

Vijay(Book5) 04/08/2023

```
In [1]: 1 import numpy as np
        2 import pandas as pd
        3 import matplotlib.pyplot as plt
        4 import seaborn as sns
        5 from sklearn.linear_model import LogisticRegression
        6 from sklearn.preprocessing import StandardScaler
        7 import re
        8 from sklearn.datasets import load_digits
        9 from sklearn.model_selection import train_test_split
```

```
In [2]: 1 a=pd.read_csv(r"C:\Users\user\Downloads\Book5.csv")
        2 a
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	3427	TAMIL NADU	1901	24.5	39.1	21.7	36.0	74.0	41.8	49.3	67.9	191.1	122.3
1	3428	TAMIL NADU	1902	67.2	9.8	25.1	21.9	84.7	39.3	55.1	113.8	98.6	282.2
2	3429	TAMIL NADU	1903	19.3	7.8	1.7	18.2	128.5	58.5	72.6	115.0	210.4	128.1
3	3430	TAMIL NADU	1904	35.2	0.1	0.7	19.5	121.9	34.9	89.0	40.4	85.7	163.2
4	3431	TAMIL NADU	1905	6.5	7.5	17.2	64.8	83.7	49.8	39.0	101.8	73.5	250.4
...
110	3537	TAMIL NADU	2011	4.3	11.2	8.0	91.5	33.4	56.0	45.5	128.9	76.0	200.4
111	3538	TAMIL NADU	2012	3.0	0.1	2.5	35.5	41.9	30.1	46.5	98.0	84.9	235.2
112	3539	TAMIL NADU	2013	3.9	30.9	30.0	20.3	42.0	54.6	42.7	110.7	113.5	127.9
113	3540	TAMIL NADU	2014	7.4	6.1	8.1	8.3	139.1	47.8	50.6	117.7	98.9	252.2
114	3541	TAMIL NADU	2015	8.3	2.3	21.7	108.8	112.4	62.4	43.5	81.6	98.4	132.6

115 rows × 20 columns



In [3]: 1 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]: 1 b=a.fillna(method='ffill')
2 b

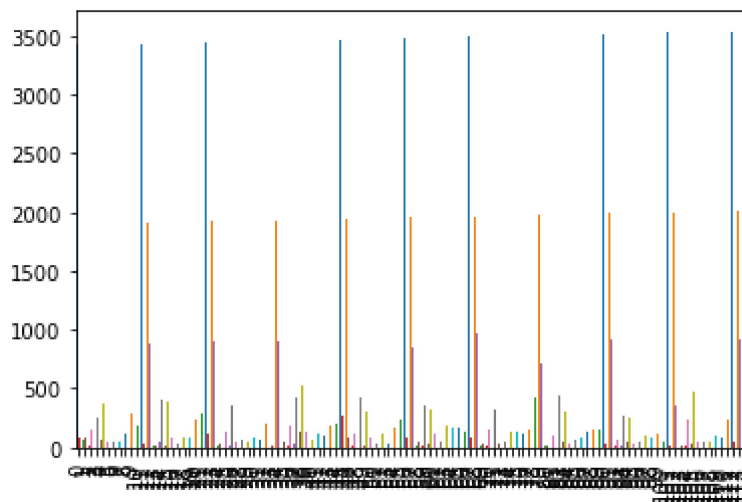
Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	3427	TAMIL NADU	1901	24.5	39.1	21.7	36.0	74.0	41.8	49.3	67.9	191.1	122.3
1	3428	TAMIL NADU	1902	67.2	9.8	25.1	21.9	84.7	39.3	55.1	113.8	98.6	282.2
2	3429	TAMIL NADU	1903	19.3	7.8	1.7	18.2	128.5	58.5	72.6	115.0	210.4	128.1
3	3430	TAMIL NADU	1904	35.2	0.1	0.7	19.5	121.9	34.9	89.0	40.4	85.7	163.2
4	3431	TAMIL NADU	1905	6.5	7.5	17.2	64.8	83.7	49.8	39.0	101.8	73.5	250.4
...
110	3537	TAMIL NADU	2011	4.3	11.2	8.0	91.5	33.4	56.0	45.5	128.9	76.0	200.4
111	3538	TAMIL NADU	2012	3.0	0.1	2.5	35.5	41.9	30.1	46.5	98.0	84.9	235.2
112	3539	TAMIL NADU	2013	3.9	30.9	30.0	20.3	42.0	54.6	42.7	110.7	113.5	127.9
113	3540	TAMIL NADU	2014	7.4	6.1	8.1	8.3	139.1	47.8	50.6	117.7	98.9	252.2
114	3541	TAMIL NADU	2015	8.3	2.3	21.7	108.8	112.4	62.4	43.5	81.6	98.4	132.6

115 rows × 20 columns

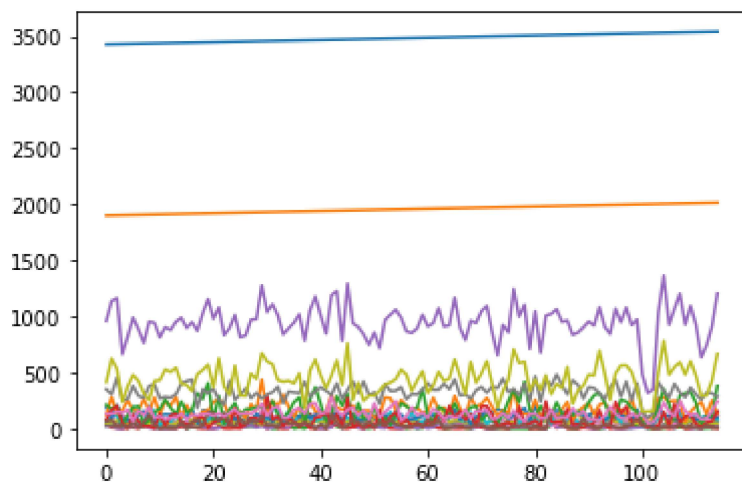
```
In [5]: 1 b.plot.bar(legend=None)
```

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



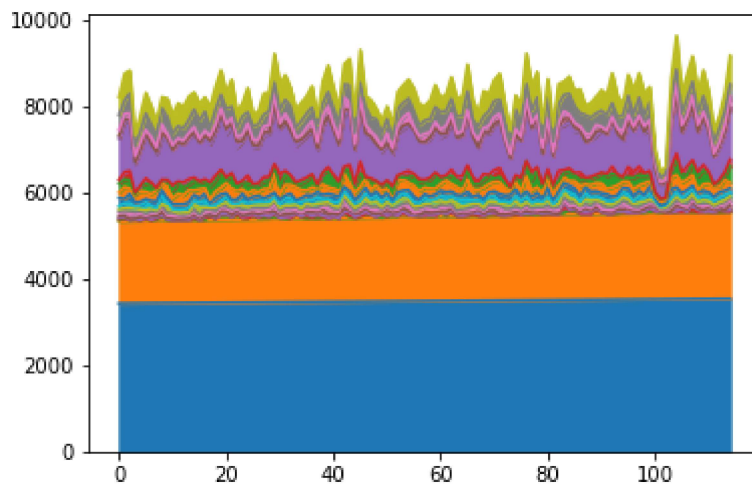
```
In [6]: 1 b.plot.line(legend=None)
```

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



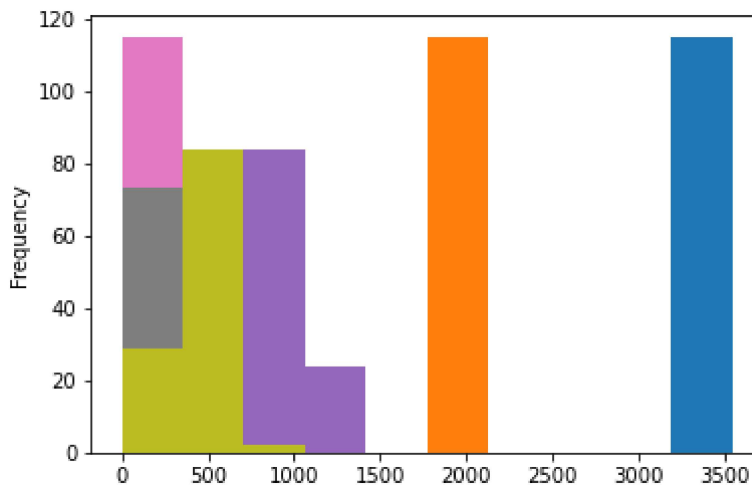
```
In [7]: 1 b.plot.area(legend=None)
```

