

## 04/08/2023 (BOOK-7)

```
In [56]: 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 [57]: a=pd.read_csv(r"C:\Users\user\Downloads\Book7.csv")
a
```

Out[57]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	3312	RAYALSEEMA	1901	7.0	50.2	0.0	12.1	38.9	53.0	73.4	60.3	109.0	81.6
1	3313	RAYALSEEMA	1902	10.0	0.2	1.7	11.0	36.8	73.6	41.3	148.3	181.7	188.5
2	3314	RAYALSEEMA	1903	30.0	0.1	0.0	3.6	80.5	67.5	127.5	140.6	219.7	95.3
3	3315	RAYALSEEMA	1904	14.8	0.0	1.7	7.1	58.8	39.8	75.1	19.4	84.7	111.5
4	3316	RAYALSEEMA	1905	6.5	6.8	17.0	18.3	44.2	66.1	50.9	219.3	36.5	180.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...
110	3422	RAYALSEEMA	2011	0.8	12.1	0.0	34.6	33.0	44.5	128.9	163.6	71.2	107.5
111	3423	RAYALSEEMA	2012	2.7	0.0	2.5	32.7	38.8	47.0	139.7	120.0	69.5	113.7
112	3424	RAYALSEEMA	2013	1.3	30.6	11.5	26.8	38.9	73.8	95.7	110.3	163.2	169.3
113	3425	RAYALSEEMA	2014	0.2	0.7	12.5	5.1	46.7	66.3	68.7	115.1	81.4	104.6
114	3426	RAYALSEEMA	2015	1.9	0.0	13.4	73.4	39.7	73.0	43.1	123.6	136.3	106.7

115 rows × 14 columns



In [58]: 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 [59]: b=a.fillna(method='ffill')  
b

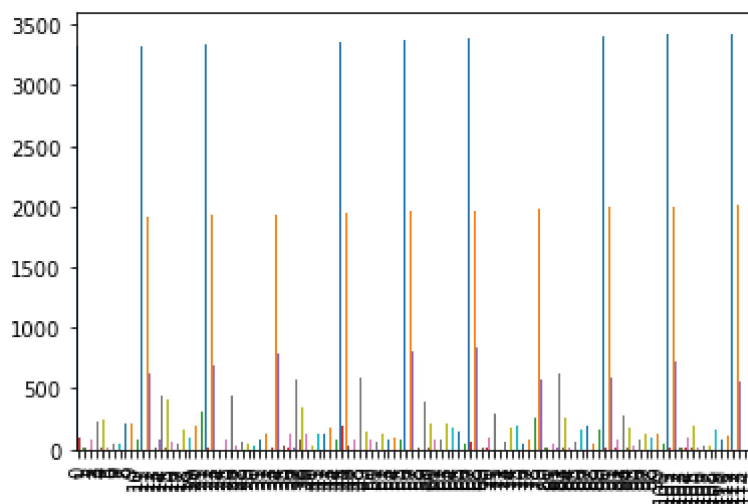
Out[59]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	3312	RAYALSEEMA	1901	7.0	50.2	0.0	12.1	38.9	53.0	73.4	60.3	109.0	81.6
1	3313	RAYALSEEMA	1902	10.0	0.2	1.7	11.0	36.8	73.6	41.3	148.3	181.7	188.5
2	3314	RAYALSEEMA	1903	30.0	0.1	0.0	3.6	80.5	67.5	127.5	140.6	219.7	95.3
3	3315	RAYALSEEMA	1904	14.8	0.0	1.7	7.1	58.8	39.8	75.1	19.4	84.7	111.5
4	3316	RAYALSEEMA	1905	6.5	6.8	17.0	18.3	44.2	66.1	50.9	219.3	36.5	180.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...
110	3422	RAYALSEEMA	2011	0.8	12.1	0.0	34.6	33.0	44.5	128.9	163.6	71.2	107.5
111	3423	RAYALSEEMA	2012	2.7	0.0	2.5	32.7	38.8	47.0	139.7	120.0	69.5	113.7
112	3424	RAYALSEEMA	2013	1.3	30.6	11.5	26.8	38.9	73.8	95.7	110.3	163.2	169.3
113	3425	RAYALSEEMA	2014	0.2	0.7	12.5	5.1	46.7	66.3	68.7	115.1	81.4	104.6
114	3426	RAYALSEEMA	2015	1.9	0.0	13.4	73.4	39.7	73.0	43.1	123.6	136.3	106.7

115 rows × 20 columns

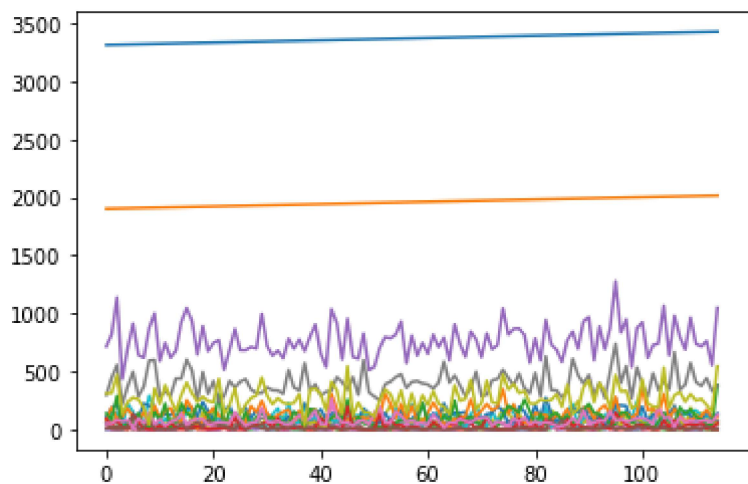
```
In [60]: b.plot.bar(legend=None)
```

```
Out[60]: <AxesSubplot:>
```



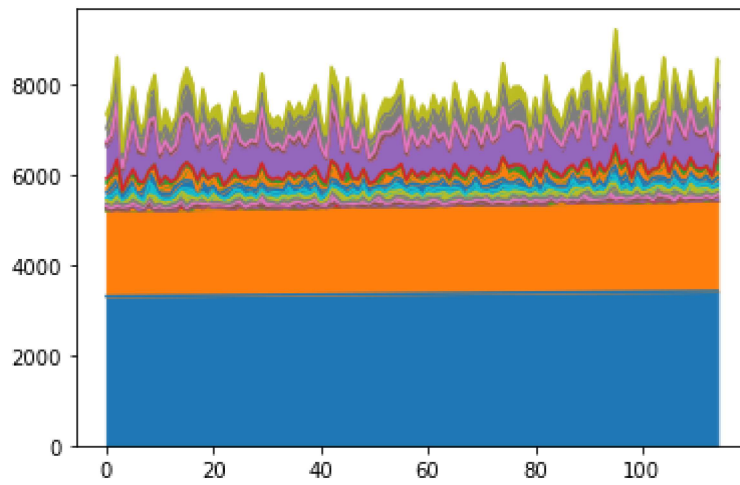
```
In [61]: b.plot.line(legend=None)
```

```
Out[61]: <AxesSubplot:>
```



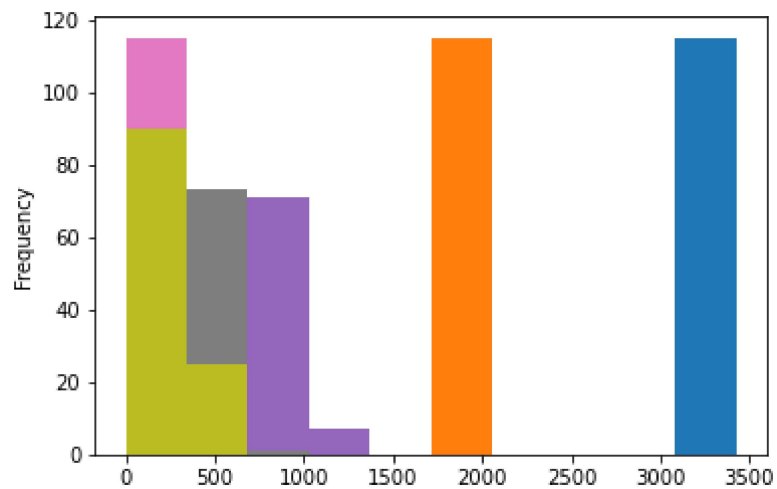
```
In [62]: b.plot.area(legend=None)
```

```
Out[62]: <AxesSubplot:>
```



```
In [63]: b.plot.hist(legend=None)
```

```
Out[63]: <AxesSubplot:ylabel='Frequency'>
```



```
In [64]: b.plot.pie(y='YEAR',figsize=(8,8),labels=None,legend=None)
```

```
Out[64]: <AxesSubplot:ylabel='YEAR'>
```

