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

In [2]:
```
df=pd.read_csv("cleaned_rainfall")
df
```

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

In [3]:
```
df["SUBDIVISION"].value_counts()
```

Out[3]:
```
KONKAN & GOA                          115
EAST RAJASTHAN                        115
WEST MADHYA PRADESH                   115
SAURASHTRA & KUTCH                    115
EAST UTTAR PRADESH                    115
UTTARAKHAND                           115
WEST UTTAR PRADESH                    115
SOUTH INTERIOR KARNATAKA              115
COASTAL ANDHRA PRADESH                115
BIHAR                                 115
JAMMU & KASHMIR                       115
KERALA                                115
GUJARAT REGION                        115
ORISSA                                115
VIDARBHA                              115
PUNJAB                                115
NORTH INTERIOR KARNATAKA              115
MATATHWADA                            115
HIMACHAL PRADESH                      115
CHHATTISGARH                          115
COASTAL KARNATAKA                     115
NAGA MANI MIZO TRIPURA                115
GANGETIC WEST BENGAL                  115
TELANGANA                             115
HARYANA DELHI & CHANDIGARH            115
MADHYA MAHARASHTRA                    115
TAMIL NADU                            115
WEST RAJASTHAN                        115
ASSAM & MEGHALAYA                     115
SUB HIMALAYAN WEST BENGAL & SIKKIM    115
EAST MADHYA PRADESH                   115
JHARKHAND                             115
RAYALSEEMA                            115
LAKSHADWEEP                           114
ANDAMAN & NICOBAR ISLANDS             110
ARUNACHAL PRADESH                      97
Name: SUBDIVISION, dtype: int64
```

MATATHWADA

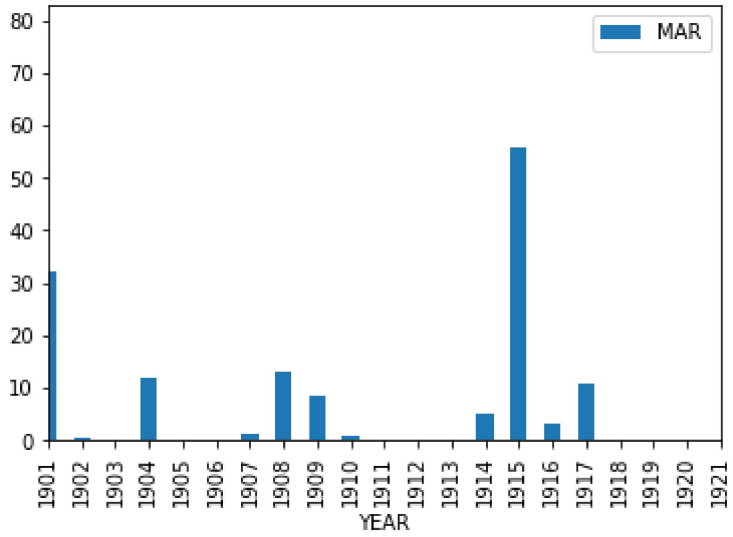
In [4]:
```
dat1=df[df["SUBDIVISION"]=="MATATHWADA"]
dat1
```

Out[4]:

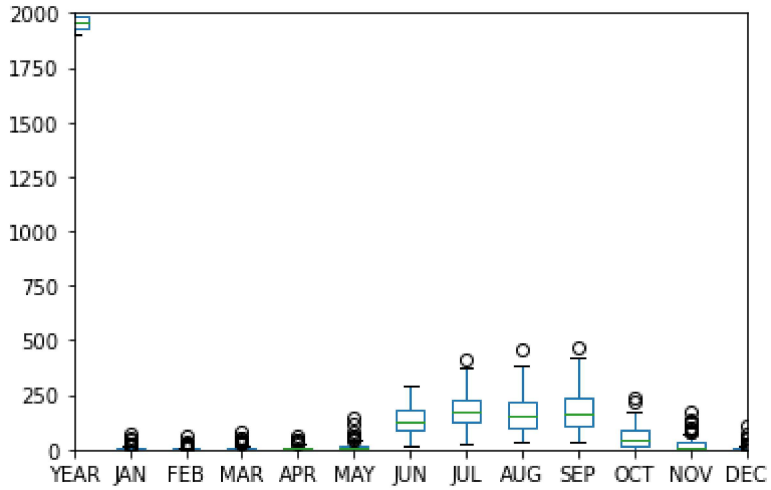
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
2737	2737	MATATHWADA	1901	15.8	3.3	32.1	48.5	26.5	193.1	184.1	249.8	74.0	81
2738	2738	MATATHWADA	1902	1.3	0.0	0.4	7.2	0.8	52.4	120.9	85.2	273.3	61
2739	2739	MATATHWADA	1903	2.6	0.8	0.0	1.7	58.3	104.4	264.2	281.9	173.3	139
2740	2740	MATATHWADA	1904	0.0	0.9	12.1	0.3	7.2	79.2	118.4	57.3	339.0	76
2741	2741	MATATHWADA	1905	1.3	2.0	0.0	6.6	4.8	84.6	94.8	137.6	157.8	15
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
2847	2847	MATATHWADA	2011	0.0	3.8	0.7	3.5	3.1	79.2	230.1	228.5	90.0	24
2848	2848	MATATHWADA	2012	0.0	0.0	0.0	0.6	2.3	72.2	161.1	101.4	120.0	68
2849	2849	MATATHWADA	2013	1.5	9.4	2.6	7.9	6.4	160.9	293.4	136.9	154.1	94
2850	2850	MATATHWADA	2014	1.4	13.4	79.0	11.9	7.0	30.4	105.0	178.9	84.5	14
2851	2851	MATATHWADA	2015	10.1	1.6	32.0	39.6	12.3	118.3	27.4	112.2	154.3	19

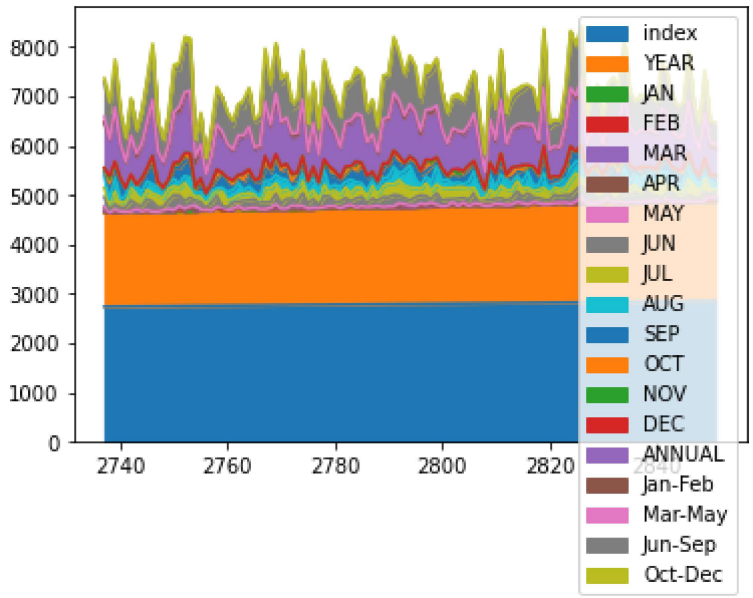
115 rows × 20 columns

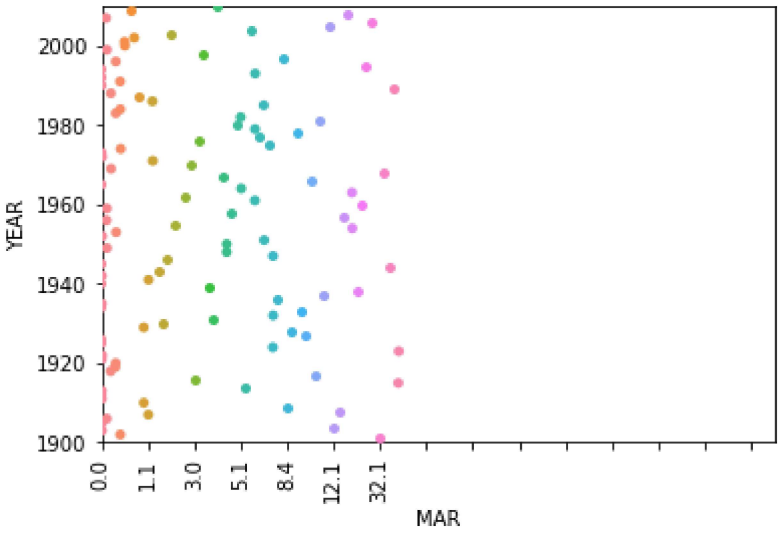
In [5]:
```
dat1.plot.bar("YEAR","MAR")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat1.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat1.plot.area()
dat1.plot.scatter("YEAR","MAR")
sns.stripplot(x=dat1["MAR"],y=dat1["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat1["MAR"],rotation="vertical")
plt.gca().xaxis.set_major_locator(tic.MultipleLocator(base=10))
plt.show()
dat1.plot.hist()
```



