

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

GUJARAT REGION

In [4]:

dat1=df[df["SUBDIVISION"]=="GUJARAT REGION"]
dat1

Out[4]:

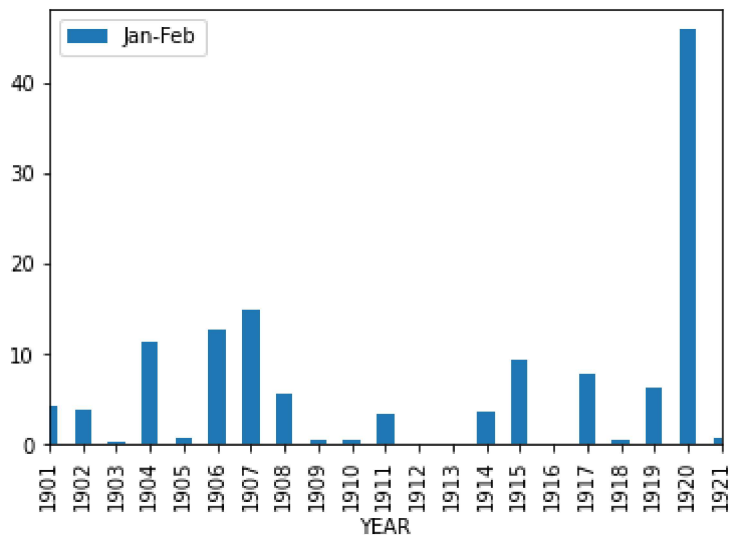
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
2277	2277	GUJARAT REGION	1901	4.2	0.0	0.6	1.6	7.0	60.3	240.2	205.4	18.1	16.6
2278	2278	GUJARAT REGION	1902	3.9	0.0	0.0	0.6	1.0	32.8	229.8	299.0	281.2	2.3
2279	2279	GUJARAT REGION	1903	0.3	0.1	1.4	0.0	12.3	30.1	452.9	202.0	183.2	5.4
2280	2280	GUJARAT REGION	1904	0.8	10.6	16.8	0.2	3.9	48.3	194.8	71.8	138.0	6.1
2281	2281	GUJARAT REGION	1905	0.1	0.7	1.1	0.3	0.0	20.1	668.3	37.9	81.3	1.4
...
2387	2387	GUJARAT REGION	2011	0.0	0.2	0.0	0.0	0.0	16.3	259.2	451.7	162.5	0.4
2388	2388	GUJARAT REGION	2012	0.1	0.0	0.0	0.0	0.0	34.4	178.2	230.3	263.8	7.1
2389	2389	GUJARAT REGION	2013	0.0	0.9	0.1	4.6	0.0	155.7	405.4	211.1	287.3	53.2
2390	2390	GUJARAT REGION	2014	5.7	0.1	0.2	1.0	1.3	11.6	307.5	138.6	235.1	3.3
2391	2391	GUJARAT REGION	2015	1.8	0.0	6.1	5.5	0.9	120.7	354.7	37.4	93.4	2.2

115 rows × 20 columns

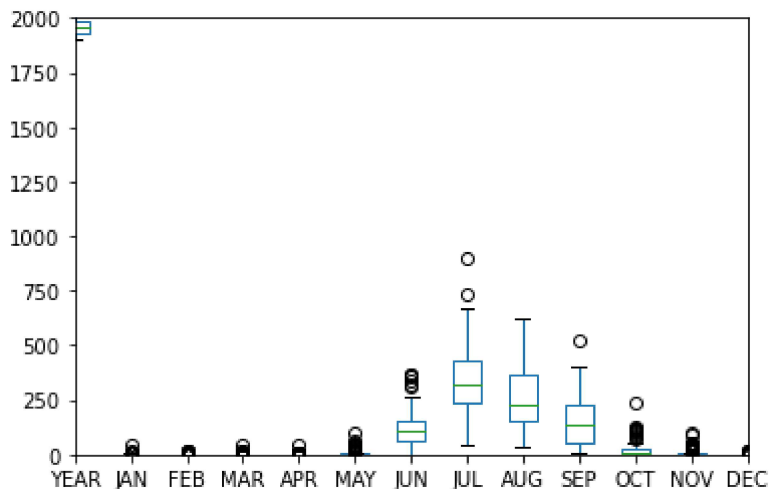
```

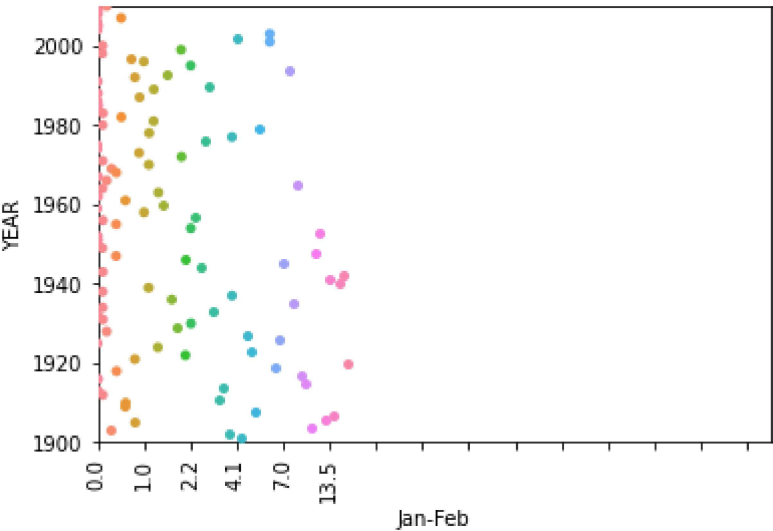
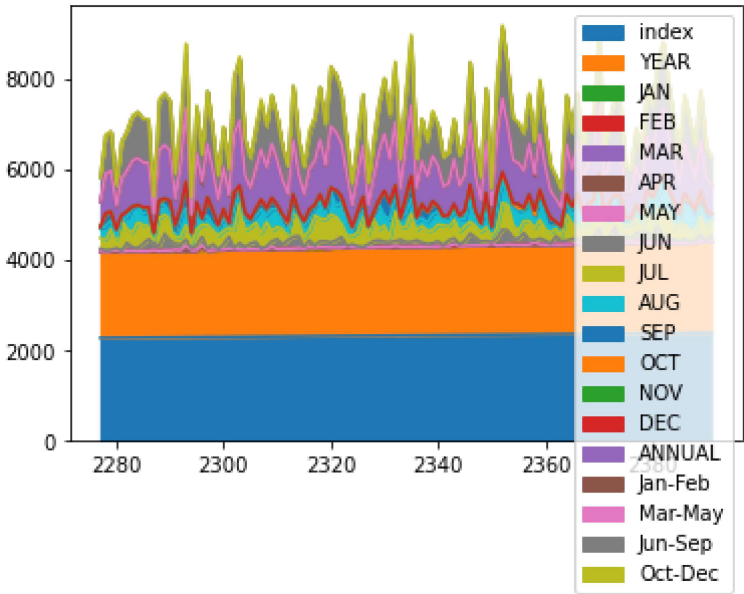
In [5]: dat1.plot.bar("YEAR", "Jan-Feb")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat1.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat1.plot.area()
dat1.plot.scatter("YEAR", "Jan-Feb")
sns.stripplot(x=dat1["Jan-Feb"],y=dat1["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)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

```

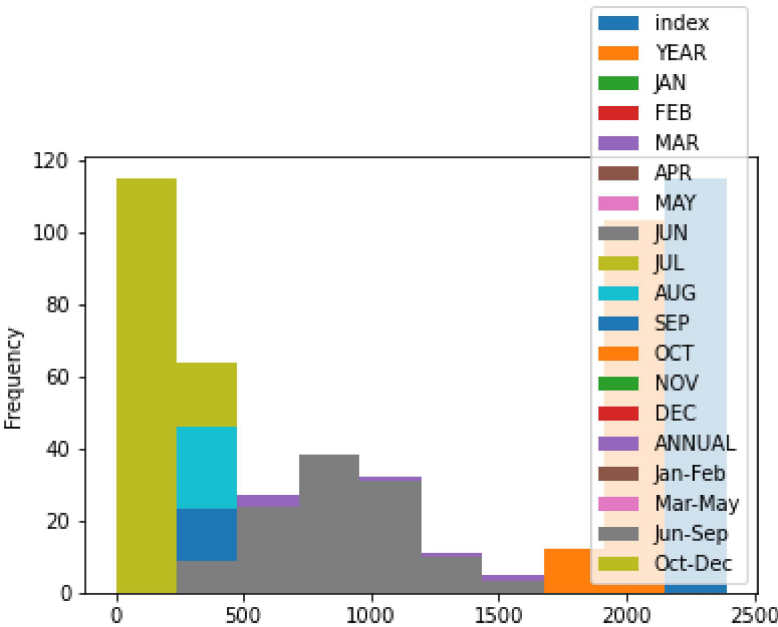


<Figure size 4320x2160 with 0 Axes>

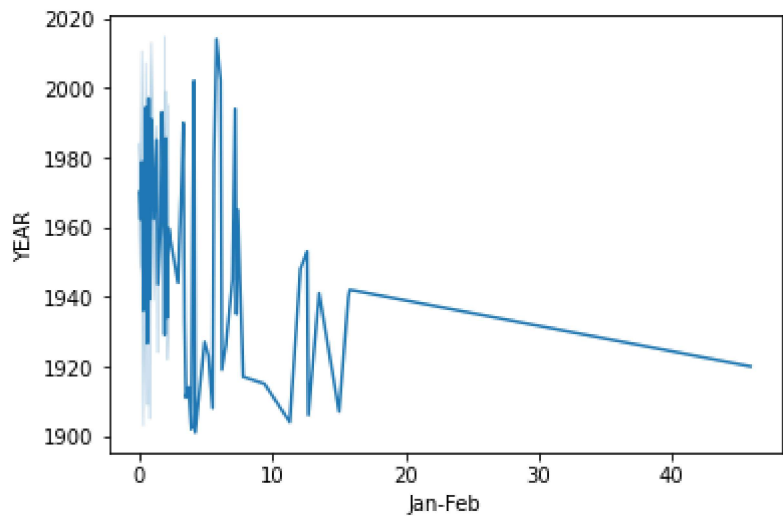




Out[5]: <AxesSubplot:ylabel='Frequency'>



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



VIDARBHA

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

Out[7]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT
2852	2852	VIDARBHA	1901	36.8	39.9	30.9	26.1	7.3	129.7	295.3	368.8	123.4	35.1
2853	2853	VIDARBHA	1902	1.6	0.1	0.0	6.5	4.1	38.0	270.7	204.7	150.9	29.0
2854	2854	VIDARBHA	1903	5.2	4.0	0.1	2.5	37.8	121.2	475.5	325.5	154.8	100.1
2855	2855	VIDARBHA	1904	4.3	2.4	12.9	0.2	14.8	148.9	158.3	151.8	196.9	61.7
2856	2856	VIDARBHA	1905	7.3	12.7	12.4	16.2	14.0	81.0	254.5	216.3	321.3	6.0
...
2962	2962	VIDARBHA	2011	0.0	1.2	0.1	7.7	0.6	137.9	247.1	302.8	191.0	4.7
2963	2963	VIDARBHA	2012	3.1	0.1	0.0	0.6	0.2	125.5	370.5	316.2	249.4	34.1
2964	2964	VIDARBHA	2013	6.6	13.0	3.8	2.8	0.5	366.7	535.5	326.1	131.7	133.1
2965	2965	VIDARBHA	2014	1.2	18.3	49.6	2.6	4.0	63.3	337.6	191.7	224.9	17.1
2966	2966	VIDARBHA	2015	26.3	4.7	66.3	28.1	12.8	254.6	137.2	288.9	167.5	7.0

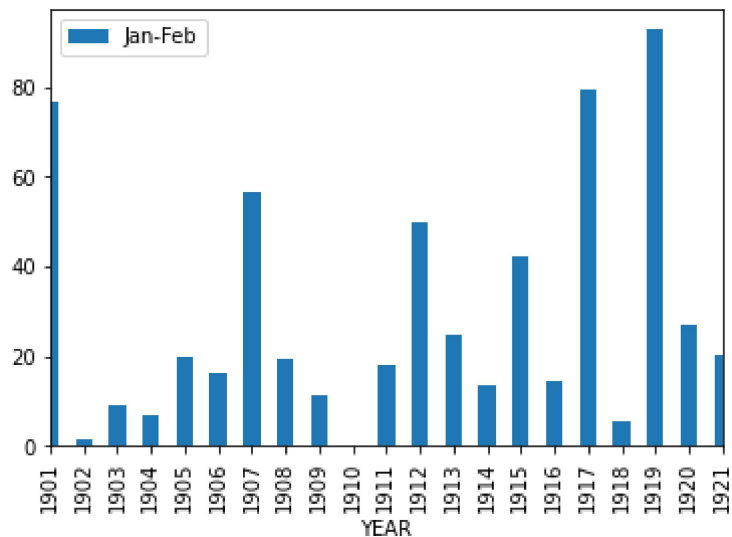
115 rows × 14 columns

In [8]:

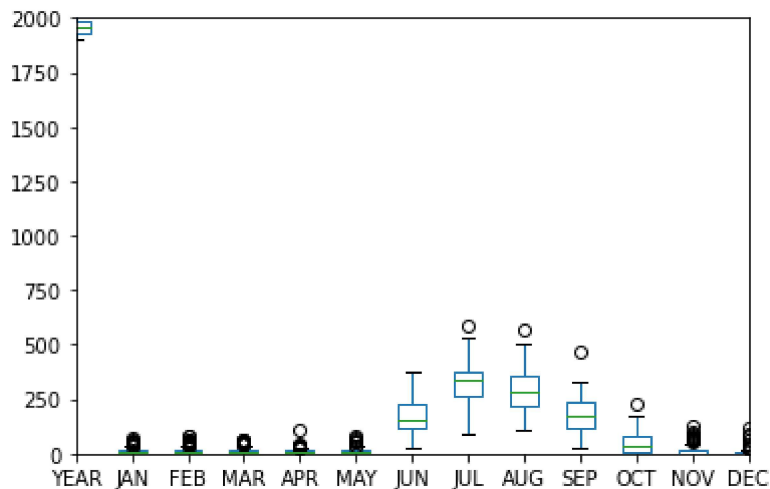
```

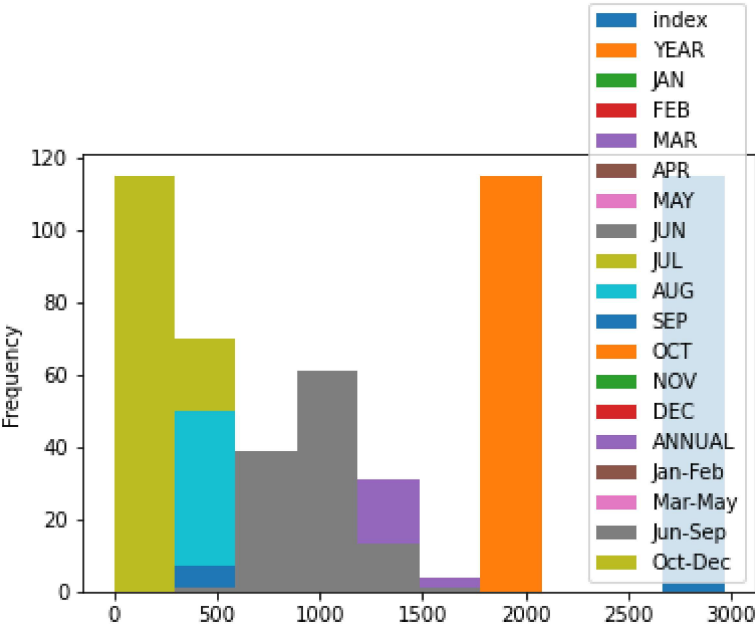
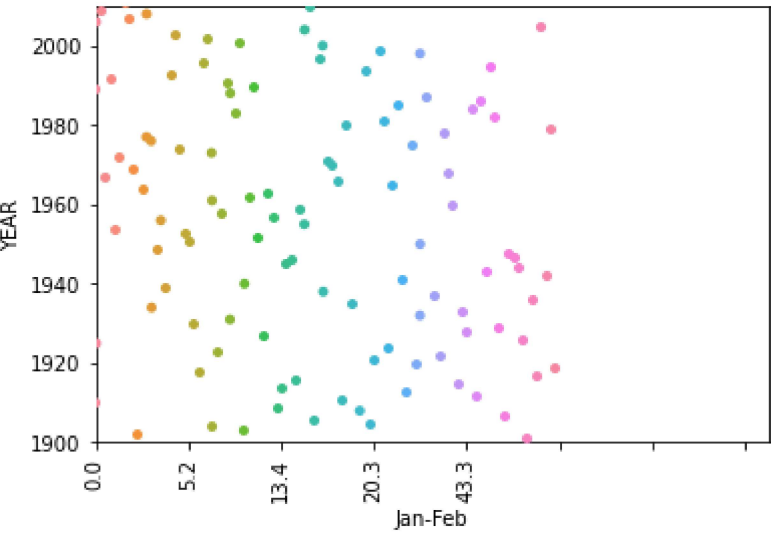
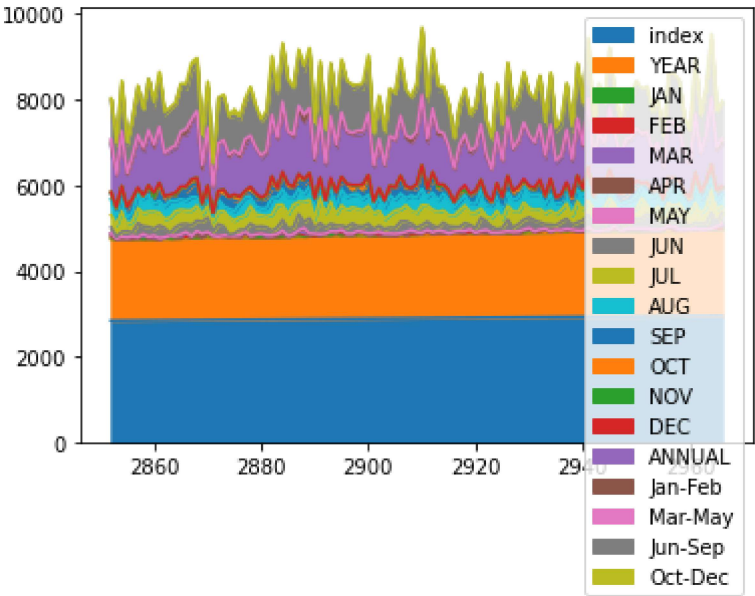
dat2.plot.bar("YEAR", "Jan-Feb")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat2.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat2.plot.area()
dat2.plot.scatter("YEAR", "Jan-Feb")
sns.stripplot(x=dat2["Jan-Feb"],y=dat2["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat2["Jan-Feb"],rotation="vertical")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

