INF 552 Homework 2

Task 2:

Linear Model: $\hat{y} = \beta_0 + \beta_1 x$

$$\beta_1 = \frac{SS_{xy}}{SS_x} \qquad \beta_0 = \bar{y} - \beta_1 \bar{x}$$

Independent Variable (x): Odometer
Dependent Variable (y): Price

Mean of Odometer ($mean_x$): 36040.05155
Mean of Price ($mean_y$): 5988.14433

Sum of Odometer: 3495885
Sum of Price: 580850

Sum of Squares of Odometer (SS_x): 3907812841
Sum of Squares of Price (SS_y): 501731942
Sum of Squares of Odometer and Price (SS_{xy}): -58050289.72

$\beta_1 = SS_{xy} / SS_x$

= -58050289.72/3907812841

= -0.014854931

$\beta_0 = mean_y - \beta_1 * mean_x$

= 5988.14433 – (-0.014854931)* 36040.05155

= 6523.516805

Observation:

It is evident from the values of β_0 and β_1 that both are inversely proportional to each other since both are of opposite signs. So, it can be concluded that as the independent variable (Odometer) value increases, the dependent variable (Price) value decreases. Also, if the value of Odometer value decreases, the value of Price will then increase.

Task 3:

Mean of x: 25
Mean of y: 0.53264745

Sum of x: 1275
Sum of y: 27.16501996

Sum of x*x: 42925
Sum of y*y: 23.83524962
Sum of x*y: 373.6052432

SS_x: 11050
SS_y: 9.365870995
SS_{xy}: -305.5202559

B1: -0.027648892
B0: 1.223869749

Pearson Linear Correlation:

$$\frac{SS_{xy}}{\sqrt{SS_x SS_y}}$$

= -305.5202559/sqrt(11050*9.365870995)

= -0.949696384

Coefficient of Determination (R^2):

= (Pearson Linear Correlation)^2

= (-0.949696384)^2

= 0.901923222

Sum of (y-yMean)^2: 9.365870995

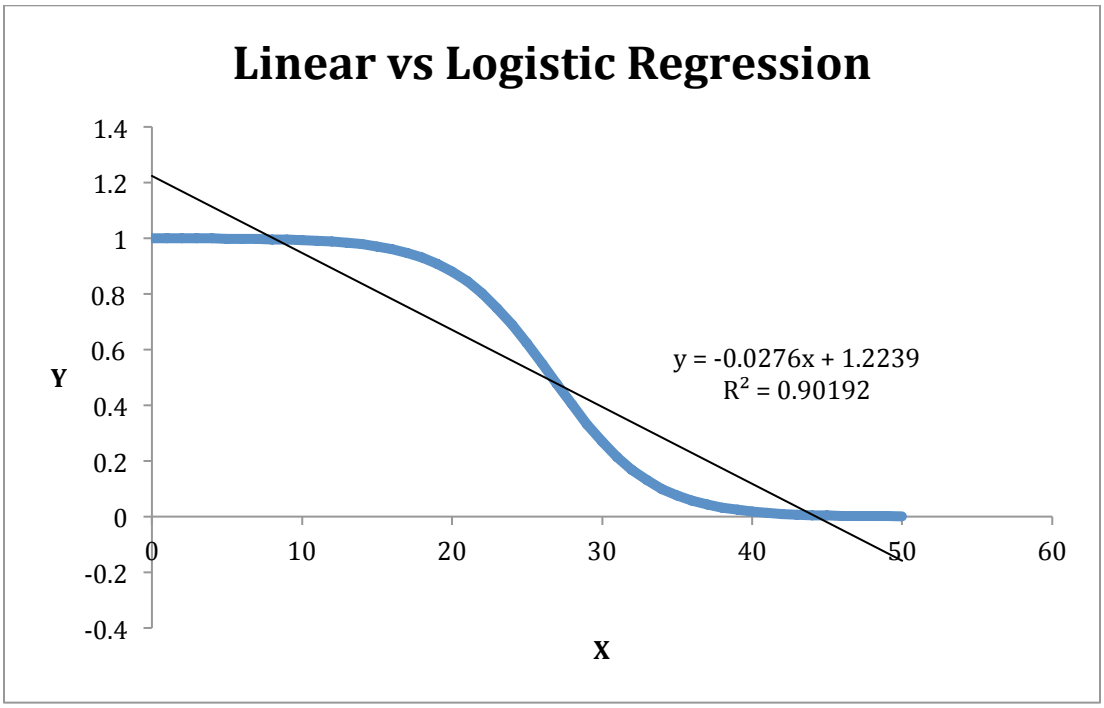
$$R^2 = 1 - \frac{SSE}{\sum (y_i - \bar{y})^2}$$

SSE = (1- R^2) * Sum of (y-yMean)^2

= (1- 0.901923222) * 9.365870995

= 0.918574454

Graph Plot:



Observation:

The linear regression model constantly decreases and the values near the end points go below 0 and beyond 1. But, the logistic regression model decreases gradually. They tend to approach 0 and 1 near the end points and the points lie within the interval [0,1]. The points in the middle do not decrease drastically unlike the ones in linear regression.