Communities and crime

Prediction of violent crime in the USA

Marie Lontsie Zanmene Yanis Bosch

19th May, 2024

Outline

The dataset

- Data sources:
  - Socio-economic data from the 1990 US Census
  - Law enforcement data from the 1990 US LEMAS survey
  - Crime data from the 1995 FBI UCR
- Creator: Michael Redmond, La Salle University, Philadelphia
- Date: 13th July, 2009

The dataset

- Size: 1994 rows, 128 columns
- Example attributes: police officers per 100K population, median rent,...
- Goal: Prediction of violent crime in the USA

Preprocessing

Column Name	Missing values(%)	Column Name	Missing values(%)
PolicReqPerOffic	84.002	PolicAveOTWorked	84.002
PolicPerPop	84.002	RacialMatchCommPol	84.002
PctPolicWhite	84.002	PctPolicBlack	84.002
PctPolicHisp	84.002	PctPolicAsian	84.002
PctPolicMinor	84.002	OfficAssgnDrugUnits	84.002
NumKindsDrugsSeiz	84.002	LemasSwFTFieldPerPop	84.002
LemasTotReqPerPop	84.002	LemasSwFTFieldOps	84.002
LemasSwFTPerPop	84.002	PolicCars	84.002
PolicOperBudg	84.002	LemasPctPolicOnPatr	84.002
LemasGangUnitDeploy	84.002	LemasSwornFT	84.002
PolicBudgPerPop	84.002	LemasTotalReq	84.002
OtherPerCap	0.0502		