In [1]:

```
#Roll No-33238
import pandas as pd
```

In [2]:

```
df=pd.read_csv('temperatures.csv')
```

In [3]:

```
df.head()
```

Out[3]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	A
0	1901	22.40	24.14	29.07	31.91	33.41	33.18	31.21	30.39	30.47	29.97	27.31	24.49	
1	1902	24.93	26.58	29.77	31.78	33.73	32.91	30.92	30.73	29.80	29.12	26.31	24.04	
2	1903	23.44	25.03	27.83	31.39	32.91	33.00	31.34	29.98	29.85	29.04	26.08	23.65	
3	1904	22.50	24.73	28.21	32.02	32.64	32.07	30.36	30.09	30.04	29.20	26.36	23.63	
4	1905	22.00	22.83	26.68	30.01	33.32	33.25	31.44	30.68	30.12	30.67	27.52	23.82	

In [4]:

```
df.shape
```

Out[4]:

```
(117, 18)
```

In [5]:

```
df.describe()
```

Out[5]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	
count	117.000000	117.000000	117.000000	117.000000	117.000000	117.000000	117.000000	11
mean	1959.000000	23.687436	25.597863	29.085983	31.975812	33.565299	32.774274	3
std	33.919021	0.834588	1.150757	1.068451	0.889478	0.724905	0.633132	
min	1901.000000	22.000000	22.830000	26.680000	30.010000	31.930000	31.100000	2
25%	1930.000000	23.100000	24.780000	28.370000	31.460000	33.110000	32.340000	3
50%	1959.000000	23.680000	25.480000	29.040000	31.950000	33.510000	32.730000	3
75%	1988.000000	24.180000	26.310000	29.610000	32.420000	34.030000	33.180000	3
max	2017.000000	26.940000	29.720000	32.620000	35.380000	35.840000	34.480000	3

In [6]:

```
df.isna().sum()
```

Out[6]:

```
YEAR       0
JAN        0
FEB        0
MAR        0
APR        0
MAY        0
JUN        0
JUL        0
AUG        0
SEP        0
OCT        0
NOV        0
DEC        0
ANNUAL     0
JAN-FEB    0
MAR-MAY    0
JUN-SEP    0
OCT-DEC    0
dtype: int64
```

In [7]:

```
from sklearn.model_selection import train_test_split
```

In [34]:

```
train,test=train_test_split(df,test_size=0.25,random_state=42)
```

In [35]:

```
train.shape
```

Out[35]:

```
(87, 18)
```

In [36]:

```
test.shape
```

Out[36]:

```
(30, 18)
```

In [37]:

```
yf=test['JAN']
```

In [38]:

```
yf.shape
```

Out[38]:

```
(30,)
```

In [39]:

```
test0=test.drop('JAN',axis=1)
```

In [40]:

```
test0.shape
```

Out[40]:

```
(30, 17)
```

In [41]:

```
trainX=train.drop('JAN',axis=1)
```

In [42]:

```
trainY=train['JAN']
```

In [43]:

```
trainX.shape
```

Out[43]:

```
(87, 17)
```

In [44]:

```
trainY.shape
```

Out[44]:

```
(87,)
```

In [45]:

```
from sklearn.linear_model import LinearRegression
```

In [46]:

```
lreg=LinearRegression().fit(trainX,trainY)
```

In [47]:

```
pred=lreg.predict(test0)
```

In [48]:

```
pred.shape
```

Out[48]:

```
(30,)
```

In [49]:

```
from sklearn.metrics import mean_squared_error
```

In [50]:

```
mean_squared_error(yf,pred)
```

Out[50]:

```
0.05728583367665706
```

In [51]:

```
from sklearn.metrics import mean_absolute_error
```

In [52]:

```
mean_absolute_error(yf,pred)
```

Out[52]:

```
0.13118655830850703
```

In [53]:

```
from sklearn.metrics import r2_score
```

In [54]:

```
r2_score(yf,pred)
```

Out[54]:

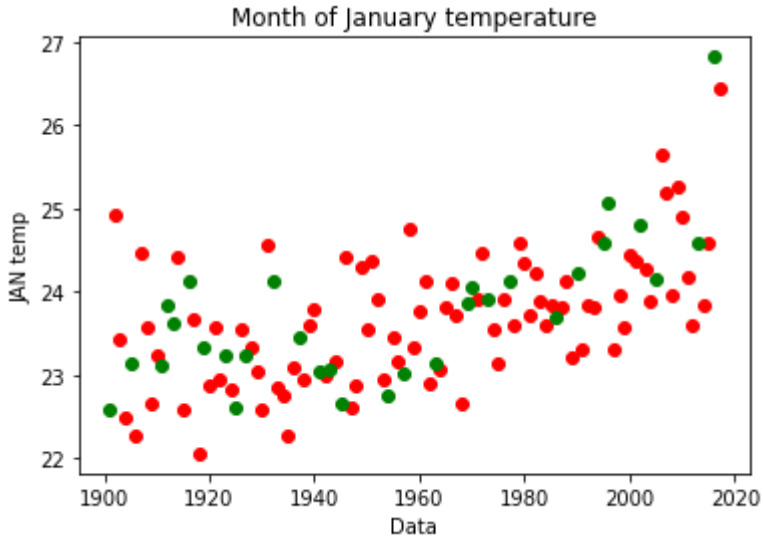
```
0.9371386510388571
```

In [55]:

```
import matplotlib.pyplot as plt
```

In [60]:

```
plt.scatter(trainX.iloc[:,0],trainY,color="red")
plt.scatter(test0.iloc[:,0],pred,color="green")
plt.title("Month of January temperature")
plt.xlabel("Data")
plt.ylabel("JAN temp")
plt.show()
```



In [111]:

```
train0,test0=train_test_split(df[['YEAR','JAN']],test_size=0.25,random_state=42)
```

In [112]:

```
train0.shape
```

Out[112]:

```
(87, 2)
```

In [113]:

```
train0x=train0['YEAR'].values
```

In [114]:

```
y0=train0['JAN'].values
```

In [115]:

```
test0x=test0['YEAR'].values
```

In [116]:

```
ytr=test0['JAN'].values
```

In [ ]:

```
X0=
```