

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\Book26.csv")
a
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	0	ANDAMAN & NICOBAR ISLANDS	1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6	386.1
1	1	ANDAMAN & NICOBAR ISLANDS	1902	0.0	159.8	12.2	0.0	446.1	537.1	228.9	753.7	666.2	159.8
2	2	ANDAMAN & NICOBAR ISLANDS	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0	187.5
3	3	ANDAMAN & NICOBAR ISLANDS	1904	9.4	14.7	0.0	202.4	304.5	495.1	502.0	160.1	820.4	228.9
4	4	ANDAMAN & NICOBAR ISLANDS	1905	1.3	0.0	3.3	26.9	279.5	628.7	368.7	330.5	297.0	269.5
...
105	105	ANDAMAN & NICOBAR ISLANDS	2011	265.9	84.8	272.8	111.4	326.5	383.2	583.2	441.5	757.1	272.8
106	106	ANDAMAN & NICOBAR ISLANDS	2012	119.9	45.6	30.9	55.8	533.9	458.2	317.3	369.6	868.9	265.9
107	107	ANDAMAN & NICOBAR ISLANDS	2013	67.1	37.6	43.0	46.3	509.3	777.0	564.8	336.7	473.6	43.0
108	108	ANDAMAN & NICOBAR ISLANDS	2014	41.9	8.6	0.0	11.1	238.0	416.6	467.6	321.6	412.9	40.0
109	109	ANDAMAN & NICOBAR ISLANDS	2015	126.8	7.6	3.1	138.2	331.9	346.4	328.9	480.0	523.3	26.9

110 rows × 20 columns



```
In [3]: a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 110 entries, 0 to 109
Data columns (total 20 columns):
#   Column          Non-Null Count  Dtype
---  -
0   index            110 non-null   int64
1   SUBDIVISION      110 non-null   object
2   YEAR             110 non-null   int64
3   JAN              110 non-null   float64
4   FEB              110 non-null   float64
5   MAR              108 non-null   float64
6   APR              108 non-null   float64
7   MAY              109 non-null   float64
8   JUN              108 non-null   float64
9   JUL              108 non-null   float64
10  AUG              108 non-null   float64
11  SEP              107 non-null   float64
12  OCT              108 non-null   float64
13  NOV              108 non-null   float64
14  DEC              107 non-null   float64
15  ANNUAL           104 non-null   float64
16  Jan-Feb          110 non-null   float64
17  Mar-May          107 non-null   float64
18  Jun-Sep          107 non-null   float64
19  Oct-Dec          107 non-null   float64
dtypes: float64(17), int64(2), object(1)
memory usage: 17.3+ KB
```

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

Out[4]:

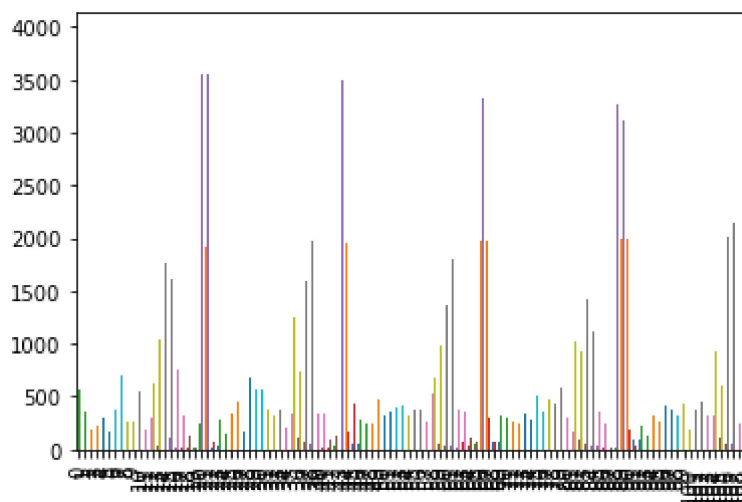
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	0	ANDAMAN & NICOBAR ISLANDS	1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6	381.1
1	1	ANDAMAN & NICOBAR ISLANDS	1902	0.0	159.8	12.2	0.0	446.1	537.1	228.9	753.7	666.2	159.8
2	2	ANDAMAN & NICOBAR ISLANDS	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0	181.1
3	3	ANDAMAN & NICOBAR ISLANDS	1904	9.4	14.7	0.0	202.4	304.5	495.1	502.0	160.1	820.4	228.9
4	4	ANDAMAN & NICOBAR ISLANDS	1905	1.3	0.0	3.3	26.9	279.5	628.7	368.7	330.5	297.0	269.5
...
105	105	ANDAMAN & NICOBAR ISLANDS	2011	265.9	84.8	272.8	111.4	326.5	383.2	583.2	441.5	757.1	272.8
106	106	ANDAMAN & NICOBAR ISLANDS	2012	119.9	45.6	30.9	55.8	533.9	458.2	317.3	369.6	868.9	265.9
107	107	ANDAMAN & NICOBAR ISLANDS	2013	67.1	37.6	43.0	46.3	509.3	777.0	564.8	336.7	473.6	43.0
108	108	ANDAMAN & NICOBAR ISLANDS	2014	41.9	8.6	0.0	11.1	238.0	416.6	467.6	321.6	412.9	40.0
109	109	ANDAMAN & NICOBAR ISLANDS	2015	126.8	7.6	3.1	138.2	331.9	346.4	328.9	480.0	523.3	26.9

110 rows × 20 columns



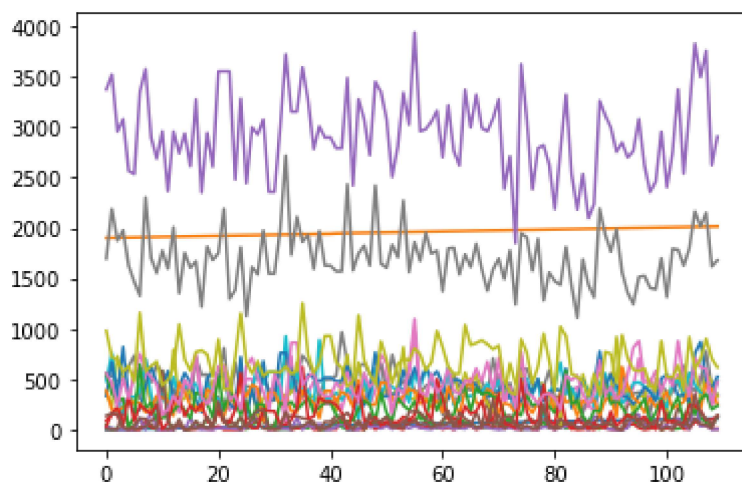
```
In [5]: b.plot.bar(legend=None)
```

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



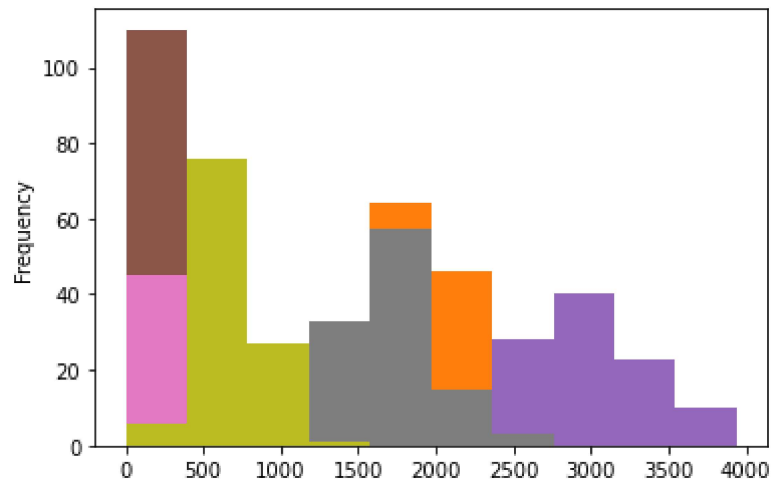
```
In [6]: b.plot.line(legend=None)
```

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



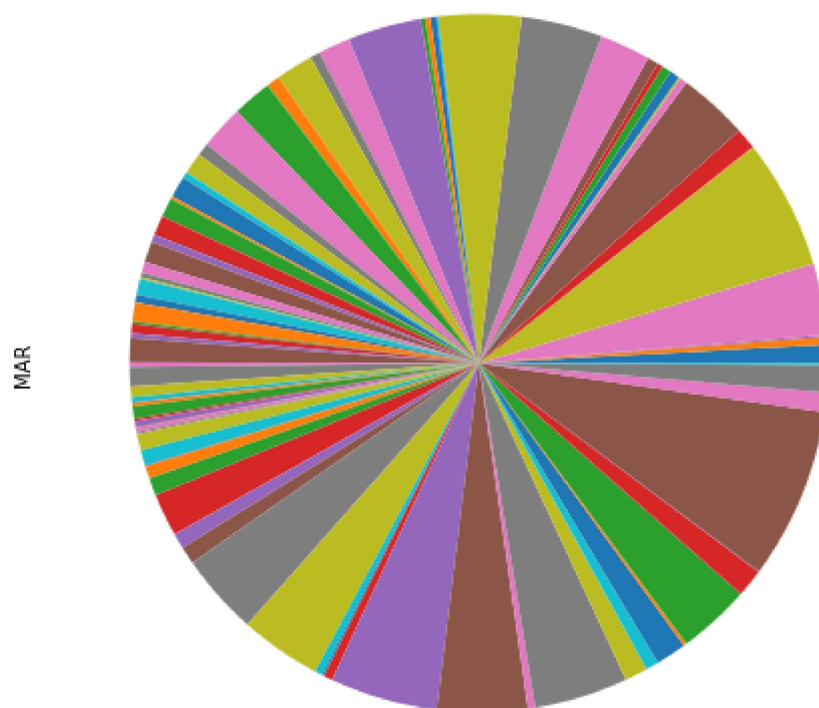
```
In [7]: b.plot.hist(legend=None)
```

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



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

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



```
In [ ]:
```

