

```
import pandas as pd
```

```
import numpy as np
```

Import CSV as DataFrame

```
df=pd.read_csv(r'https://github.com/YBI-Foundation/Dataset/raw/main/Fish.csv')

```

```
df.head()
```

	Category	Species	Weight	Height	Width	Length1	Length2	Length3
0	1	Bream	242.0	11.5200	4.0200	23.2	25.4	30.0
1	1	Bream	290.0	12.4800	4.3056	24.0	26.3	31.2
2	1	Bream	340.0	12.3778	4.6961	23.9	26.5	31.1
3	1	Bream	363.0	12.7300	4.4555	26.3	29.0	33.5
4	1	Bream	430.0	12.4440	5.1340	26.5	29.0	34.0

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 159 entries, 0 to 158
Data columns (total 8 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Category    159 non-null    int64
1   Species     159 non-null    object
2   Weight      159 non-null    float64
3   Height      159 non-null    float64
4   Width       159 non-null    float64
5   Length1     159 non-null    float64
6   Length2     159 non-null    float64
7   Length3     159 non-null    float64
dtypes: float64(6), int64(1), object(1)
memory usage: 10.1+ KB
```

```
df.describe()
```

	Category	Weight	Height	Width	Length1	Length2	Length3
count	159.000000	159.000000	159.000000	159.000000	159.000000	159.000000	159.000000
mean	3.264151	398.326415	8.970994	4.417486	26.247170	28.415723	31.227000
std	1.704249	357.978317	4.286208	1.685804	9.996441	10.716328	11.610000

```
df.shape
```

```
(159, 8)
```

```
df.columns
```

```
Index(['Category', 'Species', 'Weight', 'Height', 'Width', 'Length1',
       'Length2', 'Length3'],
      dtype='object')
```

```
y=df['Weight']
```

```
y.shape
```

```
(159,)
```

```
y
```

```
0      242.0
1      290.0
2      340.0
3      363.0
4      430.0
```

```
...
154     12.2
155     13.4
156     12.2
157     19.7
158     19.9
```

```
Name: Weight, Length: 159, dtype: float64
```

```
X=df[['Height', 'Width', 'Length1', 'Length2', 'Length3']]
```

```
X=df.drop(['Category', 'Species', 'Weight'], axis=1)
```

```
X.shape
```

```
(159, 5)
```

```
X
```

	Height	Width	Length1	Length2	Length3
0	11.5200	4.0200	23.2	25.4	30.0
1	12.4800	4.3056	24.0	26.3	31.2
2	12.3778	4.6961	23.9	26.5	31.1
3	12.7300	4.4555	26.3	29.0	33.5
4	12.4440	5.1340	26.5	29.0	34.0
...
154	2.0904	1.3936	11.5	12.2	13.4
155	2.4300	1.2690	11.7	12.4	13.5
156	2.2770	1.2558	12.1	13.0	13.8
157	2.8728	2.0672	13.2	14.3	15.2

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

```
X_train,X_test,y_train,y_test=train_test_split(X,y,train_size=0.3,random_state=2529)
```

```
X_train.shape,X_test.shape,y_train.shape,y_test.shape
```

```
((47, 5), (112, 5), (47,), (112,))
```

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

```
model=LinearRegression()
```

```
model.fit(X_train,y_train)
```

```
LinearRegression()
```

```
y_pred=model.predict(X_test)
```

```
y_pred.shape
```

```
(112,)
```

```
y_pred
```

```
array([ 454.55852783,  501.62065542,  105.01448954,  966.67007124,
        189.7938086 ,  222.59152861,  316.95390496,  388.02371242,
        367.33038424, -145.16437982, -116.63030886,  465.77132212,
        152.22389181,  823.88712795,  657.70503072,  303.60288735,
        574.51835823,  217.46478227,  697.45053723,   96.96099227,
        160.319581 , -93.24609517,  705.84753184,  771.68658418,
        789.84354787, -167.10961488,  322.26711645,  685.54620288,
```

```

348.66042548, 810.69556254, 527.46240123, 566.46955442,
482.37581402, 631.18394845, -166.31662785, 925.36056971,
790.50941397, -50.67435328, 301.28230598, 904.1537953 ,
610.0083557 , 952.80231564, 669.15043599, 708.23951091,
612.43546859, 398.23590774, 516.69860211, -63.91448611,
626.15057621, 392.80144848, 15.93709478, 878.92950245,
-192.67436056, 838.44578247, 849.5704718 , 547.9151785 ,
723.1722745 , 128.84755783, 174.6204508 , 837.46665875,
203.33530662, 22.61692047, 657.69570274, -123.42764091,
384.71311997, 181.3504614 , 605.56994951, 900.48300124,
129.04690521, 651.85457228, 899.9223446 , 620.15053308,
301.39898719, 273.83273043, 974.02611645, 360.50357818,
435.57154117, 799.30714783, 896.8931598 , 650.51137858,
194.07227238, 405.80635108, 257.48745646, 106.07592508,
806.73760198, 680.15718202, 672.80808683, 314.70090367,
235.76707604, 131.9914235 , 168.92348598, 208.3385226 ,
-88.33481149, 385.28299261, 486.05211509, 41.40108965,
-151.94209275, 92.00550969, 254.32222316, 458.87851844,
595.41604675, 230.03800049, 498.518774 , 571.37905011,
747.97493448, 184.42584718, 34.00800219, 1357.07163051,
721.97660211, 28.72675985, -31.9955029 , 625.86505533])

```

```
from sklearn.metrics import mean_squared_error,mean_absolute_error,mean_absolute_percentag
```

```
mean_squared_error(y_test,y_pred)
```

```
18825.739815999874
```

```
mean_absolute_error(y_test,y_pred)
```

```
104.80397583097229
```

```
mean_absolute_percentage_error(y_test,y_pred)
```

```
1.6878458422011842
```

```
r2_score(y_test,y_pred)
```

```
0.830889637900902
```

```
df_new = df.sample(1)
```

```
df_new
```

	Category	Species	Weight	Height	Width	Length1	Length2	Length3
96	3	Perch	225.0	7.293	3.723	22.0	24.0	25.5

```
X_new=df_new[['Height','Width','Length1','Length2','Length3']]
```

```
X_new.shape
```

```
(1, 5)
```

```
y_pred_new=model.predict(X_new)
```

```
y_pred_new
```

```
array([283.2819186])
```

