null    float64
6   APR                    112 non-null    float64
7   MAY                    112 non-null    float64
8   JUN                    112 non-null    float64
9   JUL                    111 non-null    float64
10  AUG                    112 non-null    float64
11  SEP                    111 non-null    float64
12  OCT                    111 non-null    float64
13  NOV                    108 non-null    float64
14  DEC                    110 non-null    float64
15  ANNUAL                 103 non-null    float64
16  Jan-Feb               111 non-null    float64
17  Mar-May               110 non-null    float64
18  Jun-Sep               110 non-null    float64
19  Oct-Dec               108 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 17.9+ KB
```

```
In [8]: b=a.fillna(method='ffill')
b
```

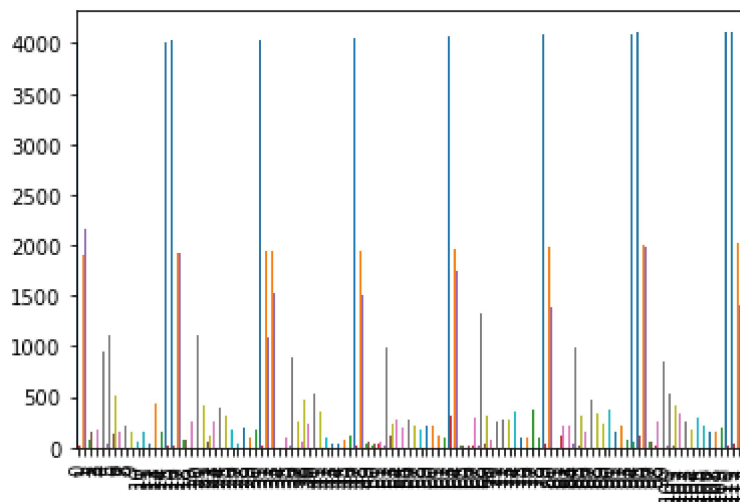
Out[8]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	4002	LAKSHADWEEP	1901	22.6	86.4	114.8	263.8	37.3	459.0	0.0	0.0	46.7	18.0
1	4003	LAKSHADWEEP	1902	99.3	9.6	32.6	40.4	179.1	374.2	413.3	170.0	214.3	38.0
2	4004	LAKSHADWEEP	1903	63.5	95.0	0.0	29.5	144.1	212.4	261.8	202.0	292.1	7.0
3	4005	LAKSHADWEEP	1904	0.0	0.0	13.5	13.2	143.3	261.3	256.0	38.9	219.9	18.0
4	4006	LAKSHADWEEP	1905	62.4	0.0	0.0	0.0	166.7	400.7	68.7	377.5	107.5	22.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
109	4111	LAKSHADWEEP	2011	5.1	2.8	3.1	85.9	107.2	153.6	350.2	254.0	255.2	17.0
110	4112	LAKSHADWEEP	2012	19.2	0.1	1.6	76.8	21.2	327.0	231.5	381.2	179.8	14.0
111	4113	LAKSHADWEEP	2013	26.2	34.4	37.5	5.3	88.3	426.2	296.4	154.4	180.0	7.0
112	4114	LAKSHADWEEP	2014	53.2	16.1	4.4	14.9	57.4	244.1	116.1	466.1	132.2	16.0
113	4115	LAKSHADWEEP	2015	2.2	0.5	3.7	87.1	133.1	296.6	257.5	146.4	160.4	16.0

114 rows × 20 columns

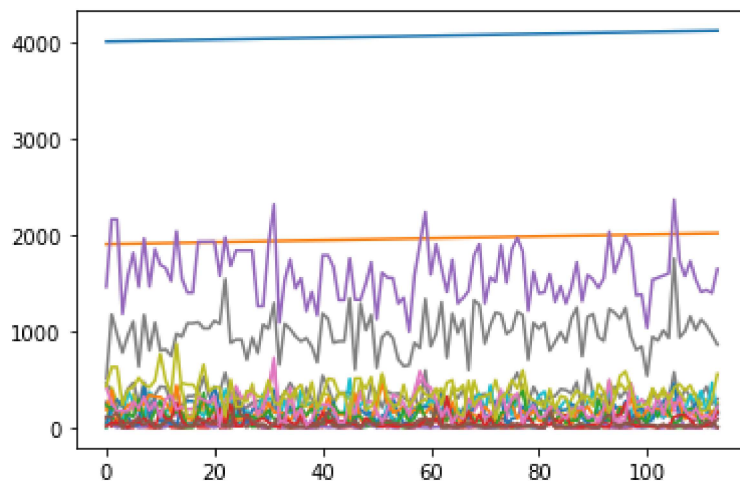
```
In [9]: b.plot.bar(legend=None)
```