```
Out[7]: <AxesSubplot:>
```



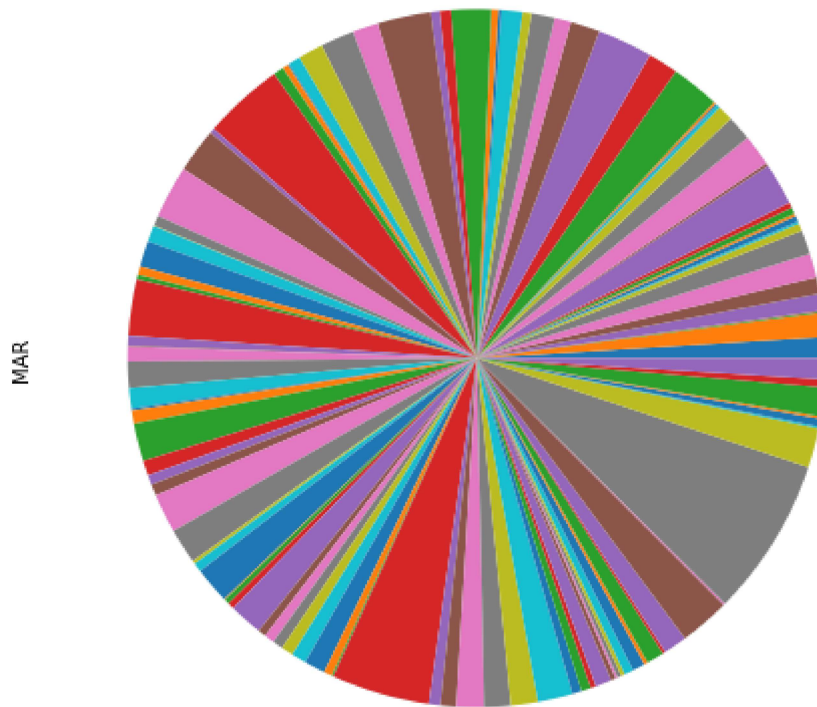
```
In [8]: 1 b.plot.hist(legend=None)
```

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



```
In [10]: 1 b.plot.pie(y='MAR',figsize=(8,8),labels=None,legend=None)
```

```
Out[10]: <AxesSubplot:ylabel='MAR'>
```



```
In [ ]: 1
```