

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 [2]: a=pd.read_csv(r"C:\Users\user\Downloads\Book32.csv")
a
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	667	ORISSA	1901	39.5	65.1	16.1	51.6	79.0	78.2	288.4	307.7	185.3	76.6
1	668	ORISSA	1902	3.4	0.2	14.2	101.1	56.7	108.3	437.4	349.1	202.7	33.7
2	669	ORISSA	1903	19.7	18.9	10.5	34.6	73.3	154.3	410.4	295.2	265.6	228.1
3	670	ORISSA	1904	0.2	12.2	20.6	10.1	100.2	342.9	336.7	350.4	227.8	111.1
4	671	ORISSA	1905	24.3	17.2	66.3	56.9	107.5	92.0	330.1	281.4	344.1	36.4
...
110	777	ORISSA	2011	3.7	16.2	4.9	58.2	75.6	210.1	199.6	358.6	398.7	20.7
111	778	ORISSA	2012	50.8	3.6	0.9	34.8	21.3	169.6	324.3	417.0	242.4	66.6
112	779	ORISSA	2013	3.3	7.8	2.1	53.6	57.7	272.6	380.0	254.9	208.1	391.6
113	780	ORISSA	2014	0.0	17.6	25.1	11.7	111.9	92.2	496.2	386.3	281.1	111.1
114	781	ORISSA	2015	15.1	3.3	10.5	67.6	32.6	238.6	294.8	264.0	237.0	24.7

115 rows × 20 columns



In [3]: 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 [4]: b=a.fillna(method='ffill')
b

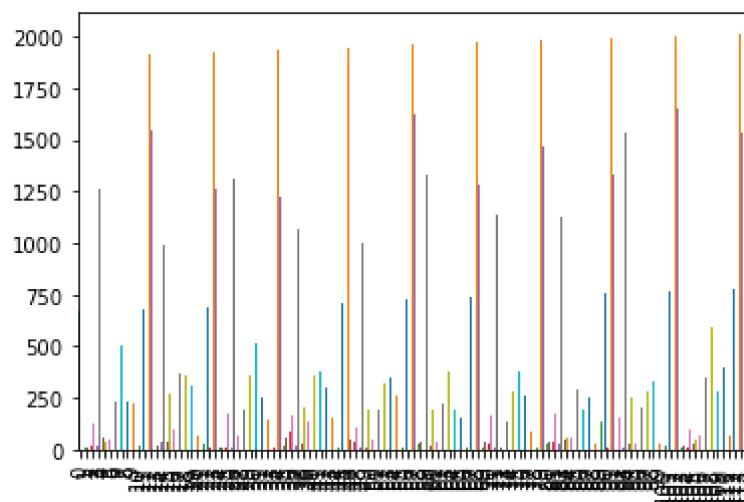
Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	667	ORISSA	1901	39.5	65.1	16.1	51.6	79.0	78.2	288.4	307.7	185.3	76.0
1	668	ORISSA	1902	3.4	0.2	14.2	101.1	56.7	108.3	437.4	349.1	202.7	33.0
2	669	ORISSA	1903	19.7	18.9	10.5	34.6	73.3	154.3	410.4	295.2	265.6	228.0
3	670	ORISSA	1904	0.2	12.2	20.6	10.1	100.2	342.9	336.7	350.4	227.8	111.0
4	671	ORISSA	1905	24.3	17.2	66.3	56.9	107.5	92.0	330.1	281.4	344.1	36.0
...
110	777	ORISSA	2011	3.7	16.2	4.9	58.2	75.6	210.1	199.6	358.6	398.7	20.0
111	778	ORISSA	2012	50.8	3.6	0.9	34.8	21.3	169.6	324.3	417.0	242.4	66.0
112	779	ORISSA	2013	3.3	7.8	2.1	53.6	57.7	272.6	380.0	254.9	208.1	391.0
113	780	ORISSA	2014	0.0	17.6	25.1	11.7	111.9	92.2	496.2	386.3	281.1	111.0
114	781	ORISSA	2015	15.1	3.3	10.5	67.6	32.6	238.6	294.8	264.0	237.0	24.0

