Proposal 1) Prescribing of Opiod among Medical Professionals.

Opioids, often prescribed as pain killers, are drugs frequently abused due to their addictive nature. Using the opioid prescriber summary filed in 2013, I could categorized each doctors/clients by speciality based on the percentages of drug prescription that are opoids. It is expected doctors/clients under the same speciality would have similar percentages of opoid prescripts. I could create categorize the profession of doctors based on level of opoid prescription. I could then determine if there are doctors who's percentage of opoid prescription was misaligned with their profession. This may indicate opoid abuse of the doctors or they're patients.

Data Set:
https://data.cms.gov/Medicare-Claims/Medicare-Part-D-Opioid-Prescriber-Summary-File-201/yb2j-f3fp

Proposal 2) Factors that Correlate with Heart Disease

Heart Disease is a serious problem in the US which is only getting worse. The link below has a database of heart disease set an integer between 0-4 indicating degree. A 0 indicates the absence of heart disease. This data set also contains multiple other information such as age, gender, whether they smoke, and level of exercise. Using this information, make a probabilistic predictive model to determine whether a person has heart disease. Possible a logical regression where it is only consider if a person has heart disease rather than the degree.

Data Set:
http://archive.ics.uci.edu/ml/datasets/Heart+Disease

Proposal 3) Exampling The Qualities of Good Movies

In this project, the movies in the data set will already be seperated in positive vs negative reviews. The purpose to this project is to determine factors which lead to a reviewer giving a positive or negative review. I could track keywords and phases which are more common among possible and negative reviews. Other than expected adjectives such as "sucks", "terrible", or "terrific", we can determine which aspects of a movie reviews tend to focus and which are more important.

Data Set:
http://ai.stanford.edu/~amaas/data/sentiment/