

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

TAMIL NADU

```
In [4]: dat1=df[df["SUBDIVISION"]=="TAMIL NADU"]  
dat1
```

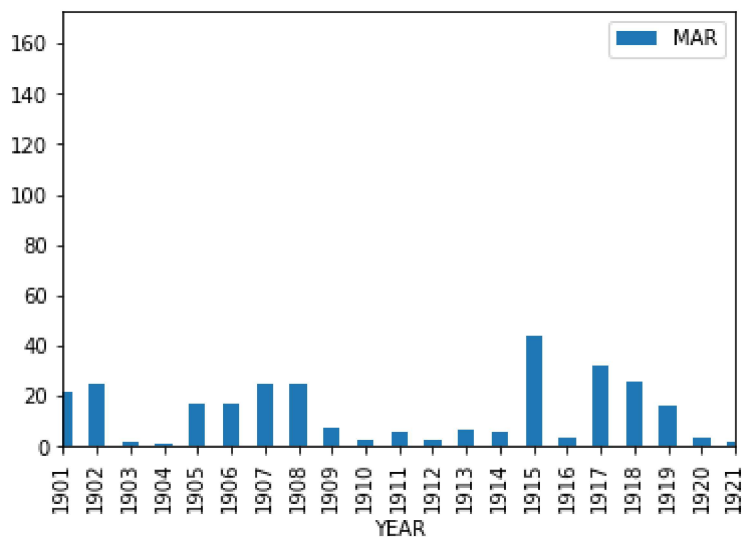
Out[4]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
3427	3427	TAMIL NADU	1901	24.5	39.1	21.7	36.0	74.0	41.8	49.3	67.9	191.1	122.3
3428	3428	TAMIL NADU	1902	67.2	9.8	25.1	21.9	84.7	39.3	55.1	113.8	98.6	282.2
3429	3429	TAMIL NADU	1903	19.3	7.8	1.7	18.2	128.5	58.5	72.6	115.0	210.4	128.1
3430	3430	TAMIL NADU	1904	35.2	0.1	0.7	19.5	121.9	34.9	89.0	40.4	85.7	163.2
3431	3431	TAMIL NADU	1905	6.5	7.5	17.2	64.8	83.7	49.8	39.0	101.8	73.5	250.4
...
3537	3537	TAMIL NADU	2011	4.3	11.2	8.0	91.5	33.4	56.0	45.5	128.9	76.0	200.4
3538	3538	TAMIL NADU	2012	3.0	0.1	2.5	35.5	41.9	30.1	46.5	98.0	84.9	235.2
3539	3539	TAMIL NADU	2013	3.9	30.9	30.0	20.3	42.0	54.6	42.7	110.7	113.5	127.9
3540	3540	TAMIL NADU	2014	7.4	6.1	8.1	8.3	139.1	47.8	50.6	117.7	98.9	252.2
3541	3541	TAMIL NADU	2015	8.3	2.3	21.7	108.8	112.4	62.4	43.5	81.6	98.4	132.6

