In [ ]:
```
#import libraries
import numpy as np
import pandas as pd

#import dataset
dataset = pd.read_csv('datasets/studentscores.csv')
X = dataset.iloc[ : , :1].values
Y = dataset.iloc[ : , 1].values

#check for missing data
from sklearn.impute import SimpleImputer

imputer = SimpleImputer(missing_values=np.nan, strategy='mean')
imputer = imputer.fit(X)
X = imputer.transform(X)

#split dataset
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=1/4, random_stat
```

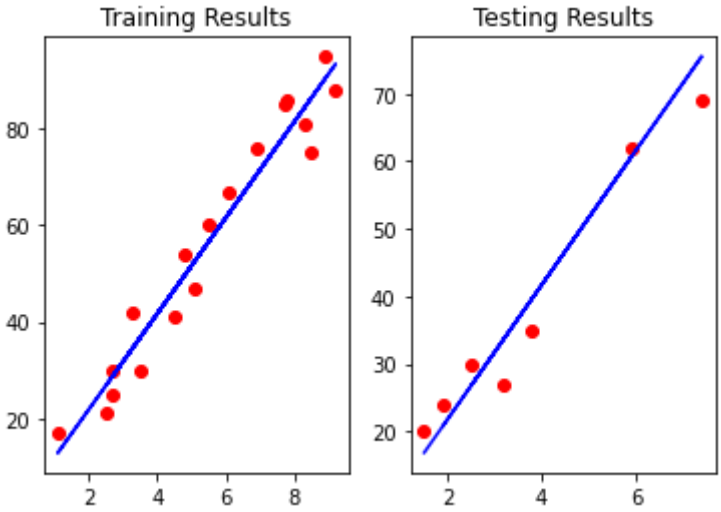
In [ ]:
```
#fit regressor
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor = regressor.fit(X_train, Y_train)
```

In [ ]:
```
#predict result
Y_pred = regressor.predict(X_test)
```

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

#visualise results
fit, (ax1, ax2) = plt.subplots(1, 2)
#training
ax1.title.set_text('Training Results')
ax1.scatter(X_train, Y_train, color = 'red')
ax1.plot(X_train, regressor.predict(X_train), color = 'blue')
#testing
ax2.title.set_text('Testing Results')
ax2.scatter(X_test, Y_test, color = 'red')
ax2.plot(X_test, regressor.predict(X_test), color = 'blue')
```

Out[ ]:
```
[<matplotlib.lines.Line2D at 0x16a21749220>]
```


Training Results
Testing Results
80
60
40
20
2
4
6
8
70
60
50
40
30
20
2
4
6