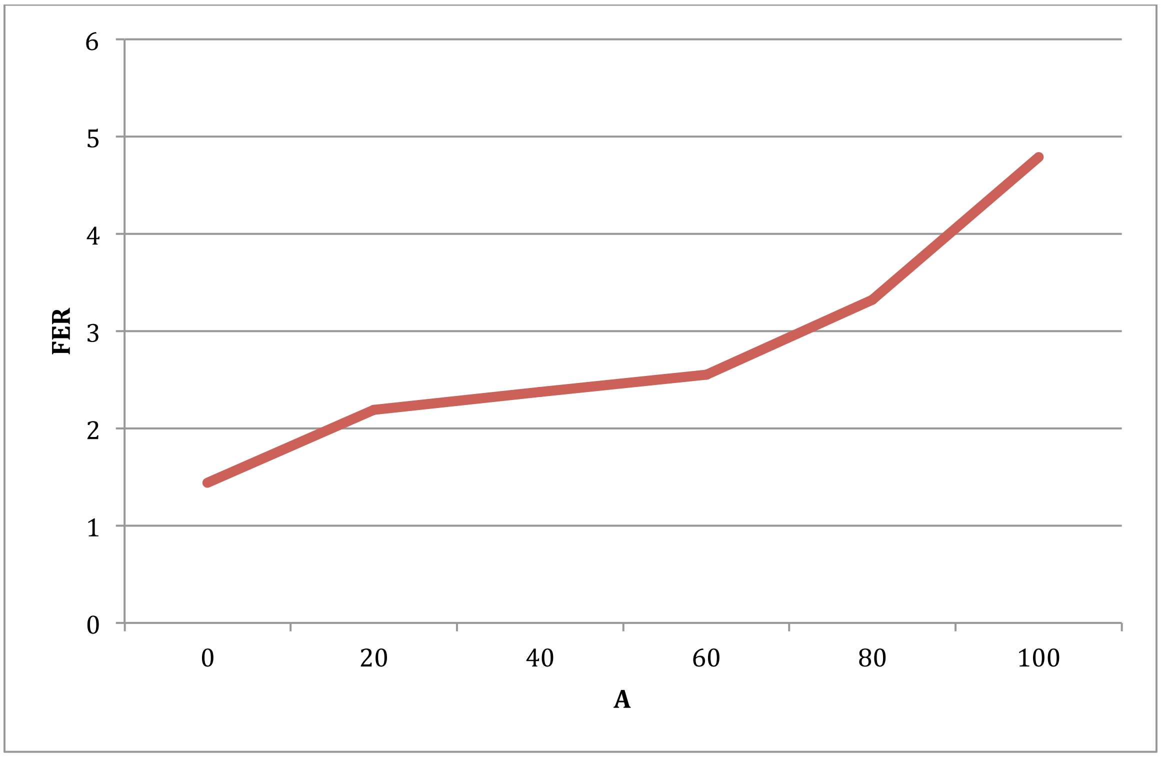
Name: Akond Rahman
Unity ID: aarahman
Date: March 22, 2016
Section: 001B

Answer to Question 12.13 (a):



Answer to Question 12.13 (b):
Adjusted R^2= 0.9715 for cubic model; from the plot it seems non-linear and adjusted R^2 is the best for the cubic model

Answer to Question 12.13 (c):
Both R^2 and Adjusted R^2 are the highest for cubic models, so cubic model will be the best choice

Answer to Question 12.13 (d):
The increase is not constant; in between 20 and 60 the rate of increase is different to that of 60-80, and also for 80-100. Weight might not be constant and depends on other factors as well (body type, breed)

Answer to Question 12.30 (a):
The three predictors account for 97.95% of the variation in ratings

Answer to Question 12.30 (b):
(R2/(1-R2)) * ($df_{error}/df_{regression}$) = $((0.979)^2 /(1-0.979^2)) * (496/3) = 3812.946$, this is smaller to the overall F statistics = 7925.829

Answer to Question 12.30 (c):
The F-statistic along with the measure of 12.30(b) tells that the three independent vriables are good predictors as they have large values.

Answer to Question 12.31 (a):
0.979566-0.895261 = 0.084305 , 8% reduced

Answer to Question 12.31 (b):
Age is not statistically significant, debt fraction is statistically significant

Answer to NBP-1 (a):
Model 1 and 5

Answer to NBP-1 (b):
Model 2, 3 and 4

Answer to NBP-1 (c):
Model 2 and 4

Answer to NB-SAS Problem -1(a):
[See attached output and SAS code for reference of the values used in answers]

54.5538
This value represents the regression sum square that is added for $\boldsymbol{\beta}_3$ if $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_4$, $\boldsymbol{\beta}_5$ are already added to the model

Answer to NB-SAS Problem -1(b):
$R(\boldsymbol{\beta}_4, \boldsymbol{\beta}_5 | \ \boldsymbol{\beta}_1 \ \boldsymbol{\beta}_2 \ \boldsymbol{\beta}_3)$
=1.812086+ 24.29437
This value represents the regression sum square that is added for $\boldsymbol{\beta}_4$ and $\boldsymbol{\beta}_5$ if $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$ are already added to the model

Answer to NB-SAS Problem -1(c):
Model 4 is nested in model 3
Model 3 is nested in model 1
Model 2 is nested in model 1
Model 4 is nested in model 1

The matrix **β−hat** for model 3 can be found from the matrix **β** of model 1 by adding zeros for co-efficient for $\boldsymbol{\beta}_4$ and **β**5. Similarly, to get model 2 and model 4 from model 1, zeros should be added for $\boldsymbol{\beta}_3$, and $\boldsymbol{\beta}_3$, $\boldsymbol{\beta}_4$, and $\boldsymbol{\beta}_5$ respectively for the matrix $\boldsymbol{\beta}$-hat for model 1.

Answer to NB-SAS Problem -1(d):
TS= 9.46, p-value < 0.0001
Decision = reject H0, at least one of $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, $\boldsymbol{\beta}_4$, and $\boldsymbol{\beta}_5$ is not null
At least one of the predictors is not null for y.
Individually, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$ are non-zero predictors for y, whereas, $\boldsymbol{\beta}_4$ and $\boldsymbol{\beta}_5$ is not. The problem with the individual analysis is the joint effect of two variables will not be observed.

Answer to NB-SAS Problem -1(e):
T.S = 14.50, p-value < 0.0001
Reject H0, at least one of the estimators $\boldsymbol{\beta}_4$, and $\boldsymbol{\beta}_5$ is non-zero.

Answer to NB-SAS Problem -1(f):
TS=16.75, p-value < 0.0001, reject H0, at least one of the estimators $\boldsymbol{\beta}_3$, $\boldsymbol{\beta}_4$, and $\boldsymbol{\beta}_5$ is non-zero

Answer to NB-SAS Problem -1(g):
$Y = -21.542883 + 3.418123x_1 - 0.024269x_2^2$

Answer to NB-SAS Problem -1(h):

Model adequacy criteria:
Global F-test: TS = 13.37, p-value < 0.0001
95% CIs include zero, so not statistically significant
R^2: 0.462, the model only considers 46.30% of the overall variability

Overall, does not satisfy all the adequacy criteria

Answer to NB-SAS Problem -1(i):
$Y = 81.7035 - 2.487 * x_1 + 0.0544x_1x_2$

Answer to NB-SAS Problem -1(j):
The models are similar because x_1 and x_2 are negatively correlated. This implies the increase in one will lead to decrease in another; so even though the models look different they imply the same correspondence.

The backward selection method is preferable as it has higher R^2 than the forward selection method. The R^2 value = 0.609. The 95% CIs for the backward method does not include zero, so the estimators are significant.