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	OCT
0	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
1	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
2	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
3	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
4	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...
4111	4111	LAKSHADWEEP	2011	5.1	2.8	3.1	85.9	107.2	153.6	350.2	254.0	255.2	117.4
4112	4112	LAKSHADWEEP	2012	19.2	0.1	1.6	76.8	21.2	327.0	231.5	381.2	179.8	145.9
4113	4113	LAKSHADWEEP	2013	26.2	34.4	37.5	5.3	88.3	426.2	296.4	154.4	180.0	72.8
4114	4114	LAKSHADWEEP	2014	53.2	16.1	4.4	14.9	57.4	244.1	116.1	466.1	132.2	169.2
4115	4115	LAKSHADWEEP	2015	2.2	0.5	3.7	87.1	133.1	296.6	257.5	146.4	160.4	165.4

4116 rows × 20 columns

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

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

RAYALSEEMA

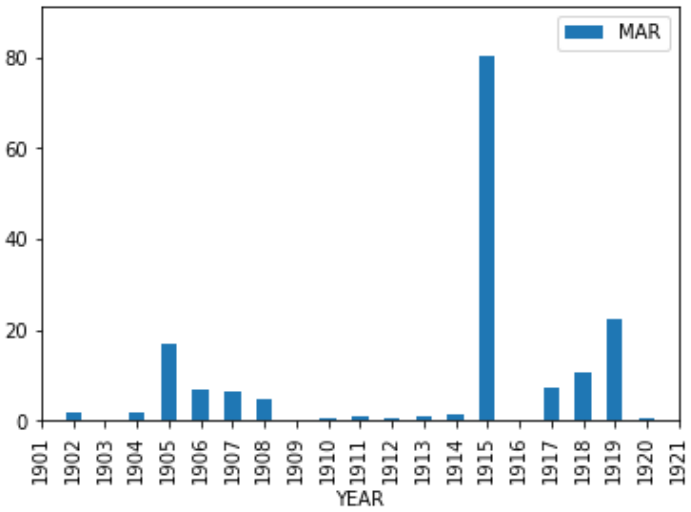
In [4]:
```
dat1=df[df["SUBDIVISION"]=="RAYALSEEMA"]
dat1
```

Out[4]:

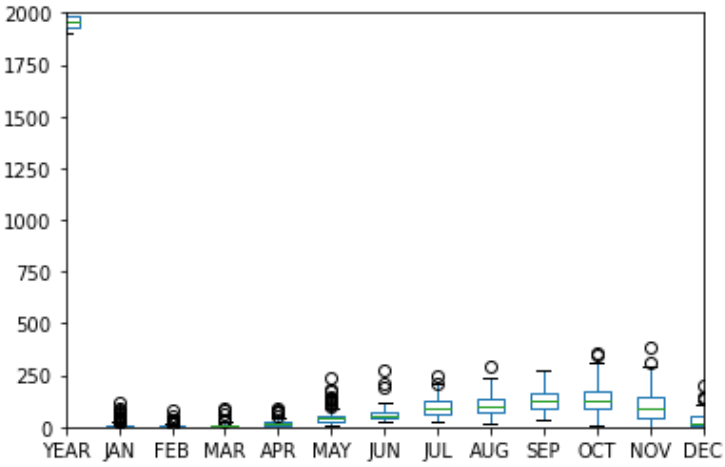
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
3312	3312	RAYALSEEMA	1901	7.0	50.2	0.0	12.1	38.9	53.0	73.4	60.3	109.0	81.6	137.2
3313	3313	RAYALSEEMA	1902	10.0	0.2	1.7	11.0	36.8	73.6	41.3	148.3	181.7	188.5	88.9
3314	3314	RAYALSEEMA	1903	30.0	0.1	0.0	3.6	80.5	67.5	127.5	140.6	219.7	95.3	289.4
3315	3315	RAYALSEEMA	1904	14.8	0.0	1.7	7.1	58.8	39.8	75.1	19.4	84.7	111.5	4.4
3316	3316	RAYALSEEMA	1905	6.5	6.8	17.0	18.3	44.2	66.1	50.9	219.3	36.5	180.2	55.4
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3422	3422	RAYALSEEMA	2011	0.8	12.1	0.0	34.6	33.0	44.5	128.9	163.6	71.2	107.5	106.9
3423	3423	RAYALSEEMA	2012	2.7	0.0	2.5	32.7	38.8	47.0	139.7	120.0	69.5	113.7	86.6
3424	3424	RAYALSEEMA	2013	1.3	30.6	11.5	26.8	38.9	73.8	95.7	110.3	163.2	169.3	38.6
3425	3425	RAYALSEEMA	2014	0.2	0.7	12.5	5.1	46.7	66.3	68.7	115.1	81.4	104.6	37.8
3426	3426	RAYALSEEMA	2015	1.9	0.0	13.4	73.4	39.7	73.0	43.1	123.6	136.3	106.7	383.8

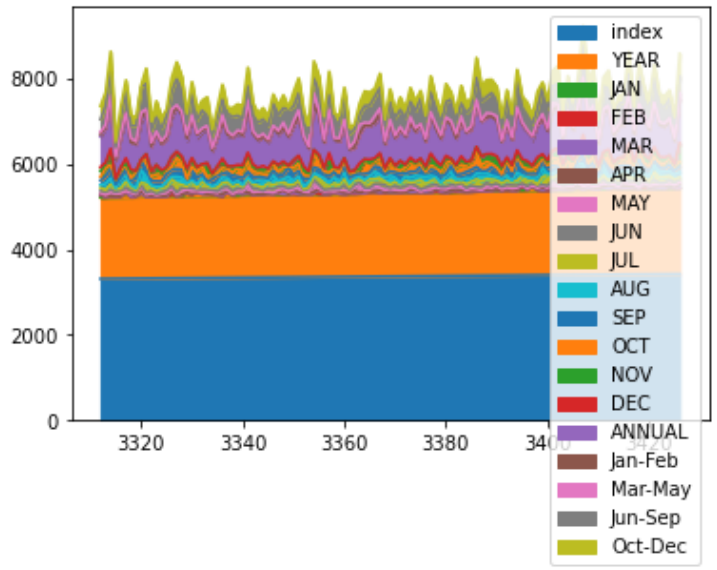
115 rows × 20 columns

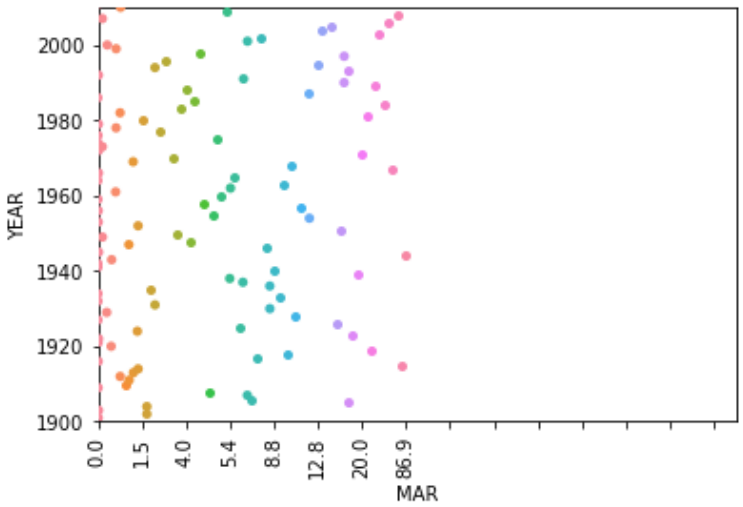
In [5]:
```
dat1.plot.bar("YEAR","MAR")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat1.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat1.plot.area()
dat1.plot.scatter("YEAR","MAR")
sns.stripplot(x=dat1["MAR"],y=dat1["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat1["MAR"],rotation="vertical")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



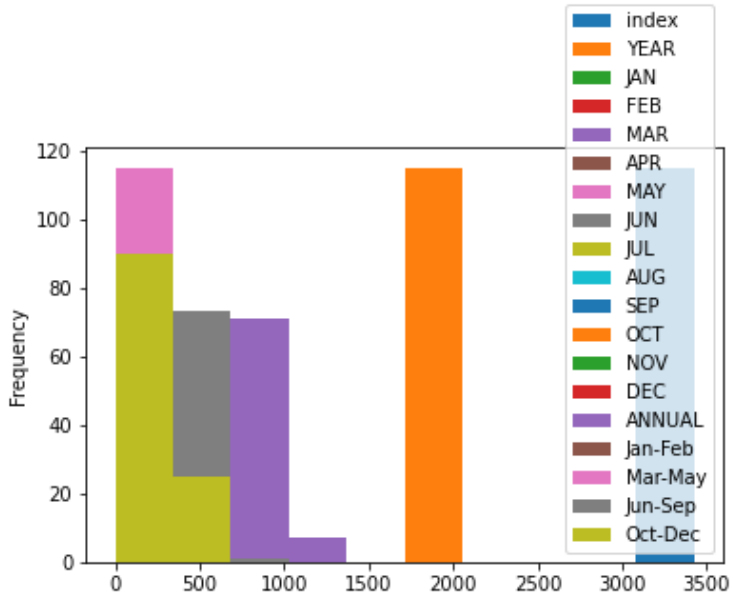

