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_coastal andhra pradesh.csv"
        df
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep
0	3082	COASTAL ANDHRA PRADESH	1901	18.8	80.9	7.2	28.7	68.7	77.7	113.0	133.7	125.3	173.4	164.8	1.5	993.8	99.7	104.6	449.7
1	3083	COASTAL ANDHRA PRADESH	1902	2.0	0.0	2.8	23.9	37.6	72.6	144.5	236.1	204.5	262.0	50.4	27.1	1063.6	2.0	64.4	657.7
2	3084	COASTAL ANDHRA PRADESH	1903	0.8	13.3	0.2	6.2	73.4	154.0	248.6	258.0	216.5	159.1	173.9	12.1	1316.2	14.2	79.8	877.1
3	3085	COASTAL ANDHRA PRADESH	1904	1.3	0.0	5.4	3.0	136.3	107.8	120.2	117.7	116.8	240.9	0.0	10.7	860.2	1.3	144.7	462.6
4	3086	COASTAL ANDHRA PRADESH	1905	1.1	16.7	68.0	37.0	68.8	84.4	64.6	210.8	170.2	66.0	7.4	0.0	795.2	17.8	173.8	530.1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	..
110	3192	COASTAL ANDHRA PRADESH	2011	0.0	17.9	0.9	62.3	67.9	86.8	196.0	215.8	129.7	74.6	4.9	5.0	861.9	17.9	131.2	628.4
111	3193	COASTAL ANDHRA PRADESH	2012	37.6	0.0	2.7	24.0	39.3	95.4	221.9	221.2	246.5	140.0	289.7	0.0	1318.4	37.6	66.1	785.0
112	3194	COASTAL ANDHRA PRADESH	2013	2.0	29.6	0.2	48.0	28.2	127.5	162.4	123.1	132.0	411.5	53.1	2.8	1120.5	31.7	76.4	545.0
113	3195	COASTAL ANDHRA PRADESH	2014	0.4	1.2	9.1	6.0	112.9	45.7	151.8	177.8	144.5	195.6	23.7	6.4	874.9	1.5	128.0	519.7
114	3196	COASTAL ANDHRA PRADESH	2015	2.0	0.6	5.5	32.3	34.1	283.8	116.0	192.0	201.8	59.7	81.2	2.0	1010.9	2.5	71.9	793.6

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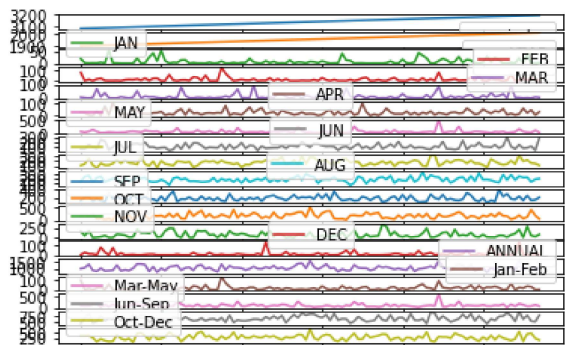
In [5]: `df.info()`

```
<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:>,
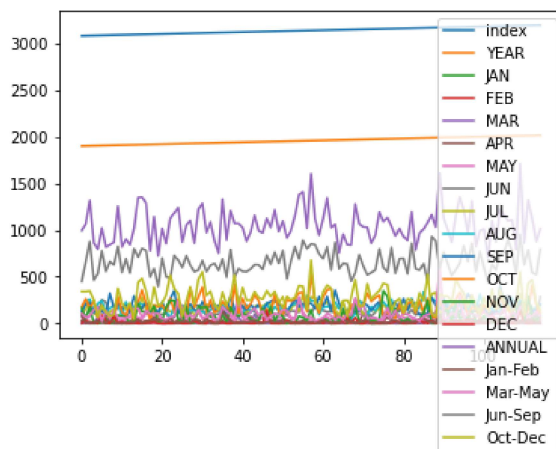   <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
        <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
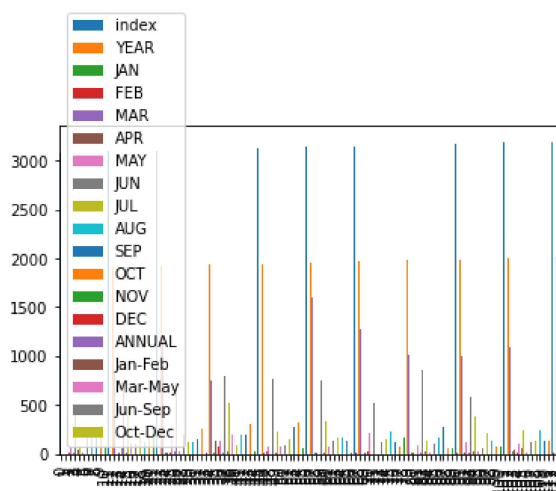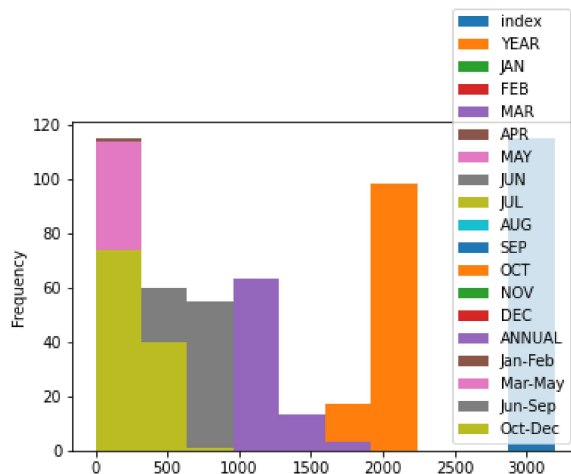   <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>], dtype=object)

Line chart

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

Out[7]: `<AxesSubplot:>`

Bar chart

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

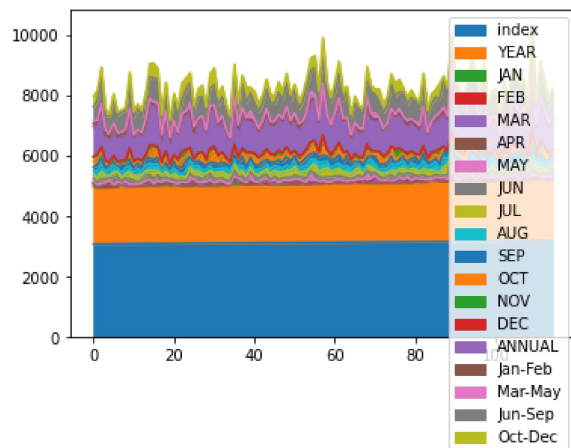Out[8]: `<AxesSubplot:>`

Histogram

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

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

Area chart

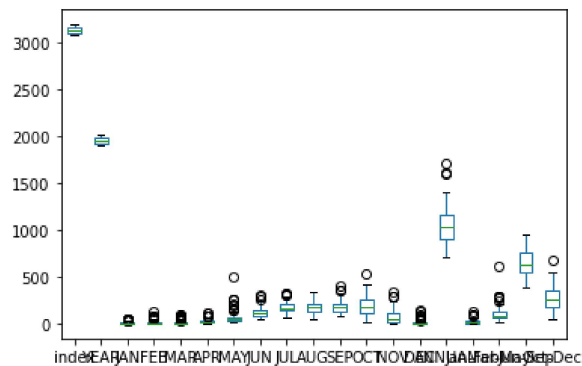In [10]: `df.plot.area()`

Out[10]: `<AxesSubplot:>`
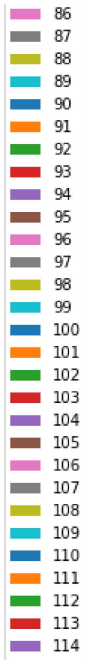

Box chart

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

Out[11]: `<AxesSubplot:>`
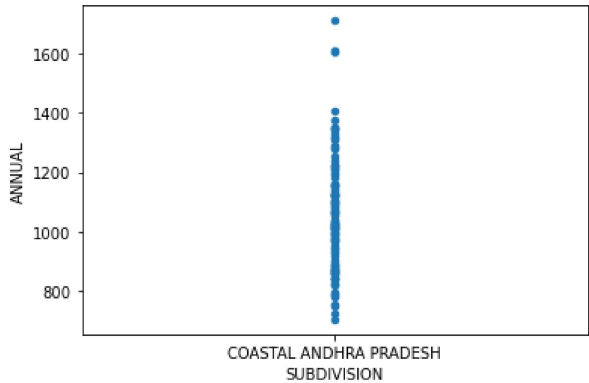
Pie chart

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

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

	86
	87
	88
	89
	90
	91
	92
	93
	94
	95
	96
	97
	98
	99
	100
	101
	102
	103
	104
	105
	106
	107
	108
	109
	110
	111
	112
	113
	114

Scatter chart

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

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

Seaborn

In [14]: `sns.pairplot(df)`
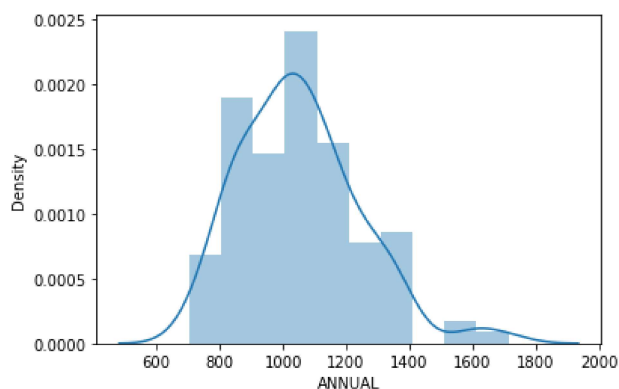
Out[14]: `<seaborn.axisgrid.PairGrid at 0x1fcdf5602b0>`
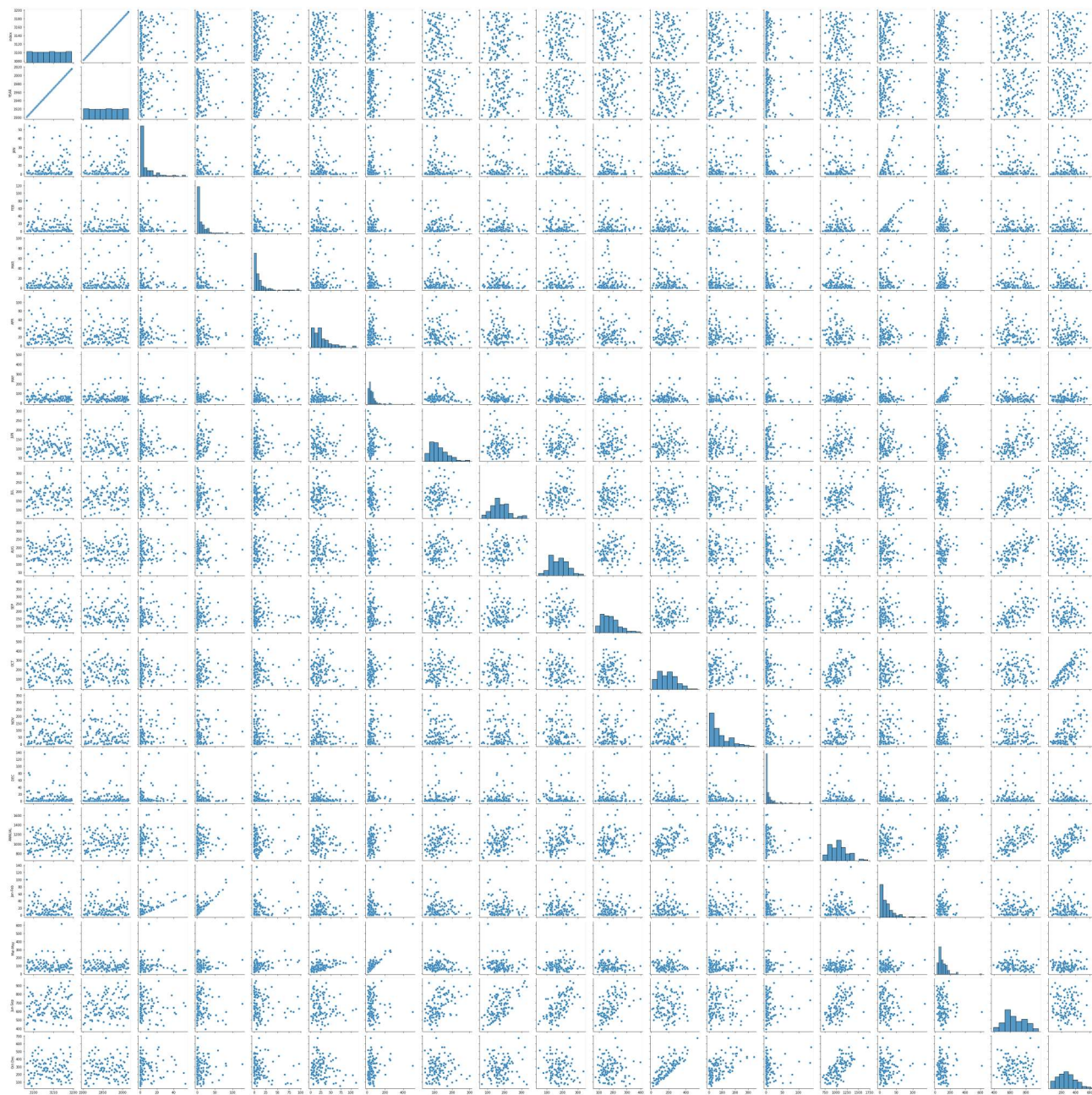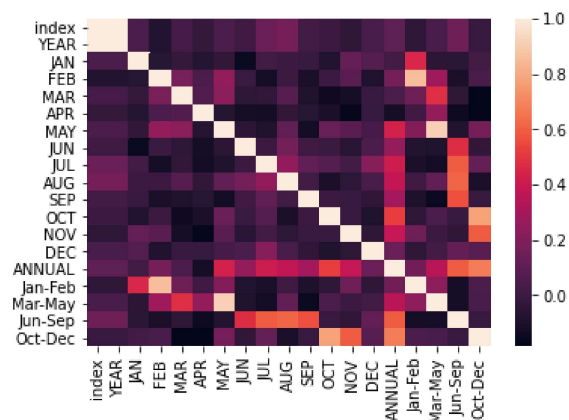
In [15]:
```python
sns.distplot(df['ANNUAL'])
```

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprec
ated 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]:
```python
sns.heatmap(df.corr())
```

Out[16]: <AxesSubplot:>

In [ ]: