**Variable Definitions for the US Crime dataset for DATA473**

-- **state**: US state (identified by number)

-- **communityname**: community name for information only (string)

-- **population**: population for community: (numeric - decimal)

-- **householdsize**: mean people per household (numeric - decimal)

-- **agePct12t21**: percentage of population that is 12-21 in age (numeric - decimal)

-- **agePct12t29**: percentage of population that is 12-29 in age (numeric - decimal)

-- **agePct16t24**: percentage of population that is 16-24 in age (numeric - decimal)

-- **agePct65up**: percentage of population that is 65 and over in age (numeric - decimal)

-- **pctUrban**: percentage of people living in areas classified as urban (numeric - decimal)

-- **medIncome**: median household income (numeric - decimal)

-- **pctWWage**: percentage of households with wage or salary income in 1989 (numeric - decimal)

-- **pctWFarmSelf**: percentage of households with farm or self-employment income in 1989 (numeric - decimal)

-- **pctWInvInc**: percentage of households with investment / rent income in 1989 (numeric - decimal)

-- **pctWSocSec**: percentage of households with social security income in 1989 (numeric - decimal)

-- **pctWPubAsst**: percentage of households with public assistance income in 1989 (numeric - decimal)

-- **pctWRetire**: percentage of households with retirement income in 1989 (numeric - decimal)

-- **medFamInc**: median family income (differs from household income for non-family households) (numeric - decimal)

-- **perCapInc**: per capita income (numeric - decimal)

-- **PctPopUnderPov**: percentage of people under the poverty level (numeric - decimal)

-- **PctLess9thGrade**: percentage of people 25 and over with less than a 9th grade education (numeric - decimal)

-- **PctNotHSGrad**: percentage of people 25 and over that are not high school graduates (numeric - decimal)

-- **PctBSorMore**: percentage of people 25 and over with a bachelors degree or higher education (numeric - decimal)

-- **PctUnemployed**: percentage of people 16 and over, in the labor force, and unemployed (numeric - decimal)

-- **PctEmploy**: percentage of people 16 and over who are employed (numeric - decimal)

-- **PctEmplManu**: percentage of people 16 and over who are employed in manufacturing (numeric - decimal)

-- **PctEmplProfServ**: percentage of people 16 and over who are employed in professional services (numeric - decimal)

-- **TotalPctDiv**: percentage of population who are divorced (numeric - decimal)

-- **PersPerFam**: mean number of people per family (numeric - decimal)

-- **PctFam2Par**: percentage of families (with kids) that are headed by two parents (numeric - decimal)

-- **PctKids2Par**: percentage of kids in family housing with two parents (numeric - decimal)

-- **PctYoungKids2Par**: percent of kids 4 and under in two parent households (numeric - decimal)

-- **PctTeen2Par**: percent of kids age 12-17 in two parent households (numeric - decimal)

-- **PctWorkMomYoungKids**: percentage of moms of kids 6 and under in labor force (numeric - decimal)

-- **PctWorkMom**: percentage of moms of kids under 18 in labor force (numeric - decimal)

-- **PctLargHouseFam**: percent of family households that are large (6 or more) (numeric - decimal)

-- **PctLargHouseOccup**: percent of all occupied households that are large (6 or more people) (numeric - decimal)

-- **PersPerOccupHous**: mean persons per household (numeric - decimal)

-- **PersPerOwnOccHous**: mean persons per owner occupied household (numeric - decimal)

-- **PersPerRentOccHous**: mean persons per rental household (numeric - decimal)

-- **PctPersOwnOccup**: percent of people in owner occupied households (numeric - decimal)

-- **PctPersDenseHous**: percent of persons in dense housing (more than 1 person per room) (numeric - decimal)

-- **PctHousOccup**: percent of housing occupied (numeric - decimal)

-- **PctHousOwnOcc**: percent of households owner occupied (numeric - decimal)

-- **PctVacantBoarded**: percent of vacant housing that is boarded up (numeric - decimal)

-- **PctVacMore6Mos**: percent of vacant housing that has been vacant more than 6 months (numeric - decimal)

-- **MedYrHousBuilt**: median year housing units built (numeric - decimal)

-- **PctHousNoPhone**: percent of occupied housing units without phone (in 1990, this was rare!) (numeric - decimal)

-- **PctWOFullPlumb**: percent of housing without complete plumbing facilities (numeric - decimal)

-- **OwnOccLowQuart**: owner occupied housing - lower quartile value (numeric - decimal)

-- **OwnOccMedVal**: owner occupied housing - median value (numeric - decimal)

-- **OwnOccHiQuart**: owner occupied housing - upper quartile value (numeric - decimal)

-- **RentLowQ**: rental housing - lower quartile rent (numeric - decimal)

-- **RentMedian**: rental housing - median rent (Census variable H32B from file STF1A) (numeric - decimal)

-- **RentHighQ**: rental housing - upper quartile rent (numeric - decimal)

-- **MedRent**: median gross rent (Census variable H43A from file STF3A - includes utilities) (numeric - decimal)

-- **MedRentPctHousInc**: median gross rent as a percentage of household income (numeric - decimal)

-- **MedOwnCostPctInc**: median owners cost as a percentage of household income - for owners with a mortgage (numeric - decimal)

-- **MedOwnCostPctIncNoMtg**: median owners cost as a percentage of household income - for owners without a mortgage (numeric - decimal)

-- **NumInShelters**: number of people in homeless shelters (numeric - decimal)

-- **NumStreet**: number of homeless people counted in the street (numeric - decimal)

-- **PctSameHouse85**: percent of people living in the same house as in 1985 (5 years before) (numeric - decimal)

-- **PctSameCity85**: percent of people living in the same city as in 1985 (5 years before) (numeric - decimal)

-- **PctSameState85**: percent of people living in the same state as in 1985 (5 years before) (numeric - decimal)

-- **LandArea**: land area in square miles (numeric - decimal)

-- **PopDens**: population density in persons per square mile (numeric - decimal)

-- **PctUsePubTrans**: percent of people using public transit for commuting (numeric - decimal)

-- **ViolentCrimesPerPop**: total number of violent crimes per 100K popuation (numeric - decimal) GOAL attribute (to be predicted)