```
Out[9]: <AxesSubplot:>
```



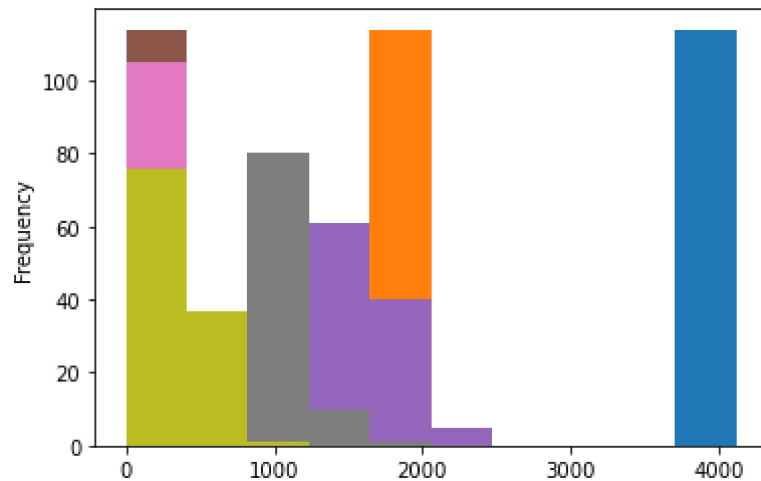
```
In [10]: b.plot.line(legend=None)
```

```
Out[10]: <AxesSubplot:>
```



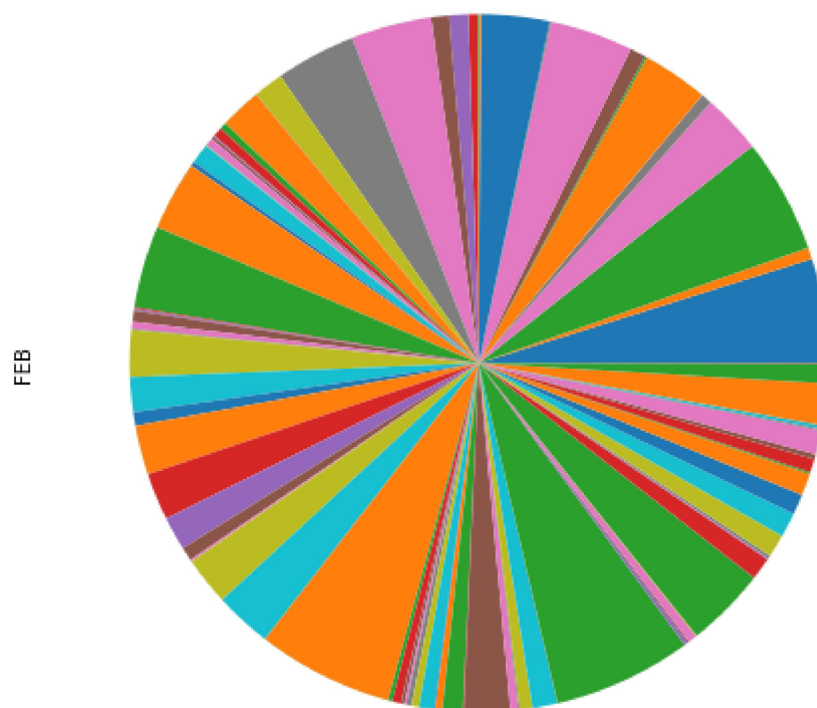
```
In [12]: b.plot.hist(legend=None)
```

```
Out[12]: <AxesSubplot:ylabel='Frequency'>
```



```
In [13]: a.plot.pie(y='FEB',figsize=(8,8),labels=None,legend=None)
```

```
Out[13]: <AxesSubplot:ylabel='FEB'>
```



```
In [ ]:
```

