

Type *Markdown* and LaTeX: α^2

Importing Libraries

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

Importing Datasets

```
In [2]: df=pd.read_csv(r"C:\Users\user\Downloads\drive-download-20230804T043023Z-001\rainfall_west_rajasthan.csv")
df
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb
0	1242	UTTARAKHAND	1901	134.5	81.4	44.5	5.9	60.8	33.6	381.1	612.3	167.1	16.3	0.0	24.9	1562.5	215.9
1	1243	UTTARAKHAND	1902	0.0	17.0	52.2	63.7	52.1	113.1	444.1	327.5	220.4	31.9	2.1	0.0	1324.2	17.1
2	1244	UTTARAKHAND	1903	68.0	7.9	87.6	10.3	37.5	83.0	251.6	442.7	249.3	57.5	0.0	11.3	1306.5	75.9
3	1245	UTTARAKHAND	1904	40.0	5.2	78.3	13.6	61.1	180.1	449.6	417.2	174.1	6.3	35.6	31.0	1492.0	45.2
4	1246	UTTARAKHAND	1905	115.4	80.7	99.8	26.1	70.3	111.5	299.9	349.5	129.5	0.0	1.0	18.5	1302.1	196.0
...
110	1352	UTTARAKHAND	2011	30.9	65.2	18.0	30.9	84.2	223.1	433.3	523.7	148.4	3.4	1.2	2.3	1564.7	96.1
111	1353	UTTARAKHAND	2012	38.8	11.9	28.1	39.2	9.1	46.0	387.1	419.5	220.6	4.7	3.4	15.5	1223.9	50.8
112	1354	UTTARAKHAND	2013	73.0	188.3	22.0	24.7	18.2	488.9	413.4	359.4	111.3	29.1	3.2	3.8	1735.4	261.3
113	1355	UTTARAKHAND	2014	45.9	99.9	68.4	37.6	52.9	62.9	462.7	264.2	107.9	40.8	0.0	44.3	1287.4	145.8
114	1356	UTTARAKHAND	2015	54.5	62.6	127.3	57.3	38.0	186.6	337.0	305.3	52.6	16.8	2.4	7.2	1247.6	117.0

115 rows × 20 columns

Data Cleaning and Data Preprocessing

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

```
In [4]: df.columns
```

```
Out[4]: Index(['index', 'SUBDIVISION', 'YEAR', 'JAN', 'FEB', 'MAR', 'APR', 'MAY',
              'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC', 'ANNUAL', 'Jan-Feb',
              'Mar-May', 'Jun-Sep', 'Oct-Dec'],
              dtype='object')
```

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

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 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: 18.9+ KB

```

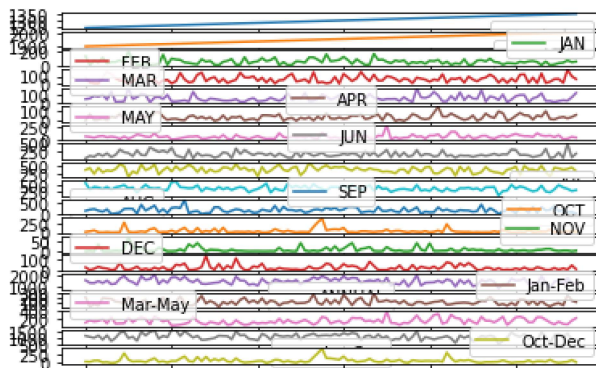
Line chart

In [6]: `df.plot.line(subplots=True)`

```

Out[6]: array([<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>,
<AxesSubplot:~>, <AxesSubplot:~>, <AxesSubplot:~>], dtype=object)

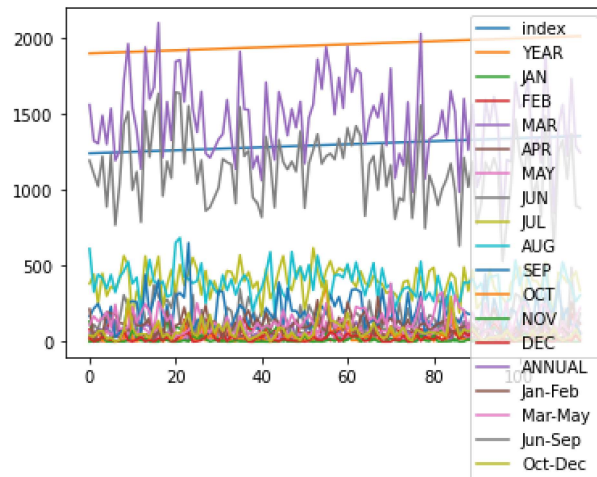
```



Line chart

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

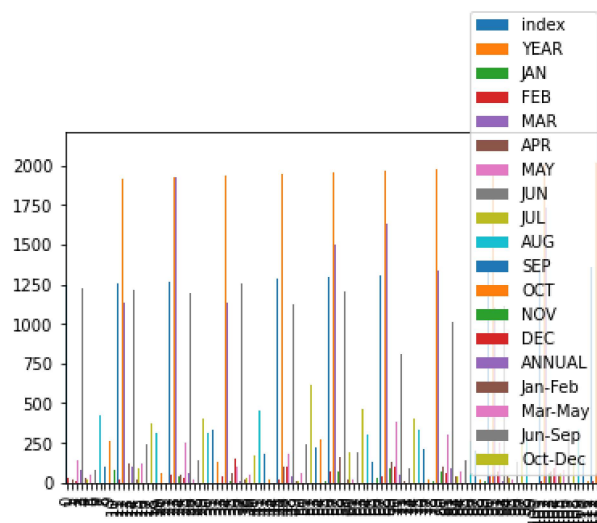
```
Out[7]: <AxesSubplot:>
```



Bar chart

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

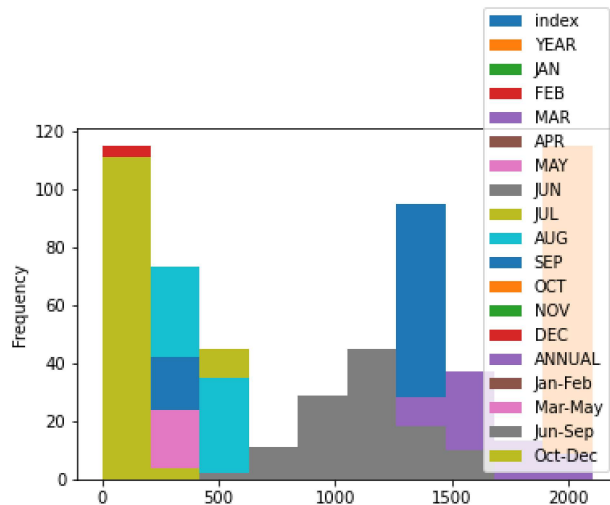
```
Out[8]: <AxesSubplot:>
```



Histogram

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

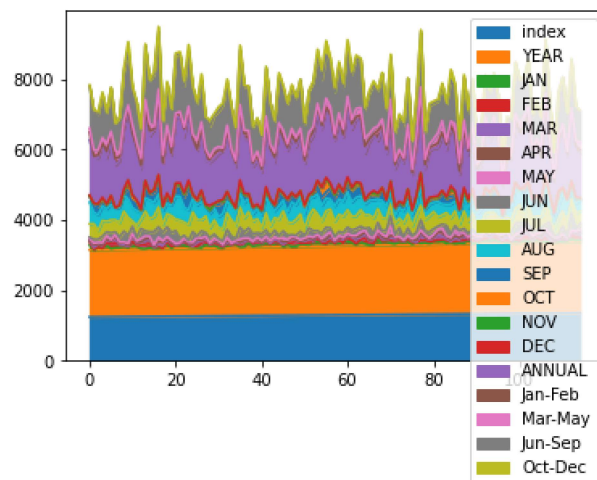
```
Out[9]: <AxesSubplot:ylabel='Frequency'>
```



Area chart

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

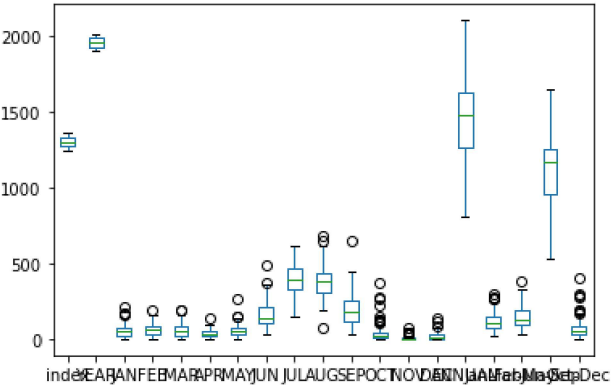
```
Out[10]: <AxesSubplot:>
```



Box chart

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

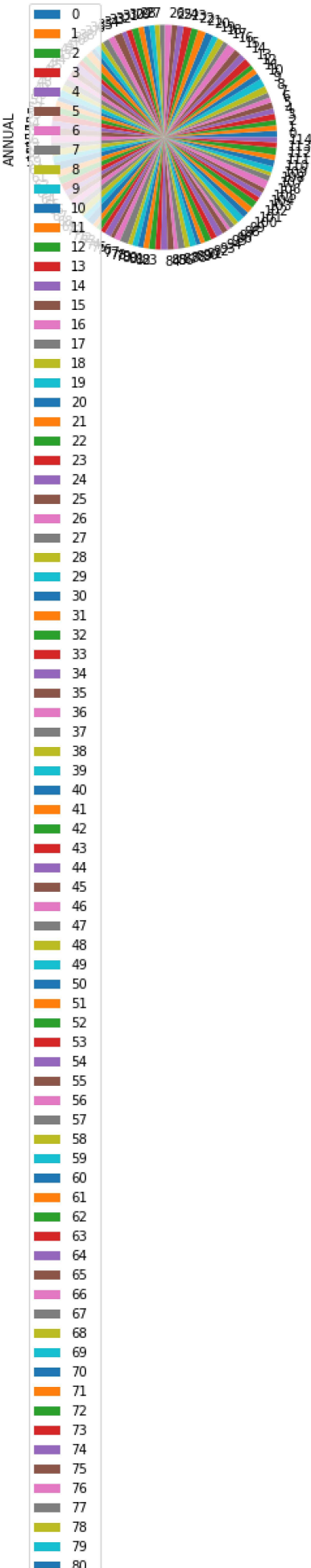
Out[11]: <AxesSubplot:>

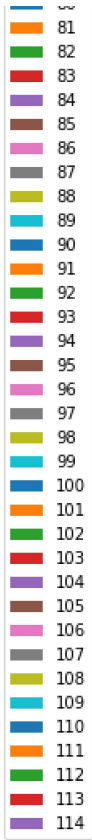


Pie chart

```
In [12]: df.plot.pie(y='ANNUAL' )
```

```
Out[12]: <AxesSubplot:ylabel='ANNUAL'>
```

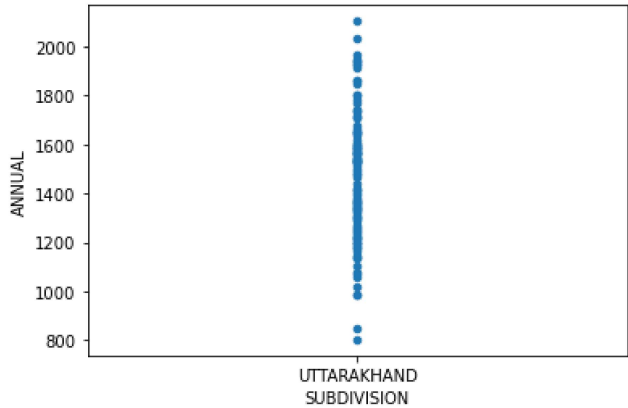





Scatter chart

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

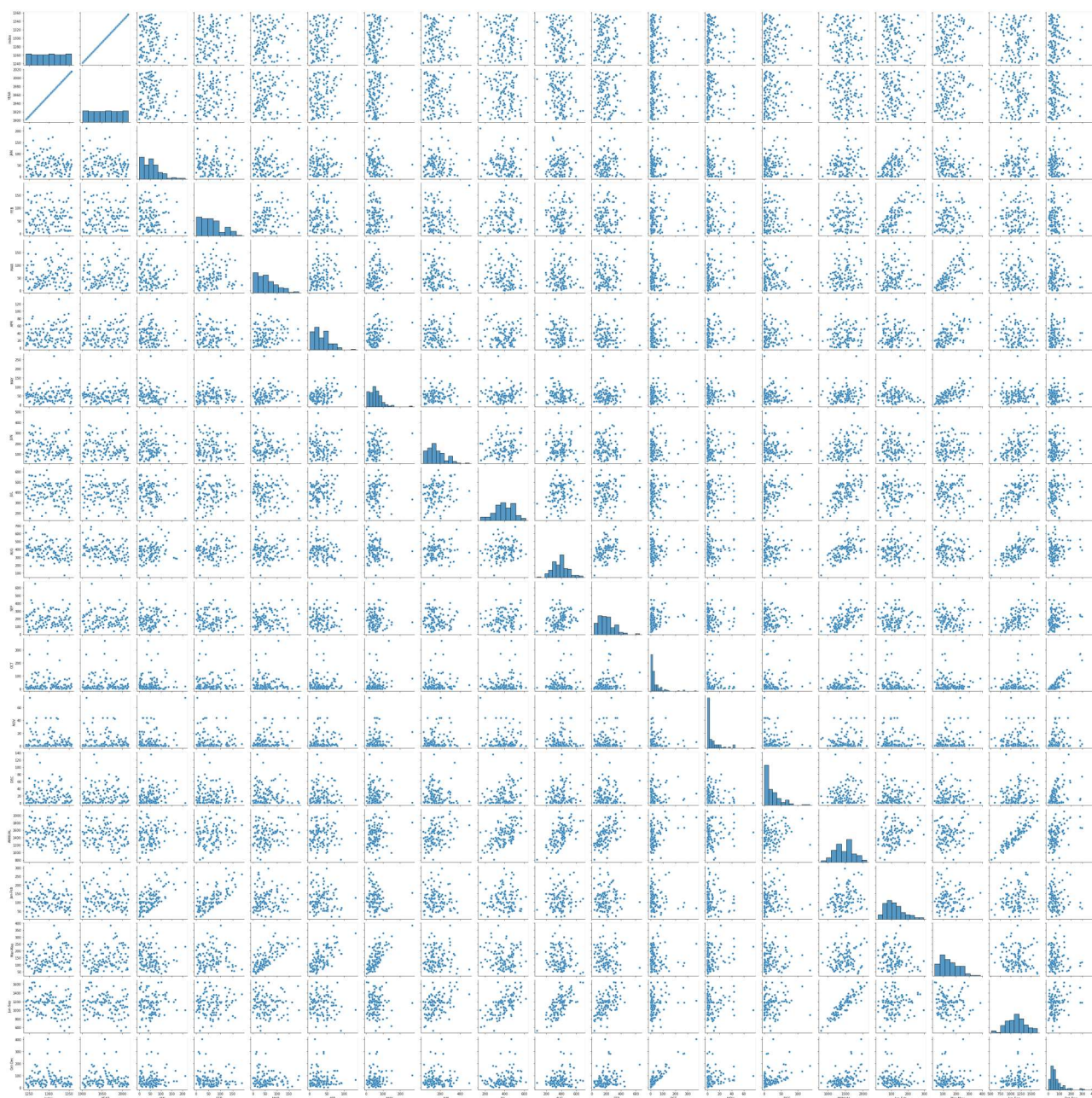
```
Out[13]: <AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL '>
```



Seaborn

```
In [14]: sns.pairplot(df)
```

```
Out[14]: <seaborn.axisgrid.PairGrid at 0x182c1035cd0>
```

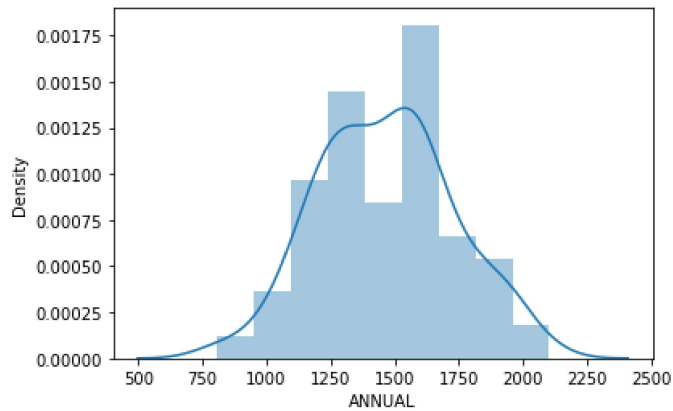


```
In [15]: sns.distplot(df['ANNUAL'])
```

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

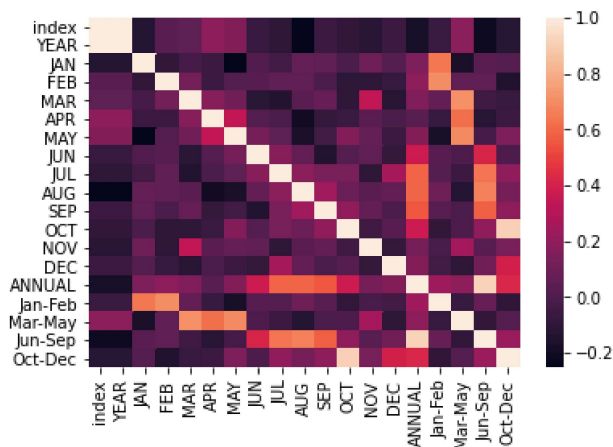
warnings.warn(msg, FutureWarning)

```
Out[15]: <AxesSubplot:xlabel='ANNUAL', ylabel='Density'>
```



```
In [16]: sns.heatmap(df.corr())
```

```
Out[16]: <AxesSubplot:>
```



```
In [ ]:
```