

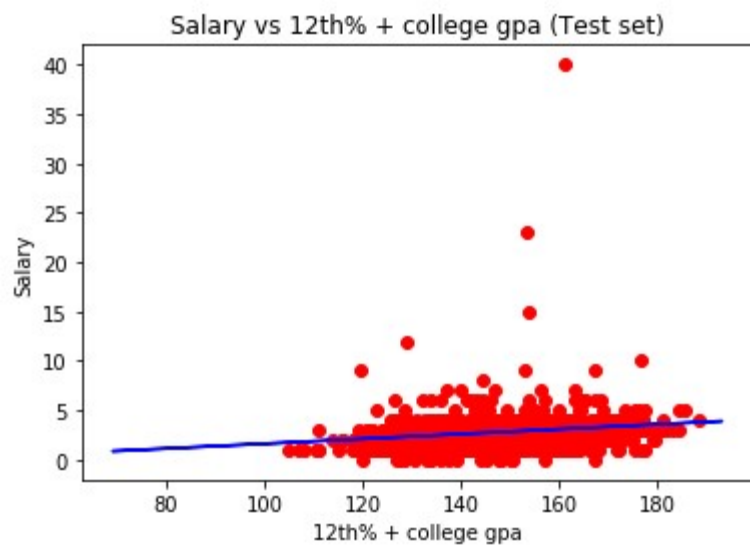
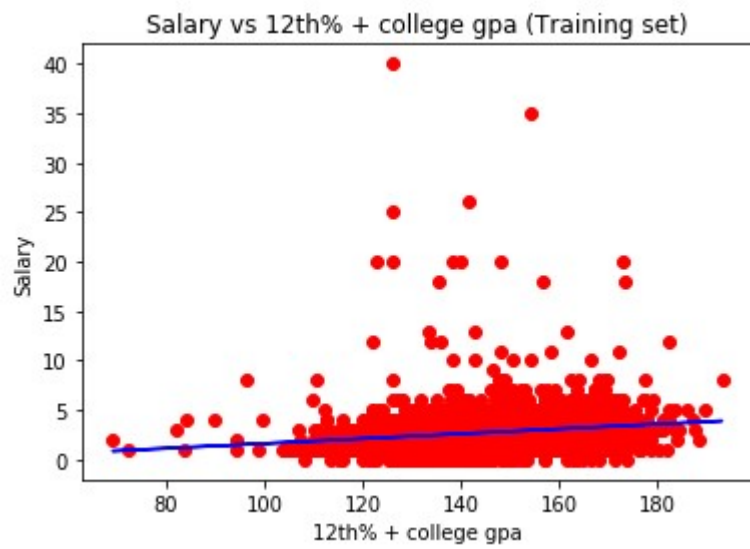
```

In [103]: import matplotlib.pyplot as plt
...: import pandas as pd
...:
...: df=pd.read_csv('G:\Data Analysis\output.csv')
...: df=df.dropna()
...:
...: df["normalised gpa"]=df.t12percentage + df.collegeGPA
...: X = df.iloc[:,[38]].values
...: y = df.iloc[:, 1].values
...: for i in range(len(y)):
...:     y[i]=int(y[i])
...:
...:
...:
...: 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_state = 0)
...:
...:
...: from sklearn.linear_model import LinearRegression
...: regressor = LinearRegression()
...: regressor.fit(X_train, y_train)
...: c=0
...: y_pred = regressor.predict(X_test)
...: for i in range(len(y_pred)):
...:     y_pred[i]=int(y_pred[i])
...:     y_pred[i]=(y_pred[i]<=y_test[i]+2 and y_pred[i]>=y_test[i]-2)
...:     if(y_pred[i]):
...:         c+=1
...:
...:
...: acc=float(c/len(y_test))

In [104]: acc
Out[104]: 0.9325153374233128

In [105]: plt.scatter(X_train, y_train, color = 'red')
...: plt.plot(X_train, regressor.predict(X_train), color = 'blue')
...: plt.title('Salary vs 12th% + college gpa (Training set)')
...: plt.xlabel('12th% + college gpa')
...: plt.ylabel('Salary')
...: plt.show()
...:
...:
...: plt.scatter(X_test, y_test, color = 'red')
...: plt.plot(X_train, regressor.predict(X_train), color = 'blue')
...: plt.title('Salary vs 12th% + college gpa (Test set)')
...: plt.xlabel('12th% + college gpa')
...: plt.ylabel('Salary')
...: plt.show()

```



In [106]: