Importing Libraries

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Importing Datasets

In [2]: df=pd.read_csv(r'C:\Users\user\Downloads\Rainfall\COASTAL KARNATAKA.csv')
 df

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	(
0	3543	COASTAL KARNATAKA	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	1
1	3544	COASTAL KARNATAKA	1903	0.7	0.0	0.0	4.1	202.8	536.5	1405.5	593.8	304.4	1
2	3545	COASTAL KARNATAKA	1904	2.4	0.0	4.8	23.7	93.2	1108.2	1070.0	465.6	245.3	1
3	3546	COASTAL KARNATAKA	1905	0.0	0.2	0.0	6.4	83.1	767.3	777.3	586.9	172.9	2
4	3547	COASTAL KARNATAKA	1906	23.0	0.0	0.0	0.5	29.8	593.6	1173.4	535.0	273.3	1
109	3652	COASTAL KARNATAKA	2011	4.8	3.8	8.7	66.1	49.3	1018.4	1080.5	861.3	545.2	1
110	3653	COASTAL KARNATAKA	2012	NaN	11.4	5.1	77.0	22.9	650.9	754.6	1027.6	382.0	1
111	3654	COASTAL KARNATAKA	2013	2.4	19.6	19.0	28.5	100.4	1153.0	1515.3	680.2	379.1	2
112	3655	COASTAL KARNATAKA	2014	0.0	0.3	1.9	40.5	181.9	507.0	1155.4	1121.0	379.3	2
113	3656	COASTAL KARNATAKA	2015	1.4	1.0	32.3	72.2	150.3	735.3	930.9	575.2	260.3	2
114 r	114 rows × 20 columns												

Data Cleaning and Data Preprocessing

```
In [3]: df=df.dropna()
df
```

Out[3]:

