EDA strategy

* Describe data
  + Total count of incidents (by race, time)
    - Table
    - Incident by race, by year (Bar plot)
  + Sample means, standard deviations
    - Aggregate
    - By Year
    - By Geography
  + % of incidents (by threat level, allegedly armed vs. unarmed, by race, by fraction of US population)
* Distribution of each variable
  + Make sure the data makes sense
  + Make sure the variables are ach informative (i.e., appropriate distribution of observations across values)
* Correlations
  + By race, gender, state, age
    - With and without threat level, flee, signs of mental illness, unarmed
    - Allegedly armed/unarmed and threat level
    - Sns.pairplot
    - Scatterplots
    - Violinplots
  + Outcomes, by race (criminal charges, official disposition)