# **Importing Libraries**

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

In [2]: df=pd.read\_csv("rainfall in india 1901-2015.csv")
df

#### Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
0	0	ANDAMAN & NICOBAR ISLANDS	1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6
1	1	ANDAMAN & NICOBAR ISLANDS	1902	0.0	159.8	12.2	0.0	446.1	537.1	228.9	753.7	666.2
2	2	ANDAMAN & NICOBAR ISLANDS	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0
3	3	ANDAMAN & NICOBAR ISLANDS	1904	9.4	14.7	0.0	202.4	304.5	495.1	502.0	160.1	820.4
4	4	ANDAMAN & NICOBAR ISLANDS	1905	1.3	0.0	3.3	26.9	279.5	628.7	368.7	330.5	297.0
4111	4111	LAKSHADWEEP	2011	5.1	2.8	3.1	85.9	107.2	153.6	350.2	254.0	255.2
4112	4112	LAKSHADWEEP	2012	19.2	0.1	1.6	76.8	21.2	327.0	231.5	381.2	179.8
4113	4113	LAKSHADWEEP	2013	26.2	34.4	37.5	5.3	88.3	426.2	296.4	154.4	180.0
4114	4114	LAKSHADWEEP	2014	53.2	16.1	4.4	14.9	57.4	244.1	116.1	466.1	132.2
4115	4115	LAKSHADWEEP	2015	2.2	0.5	3.7	87.1	133.1	296.6	257.5	146.4	160.4

4116 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: 4090 entries, 0 to 4115
        Data columns (total 20 columns):
              Column
                           Non-Null Count
                                           Dtype
                           -----
              _ _ _ _ _ _
         0
              index
                           4090 non-null
                                           int64
         1
                          4090 non-null
             SUBDIVISION
                                           object
         2
                           4090 non-null
                                           int64
             YEAR
         3
             JAN
                           4090 non-null
                                           float64
         4
             FEB
                           4090 non-null
                                           float64
         5
                                           float64
             MAR
                           4090 non-null
                                           float64
         6
                           4090 non-null
             APR
         7
                           4090 non-null
                                           float64
             MAY
         8
             JUN
                           4090 non-null
                                           float64
         9
                                           float64
             JUL
                           4090 non-null
         10 AUG
                           4090 non-null
                                           float64
                                           float64
         11
             SEP
                           4090 non-null
                           4090 non-null
                                           float64
         12 OCT
         13 NOV
                           4090 non-null
                                           float64
         14 DEC
                           4090 non-null
                                           float64
         15 ANNUAL
                           4090 non-null
                                           float64
             Jan-Feb
                           4090 non-null
                                           float64
         16
         17 Mar-May
                           4090 non-null
                                           float64
             Jun-Sep
                           4090 non-null
                                           float64
         18
                           4090 non-null
         19 Oct-Dec
                                           float64
        dtypes: float64(17), int64(2), object(1)
        memory usage: 671.0+ 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
                                   ΔDD
         500
1000
                                   MAY
                                   ILIM
         1000
2006
1000
500
500
500
500
1000
                                   IUL
                                   AUG
                                   SEP
                                   OCT
                                   NOV
                                                    DEC
                                 ANNUAL
                                                  Jan-Feb
                                  Mar-May
                                  lun-Sep
                                  Oct-Dec 2
```

#### Line chart

1000

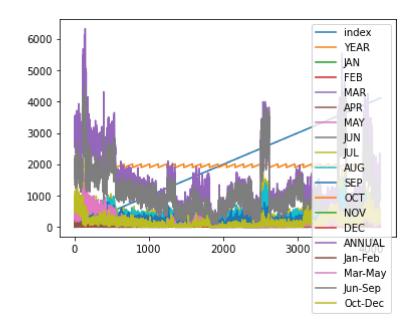
2000

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

3000

4000

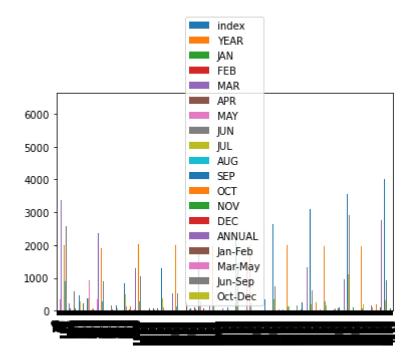
Out[7]: <AxesSubplot:>



### **Bar chart**

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

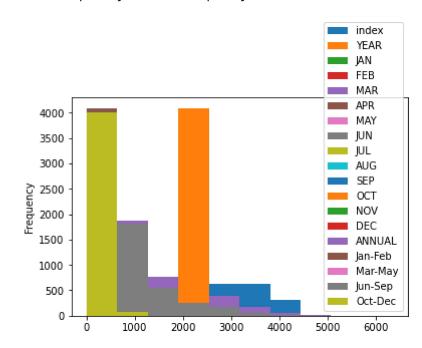
Out[8]: <AxesSubplot:>



# Histogram

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

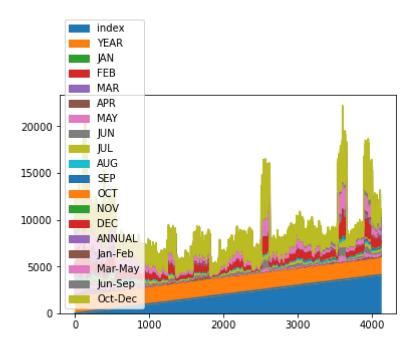
Out[9]: <AxesSubplot:ylabel='Frequency'>



### **Area chart**

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

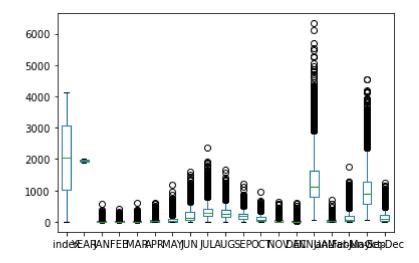
Out[10]: <AxesSubplot:>



# **Box chart**

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

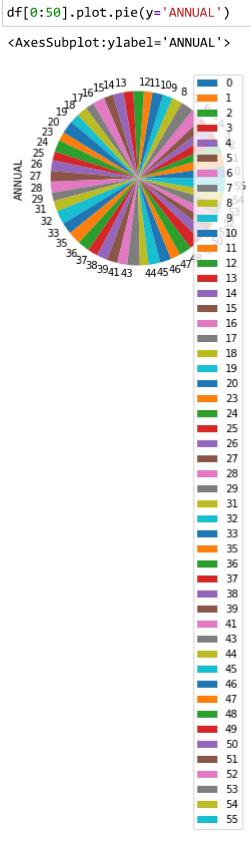
Out[11]: <AxesSubplot:>



# Pie chart

In [12]: df[0:50].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'>

