

```
In [1]: import pandas as pd
```

```
In [2]: data = pd.read_csv('temperatures.csv')
```

```
In [3]: data
```

```
Out[3]:
```

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	AI
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	
...	
112	2013	24.56	26.59	30.62	32.66	34.46	32.44	31.07	30.76	31.04	30.27	27.83	25.37	
113	2014	23.83	25.97	28.95	32.74	33.77	34.15	31.85	31.32	30.68	30.29	28.05	25.08	
114	2015	24.58	26.89	29.07	31.87	34.09	32.48	31.88	31.52	31.55	31.04	28.10	25.67	
115	2016	26.94	29.72	32.62	35.38	35.72	34.03	31.64	31.79	31.66	31.98	30.11	28.01	
116	2017	26.45	29.46	31.60	34.95	35.84	33.82	31.88	31.72	32.22	32.29	29.60	27.18	

117 rows × 18 columns



```
In [4]: x_train=data[["YEAR"]]
```

```
In [5]: y_train=data[["ANNUAL"]]
```

```
In [6]: from sklearn.linear_model import LinearRegression
```

```
In [7]: model=LinearRegression()
```

```
In [8]: model.fit(x_train,y_train)
```

```
Out[8]: LinearRegression()
```

```
In [9]: model.predict([[2050]])
```

```
Out[9]: array([[30.37543142]])
```

```
In [10]: predicted=model.predict(x_train)
```

```
In [11]: predicted
```

```
[29.5010558],  
[29.57501496],  
[29.58813654],  
[29.60125812],  
[29.6143797 ],  
[29.62750129],  
[29.64062287],  
[29.65374445],  
[29.66686603],  
[29.67998761],  
[29.69310919],  
[29.70623077],  
[29.71935236],  
[29.73247394],  
[29.74559552],  
[29.7587171 ],  
[29.77183868],  
[29.78496026],  
[29.79808184],  
[29.81120342],  
[29.82432501]
```

```
In [12]: predicted.shape
```

```
Out[12]: (117, 1)
```

```
In [13]: from sklearn.metrics import mean_squared_error
```

```
In [14]: mean_squared_error(y_train,predicted)
```

```
Out[14]: 0.10960795229110358
```

```
In [17]: from sklearn.metrics import mean_absolute_error, r2_score
```

```
In [16]: mean_absolute_error(y_train,predicted)
```

```
Out[16]: 0.22535284978630418
```

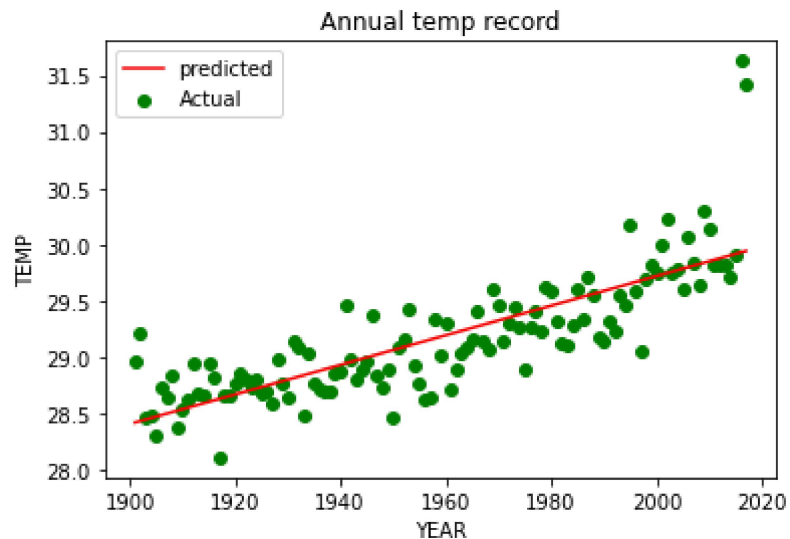
```
In [18]: r2_score(y_train,predicted)
```

```
Out[18]: 0.641807891278368
```

```
In [19]: import matplotlib.pyplot as plt
```

```
In [26]: plt.scatter(x_train,y_train, label="Actual",color="g")  
plt.plot(x_train,predicted, label="predicted",color="red")  
plt.xlabel("YEAR")  
plt.ylabel("TEMP")  
plt.title("Annual temp record")  
plt.legend()
```

Out[26]: <matplotlib.legend.Legend at 0x1c6997e1df0>



In []:

In []: