Type *Markdown* and LaTeX: $\alpha^2$

# 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_jammu _ kashmir.csv"
df
```

Out[2]:

| | index | SUBDIVISION | YEAR | JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | OCT | NOV | DEC | ANNUAL | Jan-Feb | Mar-May |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1012 | EAST UTTAR PRADESH | 1901 | 62.6 | 31.3 | 8.2 | 1.1 | 13.6 | 21.8 | 226.5 | 285.6 | 215.4 | 4.9 | 0.1 | 2.1 | 873.2 | 93.9 | 22.9 |
| **1** | 1013 | EAST UTTAR PRADESH | 1902 | 6.1 | 2.3 | 2.4 | 2.0 | 21.4 | 32.5 | 411.5 | 155.4 | 257.2 | 13.2 | 1.2 | 0.0 | 905.2 | 8.3 | 25.9 |
| **2** | 1014 | EAST UTTAR PRADESH | 1903 | 8.2 | 0.4 | 1.3 | 0.7 | 15.3 | 71.6 | 115.3 | 420.2 | 258.7 | 324.7 | 0.0 | 0.0 | 1216.4 | 8.6 | 17.3 |
| **3** | 1015 | EAST UTTAR PRADESH | 1904 | 7.3 | 1.5 | 8.3 | 0.4 | 28.7 | 148.0 | 359.4 | 328.8 | 95.0 | 50.6 | 17.0 | 26.3 | 1071.2 | 8.8 | 37.4 |
| **4** | 1016 | EAST UTTAR PRADESH | 1905 | 16.8 | 23.6 | 20.0 | 5.4 | 15.4 | 17.3 | 302.4 | 316.2 | 169.5 | 3.3 | 0.0 | 1.6 | 891.6 | 40.5 | 40.9 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **110** | 1122 | EAST UTTAR PRADESH | 2011 | 1.0 | 2.7 | 1.6 | 2.9 | 32.2 | 163.8 | 197.9 | 232.1 | 146.4 | 0.6 | 0.0 | 0.0 | 781.2 | 3.7 | 36.7 |
| **111** | 1123 | EAST UTTAR PRADESH | 2012 | 20.3 | 1.2 | 3.4 | 2.8 | 0.2 | 18.5 | 234.2 | 156.0 | 164.4 | 0.7 | 0.3 | 0.7 | 602.7 | 21.5 | 6.4 |
| **112** | 1124 | EAST UTTAR PRADESH | 2013 | 6.1 | 59.6 | 2.7 | 8.7 | 1.1 | 309.7 | 230.0 | 246.1 | 78.2 | 97.4 | 0.5 | 1.1 | 1041.4 | 65.8 | 12.6 |
| **113** | 1125 | EAST UTTAR PRADESH | 2014 | 47.4 | 25.8 | 15.4 | 1.7 | 10.7 | 47.8 | 224.5 | 138.1 | 106.7 | 74.7 | 0.0 | 8.4 | 701.2 | 73.3 | 27.7 |
| **114** | 1126 | EAST UTTAR PRADESH | 2015 | 30.0 | 4.1 | 48.2 | 23.2 | 8.6 | 95.3 | 179.0 | 175.8 | 21.9 | 11.8 | 0.5 | 4.9 | 603.3 | 34.1 | 80.0 |

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')
```

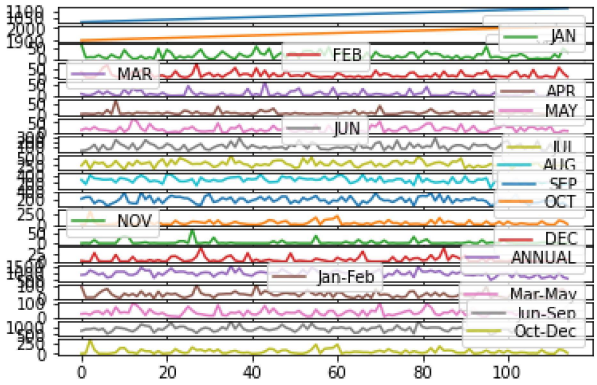
In [5]:
```
df.info()
```

```
<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)
```

