**ANALYSIS 1:**

**Introduction:**

Century 21 Ames (a real estate company) in Ames Iowa has commissioned us to answer a very important question with respect to their business. Century 21 Ames only sells houses in the NAmes, Edwards and BrkSide neighborhoods and would like to simply get an estimate of how the SalePrice of the house is related to the square footage of the living area of the house (GrLivArea) and if the SalesPrice (and its relationship to square footage) depends on which neighborhood the house is located in

**Data Description**:

The [Ames Housing dataset](http://www.amstat.org/publications/jse/v19n3/decock.pdf) was compiled by Dean De Cock for use in data science education. The data set ,train.csv describing the sale of individual residential property in Ames, Iowa from 2006 to 2010 contains 1460 observations and a large number of explanatory variables (23 nominal, 23 ordinal, 14 discrete, and 21 continuous) involved in assessing home values , with a total of 81 variables. Century 21 Ames only sells houses in the NAmes, Edwards and BrkSide neighborhoods and would like to simply get an estimate of how the SalePrice of the house is related to the square footage of the living area of the house (GrLivArea) and if the SalesPrice (and its relationship to square footage) depends on which neighborhood the house is located in. Hence the specific variables are SalesPrice( dependent variable ) is explained with living area of the house (**GrLivArea)** and **SalesPrice** (and its relationship to square footage) depends on the **neighborhood** ( three neighborhood of interest )

Further information can be found at <http://jse.amstat.org/v19n3/decock.pdf>

**Restatement of Problem:**

provide our client with a well written conclusion that quantifies the relationship between living area and sale price with respect to these three neighborhoods

**Build and Fit the Model:**

Build and fit a model that will answer this question, keeping in mind that realtors prefer to talk about living area in increments of 100 sq. ft. Provide your client with the estimate (or estimates if it varies by neighborhood) as well as confidence intervals for any estimate(s) you provide.

**Checking Assumptions :**

Residual Plots

Influential point analysis (Cook’s D and Leverage)

Make sure to address each assumption.

**Comparing Competing Models:**

Adj R2

Internal CV Press

**Parameters:**

Estimates

Interpretation

Confidence Intervals

**Conclusion:**

A short summary of the analysis.

**Appendix:**

Code and additional screen shot of plots and graphs