

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

HARYANA DELHI & CHANDIGARH

In [4]:

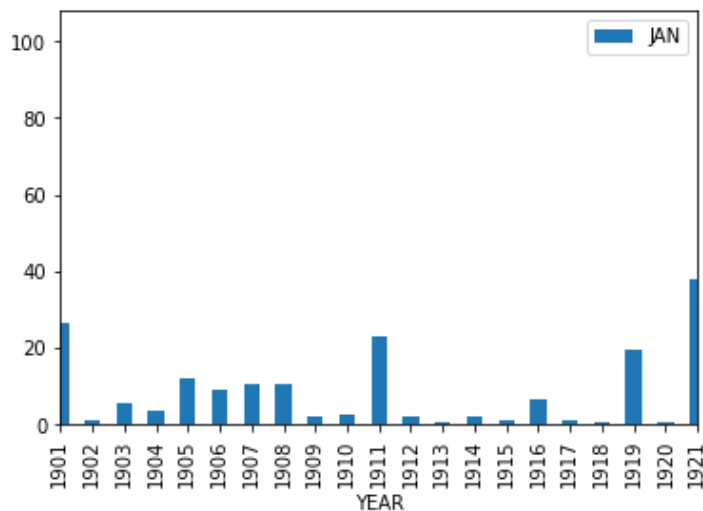
dat1=df[df["SUBDIVISION"]=="SUB HIMALAYAN WEST BENGAL & SIKKIM"]
dat1

Out[4]:

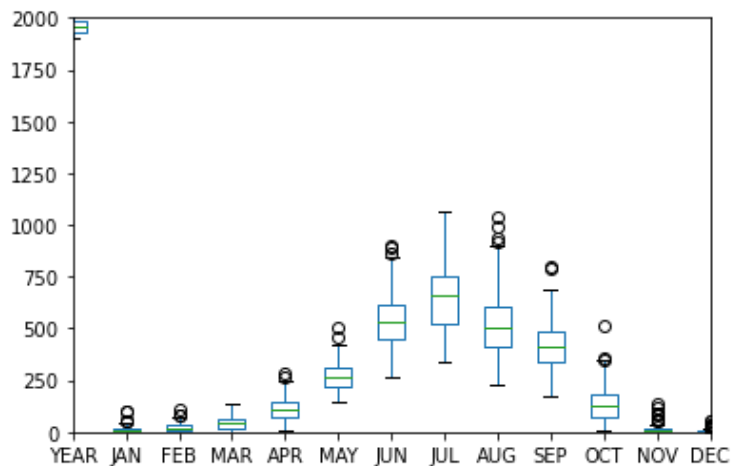
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
437	437	SUB HIMALAYAN WEST BENGAL & SIKKIM	1901	26.5	14.8	14.1	29.2	195.5	488.4	524.8	501.1	242.7	55.5	17
438	438	SUB HIMALAYAN WEST BENGAL & SIKKIM	1902	1.2	0.7	87.1	126.1	271.3	539.2	671.0	603.8	799.9	74.4	5
439	439	SUB HIMALAYAN WEST BENGAL & SIKKIM	1903	5.5	8.7	19.6	18.6	163.6	541.2	431.5	708.8	365.2	141.3	0
440	440	SUB HIMALAYAN WEST BENGAL & SIKKIM	1904	3.4	29.2	0.9	124.3	333.6	274.2	500.4	468.5	260.6	164.8	8
441	441	SUB HIMALAYAN WEST BENGAL & SIKKIM	1905	12.0	31.2	51.9	104.4	290.6	524.8	523.1	1036.6	321.1	87.9	2
...
547	547	SUB HIMALAYAN WEST BENGAL & SIKKIM	2011	8.5	19.9	71.2	135.0	247.8	419.8	612.3	470.3	356.3	46.7	26
548	548	SUB HIMALAYAN WEST BENGAL & SIKKIM	2012	15.3	13.9	45.5	159.8	202.4	604.2	684.5	332.7	434.7	119.4	12
549	549	SUB HIMALAYAN WEST BENGAL & SIKKIM	2013	3.0	23.6	32.1	114.7	296.5	404.9	588.4	416.3	308.0	199.8	16
550	550	SUB HIMALAYAN WEST BENGAL & SIKKIM	2014	0.2	26.6	37.7	47.9	308.6	543.2	384.6	563.3	371.5	31.2	5
551	551	SUB HIMALAYAN WEST BENGAL & SIKKIM	2015	15.7	15.0	64.8	149.0	304.6	508.2	393.3	626.6	354.9	53.6	20

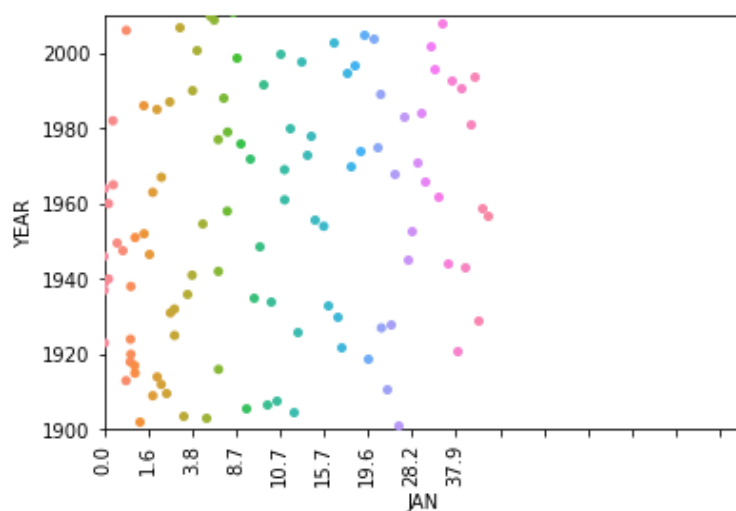
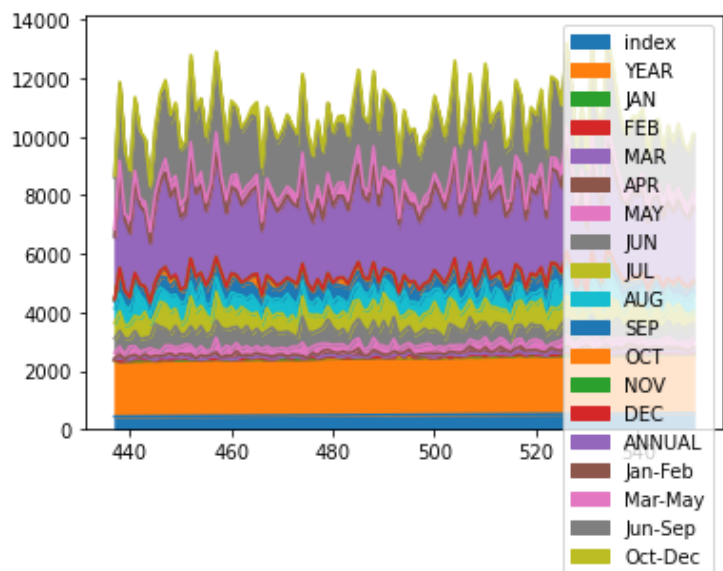
115 rows × 20 columns

```
In [5]: dat1.plot.bar("YEAR", "JAN")
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", "JAN")
sns.stripplot(x=dat1["JAN"],y=dat1["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat1["JAN"],rotation="vertical")
plt.gca().xaxis.set_major_locator(tic.MultipleLocator(base=10))
plt.show()
dat1.plot.hist()
```

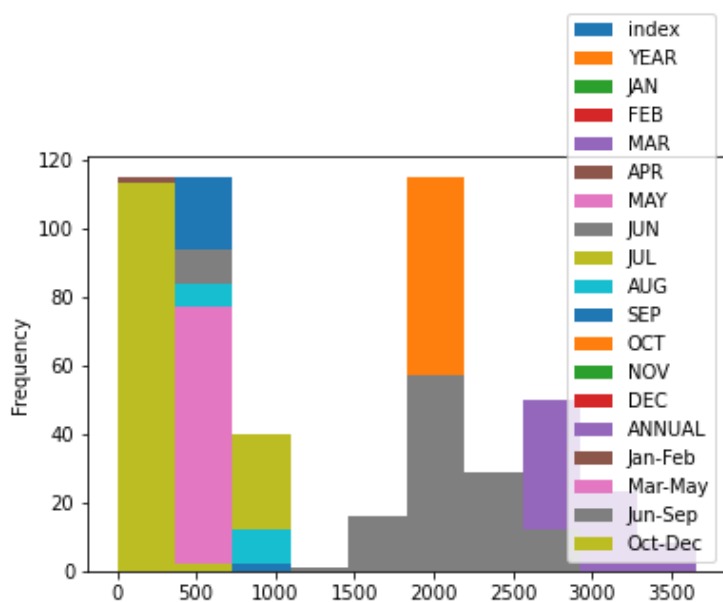


<Figure size 4320x2160 with 0 Axes>

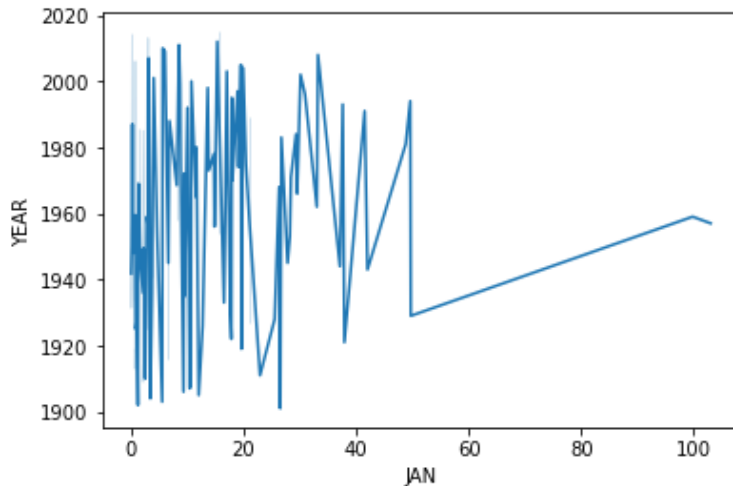




Out[5]: <AxesSubplot:ylabel='Frequency'>



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



PUNJAB

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

Out[7]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
1472	1472	PUNJAB	1901	55.7	50.1	25.2	2.1	25.2	10.4	178.2	145.0	24.4	3.7	0.0
1473	1473	PUNJAB	1902	0.0	0.8	9.9	10.9	29.6	49.9	125.6	94.9	67.2	9.0	0.0
1474	1474	PUNJAB	1903	29.5	0.5	45.0	1.3	9.2	5.2	212.2	119.1	132.5	6.9	0.0
1475	1475	PUNJAB	1904	24.2	1.7	87.8	1.2	13.8	22.0	59.9	124.0	73.8	7.4	9.8
1476	1476	PUNJAB	1905	53.0	40.3	24.3	0.5	2.2	19.2	122.6	50.3	111.1	1.2	0.0
...
1582	1582	PUNJAB	2011	3.5	35.6	8.2	17.8	18.9	162.9	120.9	193.5	140.2	0.0	1.0
1583	1583	PUNJAB	2012	62.6	3.2	1.9	31.1	1.6	11.9	120.2	135.1	112.3	2.2	0.4
1584	1584	PUNJAB	2013	9.3	50.1	11.6	3.4	3.6	120.3	117.9	217.1	24.4	16.2	6.1
1585	1585	PUNJAB	2014	21.8	20.1	30.3	24.5	20.8	20.6	76.3	41.9	105.8	6.0	0.7
1586	1586	PUNJAB	2015	17.7	31.3	68.5	29.8	16.7	48.3	130.2	88.6	69.2	9.0	0.8

115 rows × 15 columns

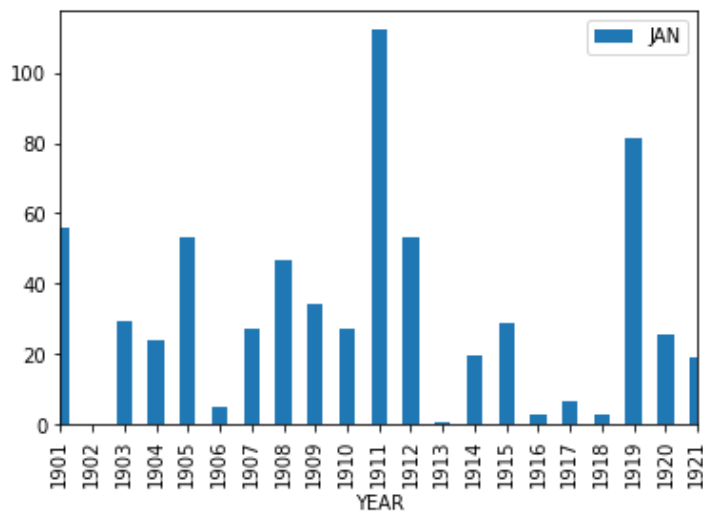


In [8]:

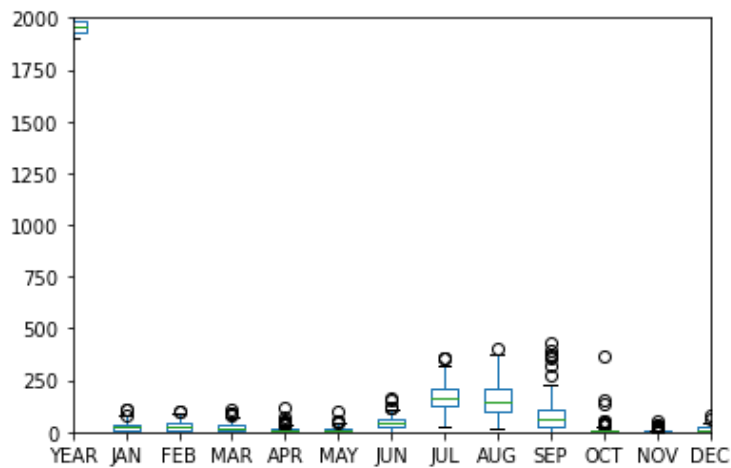
```

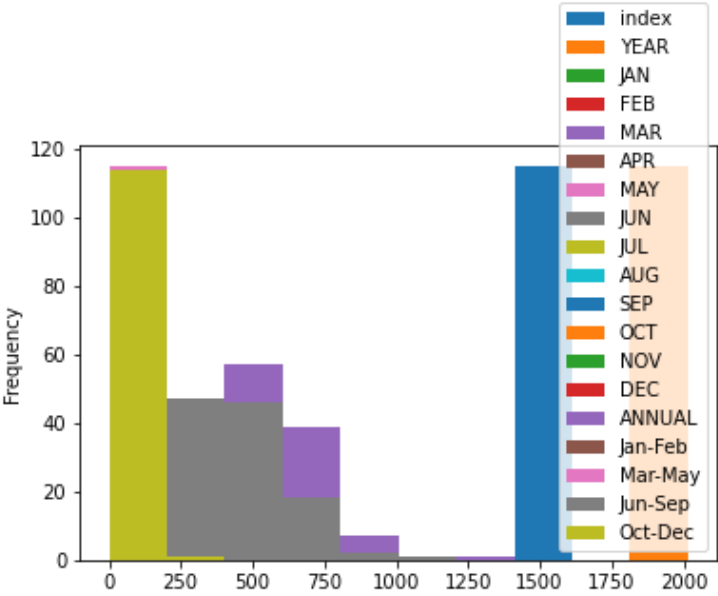
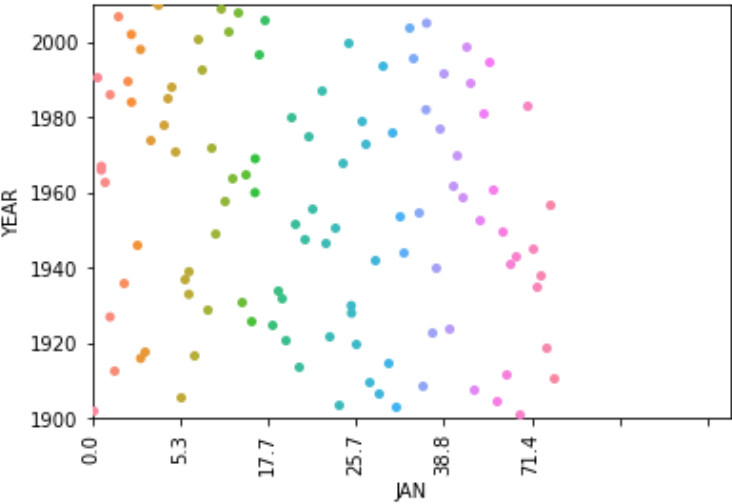
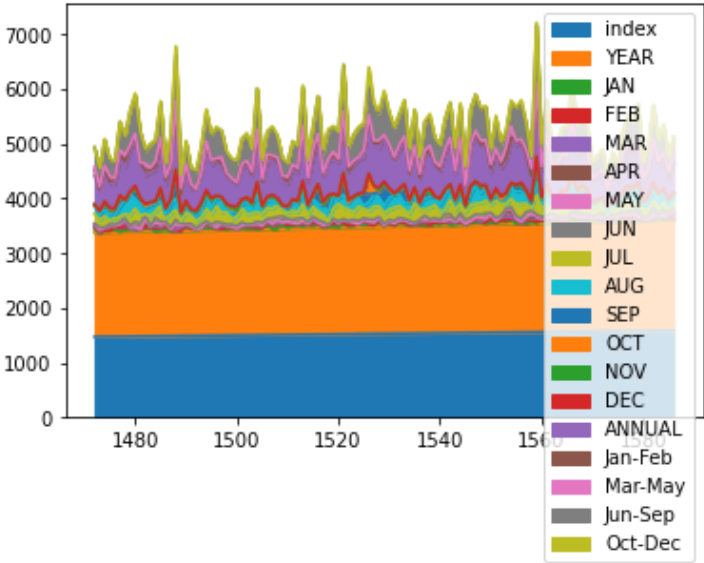
dat2.plot.bar("YEAR", "JAN")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat2.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat2.plot.area()
dat2.plot.scatter("YEAR", "JAN")
sns.stripplot(x=dat2["JAN"],y=dat2["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat2["JAN"],rotation="vertical")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

