

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

```
In [2]: data=pd.read_csv(r"C:\Users\user\Downloads\rainfall in india 1901-2015.csv")
data
```

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

Naga Mizo tripura

```
In [18]: df=data.iloc[322:437]
df
```

Out[18]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
322	322	NAGA MANI MIZO TRIPURA	1901	11.7	18.1	29.4	206.2	124.0	443.3	331.4	466.0	304.1	166.0
323	323	NAGA MANI MIZO TRIPURA	1902	4.8	0.5	36.3	297.8	215.5	480.1	392.4	312.8	318.7	107.0
324	324	NAGA MANI MIZO TRIPURA	1903	6.5	40.5	139.8	45.5	159.9	458.6	300.2	470.6	366.1	166.0
325	325	NAGA MANI MIZO TRIPURA	1904	2.3	46.9	47.5	290.3	230.5	455.3	423.5	423.6	375.8	128.0
326	326	NAGA MANI MIZO TRIPURA	1905	9.1	35.3	306.5	161.7	193.6	339.7	450.1	429.9	320.1	246.0
...
432	432	NAGA MANI MIZO TRIPURA	2011	12.6	3.6	51.4	81.1	334.9	374.2	313.3	367.6	258.3	97.0
433	433	NAGA MANI MIZO TRIPURA	2012	24.5	10.2	20.3	243.5	163.5	396.2	280.1	342.7	248.7	166.0
434	434	NAGA MANI MIZO TRIPURA	2013	0.2	5.7	19.7	60.3	348.9	206.6	255.9	291.3	241.4	128.0
435	435	NAGA MANI MIZO TRIPURA	2014	1.2	21.0	25.4	49.6	192.5	268.3	295.7	372.3	300.9	68.0
436	436	NAGA MANI MIZO TRIPURA	2015	14.4	14.2	21.6	253.5	198.3	283.9	413.6	334.2	255.9	118.0

115 rows × 20 columns

In [4]: `df.head()`

Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
321	321	ASSAM & MEGHALAYA	2015	13.4	15.5	37.5	250.9	332.5	558.5	300.1	590.9	279.9	6%
322	322	NAGA MANI MIZO TRIPURA	1901	11.7	18.1	29.4	206.2	124.0	443.3	331.4	466.0	304.1	16%
323	323	NAGA MANI MIZO TRIPURA	1902	4.8	0.5	36.3	297.8	215.5	480.1	392.4	312.8	318.7	10%
324	324	NAGA MANI MIZO TRIPURA	1903	6.5	40.5	139.8	45.5	159.9	458.6	300.2	470.6	366.1	16%
325	325	NAGA MANI MIZO TRIPURA	1904	2.3	46.9	47.5	290.3	230.5	455.3	423.5	423.6	375.8	12%



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

Out[5]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
432	432	NAGA MANI MIZO TRIPURA	2011	12.6	3.6	51.4	81.1	334.9	374.2	313.3	367.6	258.3	92
433	433	NAGA MANI MIZO TRIPURA	2012	24.5	10.2	20.3	243.5	163.5	396.2	280.1	342.7	248.7	160
434	434	NAGA MANI MIZO TRIPURA	2013	0.2	5.7	19.7	60.3	348.9	206.6	255.9	291.3	241.4	125
435	435	NAGA MANI MIZO TRIPURA	2014	1.2	21.0	25.4	49.6	192.5	268.3	295.7	372.3	300.9	69
436	436	NAGA MANI MIZO TRIPURA	2015	14.4	14.2	21.6	253.5	198.3	283.9	413.6	334.2	255.9	118



```
In [6]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 116 entries, 321 to 436
Data columns (total 20 columns):
#   Column          Non-Null Count  Dtype
---  -
0   index           116 non-null   int64
1   SUBDIVISION     116 non-null   object
2   YEAR            116 non-null   int64
3   JAN             116 non-null   float64
4   FEB             116 non-null   float64
5   MAR             116 non-null   float64
6   APR             116 non-null   float64
7   MAY             116 non-null   float64
8   JUN             116 non-null   float64
9   JUL             116 non-null   float64
10  AUG             116 non-null   float64
11  SEP             116 non-null   float64
12  OCT             116 non-null   float64
13  NOV             116 non-null   float64
14  DEC             116 non-null   float64
15  ANNUAL          116 non-null   float64
16  Jan-Feb        116 non-null   float64
17  Mar-May        116 non-null   float64
18  Jun-Sep        116 non-null   float64
19  Oct-Dec        116 non-null   float64
dtypes: float64(17), int64(2), object(1)
memory usage: 18.3+ KB
```

In [7]:

df1=data.fillna(0)
df1

Out[7]:

	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

In [8]:

df1.describe()

Out[8]:

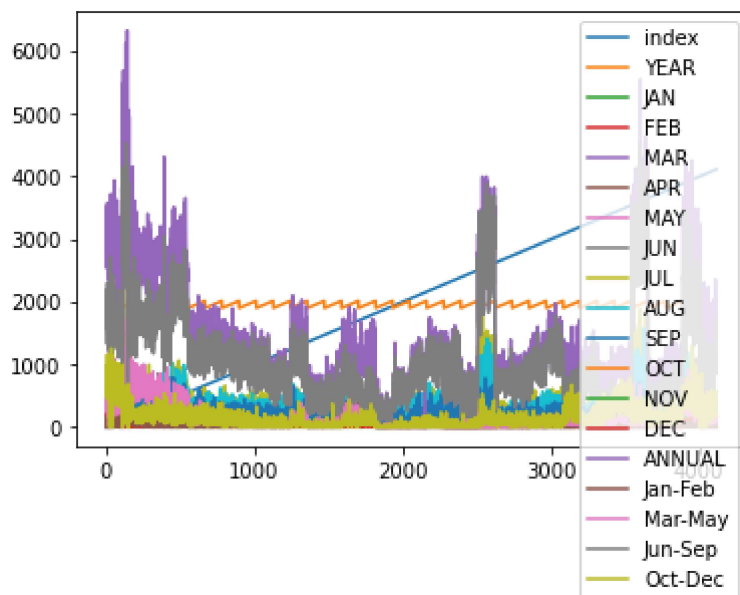
	index	YEAR	JAN	FEB	MAR	APR	MA
count	4116.000000	4116.000000	4116.000000	4116.000000	4116.000000	4116.000000	4116.000000
mean	2057.500000	1958.218659	18.938897	21.789431	27.319315	43.085520	85.682920
std	1188.331183	33.140898	33.574242	35.901220	46.936787	67.811512	123.211700
min	0.000000	1901.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	1028.750000	1930.000000	0.600000	0.600000	1.000000	3.000000	8.600000
50%	2057.500000	1958.000000	6.000000	6.700000	7.800000	15.600000	36.400000
75%	3086.250000	1987.000000	22.125000	26.800000	31.225000	49.825000	96.825000
max	4115.000000	2015.000000	583.700000	403.500000	605.600000	595.100000	1168.600000

```
In [9]: df1.columns
```

```
Out[9]: 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')
```

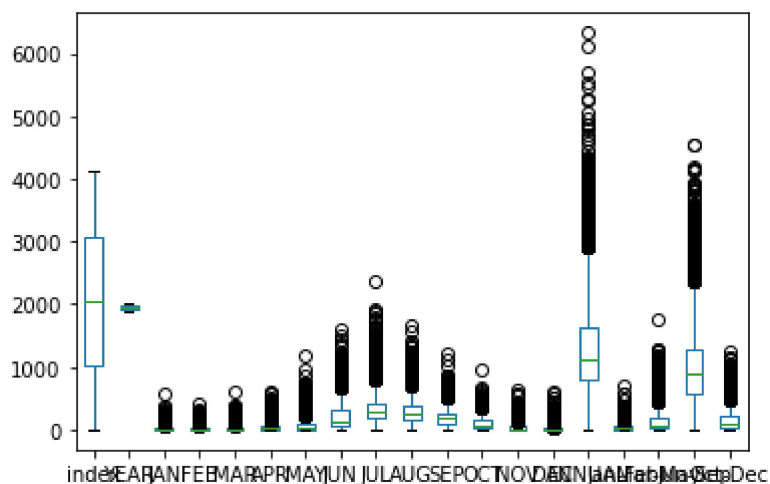
```
In [10]: df1.plot.line()
```

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



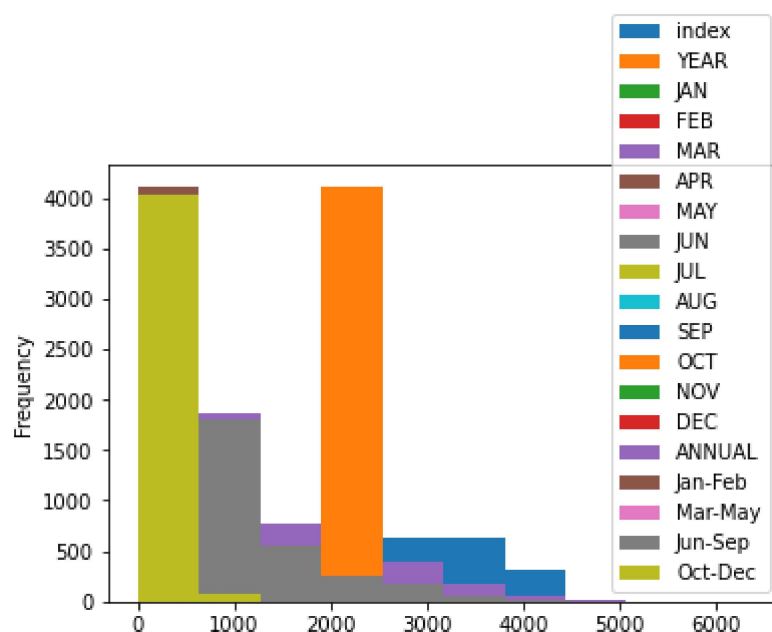
```
In [11]: df1.plot.box()
```

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



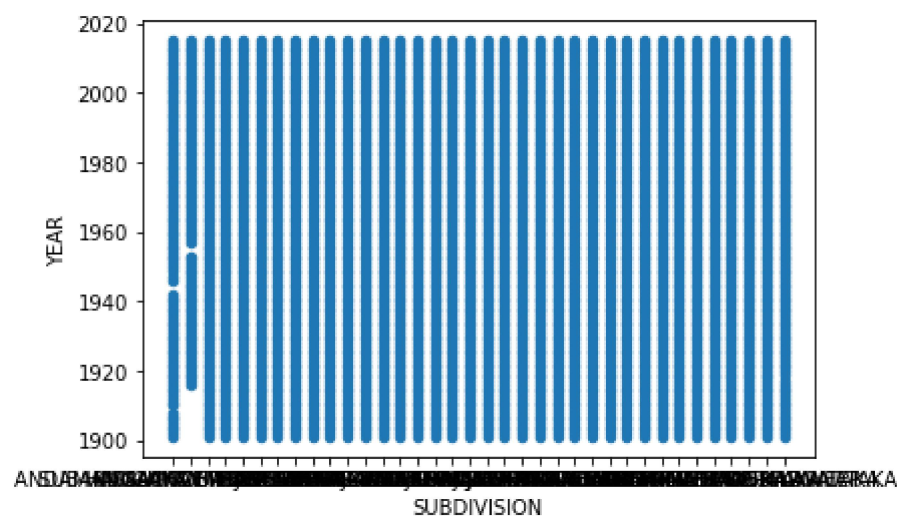
```
In [12]: df1.plot.hist()
```

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

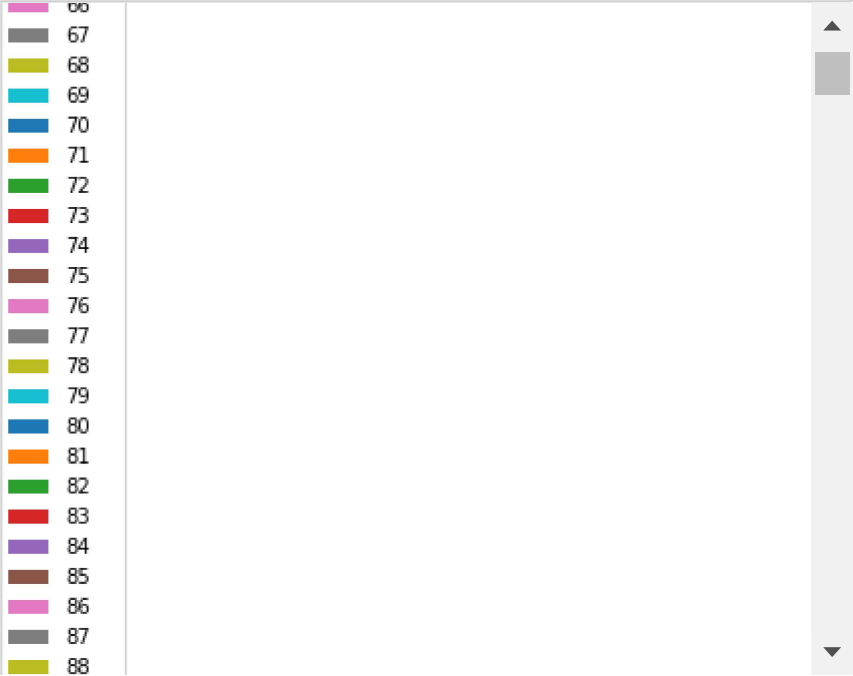


```
In [13]: df1.plot.scatter(x="SUBDIVISION",y="YEAR")
```

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

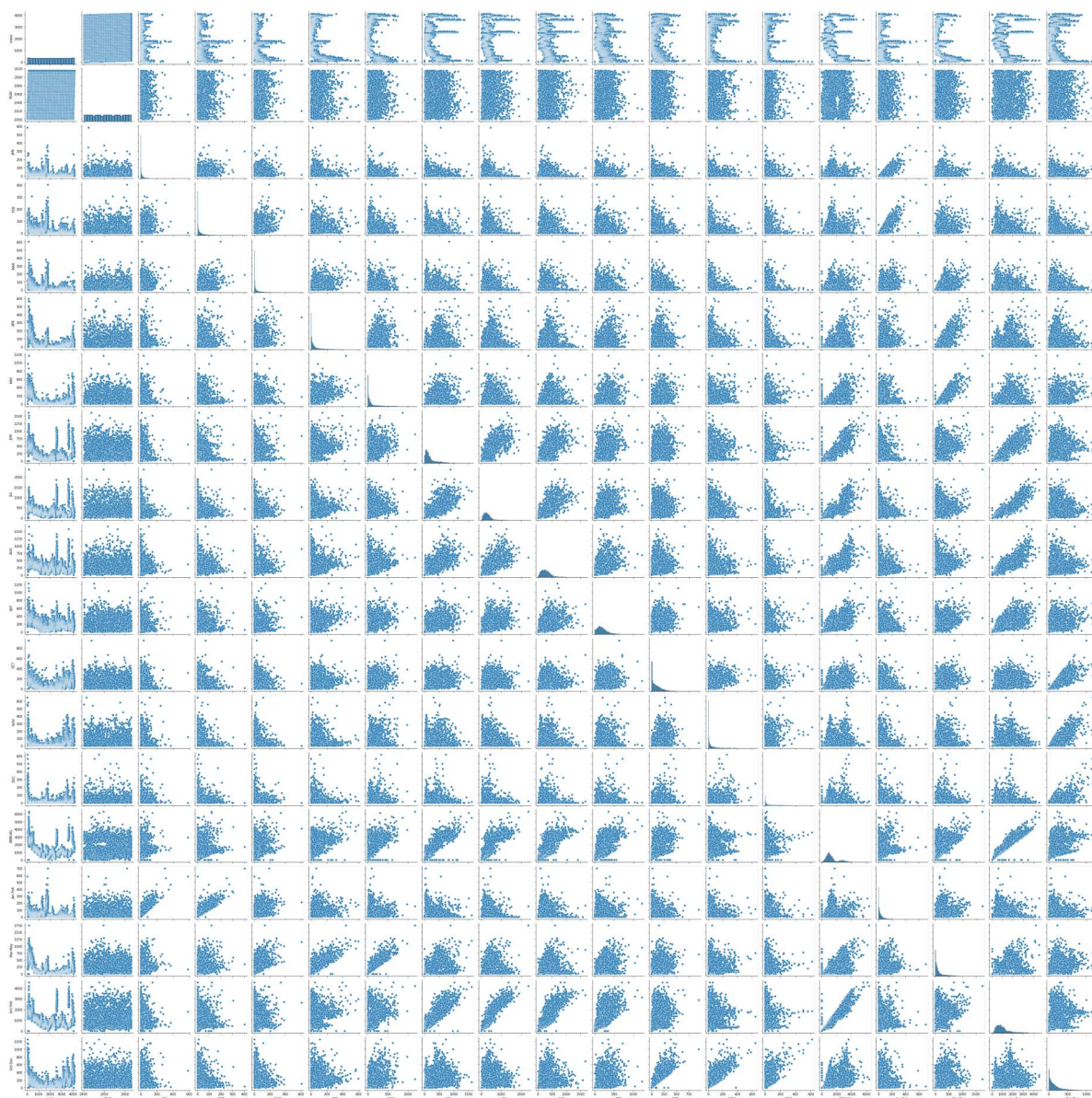


```
In [14]: df2=df1[['Mar-May']]
df2.plot.pie(subplots=True)
```



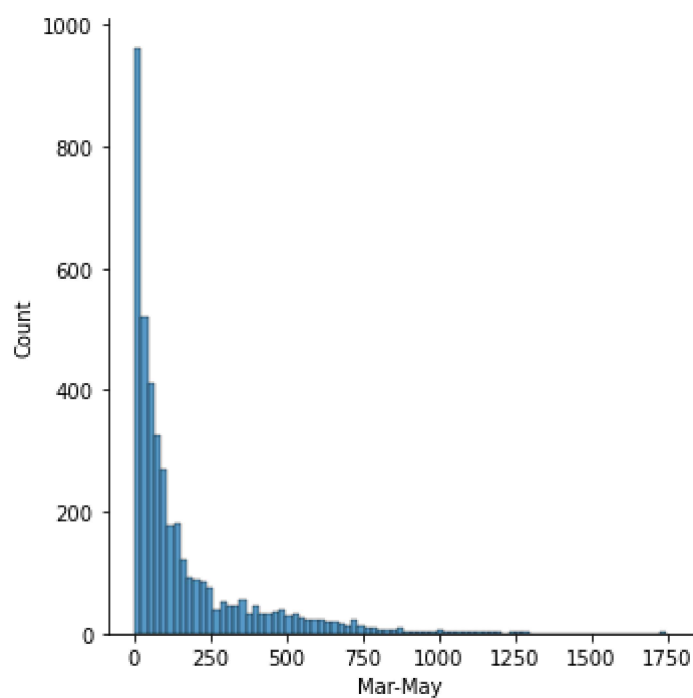

```
In [19]: sns.pairplot(df1)
```

```
Out[19]: <seaborn.axisgrid.PairGrid at 0x1c770d0c520>
```



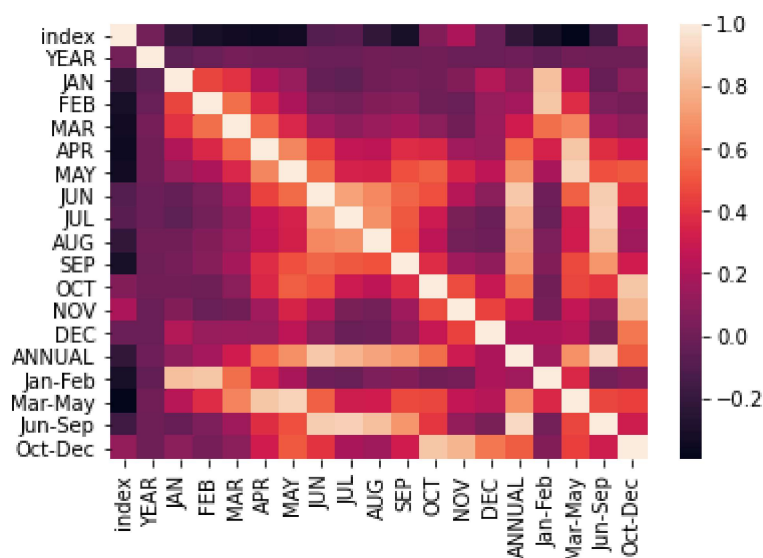
```
In [20]: sns.displot(data["Mar-May"])
```

```
Out[20]: <seaborn.axisgrid.FacetGrid at 0x1c71c5359a0>
```



```
In [21]: sns.heatmap(df1.corr())
```

```
Out[21]: <AxesSubplot:>
```



```
In [ ]:
```