Assignment 2

LINEAR REGRESSION

Linear regression is a type of predictive analysis which uses the independent variable to predict the outcome of the dependent variable. It also tells which independent variables are significant predictors of the dependent variable.

It is represented by the equation

Y = a + bX

Where,

X is the independent variable

Y is the dependent variable

a, b are the intercepts

* Simple linear regression  
  1 dependent variable 1 independent variable

* [Multiple linear regression](http://www.statisticssolutions.com/data-analysis-plan-multiple-linear-regression/)  
  1 dependent variable 2+ independent variables

LOGISTICS REGRESSION

When there are two dependent variables then it is the appropriate regression analysis to conduct. In this regression, model fit is very important. The more the independent variables in the model the more in the amount of variance. So it might also lead to overfitting, which indeed reduces the generalizability of the model beyond the data on which the model is fit.

From the reading,

With the low p-value, the model is more meaningful and the model fits the data better. A larger p-value means that the confidence level (1-p) is low, we can say it means that the coefficients in the regression model are zero and the testing data will fall into overfitting model.