

```
In [1]: import pandas as pd
```

```
In [2]: data=pd.read_csv("/home/placement/Downloads/arunachal.csv")
```

```
In [5]: data1=pd.read_csv("/home/placement/Downloads/rainfall in india 1901-2015.csv")
```

```
In [6]: data.head()
```

```
Out[6]:
```

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	NaN	629.7	333.9	NaN	NaN	NaN	NaN	117.9	811.8
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	NaN	269.6	107.9	823.8	909.1	628.4	411.5	199.3	63.5	0.0	NaN	185.9	NaN
2	112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692.6	515.8	125.2	7.8	13.7	5486.3	21.4	1196.9
3	113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286.7	628.7	948.3	40.7	8.6	4693.9	102.3	706.0
4	114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896.5	376.7	103.3	0.0	0.0	4106.7	210.3	1143.9

In [7]: data1.head()

Out[7]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
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	558.2	33.6	3373.2	136.3	560.3	1696.3	980.3
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	359.0	160.5	3520.7	159.8	458.3	2185.9	716.7
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	284.4	225.0	2957.4	156.7	236.1	1874.0	690.6
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	308.7	40.1	3079.6	24.1	506.9	1977.6	571.0
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	25.4	344.7	2566.7	1.3	309.7	1624.9	630.8

```
In [8]: list(data)
```

```
Out[8]: ['Unnamed: 0',  
        'SUBDIVISION',  
        'YEAR',  
        'JAN',  
        'FEB',  
        'MAR',  
        'APR',  
        'MAY',  
        'JUN',  
        'JUL',  
        'AUG',  
        'SEP',  
        'OCT',  
        'NOV',  
        'DEC',  
        'ANNUAL',  
        'Jan-Feb',  
        'Mar-May',  
        'Jun-Sep',  
        'Oct-Dec']
```

```
In [9]: list(data1)
```

```
Out[9]: ['SUBDIVISION',  
         'YEAR',  
         'JAN',  
         'FEB',  
         'MAR',  
         'APR',  
         'MAY',  
         'JUN',  
         'JUL',  
         'AUG',  
         'SEP',  
         'OCT',  
         'NOV',  
         'DEC',  
         'ANNUAL',  
         'Jan-Feb',  
         'Mar-May',  
         'Jun-Sep',  
         'Oct-Dec']
```

```
In [14]: data1=data.loc[(data.YEAR<2010)] #data till year 2010, can change here
```

In [15]: data1.tail()

Out[15]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	JUN
86	196	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	400.0	243.6	139.3	28.6	3.3	2335.5	216.0	604.6	134
87	197	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	227.6	263.2	77.2	69.7	21.7	2259.6	109.7	587.7	139
88	198	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	364.6	529.3	102.6	24.3	6.9	3020.7	110.8	590.9	218
89	199	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	439.3	189.7	115.1	1.7	2.6	2244.4	116.4	499.9	150
90	200	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	370.3	152.5	82.9	33.9	15.9	1749.9	110.8	394.7	111

In [17]:

Out[17]: array(['ARUNACHAL PRADESH'], dtype=object)

In [23]: data2=data1.loc[(data1,SUBDIVISION == 'ARUNACHAL PRADESH')]
data2

```
-----
NameError                                Traceback (most recent call last)
Cell In[23], line 1
----> 1 data2=data1.loc[(data1,SUBDIVISION == 'ARUNACHAL PRADESH')]
      2 data2

NameError: name 'SUBDIVISION' is not defined
```

```
In [24]: list(data1)
```

```
Out[24]: ['Unnamed: 0',  
          'SUBDIVISION',  
          'YEAR',  
          'JAN',  
          'FEB',  
          'MAR',  
          'APR',  
          'MAY',  
          'JUN',  
          'JUL',  
          'AUG',  
          'SEP',  
          'OCT',  
          'NOV',  
          'DEC',  
          'ANNUAL',  
          'Jan-Feb',  
          'Mar-May',  
          'Jun-Sep',  
          'Oct-Dec']
```

```
In [25]: data1['SUBDIVISION'].unique()
```

```
Out[25]: array(['ARUNACHAL PRADESH'], dtype=object)
```

```
In [26]: data2=data1.loc[(data1,SUBDIVISION == 'ARUNACHAL PRADESH')]  
data2
```

```
-----  
NameError                                Traceback (most recent call last)  
Cell In[26], line 1  
----> 1 data2=data1.loc[(data1,SUBDIVISION == 'ARUNACHAL PRADESH')]  
      2 data2  
  
NameError: name 'SUBDIVISION' is not defined
```

```
In [27]: data1.isna().sum()
```

```
Out[27]: Unnamed: 0      0  
SUBDIVISION      0  
YEAR             0  
JAN              1  
FEB              1  
MAR              2  
APR              0  
MAY              0  
JUN              1  
JUL              1  
AUG              0  
SEP              0  
OCT              2  
NOV              2  
DEC              2  
ANNUAL           6  
Jan-Feb          1  
Mar-May          2  
Jun-Sep          2  
Oct-Dec          3  
dtype: int64
```

```
In [28]: data1['SUBDIVISION'].unique()
```

```
Out[28]: array(['ARUNACHAL PRADESH'], dtype=object)
```

```
In [29]: data['SUBDIVISION'].unique()
```

```
Out[29]: array(['ARUNACHAL PRADESH'], dtype=object)
```

```
In [ ]:
```

