# Homework Assignment 2

## (100 points)

### Assigned Date: Thursday, February 11, 2016

### Due Date: Thursday, February 18, 2016

**Educational Goal**

Become familiar with cost function and hypothesis learning with linear regression with one variable.

**Requirements**

1. Data Description

Data downloaded at
http://www.cs.umb.edu/~ding/classes/438_638/homework/hwk2/hwk2data.txt

The file hwk2data.txt contains the population of a city (the first column) and the profit of a coffee shop in that city (the second column). A negative value for profit indicates a loss.

2. Plot the data using two given properties with X-axis for population and Y-axis for profit.
3. Fit the linear regression with one variable (two parameters $\theta_0$ and $\theta_1$), report the value of the cost function of the best model.
4. Choose two different learning rates of a very small value and a very large value, repeat Step 3. Discuss the impact of the learning rate.
5. The population of Boston in 2013 was 645,966. Please use the linear regression model you learned in Step 3 and predicate the profit a coffee shop in Boston in 2013.
6. **Provide a single runnable script for steps 2-5**. You will demonstrate your program to your TA Yahui Di in the Unix Lab from 3:00 to 4:00 PM on Monday Feb 23 or Tuesday Feb 24. You may bring your own laptop or use a machine in the Unix lab.

**Submission Requirements**

1.  Report: Prepare One PPT slide (saved as a PDF file) to explain the results of Steps 2 to 5 with your hwk2_firstname_lastname.pdf and submit the PDF file to the Blackboard.
2.  Zip your source code into one ZIP file and save it as hwk2_firstname_lastname.zip and submit the ZIP file to the Blackboard.
3.  Submit paper copies of your one slide PDF report and source code in class.
4.  In order to receive credits for this homework, you must submit hard copies (in class) and soft copies (to the Blackboard) and demonstrate your program to your TA in the Unix lab during the given time period.