

04/08/2023 (BOOK-4)

```
In [29]: 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 [30]: a=pd.read_csv(r"C:\Users\user\Downloads\Book4.csv")
a
```

Out[30]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	C
0	3542	COASTAL KARNATAKA	1901	1.8	0.6	10.7	52.4	81.6	960.9	991.2	606.4	108.0	12
1	3543	COASTAL KARNATAKA	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	18
2	3544	COASTAL KARNATAKA	1903	0.7	0.0	0.0	4.1	202.8	536.5	1405.5	593.8	304.4	18
3	3545	COASTAL KARNATAKA	1904	2.4	0.0	4.8	23.7	93.2	1108.2	1070.0	465.6	245.3	12
4	3546	COASTAL KARNATAKA	1905	0.0	0.2	0.0	6.4	83.1	767.3	777.3	586.9	172.9	22
...
110	3652	COASTAL KARNATAKA	2011	4.8	3.8	8.7	66.1	49.3	1018.4	1080.5	861.3	545.2	17
111	3653	COASTAL KARNATAKA	2012	NaN	11.4	5.1	77.0	22.9	650.9	754.6	1027.6	382.0	17
112	3654	COASTAL KARNATAKA	2013	2.4	19.6	19.0	28.5	100.4	1153.0	1515.3	680.2	379.1	26
113	3655	COASTAL KARNATAKA	2014	0.0	0.3	1.9	40.5	181.9	507.0	1155.4	1121.0	379.3	22
114	3656	COASTAL KARNATAKA	2015	1.4	1.0	32.3	72.2	150.3	735.3	930.9	575.2	260.3	20

115 rows × 20 columns



```
In [31]: 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             114 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          114 non-null    float64
16  Jan-Feb         114 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 [32]: b=a.fillna(method='ffill')
b
```

Out[32]:

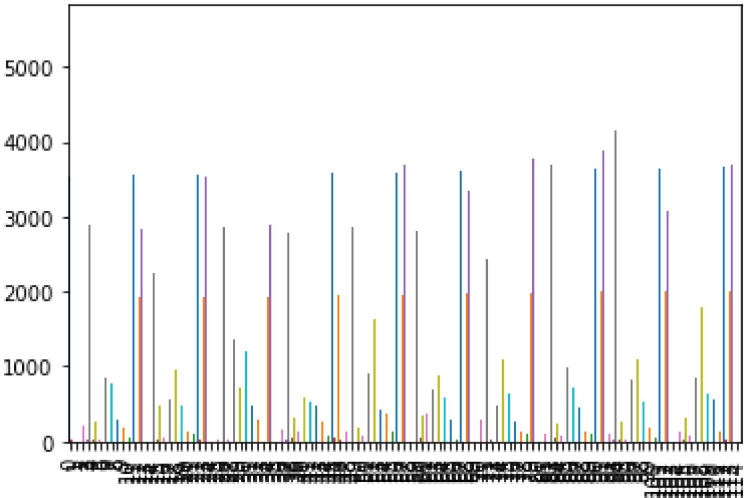
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	C
0	3542	COASTAL KARNATAKA	1901	1.8	0.6	10.7	52.4	81.6	960.9	991.2	606.4	108.0	12
1	3543	COASTAL KARNATAKA	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	18
2	3544	COASTAL KARNATAKA	1903	0.7	0.0	0.0	4.1	202.8	536.5	1405.5	593.8	304.4	18
3	3545	COASTAL KARNATAKA	1904	2.4	0.0	4.8	23.7	93.2	1108.2	1070.0	465.6	245.3	12
4	3546	COASTAL KARNATAKA	1905	0.0	0.2	0.0	6.4	83.1	767.3	777.3	586.9	172.9	22
...
110	3652	COASTAL KARNATAKA	2011	4.8	3.8	8.7	66.1	49.3	1018.4	1080.5	861.3	545.2	17
111	3653	COASTAL KARNATAKA	2012	4.8	11.4	5.1	77.0	22.9	650.9	754.6	1027.6	382.0	17
112	3654	COASTAL KARNATAKA	2013	2.4	19.6	19.0	28.5	100.4	1153.0	1515.3	680.2	379.1	26
113	3655	COASTAL KARNATAKA	2014	0.0	0.3	1.9	40.5	181.9	507.0	1155.4	1121.0	379.3	22
114	3656	COASTAL KARNATAKA	2015	1.4	1.0	32.3	72.2	150.3	735.3	930.9	575.2	260.3	20

115 rows × 20 columns



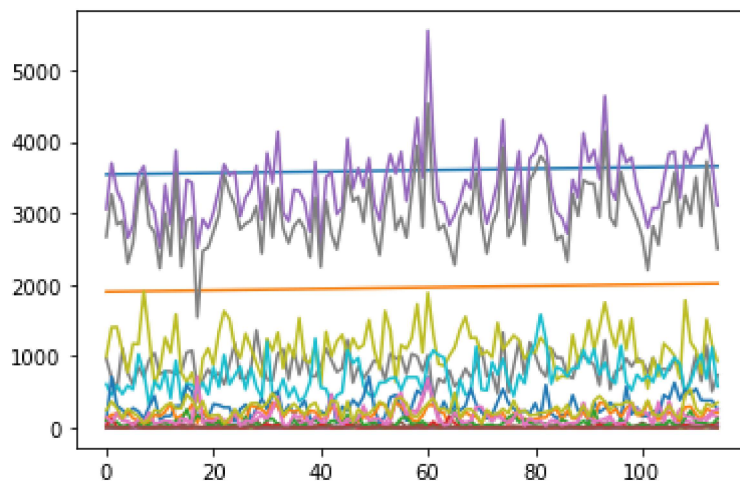
```
In [33]: b.plot.bar(legend=None)
```

Out[33]: <AxesSubplot:>



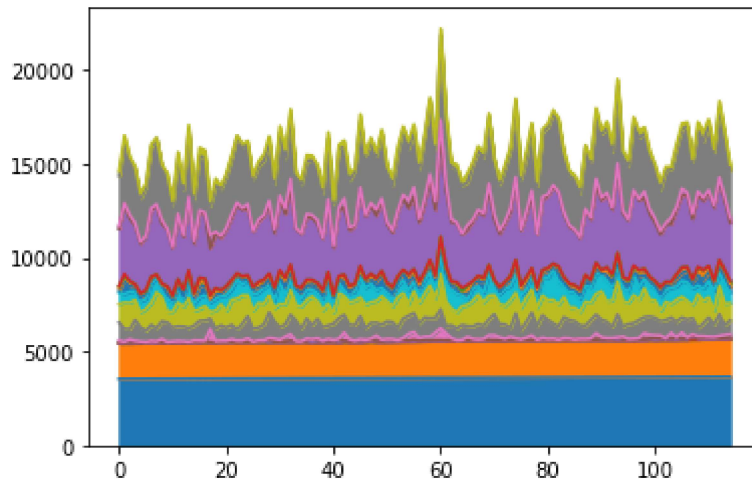
```
In [34]: b.plot.line(legend=None)
```

```
Out[34]: <AxesSubplot:>
```



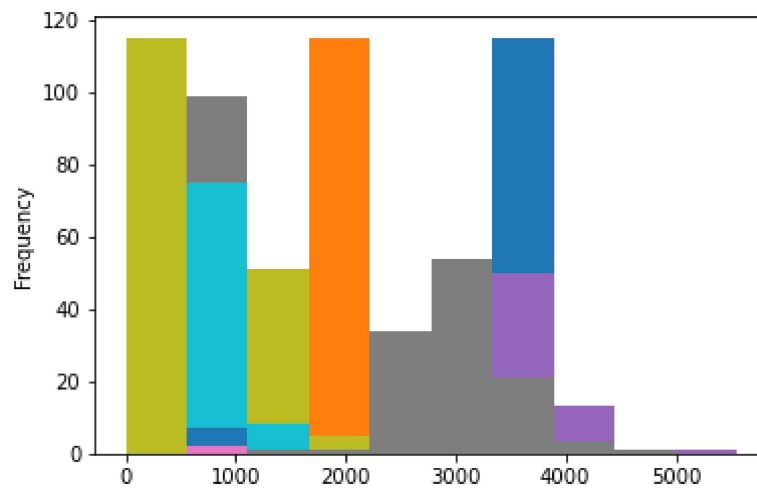
```
In [35]: b.plot.area(legend=None)
```

```
Out[35]: <AxesSubplot:>
```



```
In [36]: b.plot.hist(legend=None)
```

```
Out[36]: <AxesSubplot:ylabel='Frequency'>
```



```
In [37]: b.plot.pie(y='YEAR',figsize=(8,8),labels=None,legend=None)
```

```
Out[37]: <AxesSubplot:ylabel='YEAR'>
```