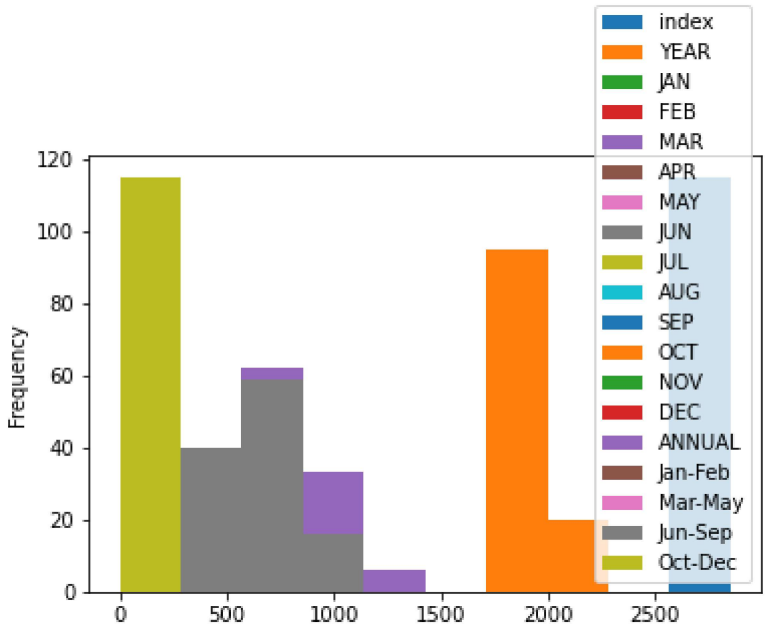

```
<Figure size 4320x2160 with 0 Axes>
```



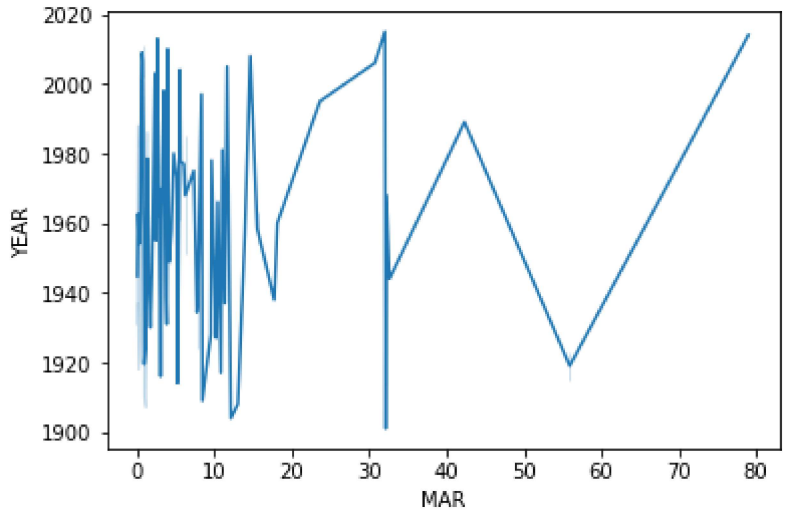




Out[5]: <AxesSubplot:ylabel='Frequency'>



In [6]:
```
sns.lineplot(x=dat1["MAR"],y=dat1["YEAR"])
plt.show()
```



WEST MADHYA PRADESH

In [7]:
```
dat2=df[df["SUBDIVISION"]=="WEST MADHYA PRADESH"]
dat2
```

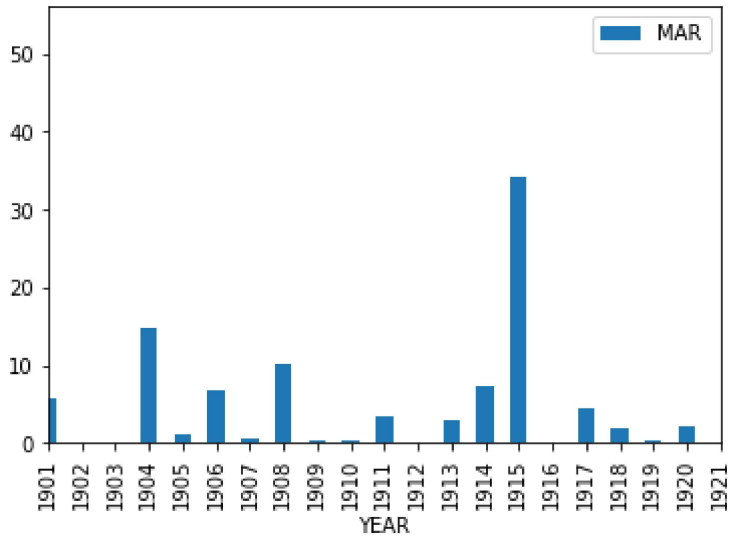
Out[7]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
2047	2047	WEST MADHYA PRADESH	1901	25.8	5.8	5.8	2.8	2.1	41.2	228.9	349.9	47.9	5.6
2048	2048	WEST MADHYA PRADESH	1902	22.1	8.4	0.0	2.0	5.9	35.9	401.9	179.4	194.1	37.9
2049	2049	WEST MADHYA PRADESH	1903	5.3	0.0	0.0	0.0	22.3	50.6	304.9	261.1	250.2	55.1
2050	2050	WEST MADHYA PRADESH	1904	3.2	15.5	14.8	0.0	12.0	96.6	273.0	218.6	125.9	3.3
2051	2051	WEST MADHYA PRADESH	1905	3.5	4.4	1.1	0.8	3.0	36.1	326.3	137.6	183.5	0.3
...	...	...	...	...	...	...	...	...	...	...	...	...	...
2157	2157	WEST MADHYA PRADESH	2011	0.0	1.7	0.1	1.8	3.6	241.5	306.7	343.3	165.0	0.2
2158	2158	WEST MADHYA PRADESH	2012	6.2	0.0	0.0	0.9	3.1	48.2	439.2	341.2	194.3	2.1
2159	2159	WEST MADHYA PRADESH	2013	1.7	31.1	8.5	2.8	0.4	263.7	485.1	432.6	98.9	68.7
2160	2160	WEST MADHYA PRADESH	2014	25.6	34.4	4.6	1.4	1.4	30.6	337.4	211.0	192.6	7.0
2161	2161	WEST MADHYA PRADESH	2015	40.2	6.4	53.5	13.3	2.0	154.1	428.2	276.6	55.6	11.0

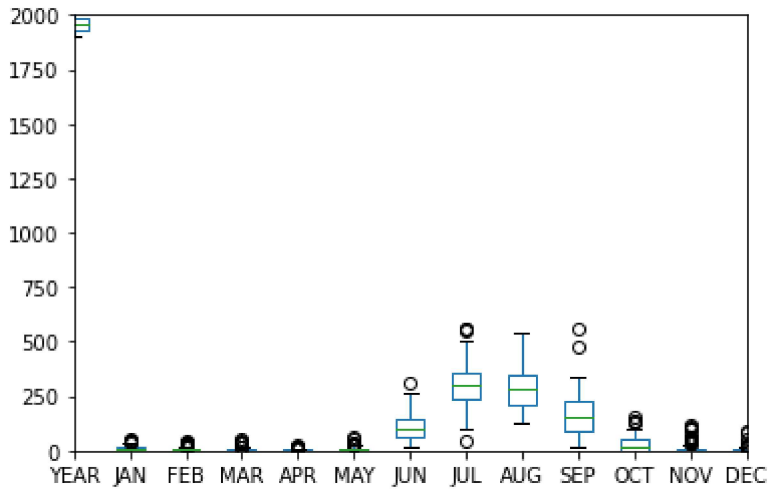
115 rows × 20 columns

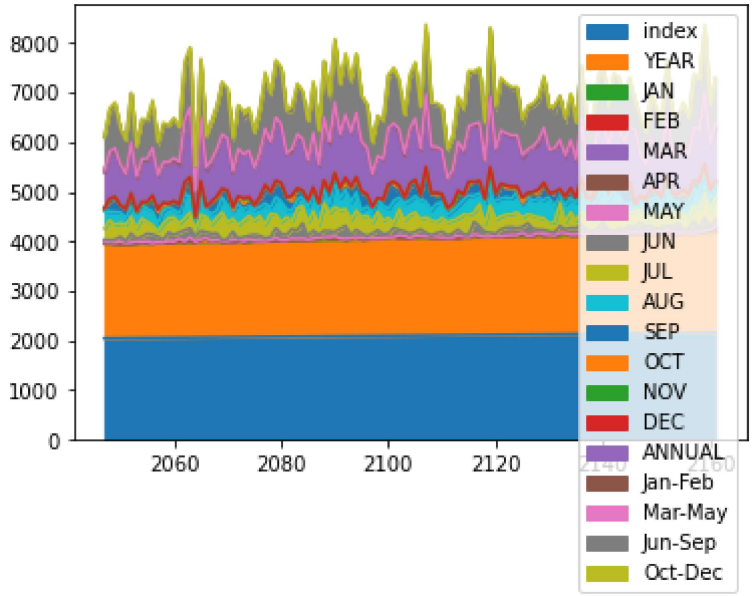
In [10]:

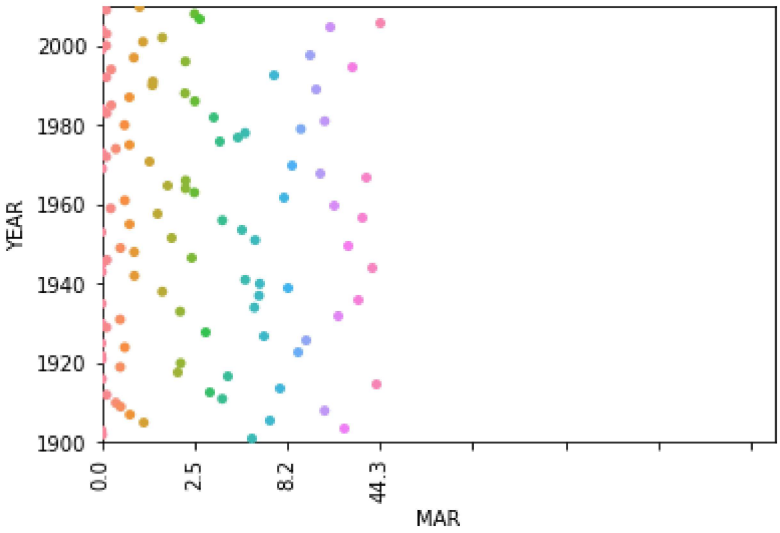
```
dat2.plot.bar("YEAR","MAR")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat2.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat2.plot.area()
dat2.plot.scatter("YEAR","MAR")
sns.stripplot(x=dat2["MAR"],y=dat2["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat2["MAR"],rotation="vertical")
plt.gca().xaxis.set_major_locator(tic.MultipleLocator(base=20))
plt.show()
dat2.plot.hist()
plt.show()
```

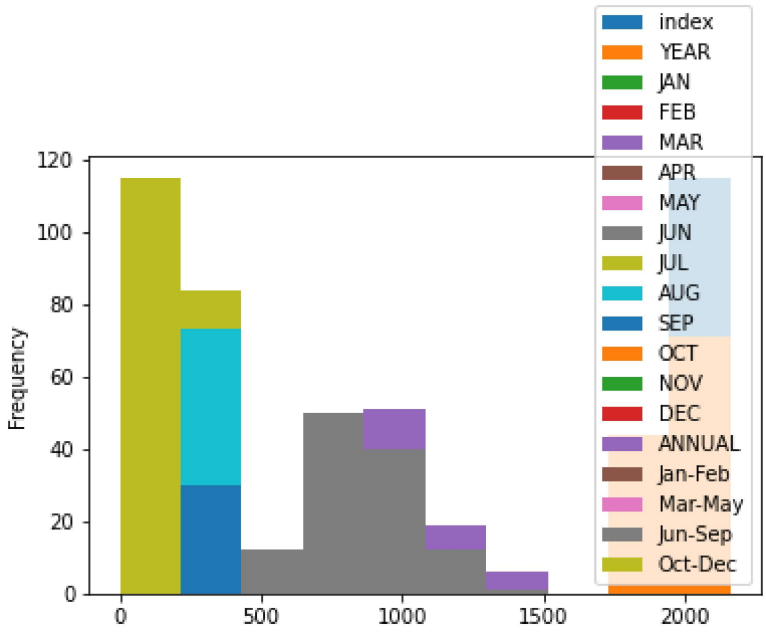


<Figure size 4320x2160 with 0 Axes>




8000
7000
6000
5000
4000
3000
2000
1000
0
2060
2080
2100
2120
index
YEAR
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
ANNUAL
Jan-Feb
Mar-May
Jun-Sep
Oct-Dec


2000
1980
1960
1940
1920
1900
YEAR
0.0
2.5
8.2
44.3
MAR


120
100
80
60
40
20
0
Frequency
0
500
1000
1500
2000
index
YEAR
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
ANNUAL
Jan-Feb
Mar-May
Jun-Sep
Oct-Dec

In [9]:
```
sns.lineplot(x=dat2["MAR"],y=dat2["YEAR"])
plt.show()
```