index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	(
3543	COASTAL KARNATAKA	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	1
3544	COASTAL KARNATAKA	1903	0.7	0.0	0.0	4.1	202.8	536.5	1405.5	593.8	304.4	1
3545	COASTAL KARNATAKA	1904	2.4	0.0	4.8	23.7	93.2	1108.2	1070.0	465.6	245.3	1:
3546	COASTAL KARNATAKA	1905	0.0	0.2	0.0	6.4	83.1	767.3	777.3	586.9	172.9	2:
3547	COASTAL KARNATAKA	1906	23.0	0.0	0.0	0.5	29.8	593.6	1173.4	535.0	273.3	1:
3651	COASTAL KARNATAKA	2010	14.4	0.4	3.5	62.2	80.2	682.7	1200.2	637.5	468.4	2
3652	COASTAL KARNATAKA	2011	4.8	3.8	8.7	66.1	49.3	1018.4	1080.5	861.3	545.2	1
3654	COASTAL KARNATAKA	2013	2.4	19.6	19.0	28.5	100.4	1153.0	1515.3	680.2	379.1	2
3655	COASTAL KARNATAKA	2014	0.0	0.3	1.9	40.5	181.9	507.0	1155.4	1121.0	379.3	2:
3656	COASTAL KARNATAKA	2015	1.4	1.0	32.3	72.2	150.3	735.3	930.9	575.2	260.3	2
	3543 3544 3545 3546 3547 3651 3652 3654 3655	3543 COASTAL KARNATAKA 3544 COASTAL KARNATAKA 3545 COASTAL KARNATAKA 3546 COASTAL KARNATAKA 3547 COASTAL KARNATAKA 3651 COASTAL KARNATAKA 3652 COASTAL KARNATAKA 3654 COASTAL KARNATAKA 3655 COASTAL KARNATAKA	3543 COASTAL 1902 3544 COASTAL 1903 3545 COASTAL 1904 3546 COASTAL 1905 3547 COASTAL 1905 3547 COASTAL 1906 3651 COASTAL 1906 3652 COASTAL 2010 3654 COASTAL 2011 3654 COASTAL 2011 3655 COASTAL 2014 3656 COASTAL 2014	3543 COASTAL 1902 3.2 3544 COASTAL 1903 0.7 3545 COASTAL 1904 2.4 3546 COASTAL 1905 0.0 3547 COASTAL 1905 0.0 3547 COASTAL 1906 23.0 3651 COASTAL 2010 14.4 3652 COASTAL 2011 4.8 3654 COASTAL 2011 4.8 3655 COASTAL 2013 2.4 3656 COASTAL 2014 0.0	3543 COASTAL 1902 3.2 0.3 3544 COASTAL 1903 0.7 0.0 3545 COASTAL 1904 2.4 0.0 3546 COASTAL 1905 0.0 0.2 3547 COASTAL 1906 23.0 0.0 3651 COASTAL KARNATAKA 2010 14.4 0.4 3652 COASTAL KARNATAKA 2011 4.8 3.8 3654 COASTAL KARNATAKA 2013 2.4 19.6 3655 COASTAL KARNATAKA 2014 0.0 0.3	3543 COASTAL 1902 3.2 0.3 4.9 3544 COASTAL 1903 0.7 0.0 0.0 3545 COASTAL 1904 2.4 0.0 4.8 3546 COASTAL 1905 0.0 0.2 0.0 3547 COASTAL 1906 23.0 0.0 0.0 3547 COASTAL 1906 23.0 0.0 0.0 3651 COASTAL KARNATAKA 2010 14.4 0.4 3.5 3652 COASTAL KARNATAKA 2011 4.8 3.8 8.7 3654 COASTAL 2011 4.8 3.8 8.7 3655 COASTAL KARNATAKA 2013 2.4 19.6 19.0 3656 COASTAL 2014 0.0 0.3 1.9	3543 COASTAL KARNATAKA 1902 3.2 0.3 4.9 10.2 3544 COASTAL KARNATAKA 1903 0.7 0.0 0.0 4.1 3545 COASTAL KARNATAKA 1904 2.4 0.0 4.8 23.7 3546 COASTAL KARNATAKA 1905 0.0 0.2 0.0 6.4 3547 COASTAL KARNATAKA 1906 23.0 0.0 0.0 0.5	3543 COASTAL 1902 3.2 0.3 4.9 10.2 54.6 3544 COASTAL 1903 0.7 0.0 0.0 4.1 202.8 3545 COASTAL 1904 2.4 0.0 4.8 23.7 93.2 3546 COASTAL 1905 0.0 0.2 0.0 6.4 83.1 3547 COASTAL 1906 23.0 0.0 0.0 0.5 29.8 3651 COASTAL 2010 14.4 0.4 3.5 62.2 80.2 3652 COASTAL 2011 4.8 3.8 8.7 66.1 49.3 3654 COASTAL 2013 2.4 19.6 19.0 28.5 100.4 3655 COASTAL 2014 0.0 0.3 1.9 40.5 181.9	3543	3543 COASTAL KARNATAKA 1902 3.2 0.3 4.9 10.2 54.6 698.4 1401.6 3544 COASTAL KARNATAKA 1903 0.7 0.0 0.0 4.1 202.8 536.5 1405.5 3645 COASTAL KARNATAKA 1904 2.4 0.0 4.8 23.7 93.2 1108.2 1070.0 3546 COASTAL KARNATAKA 1905 0.0 0.2 0.0 6.4 83.1 767.3 777.3 3547 COASTAL KARNATAKA 1906 23.0 0.0 0.0 0.5 29.8 593.6 1173.4	3543 COASTAL KARNATAKA 1902 3.2 0.3 4.9 10.2 54.6 698.4 1401.6 454.2 3544 COASTAL KARNATAKA 1903 0.7 0.0 0.0 4.1 202.8 536.5 1405.5 593.8 3545 COASTAL KARNATAKA 1904 2.4 0.0 4.8 23.7 93.2 1108.2 1070.0 465.6 COASTAL KARNATAKA 1905 0.0 0.2 0.0 6.4 83.1 767.3 777.3 586.9 3547 COASTAL KARNATAKA 1906 23.0 0.0 0.0 0.5 29.8 593.6 1173.4 535.0	3543 COASTAL KARNATAKA 1902 3.2 0.3 4.9 10.2 54.6 698.4 1401.6 454.2 708.4 3544 COASTAL KARNATAKA 1903 0.7 0.0 0.0 4.1 202.8 536.5 1405.5 593.8 304.4 3545 COASTAL KARNATAKA 1904 2.4 0.0 4.8 23.7 93.2 1108.2 1070.0 465.6 245.3 3546 COASTAL KARNATAKA 1905 0.0 0.2 0.0 6.4 83.1 767.3 777.3 586.9 172.9 3547 COASTAL KARNATAKA 1906 23.0 0.0 0.0 0.5 29.8 593.6 1173.4 535.0 273.3

113 rows × 20 columns

In [4]: df.columns

```
In [5]: df.info()
```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 113 entries, 0 to 113
Data columns (total 20 columns):

#	Column	Non-Null Count	Dtype
0	index	113 non-null	int64
1	SUBDIVISION	113 non-null	object
2	YEAR	113 non-null	int64
3	JAN	113 non-null	float64
4	FEB	113 non-null	float64
5	MAR	113 non-null	float64
6	APR	113 non-null	float64
7	MAY	113 non-null	float64
8	JUN	113 non-null	float64
9	JUL	113 non-null	float64
10	AUG	113 non-null	float64
11	SEP	113 non-null	float64
12	OCT	113 non-null	float64
13	NOV	113 non-null	float64
14	DEC	113 non-null	float64
15	ANNUAL	113 non-null	float64
16	Jan-Feb	113 non-null	float64
17	Mar-May	113 non-null	float64
18	Jun-Sep	113 non-null	float64
19	Oct-Dec	113 non-null	float64
		7), int64(2), ob	ject(1)
memo	ry usage: 18.	5+ KB	

Line Chart

```
In [6]: df.plot.line(subplots=True)
Out[6]: array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
             <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
             <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
             <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
             <AxesSubplot:>, <AxesSubplot:>], dtype=object)
                                IAN
                                FEB
                MAR
                APR 🛓
                                               MAY
         50¢
                JUN
        JUL
                SEP
                OCT
                                               NOV
                                               DEC
        100
5988
                ANNUAL
                                             lan-Feb
                                             Mar-May
                Jun-Sep
                Oct-Dec
```

```
In [7]: df.plot.line()
```

80

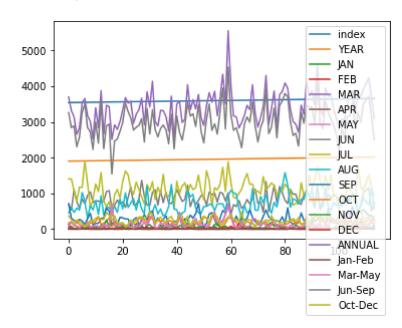
100

Out[7]: <AxesSubplot:>

20

40

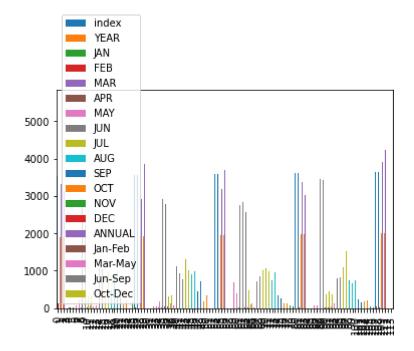
60



Bar Chart

In [8]: df.plot.bar()

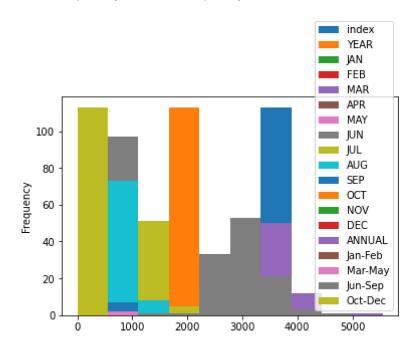
Out[8]: <AxesSubplot:>



Histogram

In [9]: df.plot.hist()

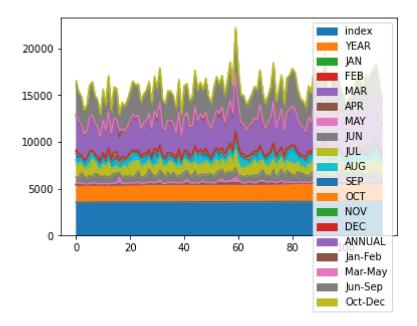
Out[9]: <AxesSubplot:ylabel='Frequency'>



Area Chart

```
In [10]: df.plot.area()
```

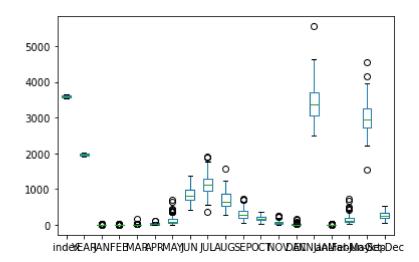
Out[10]: <AxesSubplot:>



Box Chart

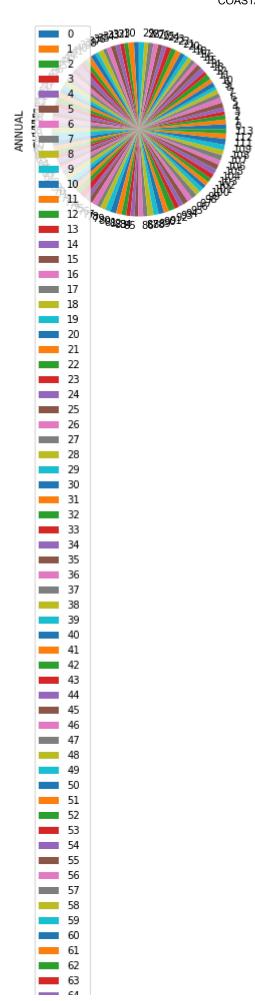
```
In [11]: df.plot.box()
```

Out[11]: <AxesSubplot:>



Pie Chart

```
In [12]: df.plot.pie(y='ANNUAL')
Out[12]: <AxesSubplot:ylabel='ANNUAL'>
```

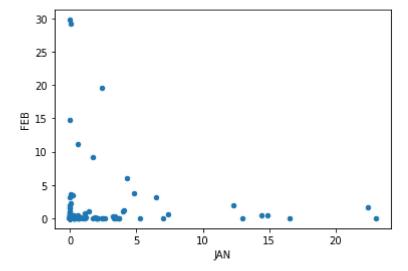




Scatter Plot

```
In [13]: df.plot.scatter(x='JAN',y='FEB')
```

Out[13]: <AxesSubplot:xlabel='JAN', ylabel='FEB'>



In [14]: df.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 113 entries, 0 to 113
Data columns (total 20 columns):

#	Column	Non-Null Count	Dtype				
0	index	113 non-null	int64				
1	SUBDIVISION	113 non-null	object				
2	YEAR	113 non-null	int64				
3	JAN	113 non-null	float64				
4	FEB	113 non-null	float64				
5	MAR	113 non-null	float64				
6	APR	113 non-null	float64				
7	MAY	113 non-null	float64				
8	JUN	113 non-null	float64				
9	JUL	113 non-null	float64				
10	AUG	113 non-null	float64				
11	SEP	113 non-null	float64				
12	OCT	113 non-null	float64				
13	NOV	113 non-null	float64				
14	DEC	113 non-null	float64				
15	ANNUAL	113 non-null	float64				
16	Jan-Feb	113 non-null	float64				
17	Mar-May	113 non-null	float64				
18	Jun-Sep	113 non-null	float64				
19	Oct-Dec	113 non-null	float64				
dtyp	es: float64(1	7), int64(2), ob	ject(1)				
memory usage: 18.5+ KB							

localhost:8888/notebooks/COASTAL KARNATAKA.ipynb

In [15]: df.describe()

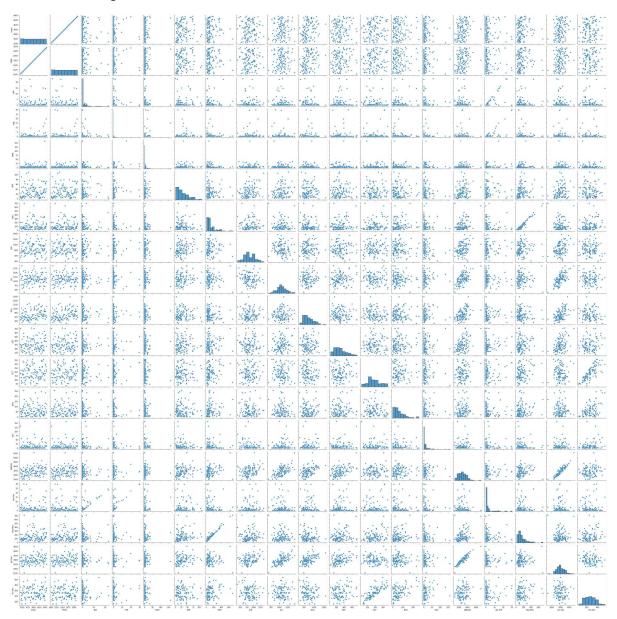
Out[15]:

	index	YEAR	JAN	FEB	MAR	APR	MAY	
count	113.000000	113.000000	113.000000	113.000000	113.000000	113.000000	113.000000	1
mean	3599.026549	1958.026549	1.938938	1.438938	6.330088	30.318584	124.036283	8
std	32.809640	32.809640	4.237132	4.677557	16.715176	23.861623	126.089062	1
min	3543.000000	1902.000000	0.000000	0.000000	0.000000	0.000000	8.400000	4
25%	3571.000000	1930.000000	0.000000	0.000000	0.200000	11.300000	44.100000	7
50%	3599.000000	1958.000000	0.100000	0.000000	1.400000	24.700000	80.200000	8
75%	3627.000000	1986.000000	2.000000	0.500000	5.700000	44.800000	164.000000	9
max	3656.000000	2015.000000	23.000000	29.800000	161.400000	110.100000	699.500000	13
4								•

EDA And Visualization

In [16]: sns.pairplot(df)

Out[16]: <seaborn.axisgrid.PairGrid at 0x1eb30fd5ee0>

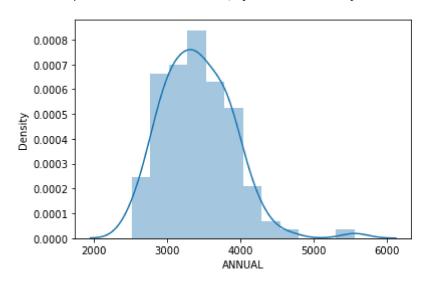


In [17]: | sns.distplot(df['ANNUAL'])

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: Fut ureWarning: `distplot` is a deprecated function and will be removed in a futu re version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)

Out[17]: <AxesSubplot:xlabel='ANNUAL', ylabel='Density'>



In [18]: sns.heatmap(df.corr())

Out[18]: <AxesSubplot:>