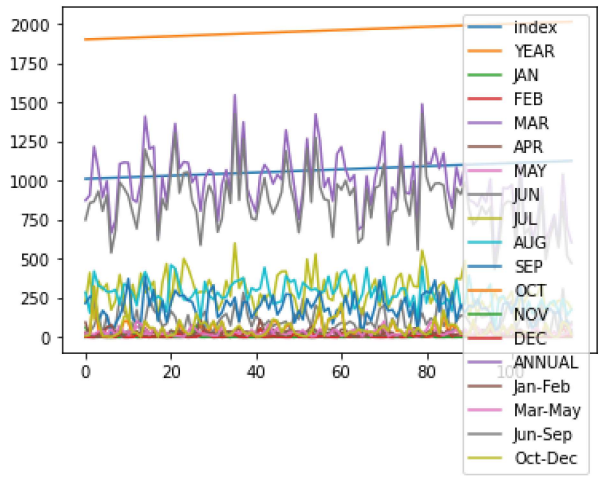
Out[6]:
```
array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
    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




# Line chart

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

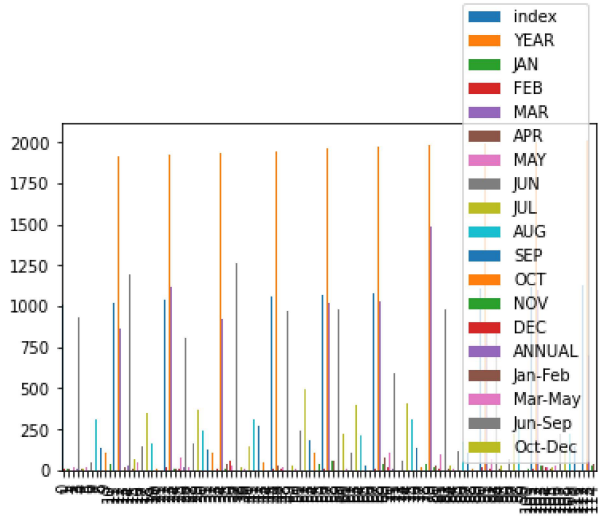
Out[7]: <AxesSubplot:>




# Bar chart

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

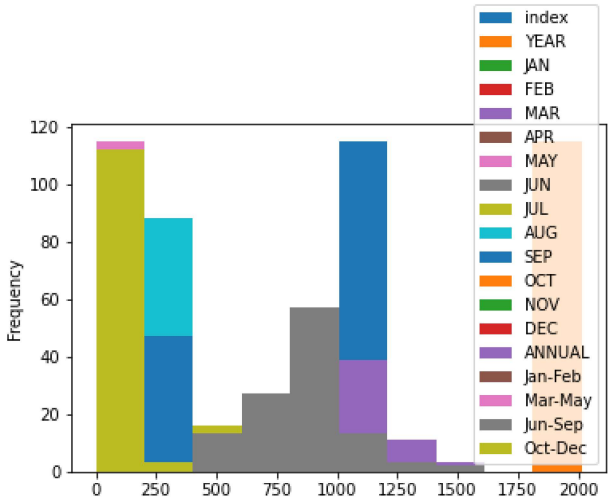
Out[8]: <AxesSubplot:>




# Histogram

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

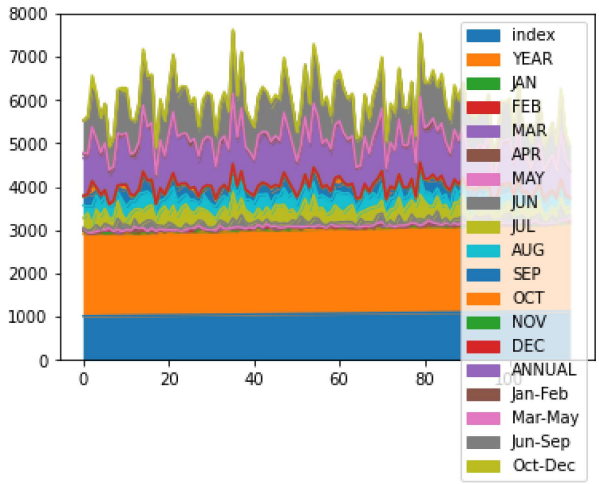
Out[9]:
```
<AxesSubplot:ylabel='Frequency'>
```




