**ENGR 2013: Engineering Analysis and Applications**

**Laboratory No. 10: Who Wants to Be a Billionaire? – Part III**

**Goal:** Understand how to use all the data available to you to fit a curve.

**Preliminary Work:** Replicate Example 5.6.4 in the textbook (Least Squares) in the textbook in Python. Also, for the stock you have selected for previous versions of this assignment, go to Yahoo Finance (or equivalent) and download the maximum amount of data available (preferably over 1,000 points). From Yahoo Finance (https://finance.yahoo.com/), type a company name, select Historical Data, set the time frame to Max, and download a csv file containing the data.

**Tasks:**

1. Implement the equation in Theorem 5.6.2 to compute a least-squares fit to your data for polynomials of degree 2, 4, 8 and 16. Plot the results and compare the accuracy of these results by computing the mean-square error (sum of the squares of the error divided by the number of points) between your fit and the actual data.
2. Using the multiple linear regression function in SKLearn, implement the same analysis as you did in (1) and show that you get the same answer. See *https://www.geeksforgeeks.org/linear-regression-python-implementation/* for more information on this. Note that your inputs should be powers of $x$ (the input signal raised to a power). For example, in the above tutorial, $x\_{2p}=x\_{2}^{p}$.
3. Determine the best order of the polynomial for this data using either (1) or (2).
4. Compare your ‘best’ least squares fit of the data to your two previous labs on this topic in a single plot. Comment on the similarities and differences. Did making use of the additional data points improve the fit of the model to the data?

**Summary:**

Are you a billionaire yet? Why not? What are the weaknesses of the least squares approach? What limits the power of these approaches on a complex problem like the prediction of a stock market price?