Linear Regression practice set questions:

Write a MatLab (or a programming language of your choice) program to perform or

answer the following tasks:

1. Use Matlab command “load patients” to load patient self evaluation dataset.

2. You can also load the data from “patients.csv” file.

3. Use variables **Age, Gender, Height, Weight, Smoker, Location,**

**SelfAssessedHealthStatus** to build a linear regression model to predict the

systolic blood pressure.

4. What are the regression coefficients (thetas)?

5. How do you interpret those numbers?

6. If you need to identify one outlier record, which record is a potential outlier? How

do you reach this conclusion?

7. If you need to identify one or few useless features (independent variables or

predictors), which one(s) will you choose? Why do you reach this conclusion?