In [15]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

In [16]:
```
data = pd.read_csv("temperatures.csv")
```

In [17]:
```
data.head()
```

Out[17]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNU
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	28
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	29
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	28
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	28
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	28

In [18]:
```
model = LinearRegression()
```

In [19]:
```
x_train = data[['YEAR']].values.reshape(117,1)
y_train = data[['ANNUAL']].values.reshape(117,1)
model.fit(x_train, y_train)
```

Out[19]:
```
LinearRegression
LinearRegression()
```

In [20]:
```
model.predict([[2304]])
```

Out[20]:
```
array([[33.70831308]])
```

In [21]:
```
predicted =  model.predict(x_train)
```

In [22]:
```
#mse
np.mean(abs(y_train-predicted))
```

Out[22]:
```
0.22535284978630413
```

In [23]:
```
#mae
np.mean((y_train-predicted)**2)
```

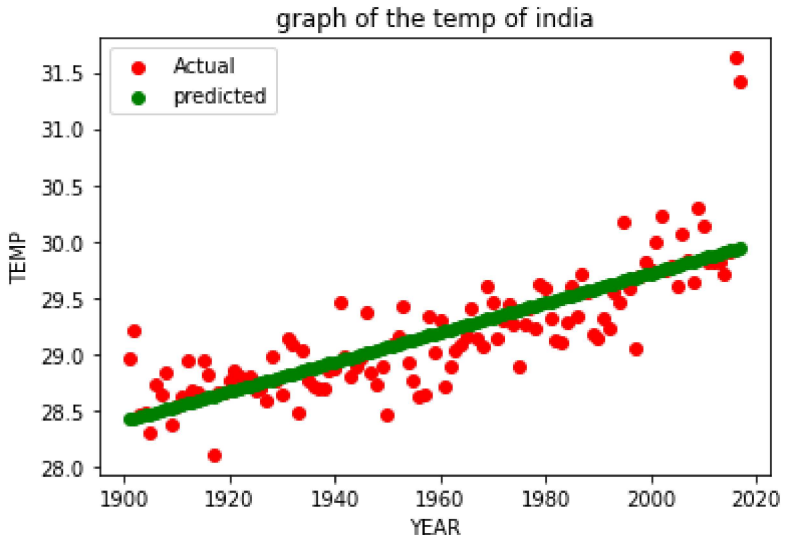
Out[23]:
```
0.10960795229110352
```

In [24]:
```
model.score(x_train, y_train)
```

Out[24]:
```
0.6418078912783682
```

In [25]:
```
plt.title("graph of the temp of india")
plt.xlabel("YEAR")
plt.ylabel('TEMP')
plt.scatter(x_train, y_train, label = 'Actual', color = 'r')
plt.scatter(x_train, predicted, label = 'predicted', color = 'g')
plt.legend()
```

Out[25]: <matplotlib.legend.Legend at 0x1b68ac35480>



In [ ]: