

9. arunachal data 17june23

#saimohan

In [21]: `import pandas as pd`

In [22]: `data=pd.read_csv('/home/placement/Desktop/saimohan/csv files/arunachal.csv')`

In [23]: `data.describe()`

Out[23]:

	Unnamed: 0	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
count	91.00000	91.000000	90.000000	90.000000	89.000000	91.000000	91.000000	90.000000	90.000000	91.000000	91.000000
mean	155.00000	1962.747253	48.598889	93.966667	154.446067	262.990110	364.651648	659.556667	711.963333	502.163736	433.273626
std	26.41338	27.695003	34.687078	46.258375	87.918484	113.395773	181.095447	311.642230	356.372598	275.716730	204.991358
min	110.00000	1916.000000	1.800000	6.100000	28.500000	94.700000	101.800000	239.400000	233.000000	172.400000	152.500000
25%	132.50000	1938.500000	20.075000	65.625000	101.700000	180.600000	237.150000	425.675000	442.150000	301.100000	282.150000
50%	155.00000	1964.000000	45.400000	87.600000	141.700000	245.400000	314.600000	545.750000	613.000000	411.600000	384.300000
75%	177.50000	1986.500000	65.150000	120.400000	189.600000	335.300000	447.050000	840.400000	922.075000	669.200000	521.150000
max	200.00000	2009.000000	164.500000	208.500000	605.600000	595.100000	1168.600000	1609.900000	2362.800000	1664.600000	1222.000000

In [24]: data.head()

Out[24]:

	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 [25]: list(data)
```

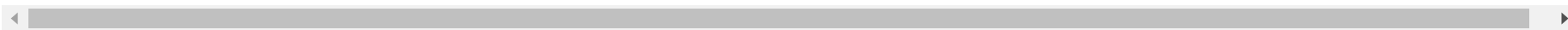
```
Out[25]: ['Unnamed: 0',  
          'SUBDIVISION',  
          'YEAR',  
          'JAN',  
          'FEB',  
          'MAR',  
          'APR',  
          'MAY',  
          'JUN',  
          'JUL',  
          'AUG',  
          'SEP',  
          'OCT',  
          'NOV',  
          'DEC',  
          'ANNUAL',  
          'Jan-Feb',  
          'Mar-May',  
          'Jun-Sep',  
          'Oct-Dec']
```

In [26]: *#we can drop the column by using below command.....*
`data1=data.drop(['Unnamed: 0'],axis=1)`
`data1`

Out[26]:

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
0	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	NaN	NaN
1	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	2772.8	262.
2	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	4121.3	146.
3	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	2888.0	997.
4	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	2649.2	103.
...
86	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	1343.7	171.
87	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	1393.5	168.
88	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	2185.1	133.
89	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	1508.7	119.
90	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	1111.8	132.

91 rows × 19 columns



```
In [27]: data.shape
```

```
Out[27]: (91, 20)
```

```
In [28]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 91 entries, 0 to 90
Data columns (total 20 columns):
 #   Column          Non-Null Count  Dtype  
---  -
 0   Unnamed: 0      91 non-null    int64  
 1   SUBDIVISION     91 non-null    object  
 2   YEAR            91 non-null    int64  
 3   JAN             90 non-null    float64 
 4   FEB             90 non-null    float64 
 5   MAR             89 non-null    float64 
 6   APR             91 non-null    float64 
 7   MAY             91 non-null    float64 
 8   JUN             90 non-null    float64 
 9   JUL             90 non-null    float64 
10  AUG             91 non-null    float64 
11  SEP             91 non-null    float64 
12  OCT             89 non-null    float64 
13  NOV             89 non-null    float64 
14  DEC             89 non-null    float64 
15  ANNUAL          85 non-null    float64 
16  Jan-Feb        90 non-null    float64 
17  Mar-May        89 non-null    float64 
18  Jun-Sep        89 non-null    float64 
19  Oct-Dec        88 non-null    float64 
dtypes: float64(17), int64(2), object(1)
memory usage: 14.3+ KB
```

```
In [29]: data.isna().sum()
```

```
Out[29]: 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
```

MODE

10,11,10,19,18,2,4,5,1,10,3,10

mode is 10

max repeated value

MEAN

sum of terms ny n

eg:10,10,10,10,20=[60/5] =12

MEDIAN

-->put in accending order

-->odd-----middle no is median

-->even----- $(a+b)/2$

eg: 13,14,18,23,40,50,51

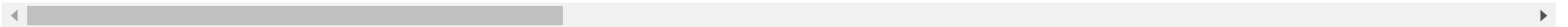
median is 23

```
In [33]: data=data.fillna(data.mean)
data
```

Out[33]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	A
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	<bound method NDFrame._add_numeric_operations....	62
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	<bound method NDFrame._add_numeric_operations....	269.6	107.9	823.8	909.1	62
2	112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	69
3	113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	28
4	114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	89
...
86	196	ARUNACHAL PRADESH	2005	48.4	167.6	229.5	195.3	179.8	269.3	430.8	40
87	197	ARUNACHAL PRADESH	2006	6.0	103.7	63.3	202.7	321.7	520.4	382.2	22
88	198	ARUNACHAL PRADESH	2007	13.4	97.4	48.1	292.4	250.4	530.2	761.0	36
89	199	ARUNACHAL PRADESH	2008	76.7	39.7	122.6	192.4	185.0	423.6	456.1	43
90	200	ARUNACHAL PRADESH	2009	18.0	92.8	72.1	132.7	189.9	259.1	329.9	37

91 rows × 20 columns



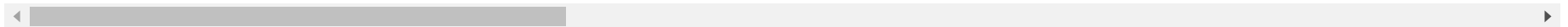

```
In [32]: data.isna().sum()
```

```
Out[32]: Unnamed: 0      0  
SUBDIVISION      0  
YEAR             0  
JAN              0  
FEB              0  
MAR              0  
APR              0  
MAY              0  
JUN              0  
JUL              0  
AUG              0  
SEP              0  
OCT              0  
NOV              0  
DEC              0  
ANNUAL           0  
Jan-Feb          0  
Mar-May          0  
Jun-Sep          0  
Oct-Dec          0  
dtype: int64
```

```
In [42]: saimohan=10
data.head(saimohan)
```

Out[42]:

	Unnamed: 0	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AU
0	110	ARUNACHAL PRADESH	1916	48.1	69.8	71.1	316.1	424.6	1124.9	<bound method NDFrame._add_numeric_operations....	629
1	111	ARUNACHAL PRADESH	1917	21.4	164.5	<bound method NDFrame._add_numeric_operations....	269.6	107.9	823.8	909.1	628
2	112	ARUNACHAL PRADESH	1918	10.4	11.0	191.2	144.6	861.1	1609.9	1303.0	692
3	113	ARUNACHAL PRADESH	1919	34.5	67.8	28.5	256.9	420.6	973.6	999.0	286
4	114	ARUNACHAL PRADESH	1920	14.0	196.3	605.6	364.7	173.6	840.6	535.4	896
5	115	ARUNACHAL PRADESH	1921	78.9	54.3	180.3	358.0	598.0	1233.2	1433.0	885
6	116	ARUNACHAL PRADESH	1922	50.7	59.4	170.4	299.5	350.5	1109.3	918.7	488
7	117	ARUNACHAL PRADESH	1923	9.4	160.8	34.0	240.9	445.4	408.6	1278.5	251
8	118	ARUNACHAL PRADESH	1924	85.7	45.1	74.1	162.4	515.7	1165.0	942.7	713
9	119	ARUNACHAL PRADESH	1925	80.6	114.0	143.3	223.0	587.2	611.6	611.0	684



In []:

