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

COASTAL ANDHRA PRADESH

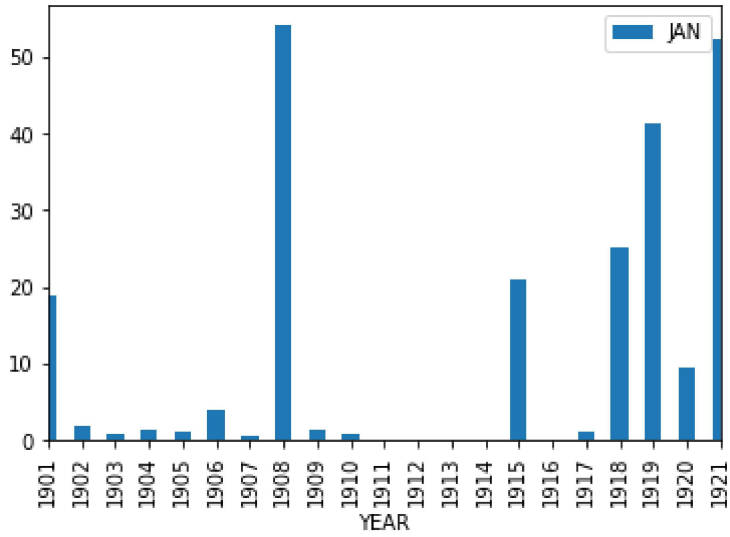
In [4]:
```
dat1=df[df["SUBDIVISION"]=="COASTAL ANDHRA PRADESH"]
dat1
```

Out[4]:

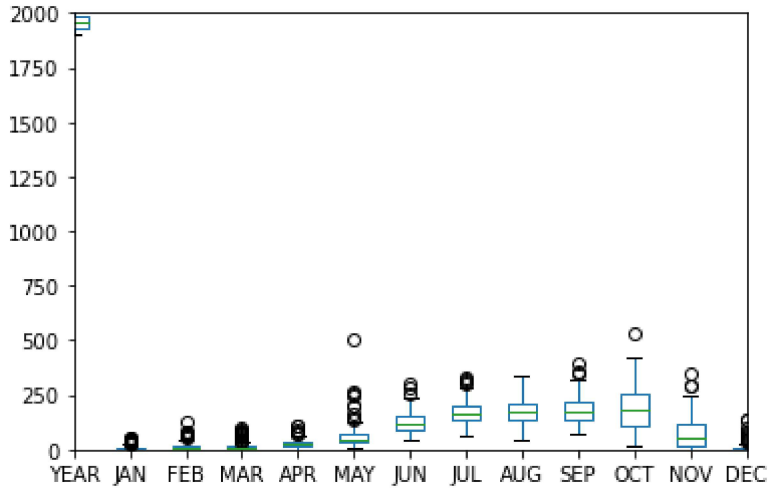
	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
3082	3082	COASTAL ANDHRA PRADESH	1901	18.8	80.9	7.2	28.7	68.7	77.7	113.0	133.7	125.3	173
3083	3083	COASTAL ANDHRA PRADESH	1902	2.0	0.0	2.8	23.9	37.6	72.6	144.5	236.1	204.5	262
3084	3084	COASTAL ANDHRA PRADESH	1903	0.8	13.3	0.2	6.2	73.4	154.0	248.6	258.0	216.5	159
3085	3085	COASTAL ANDHRA PRADESH	1904	1.3	0.0	5.4	3.0	136.3	107.8	120.2	117.7	116.8	240
3086	3086	COASTAL ANDHRA PRADESH	1905	1.1	16.7	68.0	37.0	68.8	84.4	64.6	210.8	170.2	66
...	...	...	...	...	...	...	...	...	...	...	...	...	
3192	3192	COASTAL ANDHRA PRADESH	2011	0.0	17.9	0.9	62.3	67.9	86.8	196.0	215.8	129.7	74
3193	3193	COASTAL ANDHRA PRADESH	2012	37.6	0.0	2.7	24.0	39.3	95.4	221.9	221.2	246.5	140
3194	3194	COASTAL ANDHRA PRADESH	2013	2.0	29.6	0.2	48.0	28.2	127.5	162.4	123.1	132.0	411
3195	3195	COASTAL ANDHRA PRADESH	2014	0.4	1.2	9.1	6.0	112.9	45.7	151.8	177.8	144.5	195
3196	3196	COASTAL ANDHRA PRADESH	2015	2.0	0.6	5.5	32.3	34.1	283.8	116.0	192.0	201.8	59

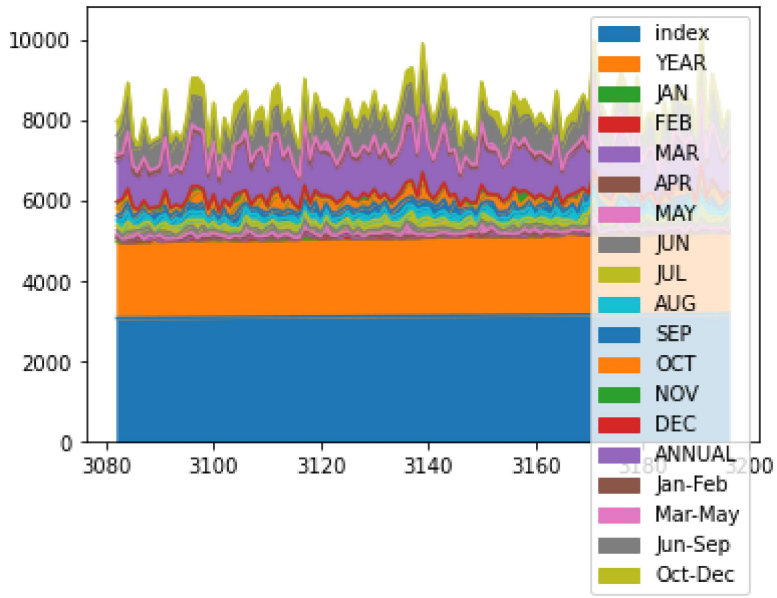
115 rows × 20 columns

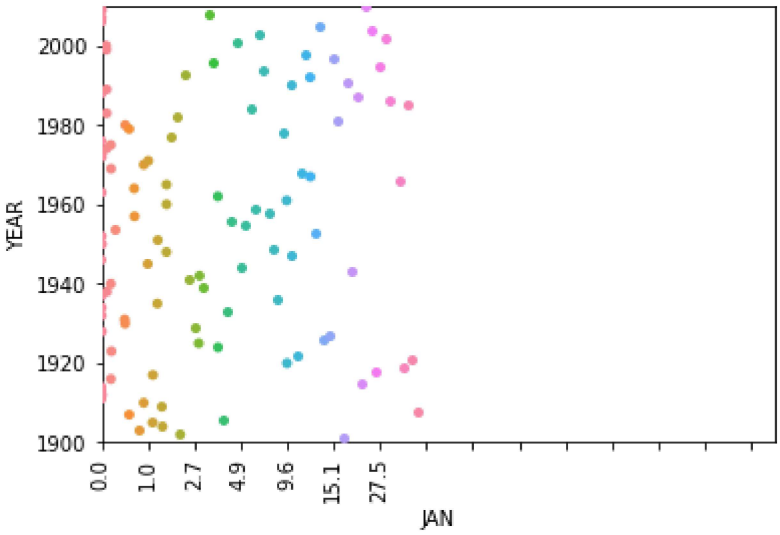
In [5]:
```
dat1.plot.bar("YEAR","JAN")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat1.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat1.plot.area()
dat1.plot.scatter("YEAR","JAN")
sns.stripplot(x=dat1["JAN"],y=dat1["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat1["JAN"],rotation="vertical")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



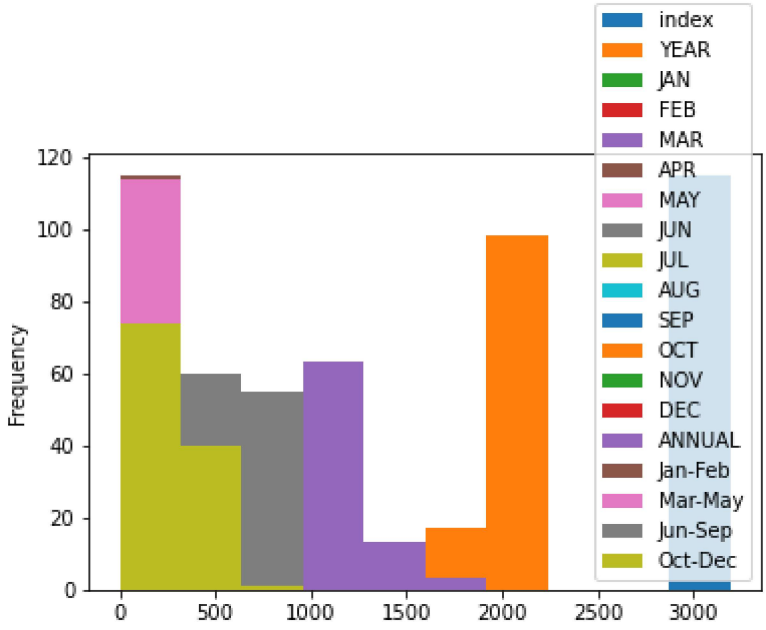

```
<Figure size 4320x2160 with 0 Axes>
```



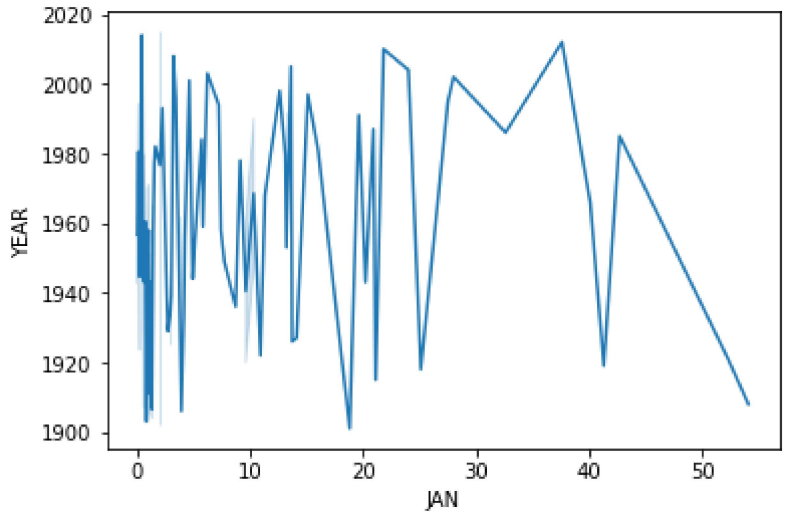




Out[5]: <AxesSubplot:ylabel='Frequency'>



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



WEST UTTAR PRADESH

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

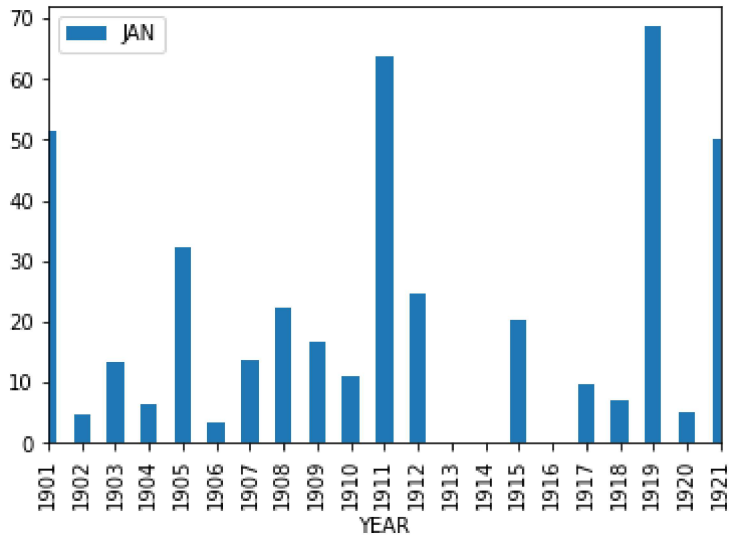
Out[7]:

	index	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OC
1127	1127	WEST UTTAR PRADESH	1901	51.4	25.6	9.5	0.7	5.6	23.8	201.9	374.3	67.7	7.(
1128	1128	WEST UTTAR PRADESH	1902	4.6	4.6	0.6	4.8	7.2	54.5	325.9	180.6	143.1	9.(
1129	1129	WEST UTTAR PRADESH	1903	13.4	0.4	1.2	0.0	8.2	32.7	145.4	279.1	150.4	177.:
1130	1130	WEST UTTAR PRADESH	1904	6.3	2.0	29.7	0.4	24.8	68.5	358.8	311.1	97.1	2.1
1131	1131	WEST UTTAR PRADESH	1905	32.3	26.6	14.8	3.6	7.1	18.9	139.8	95.0	92.2	0.:
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1237	1237	WEST UTTAR PRADESH	2011	2.1	10.4	3.9	2.8	29.6	175.9	215.9	232.3	101.7	0.1
1238	1238	WEST UTTAR PRADESH	2012	14.5	0.1	1.4	4.7	0.3	4.0	145.1	149.1	67.8	0.:
1239	1239	WEST UTTAR PRADESH	2013	20.4	69.5	3.5	1.6	2.1	190.6	233.9	287.1	52.2	61.:
1240	1240	WEST UTTAR PRADESH	2014	48.3	29.4	22.6	5.3	11.0	22.0	151.6	81.0	84.7	14.(
1241	1241	WEST UTTAR PRADESH	2015	31.6	7.2	66.8	21.0	8.1	72.0	194.2	143.5	26.5	6.:

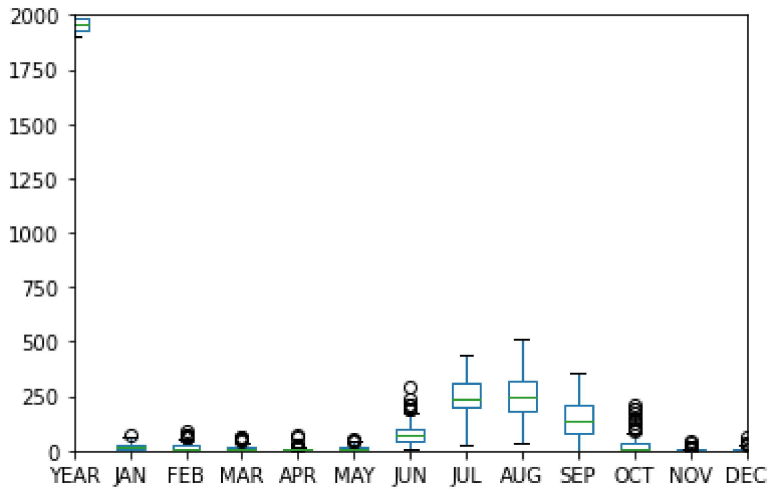
115 rows × 20 columns

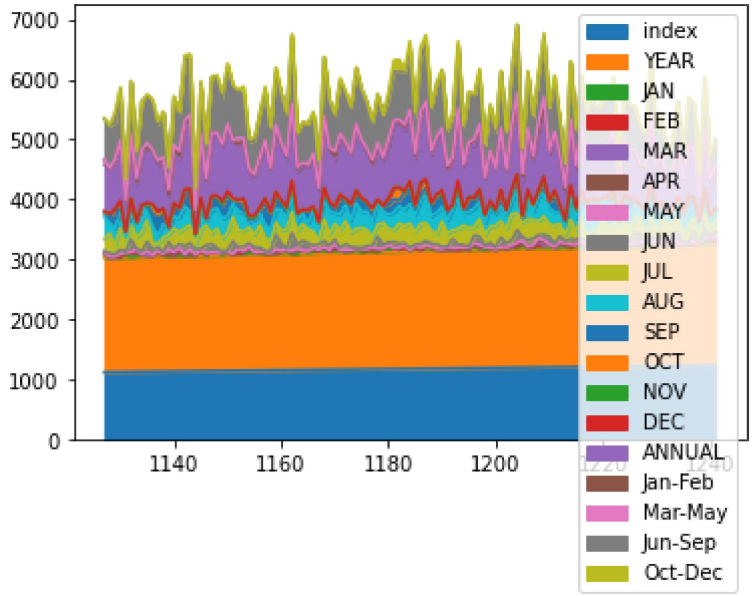
In [8]:

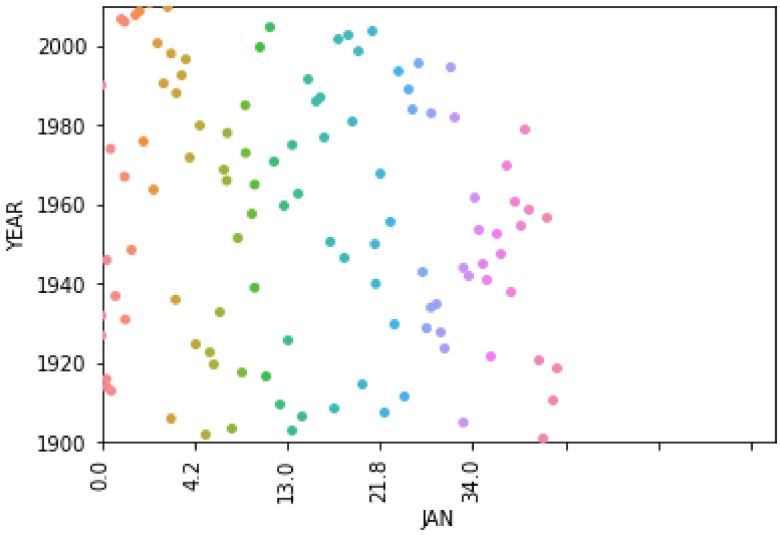
```
dat2.plot.bar("YEAR","JAN")
plt.xlim(0,20)
plt.figure(figsize=(60,30))
plt.show()
dat2.plot.box()
plt.xlim(2,14)
plt.ylim(0,2000)
plt.show()
dat2.plot.area()
dat2.plot.scatter("YEAR","JAN")
sns.stripplot(x=dat2["JAN"],y=dat2["YEAR"],jitter=True)
plt.ylim(1900,2010)
plt.xlim(0,145)
plt.xticks(dat2["JAN"],rotation="vertical")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

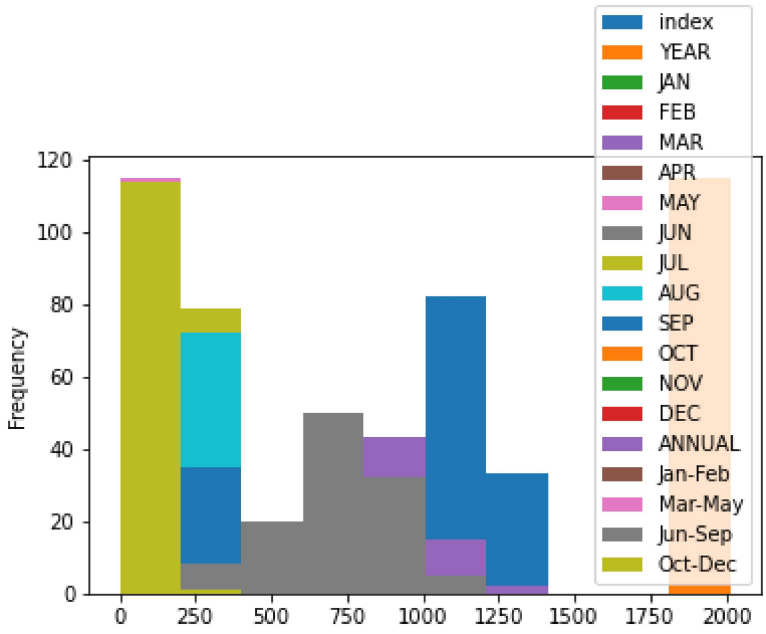


<Figure size 4320x2160 with 0 Axes>




7000
6000
5000
4000
3000
2000
1000
0
1140
1160
1180
1200
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
4.2
13.0
21.8
34.0
JAN


120
100
80
60
40
20
0
Frequency
0
250
500
750
1000
1250
1500
1750
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["JAN"],y=dat2["YEAR"])
plt.show()
```