horizontal line

**Team: Who’s Your Data?**

HW03: Tenderloin Crime

Cannady Michael, Gonzalez Sam, Lanier Aaron, Mehta Sanman, Molinari Diego, Provost Robyn, Sarmiento Linda

Tenderloin Crime

**DUE: 9th April 2017**

# ASSIGNMENT INSTRUCTIONS

Submit a single group report as a pdf file that is no longer than 10 pages; you may include an appendix if needed.

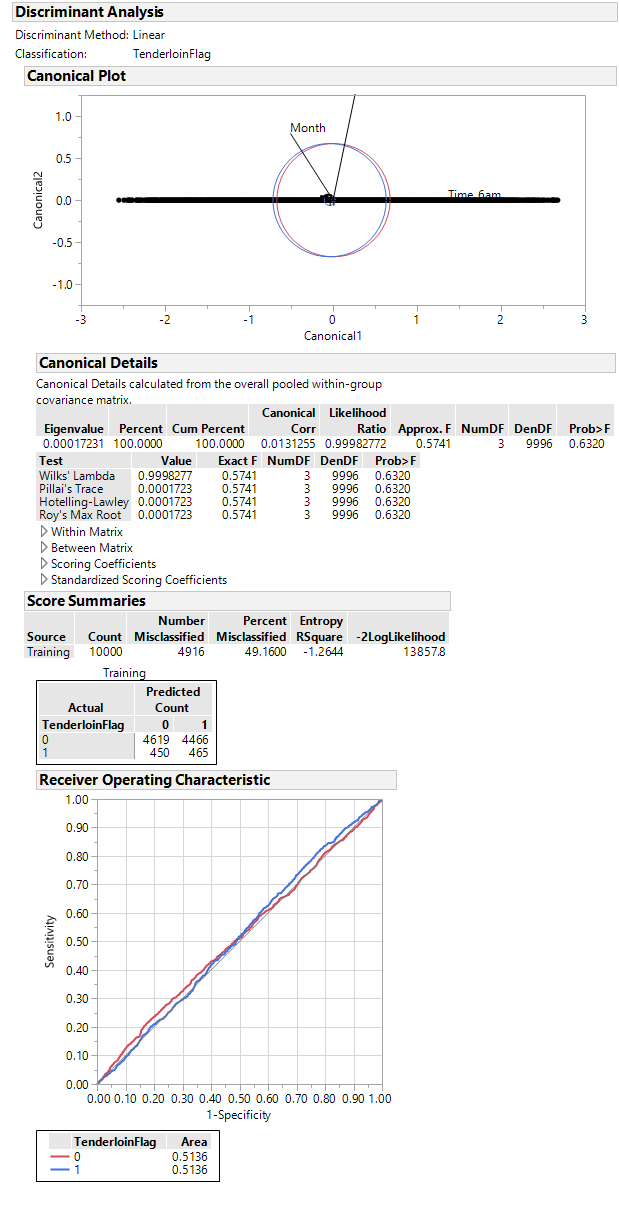
## Homework Questions

**With the San Francisco crime data:**

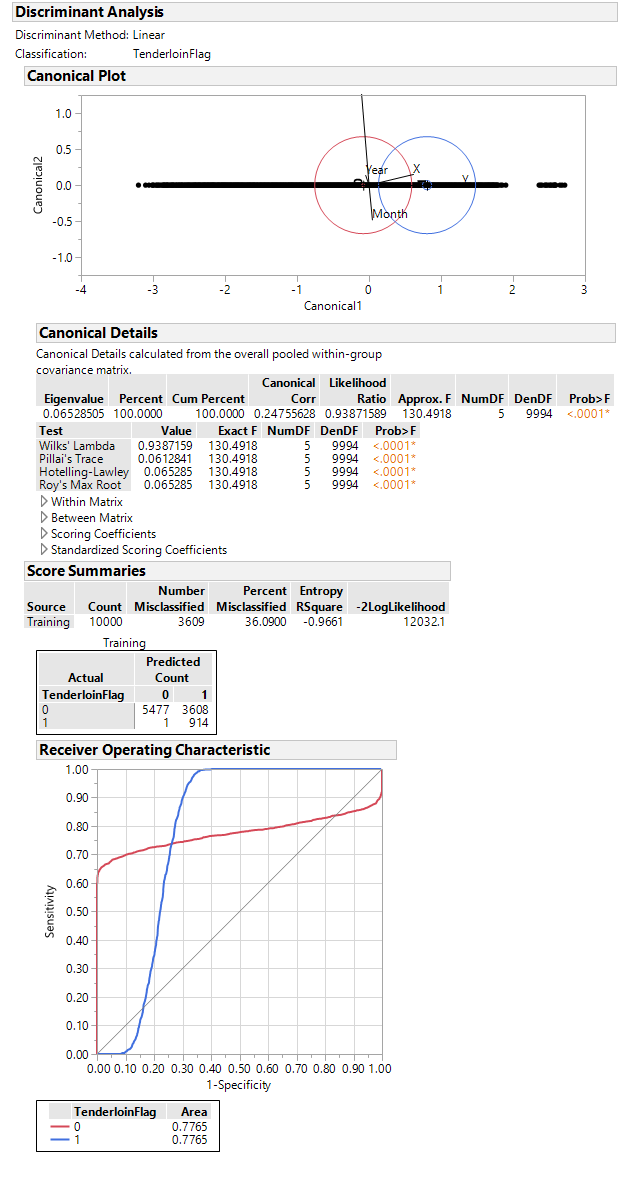
1. Use LDA to build a predictive model for whether incidents happen in the Tenderloin district or not. For predictor variables, use 'Year', 'Month', 'Time\_6am', 'X' and 'Y'.
2. According to cross-validation, what is the optimal probability threshold to use when predicting? For this optimal model, report confusion matrix, misclassification rate, sensitivity and specificity; also show an ROC curve and highlight the point on the curve that corresponds to your selected model.
3. Now use logistic regression as your predictive model, with the same numeric predictor variables as before but now also including the categorical predictor variable 'DayOfWeek'.
4. Again use cross-validation and report summaries of selected model; show ROC curve.

# Model Results

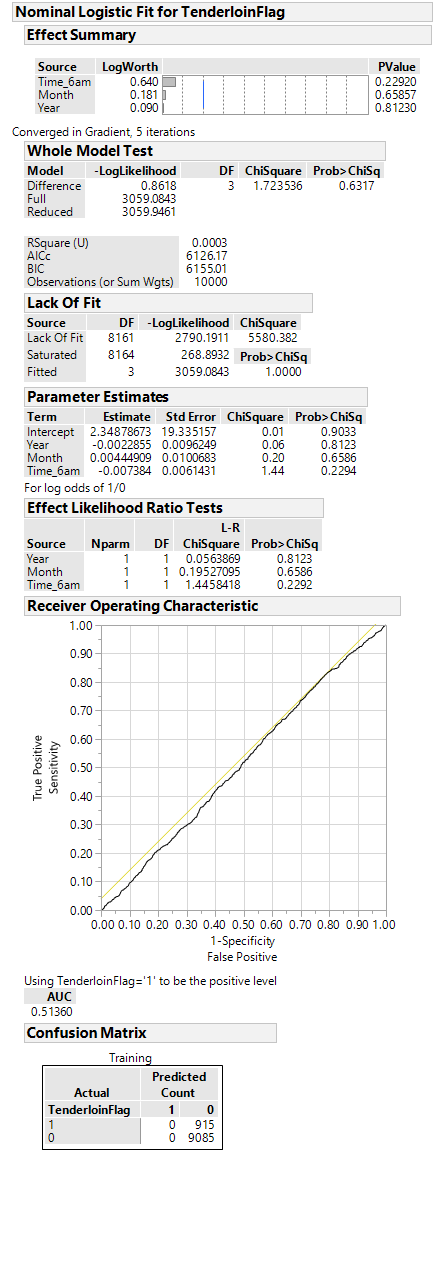
1. LDA to build predictive model for whether incidents happen in Tenderloin or not
   1. Without Location Data (X,Y)



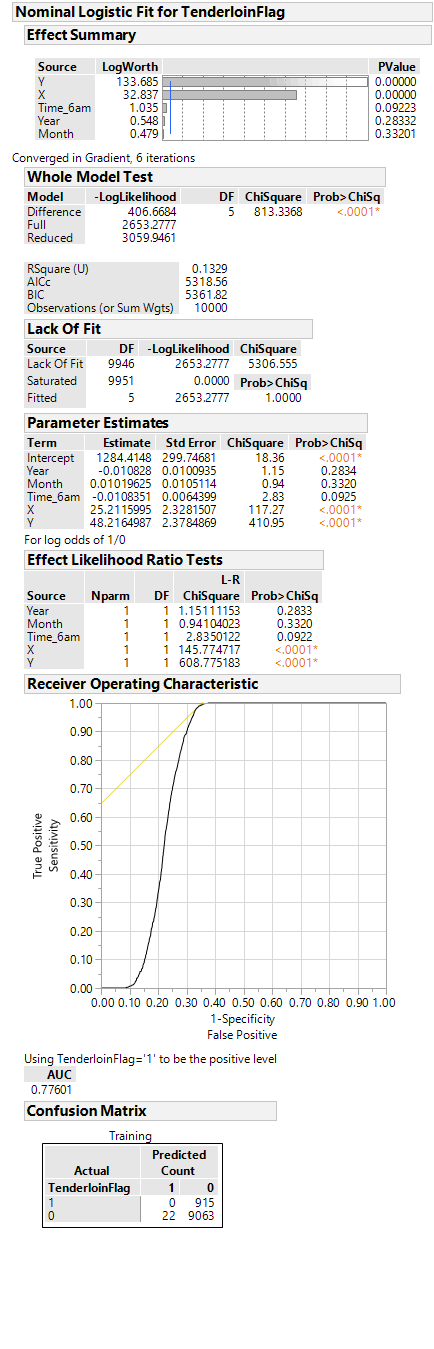
* 1. With Location Data



1. According to cross-validation, what is the optimal probability threshold to use when predicting? For this optimal model, report confusion matrix, misclassification rate, sensitivity and specificity; also show an ROC curve and highlight the point on the curve that corresponds to your selected model.
2. Now use logistic regression as your predictive model, with the same numeric predictor variables as before but now also including the categorical predictor variable 'DayOfWeek'.
   1. Without Location Data (X,Y)



* 1. With Location Data (X,Y)



1. Again use cross-validation and report summaries of selected model; show ROC curve.