Topic 1: Predicting Recidivism likelihood from State of Ohio data (main idea)

This data set gives detailed demographic information on prisoners and their crimes as well as how long/ if they went back to prison within the first 1,000 days of being released. With this information I plan to look at different racial demographics that committed the same crime and see the likelihood differences of recidivism. Additionally, I plan to create a model that will predict based on the amount of time that has passed from release the likelihood of recidivism. I plan to implement the “Customer Churn Modeling” with a twist to analyze the epidemic of recidivism. I hope to build statistical models that give insights into this issue and hopefully also shed some like on how to help prevent or decrease the rates recidivism.

Data Sources: <https://catalog.data.gov/dataset/3-year-recidivism-for-offenders-released-from-prison>

Topic 2: Predicting Chicago crime locations

This project would involve creating a predictive model for criminal patterns based off the historical and current data of all crimes (except Homicide) in Chicago. The data set has records of the Chicago crimes committed from 2001 to present, where they happened and if there was an arrest.

Data Sources: <https://catalog.data.gov/dataset/crimes-2001-to-present-398a4>

Topic 3: Bike Sharing Models based on weather data

This is a data set that gives bike sharing statistics including detailed weather at the time the bike is picked up. I would create a model that predicted likelihoods of a person using bike a share based on the different weather conditions.

Data Sources:

<https://archive.ics.uci.edu/ml/datasets/bike+sharing+dataset>

<https://www.capitalbikeshare.com/system-data>