

# 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("rainfall_andaman_nicobar_islands.csv")
df
```

Out[2]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
0	0	ANDAMAN & NICOBAR ISLANDS	1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6	388.5	558.1	100.0
1	1	ANDAMAN & NICOBAR ISLANDS	1902	0.0	159.8	12.2	0.0	446.1	537.1	228.9	753.7	666.2	197.2	359.5	100.0
2	2	ANDAMAN & NICOBAR ISLANDS	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0	181.2	284.2	100.0
3	3	ANDAMAN & NICOBAR ISLANDS	1904	9.4	14.7	0.0	202.4	304.5	495.1	502.0	160.1	820.4	222.2	308.4	100.0
4	4	ANDAMAN & NICOBAR ISLANDS	1905	1.3	0.0	3.3	26.9	279.5	628.7	368.7	330.5	297.0	260.7	25.5	100.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	100.0
105	105	ANDAMAN & NICOBAR ISLANDS	2011	265.9	84.8	272.8	111.4	326.5	383.2	583.2	441.5	757.1	212.3	150.0	100.0
106	106	ANDAMAN & NICOBAR ISLANDS	2012	119.9	45.6	30.9	55.8	533.9	458.2	317.3	369.6	868.9	209.7	300.0	100.0
107	107	ANDAMAN & NICOBAR ISLANDS	2013	67.1	37.6	43.0	46.3	509.3	777.0	564.8	336.7	473.6	455.8	354.0	100.0
108	108	ANDAMAN & NICOBAR ISLANDS	2014	41.9	8.6	0.0	11.1	238.0	416.6	467.6	321.6	412.9	402.6	201.0	100.0
109	109	ANDAMAN & NICOBAR ISLANDS	2015	126.9	7.6	3.1	138.2	331.9	346.4	328.9	480.0	523.3	252.1	236.0	100.0

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

110 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()
```

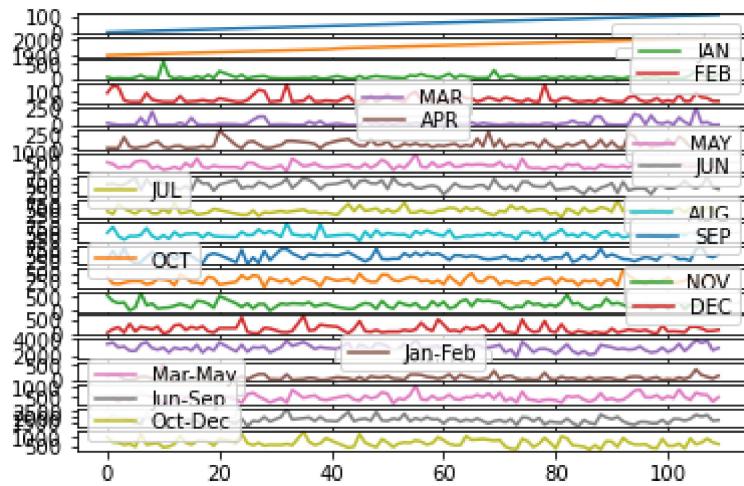
```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 104 entries, 0 to 109
Data columns (total 20 columns):
 #   Column      Non-Null Count  Dtype  
---  --  
 0   index       104 non-null    int64  
 1   SUBDIVISION 104 non-null    object  
 2   YEAR        104 non-null    int64  
 3   JAN          104 non-null    float64 
 4   FEB          104 non-null    float64 
 5   MAR          104 non-null    float64 
 6   APR          104 non-null    float64 
 7   MAY          104 non-null    float64 
 8   JUN          104 non-null    float64 
 9   JUL          104 non-null    float64 
 10  AUG          104 non-null    float64 
 11  SEP          104 non-null    float64 
 12  OCT          104 non-null    float64 
 13  NOV          104 non-null    float64 
 14  DEC          104 non-null    float64 
 15  ANNUAL       104 non-null    float64 
 16  Jan-Feb     104 non-null    float64 
 17  Mar-May     104 non-null    float64 
 18  Jun-Sep     104 non-null    float64 
 19  Oct-Dec     104 non-null    float64 
dtypes: float64(17), int64(2), object(1)
memory usage: 17.1+ 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:>], dtype=object)
```

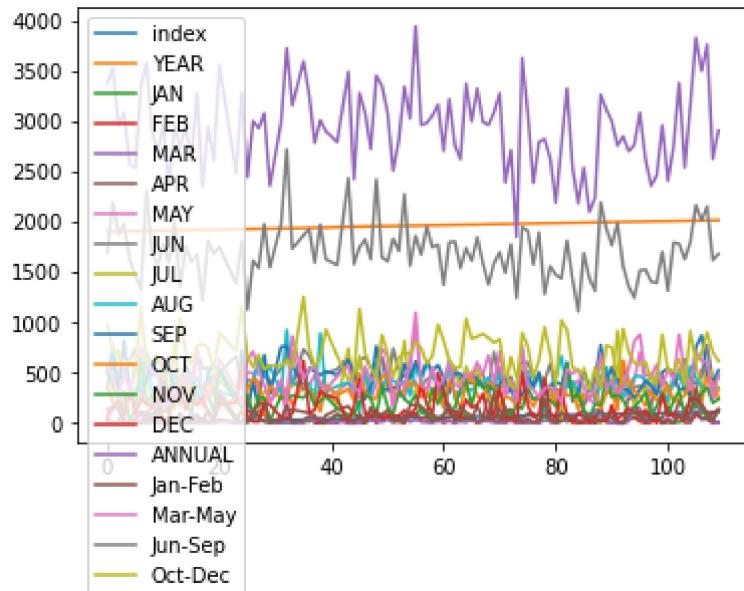
Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js



## Line chart

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

Out[7]: <AxesSubplot:>

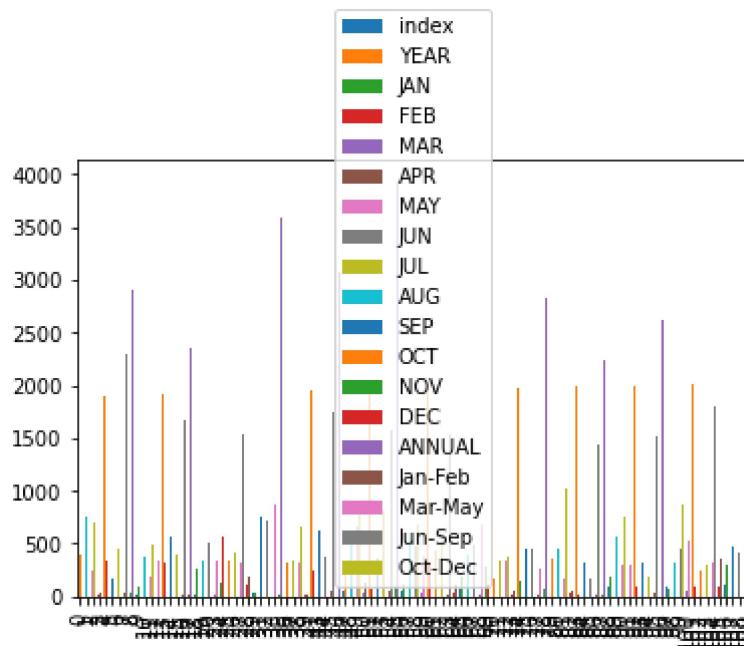


## Bar chart

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

Out[8]: <AxesSubplot:>

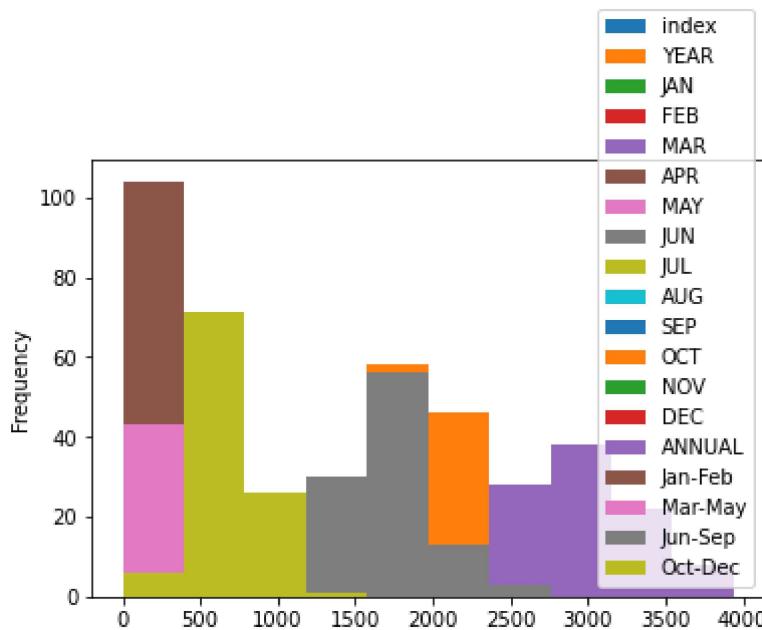
Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js



## Histogram

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

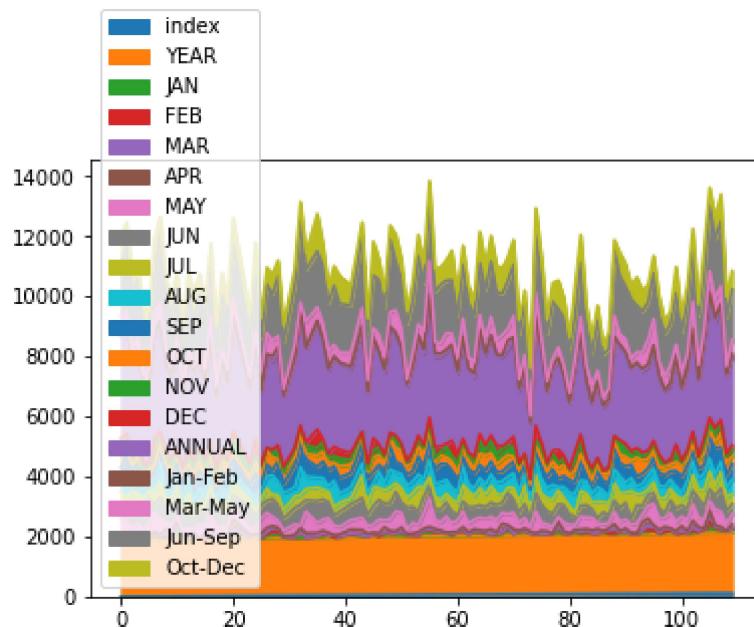
```
Out[9]: <AxesSubplot:ylabel='Frequency'>
```



## Area chart

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

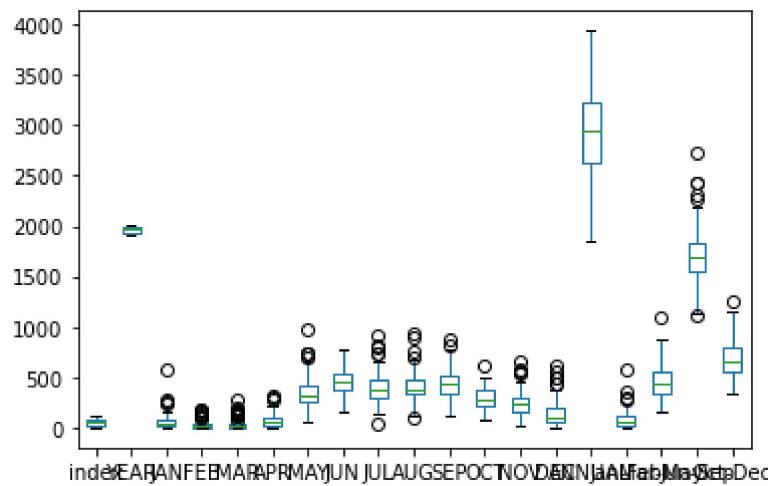
```
Out[10]: <AxesSubplot>
Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js
```



## Box chart

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

```
Out[11]: <AxesSubplot:
```

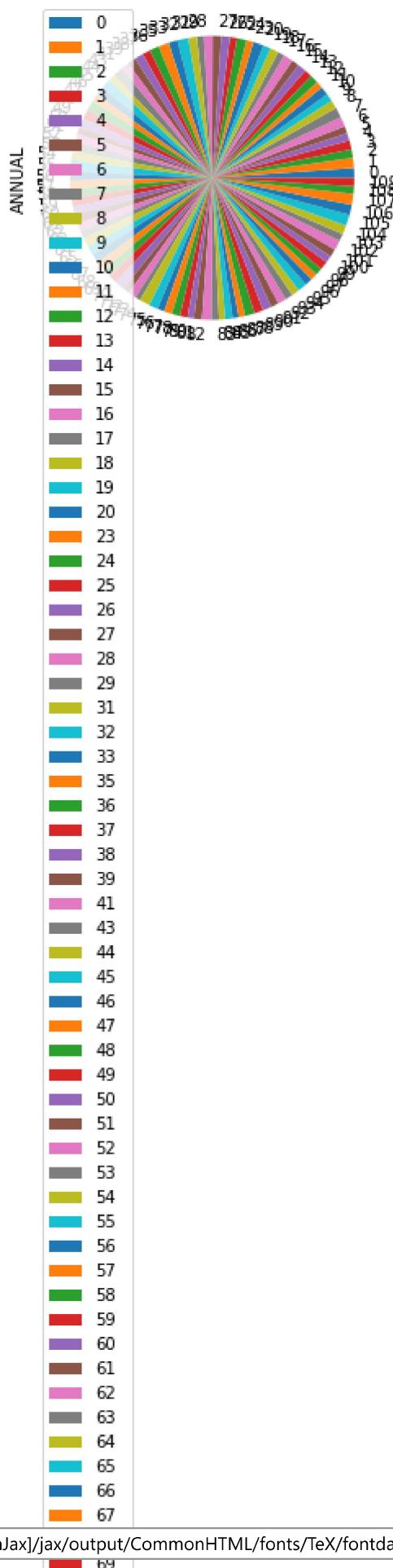


## Pie chart

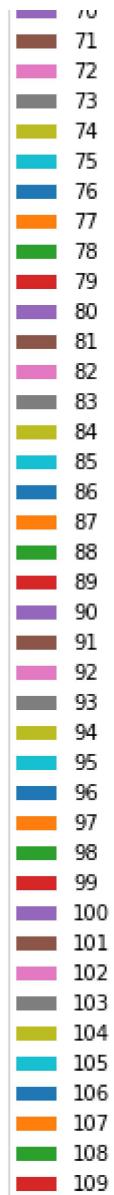
```
In [12]: df.plot.pie(y='ANNUAL' )
```

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

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js



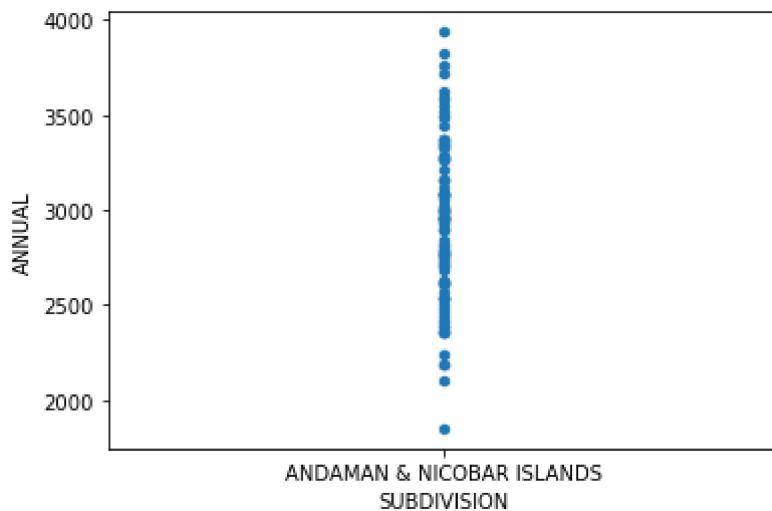
Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js



## Scatter chart

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

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



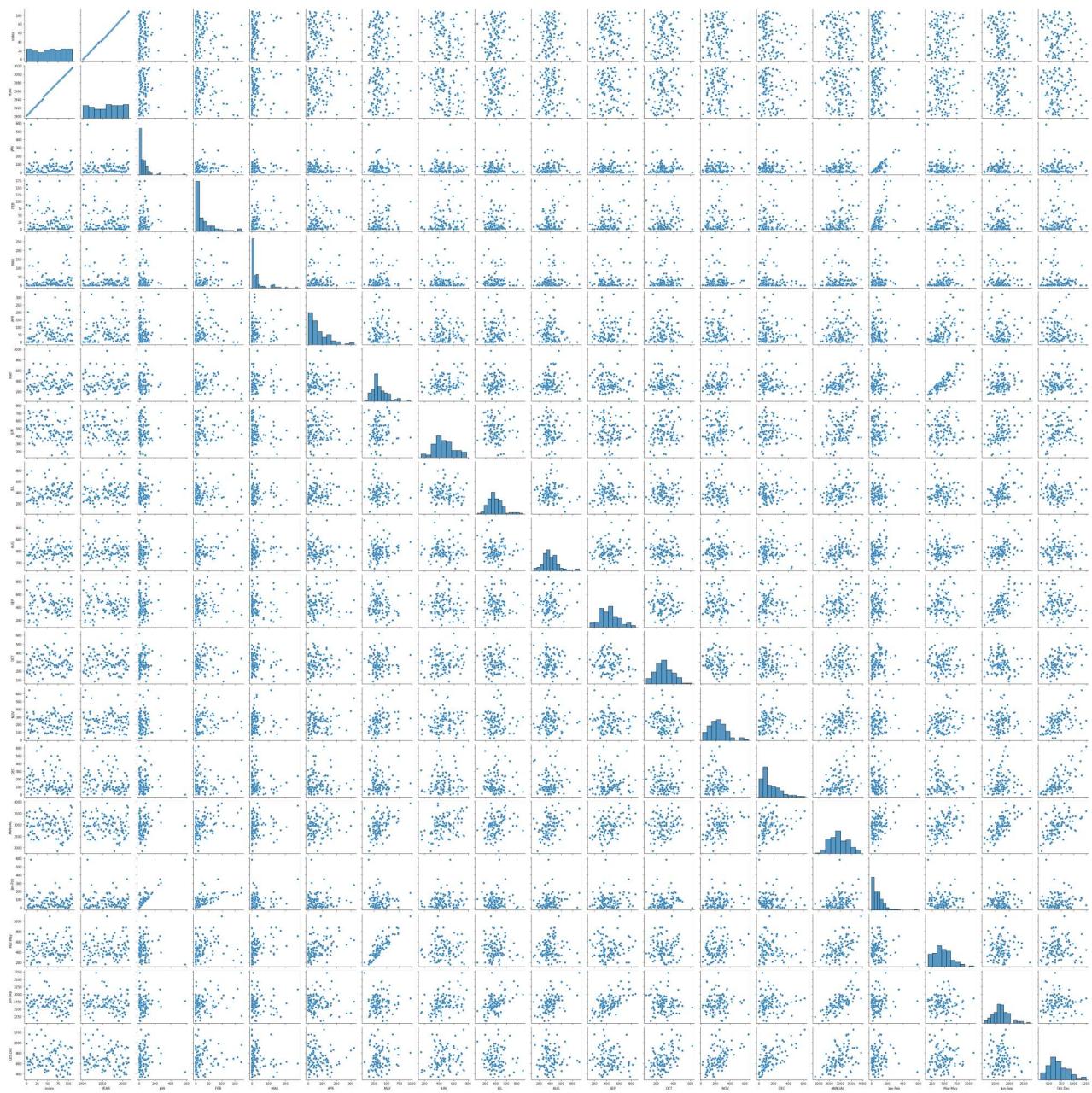
## Seaborn

In [14]:

```
sns.pairplot(df)
```

Out[14]: &lt;seaborn.axisgrid.PairGrid at 0x1e98dd12c10&gt;

## rainfall\_andaman\_nicobar islands



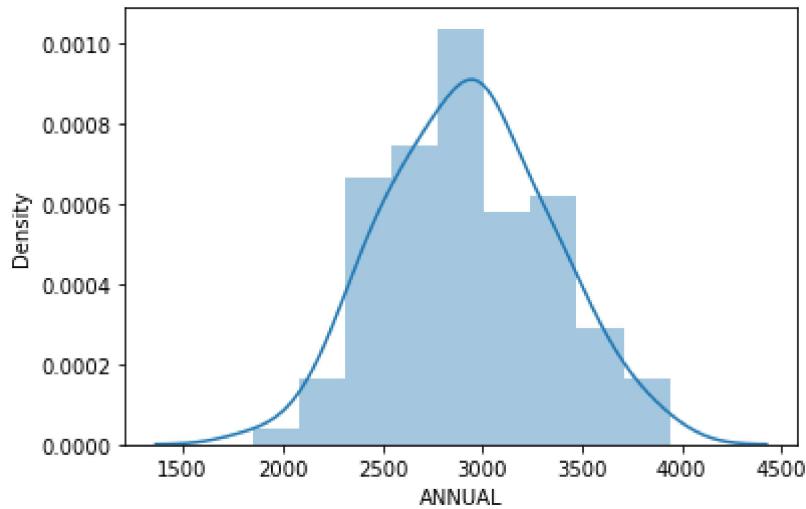
In [15]:  
`sns.distplot(df['ANNUAL'])`

```
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)
```

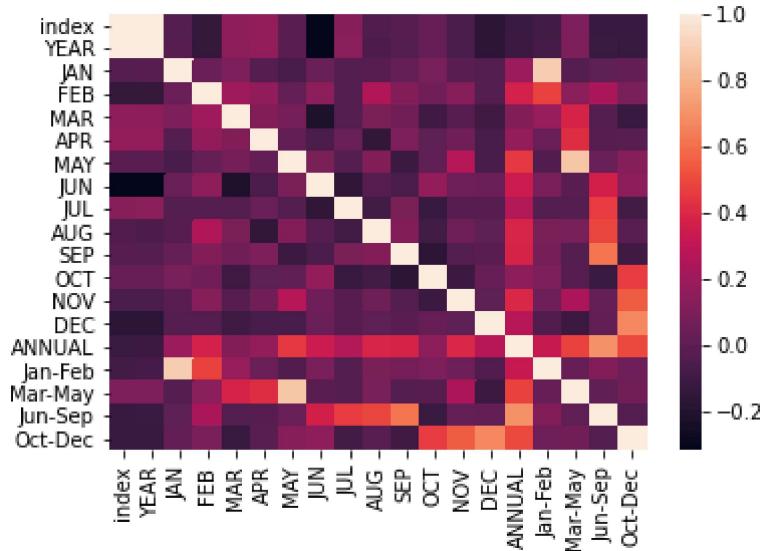
Out[15]: `<AxesSubplot:xlabel='ANNUAL', ylabel='Density'>`

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js



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

Out[16]: <AxesSubplot:>



In [ ]: