What we've tried . . .

Data Preprocessing

1. **Used training and validation set to perform classification (LR, LDA, QDA, KNN) based on dummy variables of:**
   (a) SubjectRace (removed 'U' observations)
      i. Not much improvement over guessing 'N'
   (b) SubjectRace + SubjectGender (removed 'N/A' and 'U' observations)
      i. Not much improvement . . .
   (c) SubjectRace + SubjectGender + NumberOfOfficers (replaced NA with 1)
      i. Not much improvement . . .

Looked at FullNarratives for keywords associated with fatalities and found the indexes with those keywords to change prediction to fatality after using the *Just Say No* method

Submission History

5.19.20

1. Submission 1: 0.67619
   (a) Implemented 'Just Say No' Method
2. Submission 2: 0.79761
   (a) Implemented 'Just Say No' Method and also converted 180 "fatal" indexes rows (from looking at FullNarratives for keywords associated with higher fatalities) to "yes"
3. Submission 3: 0.8
   (a) Like above, but after a more thorough inspection of some indexes, removed some false positive "fatal" indexes