115 rows × 20 columns

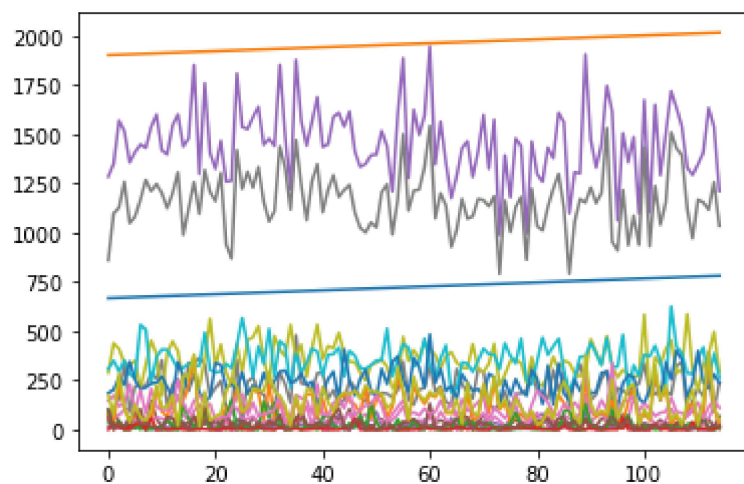
```
In [5]: b.plot.bar(legend=None)
```

```
Out[5]: <AxesSubplot:>
```



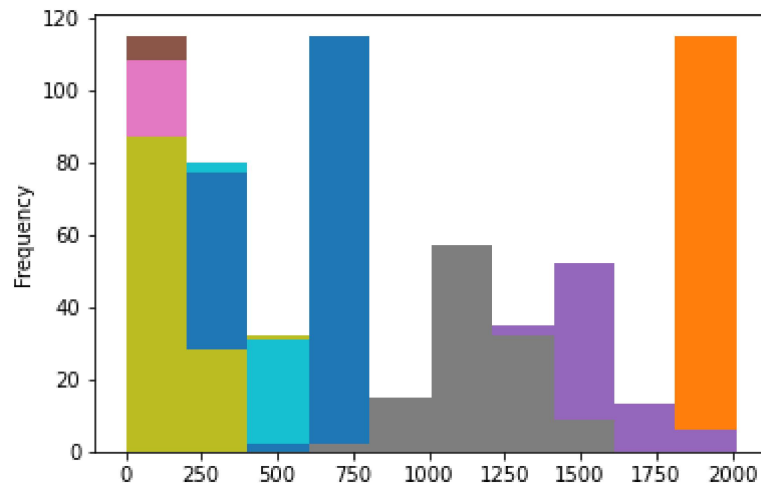
```
In [6]: b.plot.line(legend=None)
```

```
Out[6]: <AxesSubplot:>
```



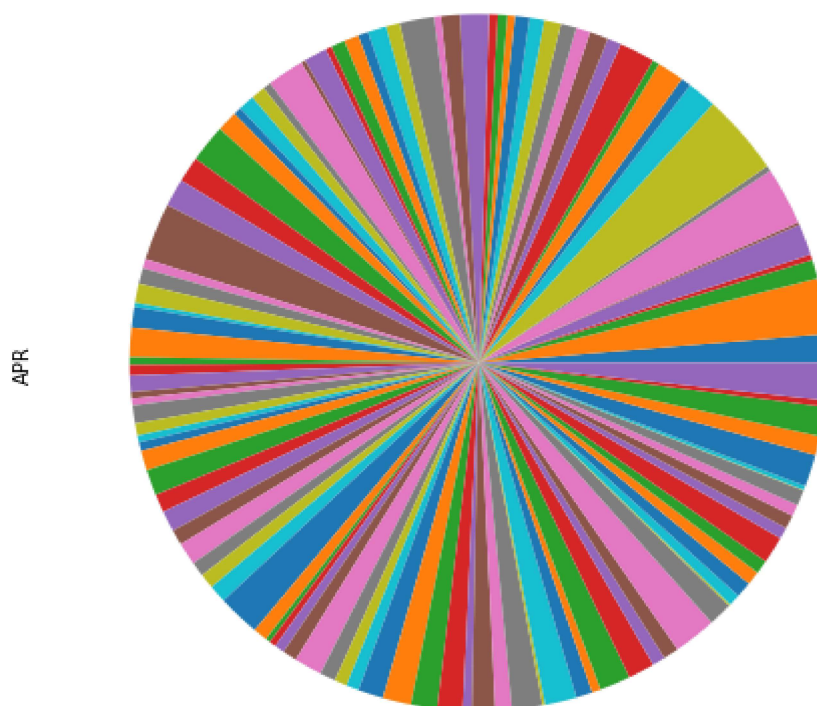
```
In [7]: b.plot.hist(legend=None)
```

```
Out[7]: <AxesSubplot:ylabel='Frequency'>
```



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

```
Out[8]: <AxesSubplot:ylabel='APR'>
```



```
In [ ]:
```