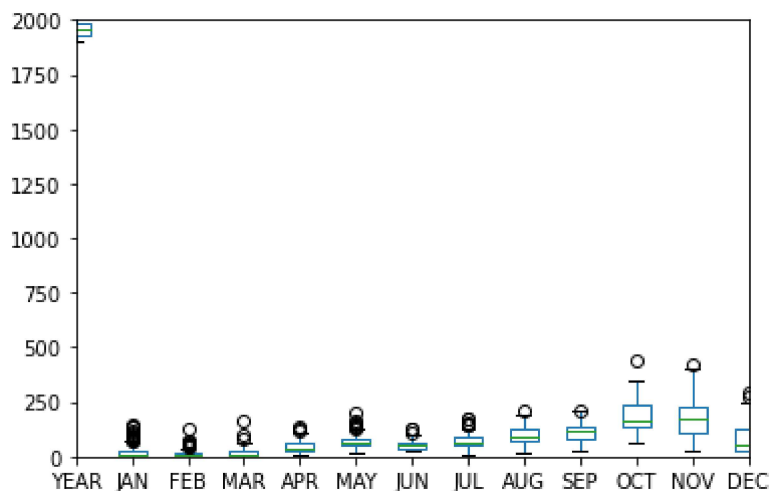
115 rows × 20 columns

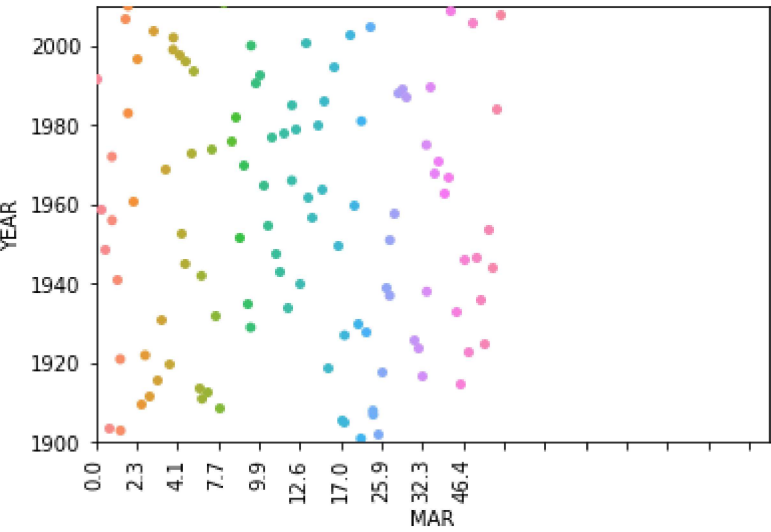
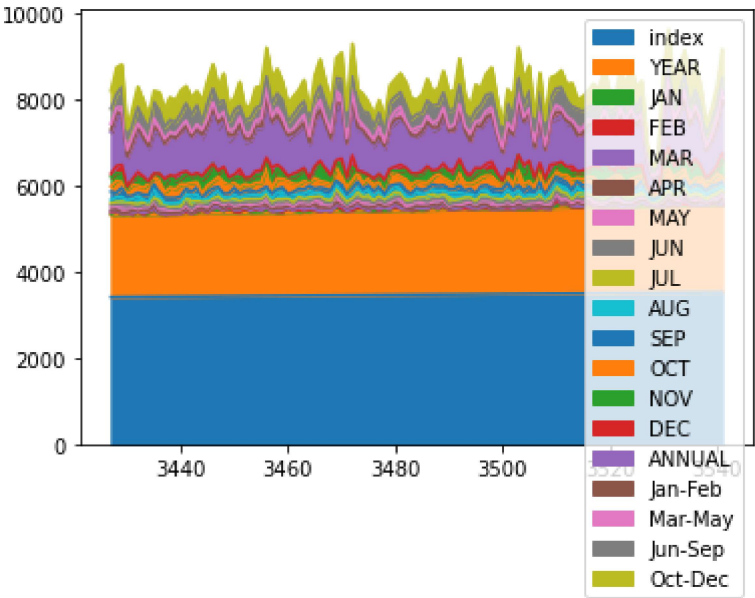


```
In [5]: dat1.plot.bar("YEAR", "MAR")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat1.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat1.plot.area()
dat1.plot.scatter("YEAR", "MAR")
sns.stripplot(x=dat1["MAR"],y=dat1["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat1["MAR"],rotation="vertical")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

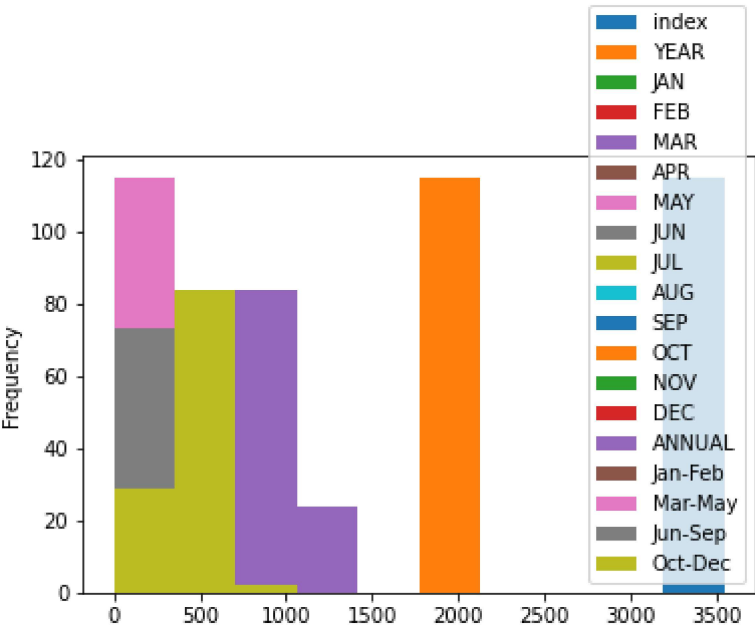


<Figure size 4320x2160 with 0 Axes>

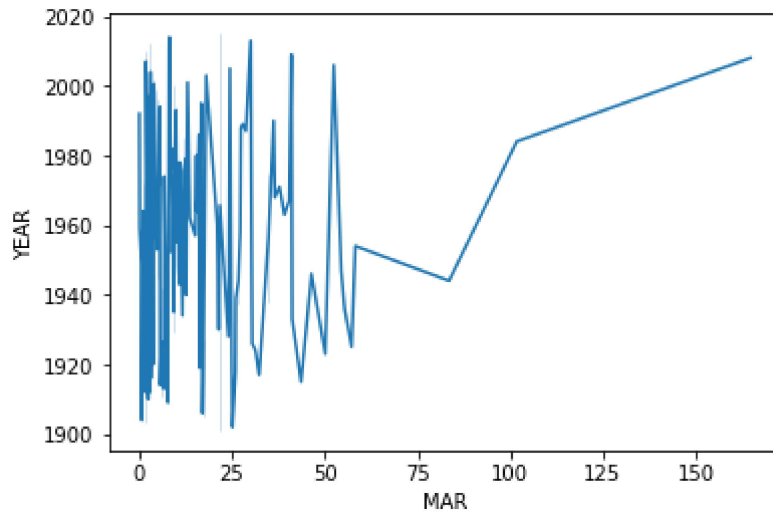




Out[5]: <AxesSubplot:ylabel='Frequency'>



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



ORISSA

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

Out[7]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
667	667	ORISSA	1901	39.5	65.1	16.1	51.6	79.0	78.2	288.4	307.7	185.3	76.