# Area chart

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

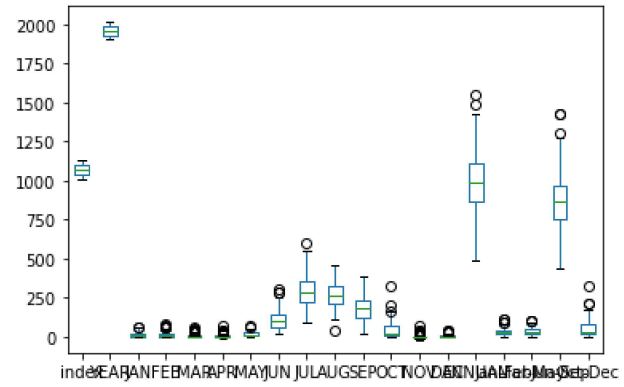
Out[10]:
```
<AxesSubplot:>
```




# Box chart

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

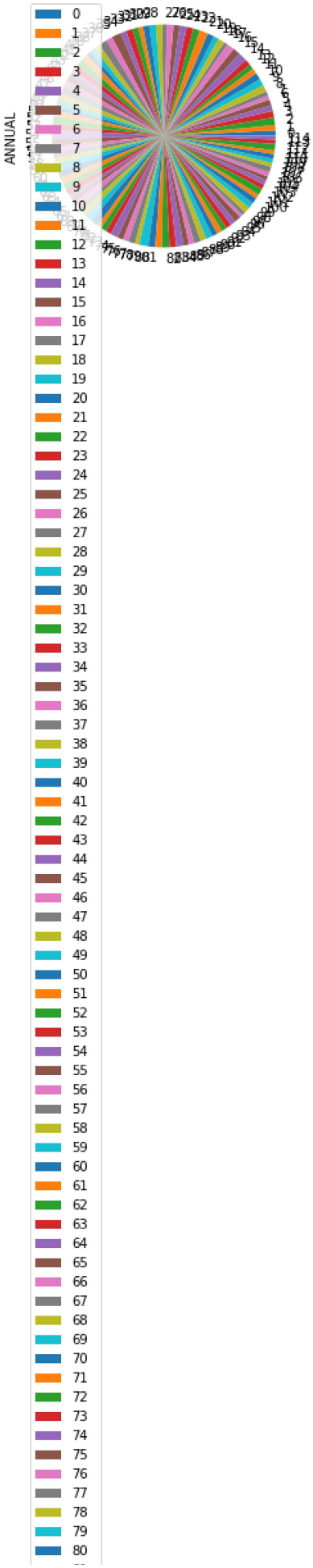
Out[11]: <AxesSubplot:>

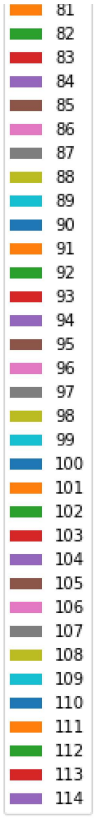



# Pie chart

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

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

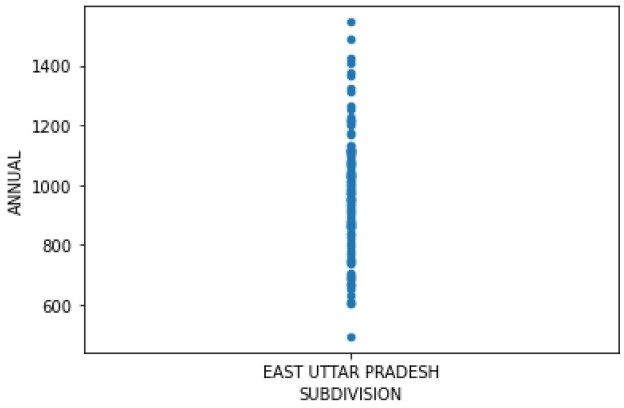

ANNUAL
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80




## Scatter chart

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

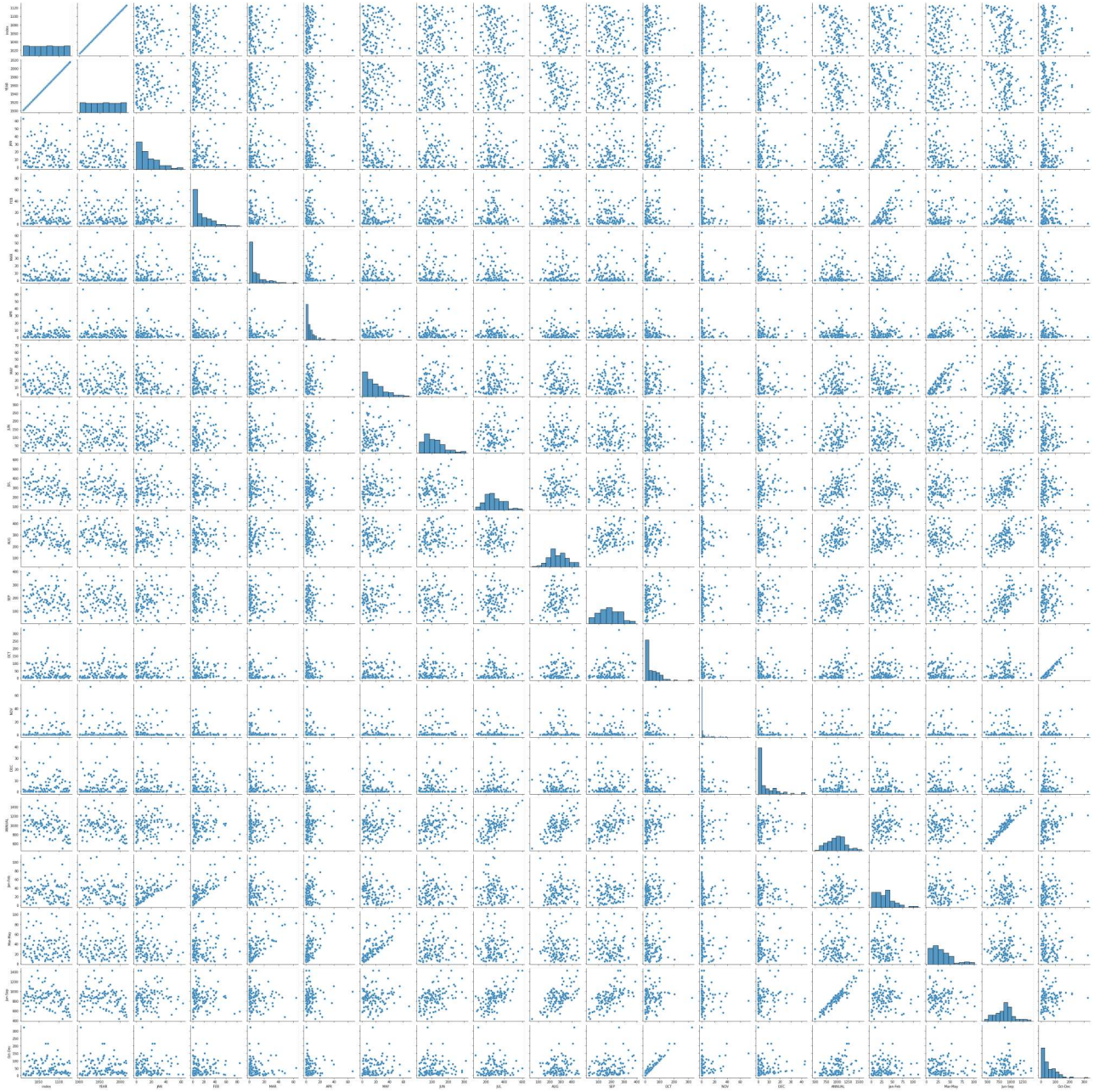
Out[13]:
```
<AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL'>
```




## Seaborn

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

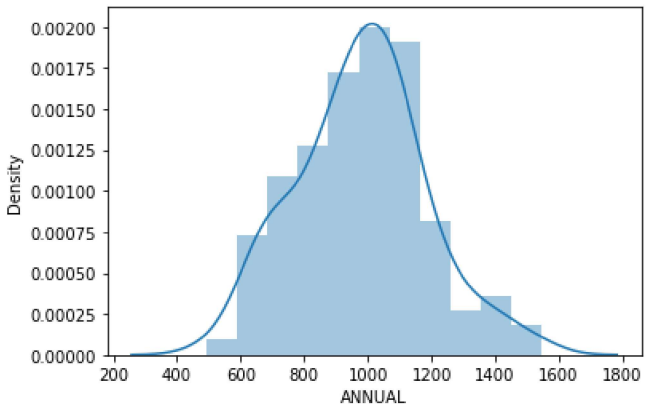
Out[14]:
```
<seaborn.axisgrid.PairGrid at 0x2c337c48820>
```



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

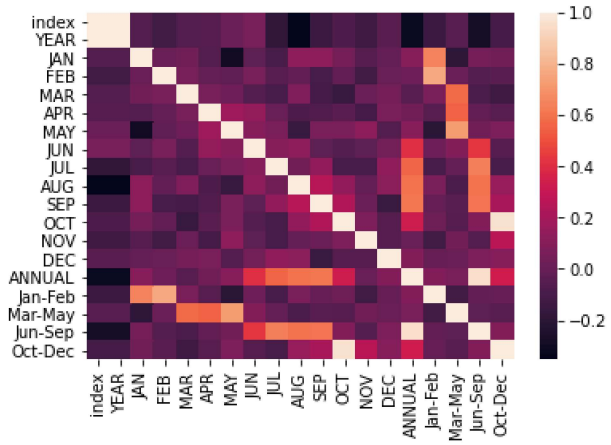
```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function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

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




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

Out[16]: <AxesSubplot:>




In [ ]:
```

```