

In [1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:

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

In [4]:

```
df.head()
```

Out[4]:

	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 [7]:

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

In [11]:

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

In [19]:

```
df.iloc[:,1: -5].values
```

Out[19]:

```
array([[22.4 , 24.14, 29.07, ..., 29.97, 27.31, 24.49],
       [24.93, 26.58, 29.77, ..., 29.12, 26.31, 24.04],
       [23.44, 25.03, 27.83, ..., 29.04, 26.08, 23.65],
       ...,
       [24.58, 26.89, 29.07, ..., 31.04, 28.1 , 25.67],
       [26.94, 29.72, 32.62, ..., 31.98, 30.11, 28.01],
       [26.45, 29.46, 31.6 , ..., 32.29, 29.6 , 27.18]])
```

In [23]:

```
df.iloc[ 1: ,0]
```

Out[23]:

```
1      1902
2      1903
3      1904
4      1905
5      1906
...
112     2013
113     2014
114     2015
115     2016
116     2017
Name: YEAR, Length: 116, dtype: int64
```

In [24]:

```
from sklearn import linear_model, metrics
```

In [26]:

```
X=df[["YEAR"]]
y=df[["JAN"]]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_s
print(X_train.shape, X_test.shape)

(93, 1) (24, 1)
```

In [31]:

```
X=df[["YEAR"]]
Y=df[["JAN"]]
```

In [32]:

```
lr = LinearRegression()
```

In [33]:

```
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_sta
```

In [34]:

```
reg = linear_model.LinearRegression()
```

In [35]:

```
lr.predict([2021])
```

```
-----
-----
NotFittedError                                Traceback (most recent call
last)
<ipython-input-35-4b2b2c5fff2e> in <module>
----> 1 lr.predict([2021])

~/local/lib/python3.8/site-packages/sklearn/linear_model/_base.py in
predict(self, X)
    360         Returns predicted values.
    361         """
--> 362         return self._decision_function(X)
    363
    364         _preprocess_data = staticmethod(_preprocess_data)

~/local/lib/python3.8/site-packages/sklearn/linear_model/_base.py in
_decision_function(self, X)
    341
    342     def _decision_function(self, X):
--> 343         check_is_fitted(self)
    344
    345         X = self._validate_data(X, accept_sparse=["csr", "csc"
, "coo"], reset=False)

~/local/lib/python3.8/site-packages/sklearn/utils/validation.py in ch
eck_is_fitted(estimator, attributes, msg, all_or_any)
   1220
   1221     if not fitted:
-> 1222         raise NotFittedError(msg % {"name": type(estimator).__
name__})
   1223
   1224
```

NotFittedError: This LinearRegression instance is not fitted yet. Call 'fit' with appropriate arguments before using this estimator.

In [36]:

```
X=df[["YEAR"]]
Y=df[["JAN"]]
```

```
-----
-----
NameError                                Traceback (most recent call
last)
<ipython-input-36-b72394811e04> in <module>
----> 1 X=trainData[["YEAR"]]
      2 Y=trainData[["JAN"]]
```

NameError: name 'trainData' is not defined

In []: