

# Importing Libraries

In [17]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

# Importing Datasets

In [18]:

```
df=pd.read_csv("rainfall_east madhya pradesh.csv")
df
```

Out[18]:

		index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
<b>0</b>	2162		EAST MADHYA PRADESH	1901	48.5	38.1	15.7	10.7	6.2	61.0	367.5	589.2	189.9	5.9	0.0
<b>1</b>	2163		EAST MADHYA PRADESH	1902	14.9	8.9	0.0	3.6	2.7	28.0	411.9	227.0	236.6	17.0	27.6
<b>2</b>	2164		EAST MADHYA PRADESH	1903	5.6	2.9	0.3	0.9	37.5	67.5	261.4	366.7	257.4	177.9	0.0
<b>3</b>	2165		EAST MADHYA PRADESH	1904	2.0	15.3	48.2	0.0	8.6	109.9	443.2	316.6	135.6	44.8	3.2
<b>4</b>	2166		EAST MADHYA PRADESH	1905	15.9	8.0	14.3	12.3	10.2	34.4	292.4	243.3	250.9	2.9	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
<b>110</b>	2272		EAST MADHYA PRADESH	2011	0.6	1.9	0.3	7.1	4.7	332.5	323.6	326.9	276.5	1.1	0.0
<b>111</b>	2273		EAST MADHYA PRADESH	2012	39.4	0.7	0.6	1.1	1.2	67.8	398.9	351.7	172.6	12.7	3.8
<b>112</b>	2274		EAST MADHYA PRADESH	2013	2.0	43.4	14.1	9.5	0.3	311.9	456.2	480.8	78.0	124.2	0.5
<b>113</b>	2275		EAST MADHYA PRADESH	2014	32.1	49.7	17.8	5.1	2.5	91.8	283.4	231.8	139.6	56.4	1.9
<b>114</b>	2276		EAST MADHYA PRADESH	2015	37.3	11.0	73.4	25.8	6.3	139.2	262.2	272.1	71.6	38.2	1.2

115 rows × 20 columns

# Data Cleaning and Data Preprocessing

```
In [19]: df=df.dropna()
```

```
In [20]: df.columns
```

```
Out[20]: 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')
```

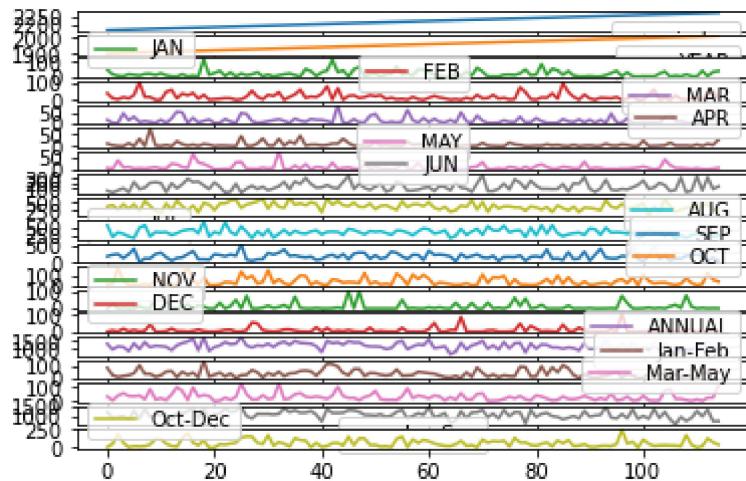
```
In [21]: df.info()
```

```
<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 [22]: df.plot.line(subplots=True)
```