```
<Figure size 4320x2160 with 0 Axes>
```



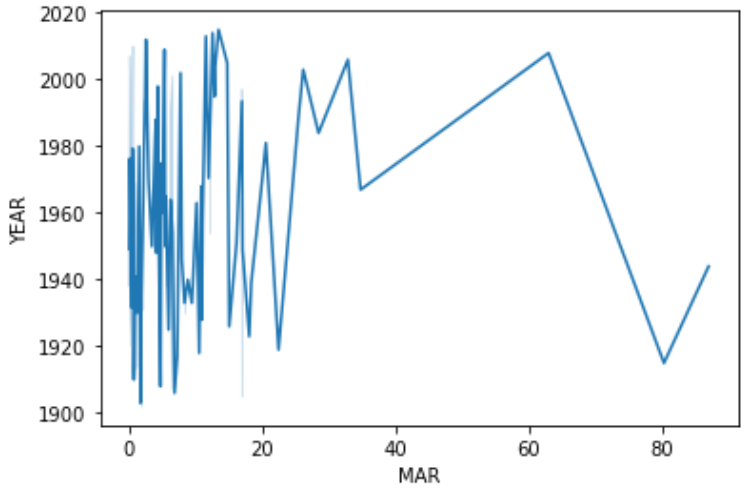




Out[5]: <AxesSubplot:ylabel='Frequency'>



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



CHHATTISGARH

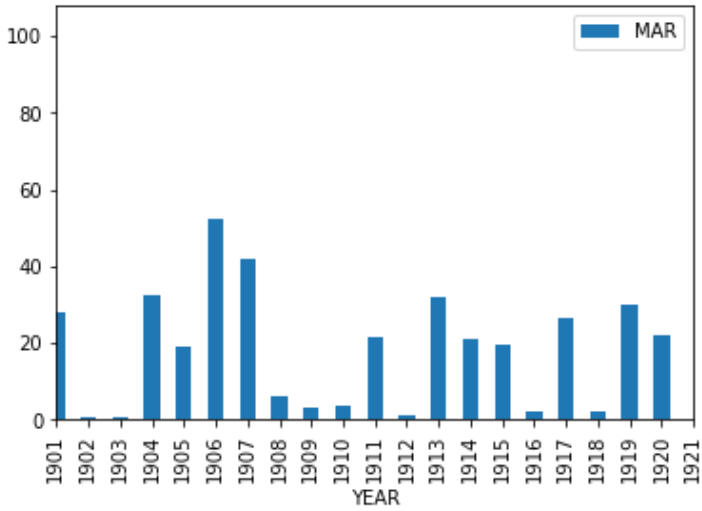
In [7]:
```
dat2=df[df["SUBDIVISION"]=="CHHATTISGARH"]
dat2
```

Out[7]:

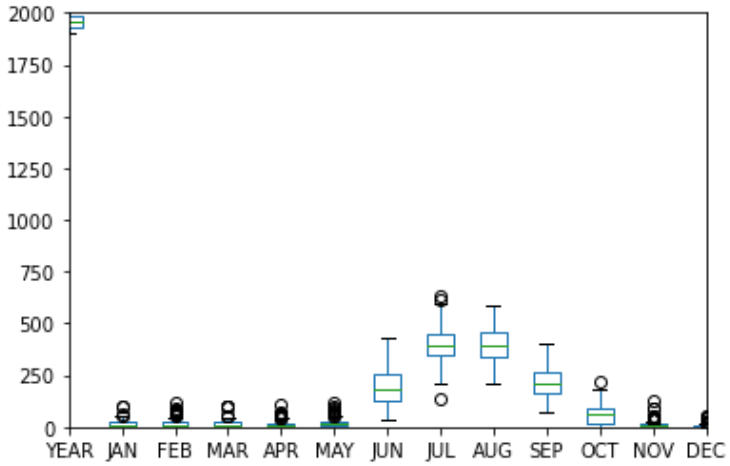
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
2967	2967	CHHATTISGARH	1901	48.9	116.5	27.8	5.5	18.4	101.6	381.0	476.7	182.8	27.3
2968	2968	CHHATTISGARH	1902	0.6	6.5	0.4	13.9	10.3	37.2	403.8	236.6	198.1	4.7
2969	2969	CHHATTISGARH	1903	6.2	13.9	0.4	6.8	51.1	110.7	365.9	396.0	212.0	168.0
2970	2970	CHHATTISGARH	1904	0.0	8.6	32.3	0.2	77.5	369.5	303.6	483.6	86.8	129.3
2971	2971	CHHATTISGARH	1905	50.3	22.6	19.0	24.6	31.8	40.4	443.7	270.8	338.8	8.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...
3077	3077	CHHATTISGARH	2011	0.3	11.5	2.6	35.0	16.8	183.5	272.6	379.8	382.2	15.5
3078	3078	CHHATTISGARH	2012	36.6	4.8	1.1	14.9	9.4	147.3	430.6	442.2	245.3	19.8
3079	3079	CHHATTISGARH	2013	2.8	19.7	4.9	45.8	5.7	263.6	418.8	336.6	140.9	180.9
3080	3080	CHHATTISGARH	2014	2.3	29.0	21.4	17.3	25.0	104.9	416.7	327.7	252.7	77.9
3081	3081	CHHATTISGARH	2015	15.8	1.2	21.2	37.0	13.0	257.6	248.6	286.6	216.9	17.7

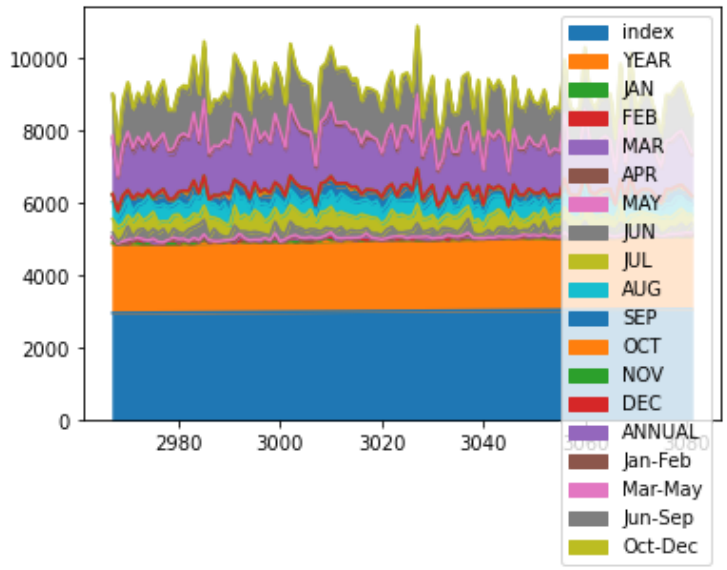
115 rows × 20 columns

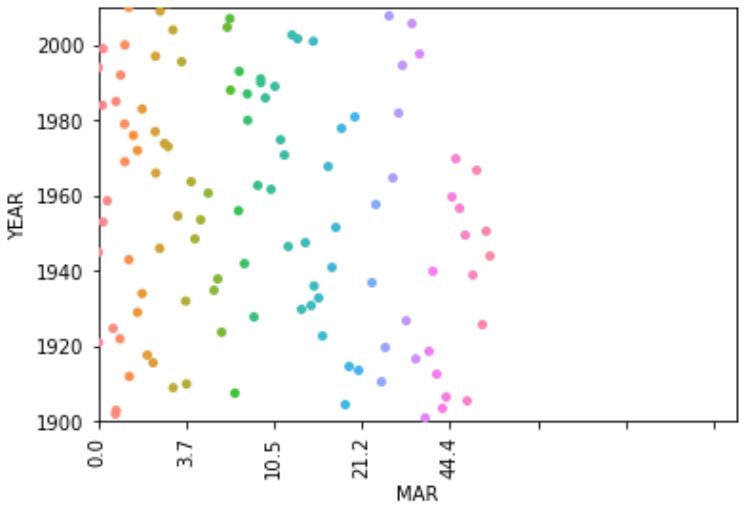
In [8]:
```
dat2.plot.bar("YEAR","MAR")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat2.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat2.plot.area()
dat2.plot.scatter("YEAR","MAR")
sns.stripplot(x=dat2["MAR"],y=dat2["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat2["MAR"],rotation="vertical")
plt.gca().xaxis.set_major_locator(tic.MultipleLocator(base=20))
plt.show()
dat2.plot.hist()
plt.show()
```

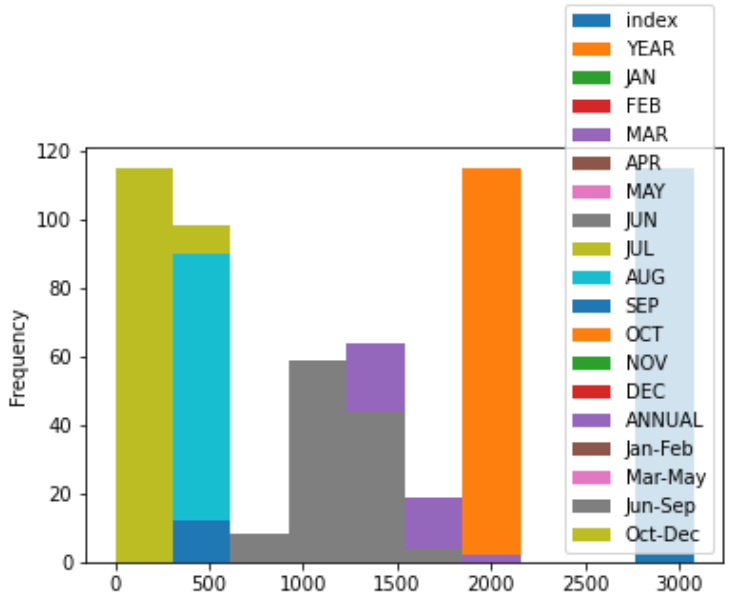



```
<Figure size 4320x2160 with 0 Axes>
```




index
YEAR
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
ANNUAL
Jan-Feb
Mar-May
Jun-Sep
Oct-Dec
10000
8000
6000
4000
2000
0
2980
3000
3020
3040


YEAR
2000
1980
1960
1940
1920
1900
0.0
3.7
10.5
21.2
44.4
MAR


index
YEAR
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
ANNUAL
Jan-Feb
Mar-May
Jun-Sep
Oct-Dec
Frequency
120
100
80
60
40
20
0
0
500
1000
1500
2000
2500
3000

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