In [1]:
```
import pandas as p
import numpy as n
import seaborn as s
import matplotlib.pyplot as plt
```

In [2]:
```
df = p.read_csv("temperatures.csv")
df
```

Out[2]:

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
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 [14]:
```
X = df[["YEAR"]].values
Y = df[["ANNUAL"]].values
type(X)
```

Out[14]:
```
numpy.ndarray
```

In [15]:
```
from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test = train_test_split(X,Y,random_state=0,test_siz
print(X_train.shape,X_test.shape,Y_train.shape,Y_test.shape)
```

```
(87, 1) (30, 1) (87, 1) (30, 1)
```

In [16]:
```
from sklearn.linear_model import LinearRegression
```

In [17]:
```
LR = LinearRegression()
```

In [18]:
```
LR.fit(X_train,Y_train)
```

Out[18]:
```
LinearRegression()
```
In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

In [22]:
```
Y_predict = LR.predict(X_test)
```

In [23]:
```
from sklearn.metrics import mean_absolute_error,mean_squared_error,r2_score
```

In [24]:
```
mean_absolute_error(Y_test,Y_predict)
```

Out[24]:
```
0.27095859373514986
```

In [25]:
```
mean_squared_error(Y_test,Y_predict)
```

Out[25]:
```
0.16530868767622497
```

r2_score(Y_test,Y_predict)

In [26]:
```
r2_score(Y_test,Y_predict)
```

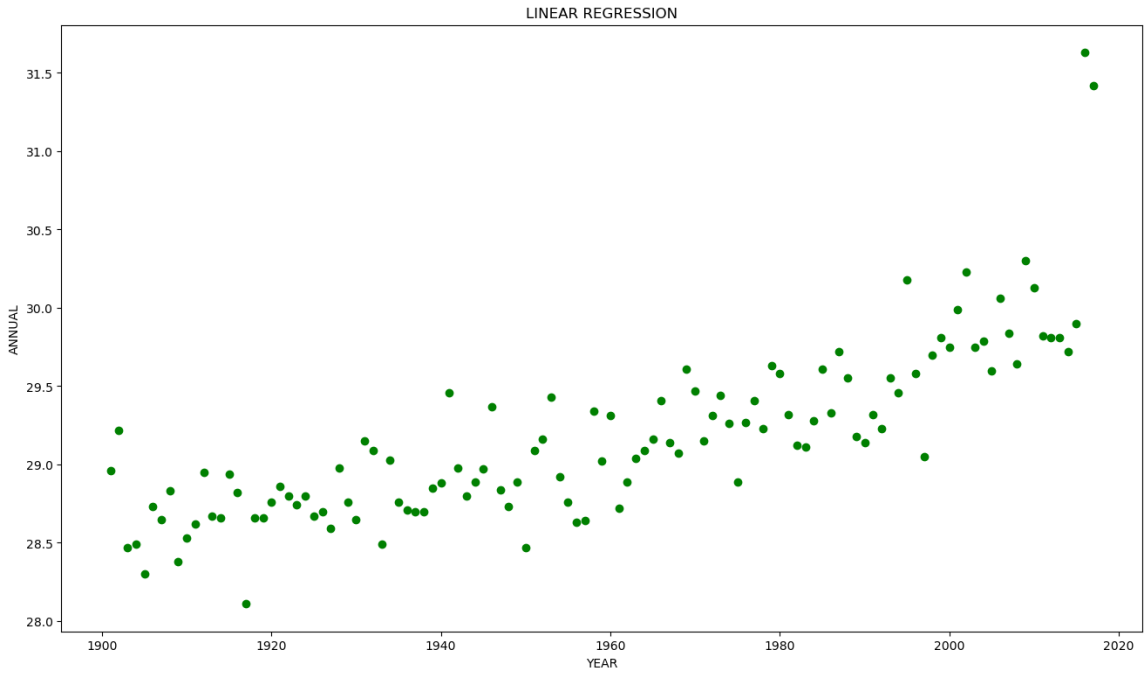
Out[26]:
```
0.6593855137850329
```

In [27]:
```
LR.predict([[3000]])
```

Out[27]:
```
array([[41.4083029]])
```

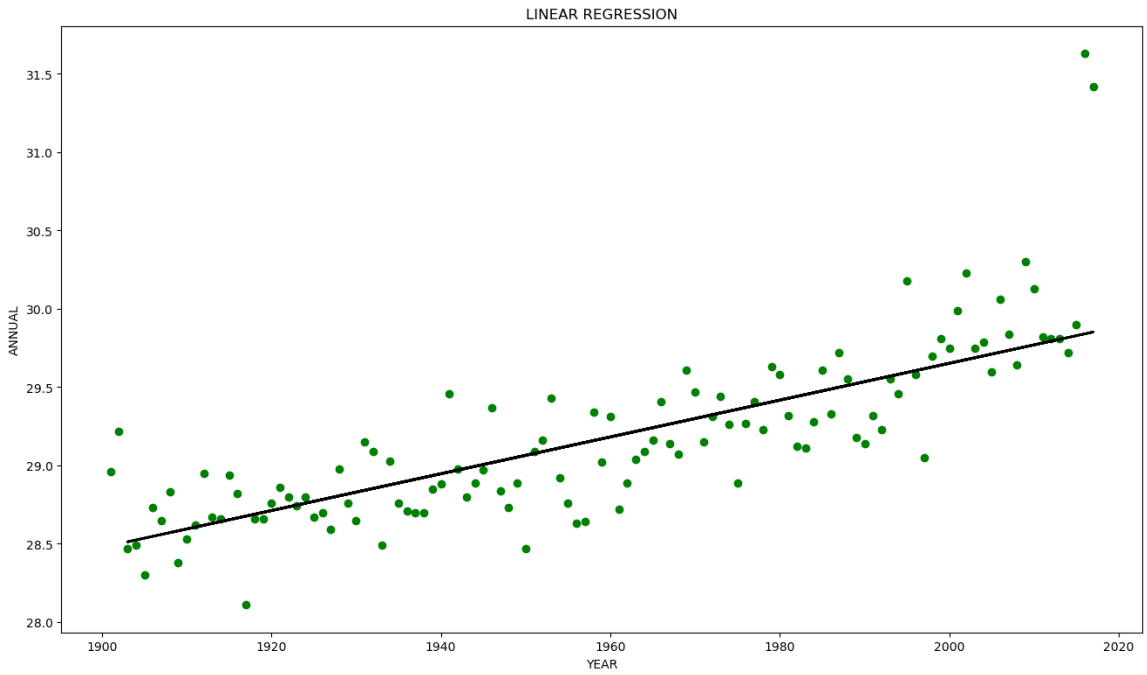
In [28]:
```
plt.figure(figsize=(16,9))
plt.xlabel("YEAR")
plt.ylabel("ANNUAL")
plt.title("LINEAR REGRESSION")
plt.plot(X,Y,'o',color="green")
```

Out[28]: [<matplotlib.lines.Line2D at 0x1b4fa15fa50>]



In [40]:
```
plt.figure(figsize=(16,9))
plt.xlabel("YEAR")
plt.ylabel("ANNUAL")
plt.title("LINEAR REGRESSION")
plt.plot(X,Y,'o',color="green")
plt.plot(X_test,Y_predict,linewidth=2,color="black")
# plt.scatter(X,Y,color="red",linewidth=1)
```

Out[40]: [<matplotlib.lines.Line2D at 0x1b4800f8a50>]



In [ ]: