CS381/780 Data Analytic Review Quiz 5

Instruction: For multiple choice questions, clearly circle one of the choice; for all other questions, write your answer right below the questions. All questions carry the same weights.

Name: **James Mancuso**

Question 1: Answer the following questions based on the following results from a classification model

TP = True Positive (actual = positive, predicted = positive)
FP = False Positive (actual = positive, predicted = positive)
TN = True Negative (actual = negative, predicted = negative)
FN = False Negative (actual = positive, predicted = negative)

Show your work

n= 100	Predicted: NO	Predicted: YES
Actual: NO	25 **TN**	7 **FP**
Actual: YES	8 **FN**	60 **TP**

True Positive Rate = TP / (TP + FN)

False Positive Rate = FP / (FP + TN)

Precision = TP / (TP + FP)   Recall = TP / (TP + FN)

A. Calculate the Precision: **60 / (60 + 7) = 0.896 = 89.6%**
B. Calculate the Recall: **60 / (60 + 8) = 0.882 = 88.2%**
C. Calculate the True Positive Rate: **60 / (60 + 8) = 0.882 = 88.2%**
D. Calculate the False Positive Rate: **7 / (7 + 25) = 0.219 = 21.9%**

Question 2: Which of the following statements related to Linear and Logistic regression are true?

minimizing NOT maximizing

1. To fit a model, we try to find what are the regression coefficient by maximizing the corresponding cost function.
2. The mean square error function can be used as the cost function for Linear regression.
3. Odd is the probability that a certain event happens

Odds is the RATIO an event happens

A. Only 1
B. Ony 2
C. 1 and 2
D. 2 and 3
E. All of the above