```
1 import pandas as pd
2 import numpy as np
3 import matplotlib.pyplot as plt
4 import seaborn as sns
```

```
1 data = pd.read_csv("airquality.csv")
2 data
```

	Unnamed: 0	Ozone	Solar.R	Wind	Temp	Month	Day
0	1	41.0	190.0	7.4	67	5	1
1	2	36.0	118.0	8.0	72	5	2
2	3	12.0	149.0	12.6	74	5	3
3	4	18.0	313.0	11.5	62	5	4
4	5	NaN	NaN	14.3	56	5	5
...	...	...	...	...	...	...	...
148	149	30.0	193.0	6.9	70	9	26
149	150	NaN	145.0	13.2	77	9	27
150	151	14.0	191.0	14.3	75	9	28
151	152	18.0	131.0	8.0	76	9	29
152	153	20.0	223.0	11.5	68	9	30

153 rows × 7 columns

```
1 data.isnull().sum()
```

```
Unnamed: 0     0
Ozone         37
Solar.R        7
Wind           0
Temp           0
Month          0
Day            0
dtype: int64
```

```
1 ozone_mean = np.mean(data['Ozone'])
2 ozone_mean
```

```
42.12931034482759
```

```
1 data['Ozone'] = data['Ozone'].fillna(ozone_mean)
2 data['Ozone']
```

```
0      41.00000
```

```
1      36.00000
2      12.00000
3      18.00000
4      42.12931
         ...
148    30.00000
149    42.12931
150    14.00000
151    18.00000
152    20.00000
Name: Ozone, Length: 153, dtype: float64
```

```
1 data.isnull().sum()
```

```
Unnamed: 0    0
Ozone         0
Solar.R       7
Wind          0
Temp          0
Month         0
Day           0
dtype: int64
```

```
1 solar_mean = np.mean(data['Solar.R'])
2 solar_mean
```

```
185.93150684931507
```

```
1 data['Solar.R'] = data['Solar.R'].fillna(solar_mean)
2 data.isnull().sum()
```

```
Unnamed: 0    0
Ozone         0
Solar.R       0
Wind          0
Temp          0
Month         0
Day           0
dtype: int64
```

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

```
1 x = data.iloc[:,3:5]
2
```

```
1 y = data['Ozone']
2 y
```

```
0      41.00000
1      36.00000
2      12.00000
3      18.00000
```

```
4      42.12931
         ...
148    30.00000
149    42.12931
150    14.00000
151    18.00000
152    20.00000
Name: Ozone, Length: 153, dtype: float64
```

```
1 x_train,x_test,y_train,y_test = train_test_split(x,y,test_size = 0.25, rando
```

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

```
1 train_model = LinearRegression().fit(x_train,y_train)
2 x_predict = train_model.predict(x_train)
3 y_predict = train_model.predict(x_test)
4 y_predict
```

```
array([67.94115909, 67.1930407 , 59.35882048, 67.21302389, 38.24705838,
       42.37468509, 38.94854934, 53.08078661, -0.99731439, 60.91500683,
       51.22415933,  2.4221603 , 36.08999015, 49.94843074, 15.40592231,
       11.26497348, 64.05402376, 34.52714275, 18.11724498, 26.40580372,
       24.10816004, 45.22658321, 24.83629525, 33.25141416, 71.34731166,
       77.90580328, 35.09471932, 55.93934579, 54.08271851, 25.69099063,
       52.09217684, 49.65465086, 47.08321049, 44.95278652, 28.262431  ,
       39.23566815, 54.64363402, -2.57348392, 53.37456649])
```

```
1 y_train
```

```
136     9.00000
117    73.00000
119    76.00000
17      6.00000
46     21.00000
         ...
8       8.00000
73     27.00000
144    23.00000
118    42.12931
99     89.00000
Name: Ozone, Length: 114, dtype: float64
```

```
1 from sklearn.metrics import mean_squared_error
2 from sklearn.metrics import r2_score
3 training_mse = mean_squared_error(y_train,x_predict)
4 print("Trainging MSE = " ,training_mse)
5 trmse = training_mse**0.5
6 print("Training RMSE = ",trmse)
7 print("Training R2 = ",r2_score(y_train,x_predict))
```

```
Trainging MSE =  428.87673991085745
Training RMSE =  20.70933943685451
Training R2 =  0.5039685716310263
```

```
1 testing_mse = mean_squared_error(y_test,y_predict)
2 print("Testing MSE = " ,testing_mse)
3 termse = testing_mse**0.5
4 print("Training RMSE = ",termse)
5 print("Training R2 = ",r2_score(y_test,y_predict))
```

```
Testing MSE =  521.3771679750005
Training RMSE =  22.833684940784316
Training R2 =  0.20484414901875603
```

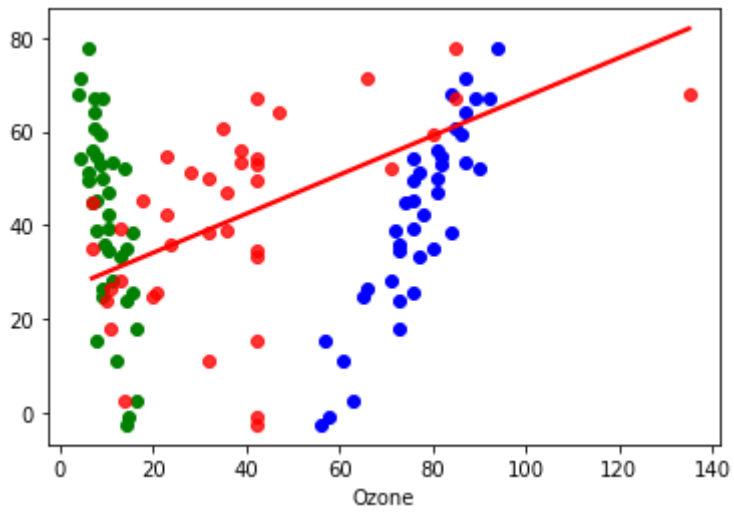
```
1 x_train.shape
```

```
(114, 2)
```

```
1
2
```

```
1 plt.scatter(x_test['Temp'],y_predict,color = 'blue')
2 plt.scatter(x_test['Wind'],y_predict,color = 'green')
3 #plt.scatter(x_test['Solar.R'],y_predict,color = 'orange')
4 sns.regplot(y_test,y_predict,ci=None,color ='red');
```

```
/home/pvg/anaconda3/lib/python3.9/site-packages/seaborn/_decorators.py:36: FutureWarn
  warnings.warn(
```




```
1
```