

Vijay(Book16) 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\Book16.csv")
        2 a
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
0	2277	GUJARAT REGION	1901	4.2	0.0	0.6	1.6	7.0	60.3	240.2	205.4	18.1	16.6
1	2278	GUJARAT REGION	1902	3.9	0.0	0.0	0.6	1.0	32.8	229.8	299.0	281.2	2.3
2	2279	GUJARAT REGION	1903	0.3	0.1	1.4	0.0	12.3	30.1	452.9	202.0	183.2	5.4
3	2280	GUJARAT REGION	1904	0.8	10.6	16.8	0.2	3.9	48.3	194.8	71.8	138.0	6.1
4	2281	GUJARAT REGION	1905	0.1	0.7	1.1	0.3	0.0	20.1	668.3	37.9	81.3	1.4
...
110	2387	GUJARAT REGION	2011	0.0	0.2	0.0	0.0	0.0	16.3	259.2	451.7	162.5	0.4
111	2388	GUJARAT REGION	2012	0.1	0.0	0.0	0.0	0.0	34.4	178.2	230.3	263.8	7.1
112	2389	GUJARAT REGION	2013	0.0	0.9	0.1	4.6	0.0	155.7	405.4	211.1	287.3	53.2
113	2390	GUJARAT REGION	2014	5.7	0.1	0.2	1.0	1.3	11.6	307.5	138.6	235.1	3.3
114	2391	GUJARAT REGION	2015	1.8	0.0	6.1	5.5	0.9	120.7	354.7	37.4	93.4	2.2

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	2277	GUJARAT REGION	1901	4.2	0.0	0.6	1.6	7.0	60.3	240.2	205.4	18.1	16.6
1	2278	GUJARAT REGION	1902	3.9	0.0	0.0	0.6	1.0	32.8	229.8	299.0	281.2	2.3
2	2279	GUJARAT REGION	1903	0.3	0.1	1.4	0.0	12.3	30.1	452.9	202.0	183.2	5.4
3	2280	GUJARAT REGION	1904	0.8	10.6	16.8	0.2	3.9	48.3	194.8	71.8	138.0	6.1
4	2281	GUJARAT REGION	1905	0.1	0.7	1.1	0.3	0.0	20.1	668.3	37.9	81.3	1.4
...
110	2387	GUJARAT REGION	2011	0.0	0.2	0.0	0.0	0.0	16.3	259.2	451.7	162.5	0.4
111	2388	GUJARAT REGION	2012	0.1	0.0	0.0	0.0	0.0	34.4	178.2	230.3	263.8	7.1
112	2389	GUJARAT REGION	2013	0.0	0.9	0.1	4.6	0.0	155.7	405.4	211.1	287.3	53.2
113	2390	GUJARAT REGION	2014	5.7	0.1	0.2	1.0	1.3	11.6	307.5	138.6	235.1	3.3
114	2391	GUJARAT REGION	2015	1.8	0.0	6.1	5.5	0.9	120.7	354.7	37.4	93.4	2.2

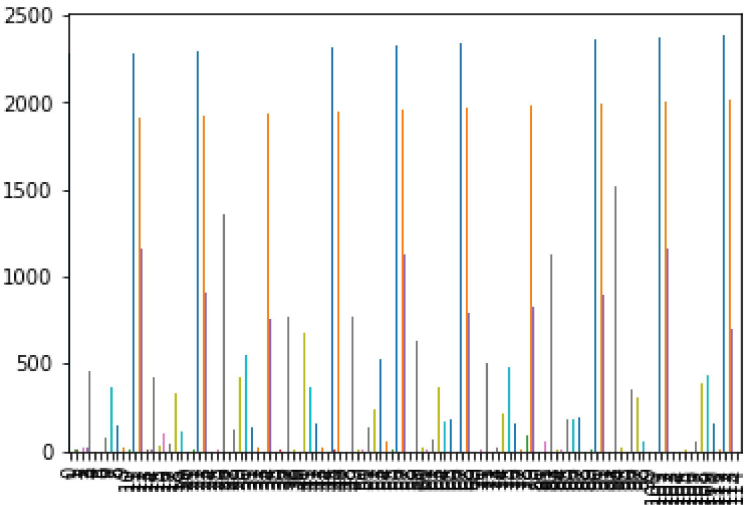
115 rows × 20 columns



In [5]:

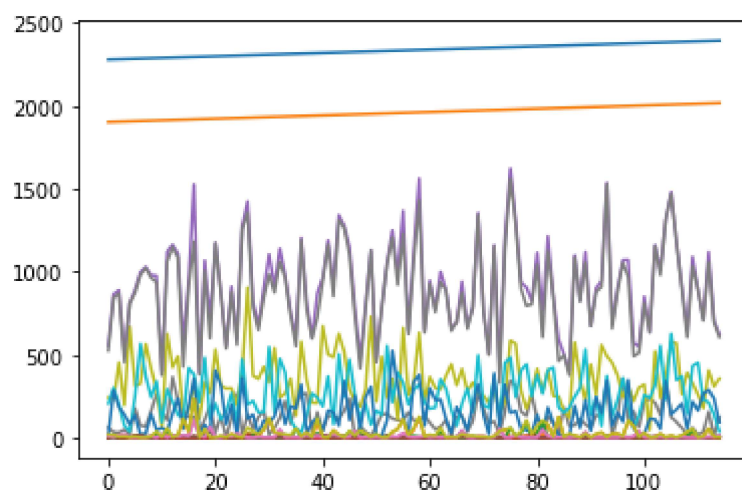
```
1 b.plot.bar(legend=None)
```

Out[5]: <AxesSubplot:>



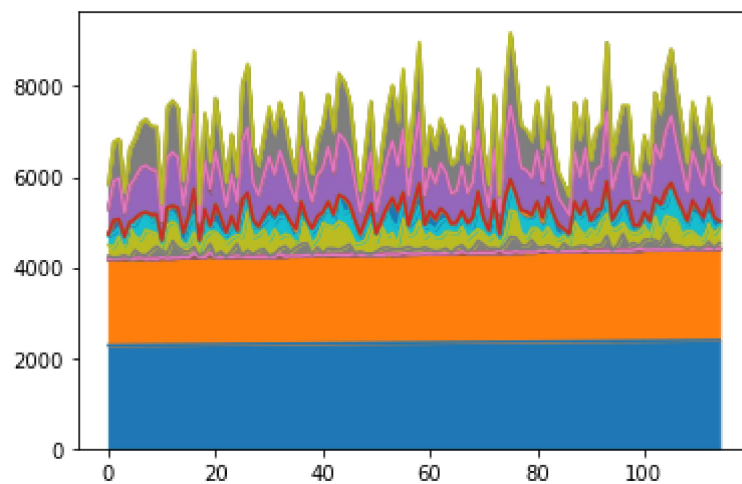
```
In [6]: 1 b.plot.line(legend=None)
```

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



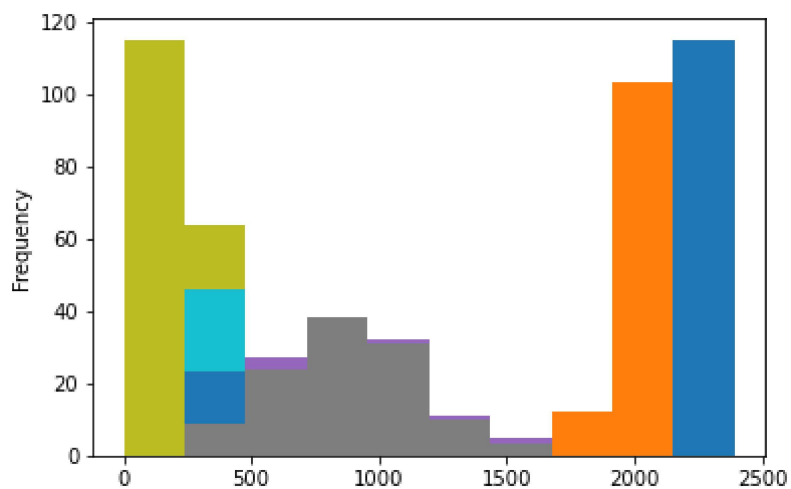
```
In [7]: 1 b.plot.area(legend=None)
```

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



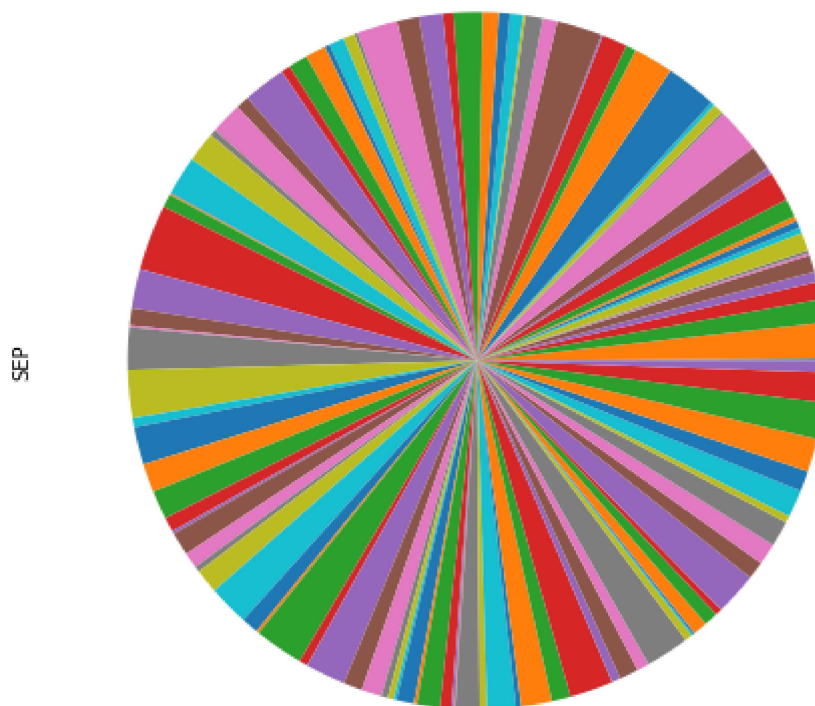
```
In [8]: 1 b.plot.hist(legend=None)
```

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



```
In [9]: 1 b.plot.pie(y='SEP',figsize=(8,8),labels=None,legend=None)
```

```
Out[9]: <AxesSubplot:ylabel='SEP'>
```



```
In [ ]: 1
```