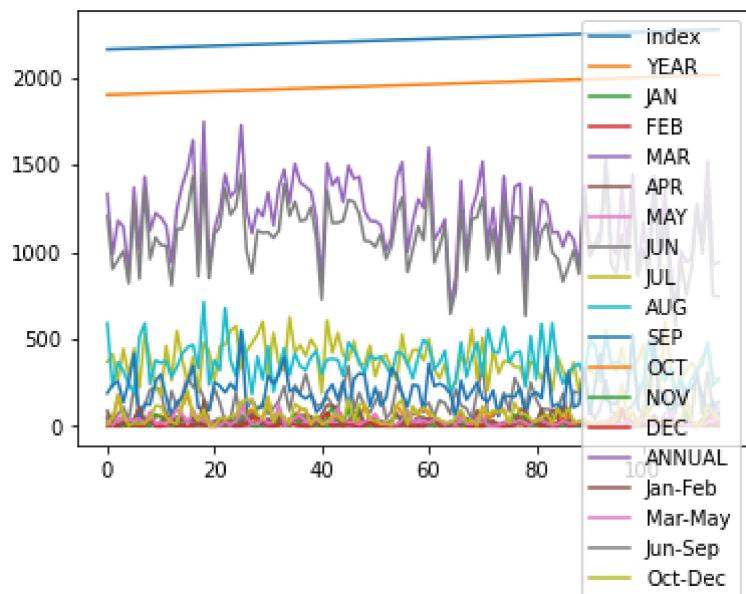
```
Out[22]: array([<AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>, <AxesSubplot:>,
       <AxesSubplot:>, <AxesSubplot:>], dtype=object)
```



## Line chart

In [23]: `df.plot.line()`

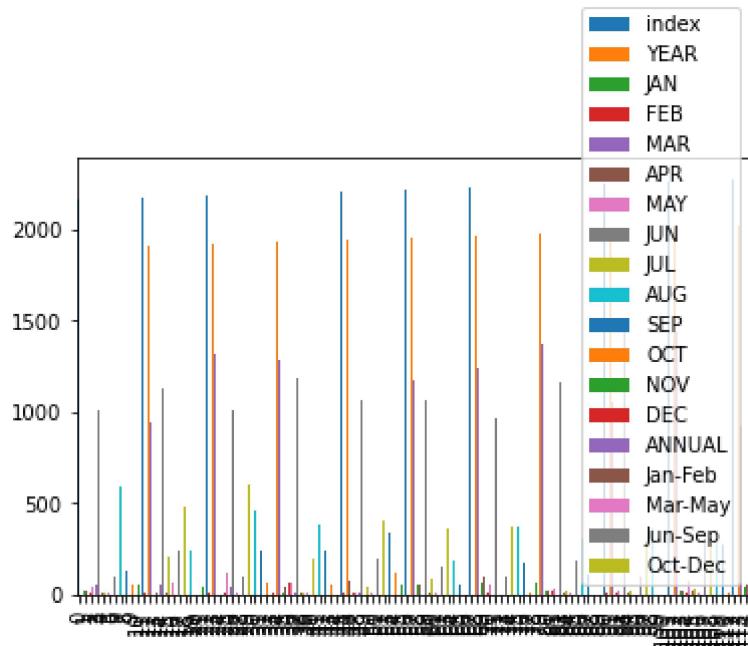
Out[23]: <AxesSubplot:>



## Bar chart

In [24]: `df.plot.bar()`

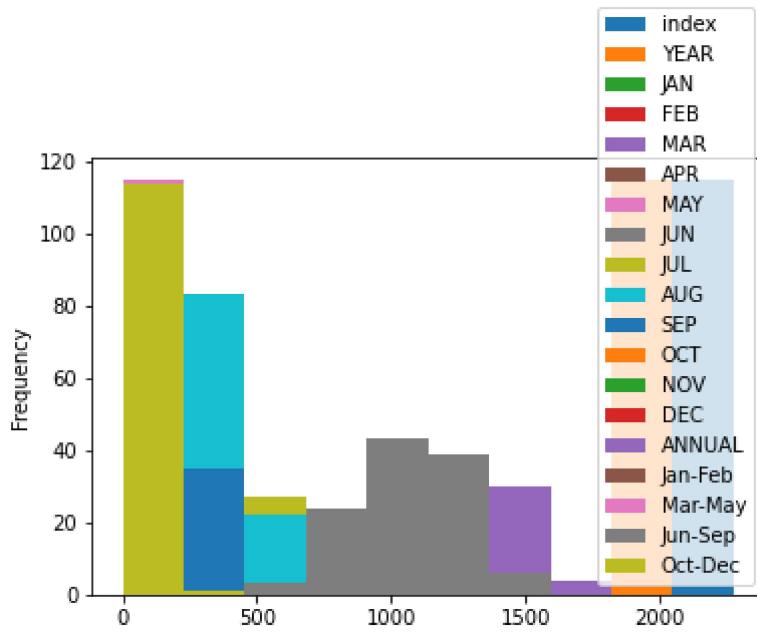
Out[24]: <AxesSubplot:>



## Histogram

```
In [25]: df.plot.hist()
```

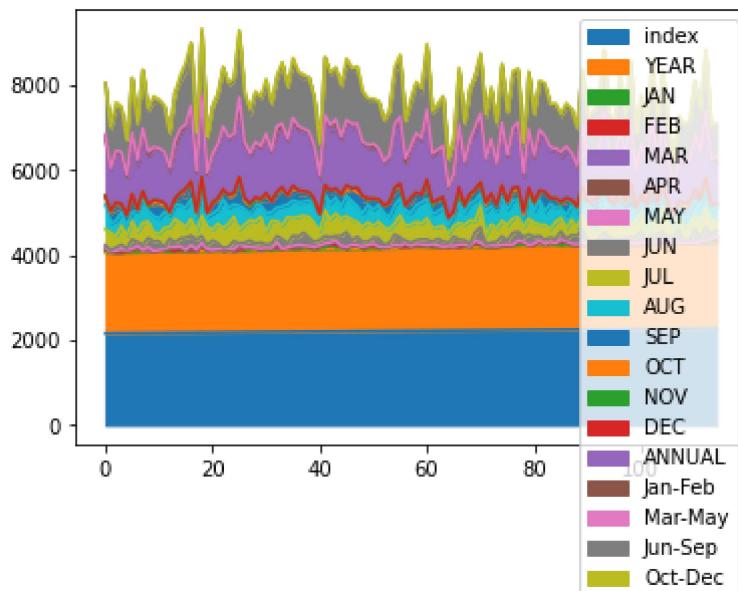
```
Out[25]: <AxesSubplot:ylabel='Frequency'>
```



## Area chart

```
In [26]: df.plot.area()
```

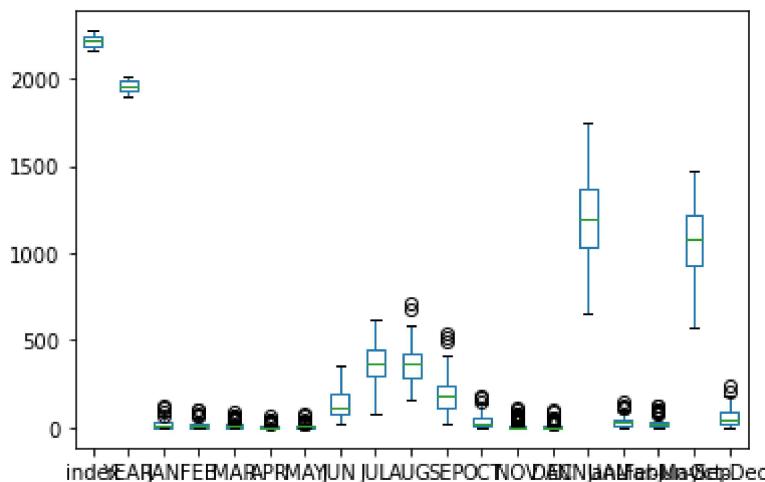
```
Out[26]: <AxesSubplot:>
```



## Box chart

```
In [27]: df.plot.box()
```

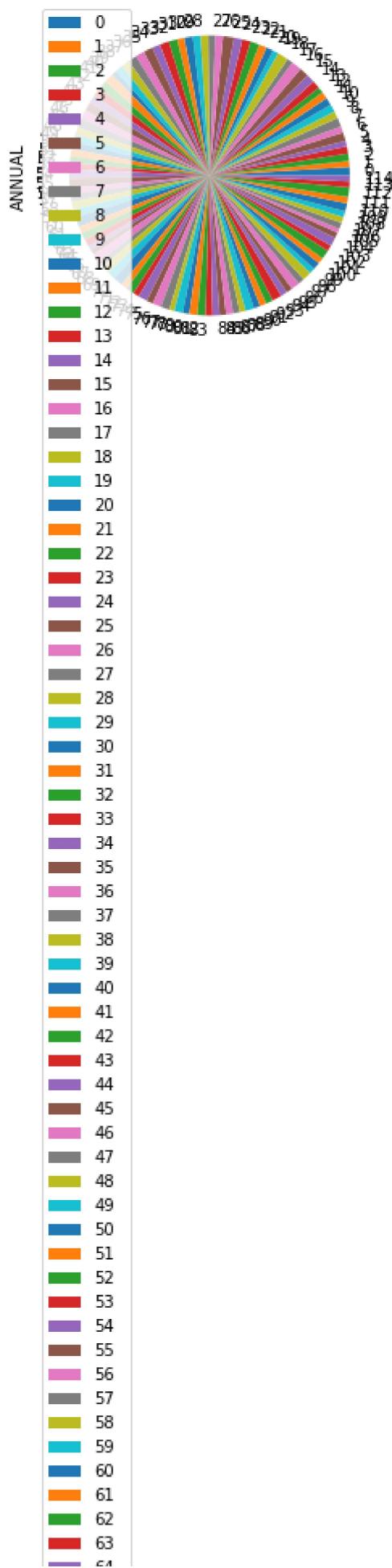
```
Out[27]: <AxesSubplot:>
```

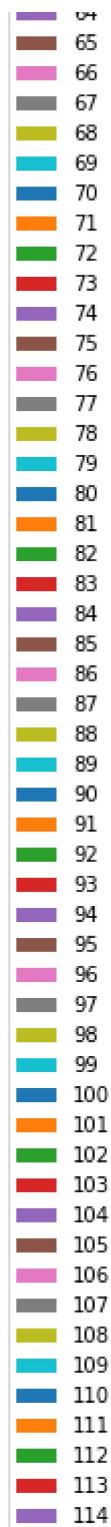


## Pie chart

```
In [28]: df.plot.pie(y='ANNUAL' )
```

```
Out[28]: <AxesSubplot:ylabel='ANNUAL'>
```





## Scatter chart

```
In [29]: df.plot.scatter(x='SUBDIVISION' ,y='ANNUAL')
```

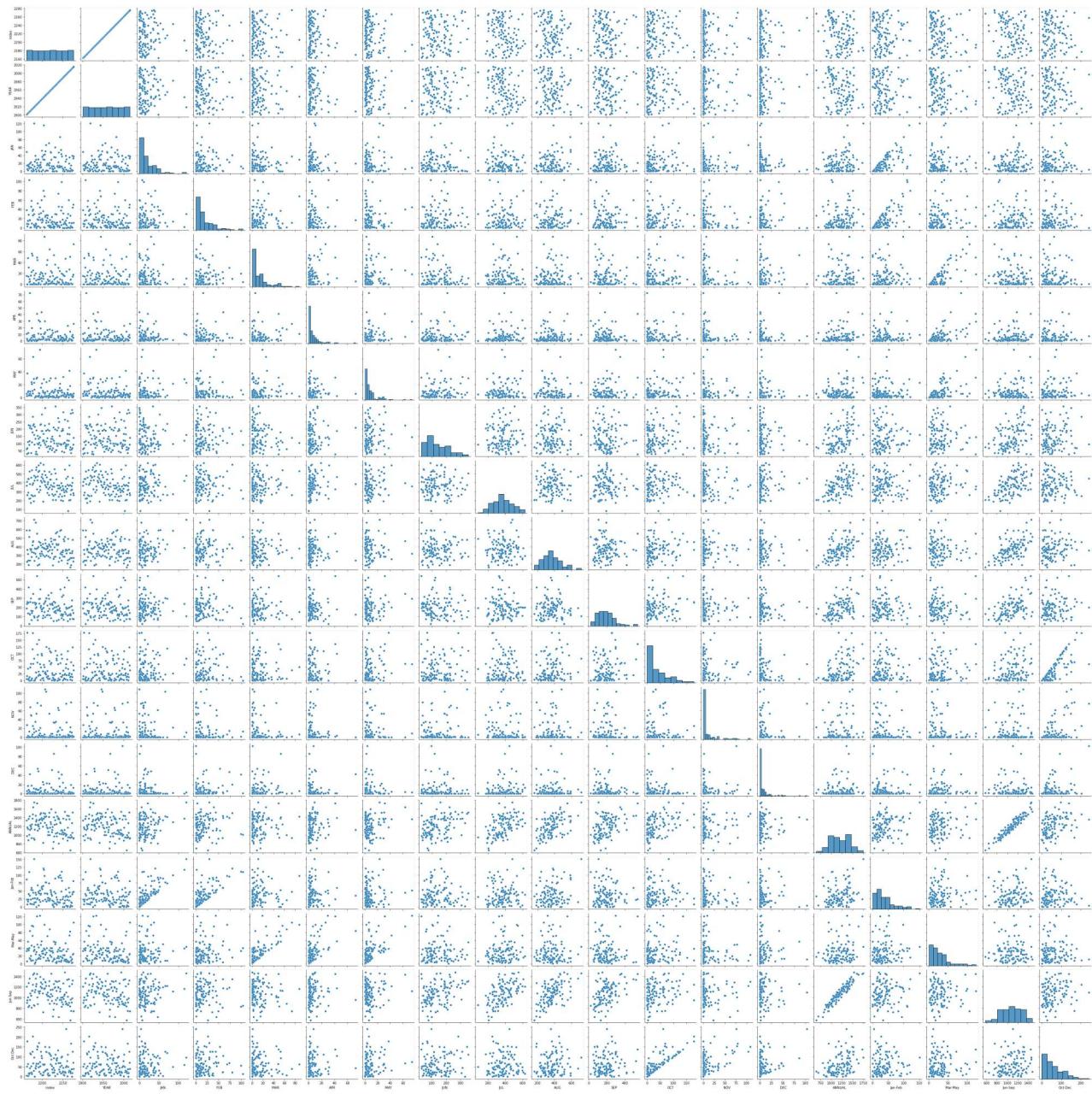
```
Out[29]: <AxesSubplot:xlabel='SUBDIVISION', ylabel='ANNUAL'>
```



## Seaborn

```
In [30]: sns.pairplot(df)
```

```
Out[30]: <seaborn.axisgrid.PairGrid at 0x22d89b4c610>
```

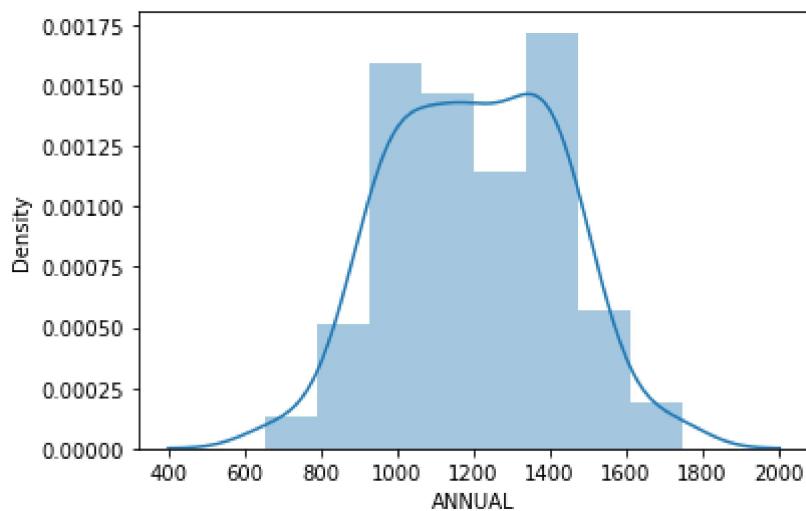


In [31]: `sns.distplot(df['ANNUAL'])`

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



```
In [32]: sns.heatmap(df.corr())
```

```
Out[32]: <AxesSubplot:
```