668	668	ORISSA	1902	3.4	0.2	14.2	101.1	56.7	108.3	437.4	349.1	202.7	33.
669	669	ORISSA	1903	19.7	18.9	10.5	34.6	73.3	154.3	410.4	295.2	265.6	228.
670	670	ORISSA	1904	0.2	12.2	20.6	10.1	100.2	342.9	336.7	350.4	227.8	111.
671	671	ORISSA	1905	24.3	17.2	66.3	56.9	107.5	92.0	330.1	281.4	344.1	36.
...
777	777	ORISSA	2011	3.7	16.2	4.9	58.2	75.6	210.1	199.6	358.6	398.7	20.
778	778	ORISSA	2012	50.8	3.6	0.9	34.8	21.3	169.6	324.3	417.0	242.4	66.
779	779	ORISSA	2013	3.3	7.8	2.1	53.6	57.7	272.6	380.0	254.9	208.1	391.
780	780	ORISSA	2014	0.0	17.6	25.1	11.7	111.9	92.2	496.2	386.3	281.1	111.
781	781	ORISSA	2015	15.1	3.3	10.5	67.6	32.6	238.6	294.8	264.0	237.0	24.

115 rows × 20 columns

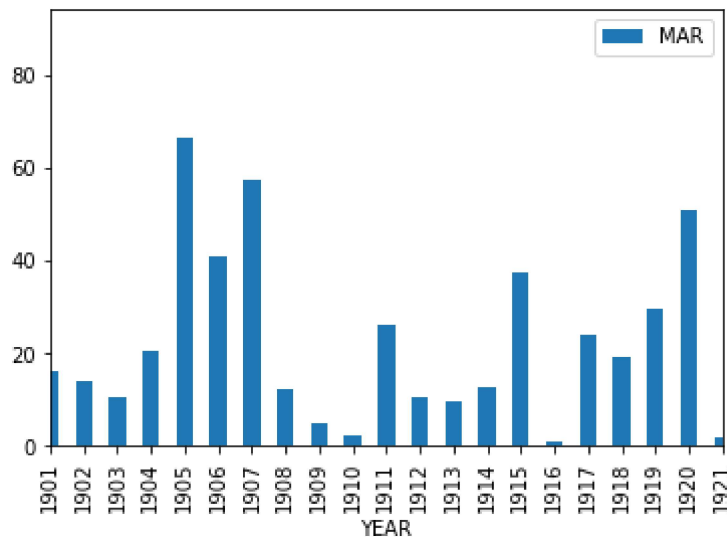


In [8]:

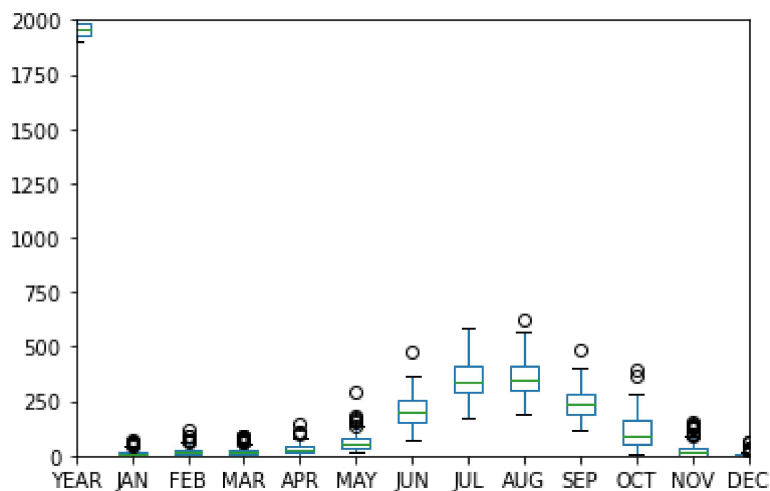
```

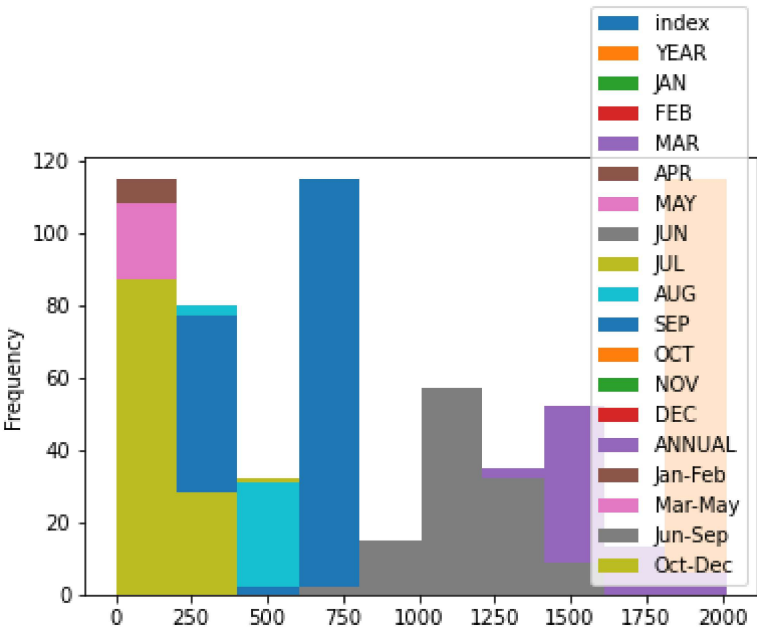
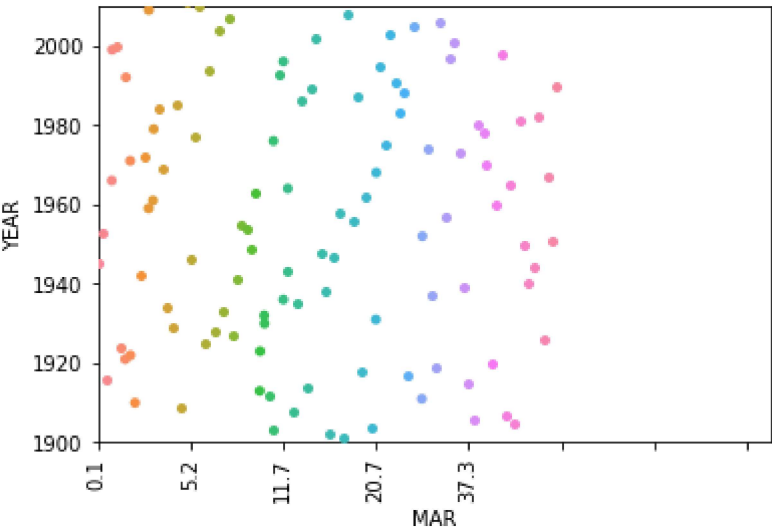
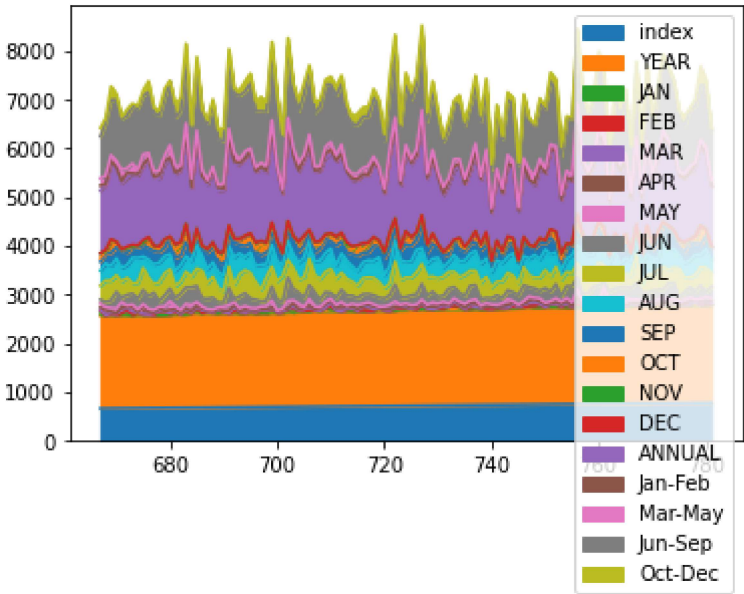
dat2.plot.bar("YEAR", "MAR")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat2.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat2.plot.area()
dat2.plot.scatter("YEAR", "MAR")
sns.stripplot(x=dat2["MAR"],y=dat2["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat2["MAR"],rotation="vertical")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

```



<Figure size 4320x2160 with 0 Axes>






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

