Metadata

Historically, crime rates in the United States have experienced many fluctuations over the years, varying by the type of crime, geographic location and other factors. However, these crime rates, with the high level of public attention and scrutiny they receive, still remain a subject of significant concern given their large impact on communities and society as a whole. The statistics carry great importance in managing public safety and security by drawing awareness towards the prevalence of criminal activities in a given area.

This report aims to draw observations regarding the correlation between crime rates and socioeconomic factors in the United States, such as income, employment rates, and education attainment, whilst factoring in the distinction between correlation and causation. Effective crime prevention and reduction strategies can be employed by addressing underlying such socio economic issues. To explore the relationship between the multiple contributing factors, we performed an analysis on four respective datasets, each containing a wide variety of data values which we filter according to their relevance to our project aim.

The datasets used in this report include:

- Dataset1 - Unemployment level data obtained from U.S.DEPARTMENT OF AGRICULTURE (USDA)
- Dataset2 - Education attainment level data obtained from the"National Center for Education Statistics". The Institute of Education Sciences (IES) is the statistics, research, and evaluation arm of the U.S. Department of Education.
- Dataset3 - Income data obtained from 'The Bureau of Economic Analysis' (BEA)
- Dataset4 - Crime data obtained from FBI: UCR.

Dataset 1: Our first data set was retrieved from the official U.S.DEPARTMENT OF AGRICULTURE (USDA), Economic Research Service. The original source of data comes from the U.S. Department of Labor, Bureau of Labor Statistics and U.S. Department of Commerce, Bureau of the Census. The Economic Research Service collected the data and prepared the final table (3277 rows x 100 columns).

Dataset 2: This dataset is obtained from the "National Center for Education Statistics". The Institute of Education Sciences (IES) is the statistics, research, and evaluation arm of the U.S. Department of Education. These statistics were originally sourced from the U.S. Department of Commerce, Census Bureau across from 2013 - 2019 American Community Survey (ACS). This dataset contains the total population of each state over the age of 25 years of age and corresponding percentage of high school completion and bachelor degrees completion.

Dataset 3: The data was originally collected through a combination of surveys, data sources, and administrative records, gathered from federal agencies such as the U.S. Census Bureau and the Bureau

of Labor Statistics (BLS). Intercensal population statistics were produced using the 'Census Bureau Das Gupta' method on the Census Bureau decennial counts for 2010 and 2020 to create consistent time series that are used to prepare per capita personal income statistics. The method was utilised to account for the unreleased official Census population data for particular years.

Dataset 4: The provided dataset contains information regarding crime rates and totals of states across the United States from 1960 to 2019. In this dataset, there are 21 variables. The crime types are divided into two main categories: property and violent crime. Property crimes refer to burglary, larceny, and motor-related crime while violent crimes refer to assault, murder, rape, and robbery.