**Business Questions to Answer:**

* What is the likelihood of arrest based on primary type of crime, beat, district, community area?
* What time of day do most crimes occur? By type?
* What are the highest crime types per location (district, community area, etc.)?
* Do crimes typically occur in more private or public areas (location description)?
* Potential: Show geographical location of crimes

**Introduction and Business Understanding:** The data is extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system.

***Business Objectives:***

**Data Understanding and Preparation:** CLEAR reported a total of 265,462 crimes in 2016. The original data set contained 23 columns.

***Unnecessary Columns:*** The “Year” column was removed from our dataset, as all records contained a value of 2016 within the column. Additionally, several other columns were removed since they would not be beneficial to our model nor help us answer our business objectives ("Unnamed: 0","Case Number", "IUCR", "X Coordinate", "Y Coordinate", "Updated On", "Block", and "FBI Code").

***Data Types:*** The default data types for "District", "Ward", and "Community Area" were float. However, after a review of the data, it was determined that an integer data type is more appropriate and was updated within the dataset. The default data type for the “Date” column was object. A review of the data shows that this column includes a date and timestamp, therefore this field was updated to a [datetime] datatype.

***Missing Values:*** 14,189 records were missing several values within the “Latitude”, “Longitude” and “Location”columns and 783 records were missing a “Location Description”. These records were dropped from our dataset, as they made up a small portion of the population (5.55%) and there was no easy way to decide what their values should be. After dropping these records, the number of rows within the dataset was reduced to 250,732.

***Binning:***

***Test/Training:***

**Initial Assumptions:**

**References:** https://www.kaggle.com/datasets/currie32/crimes-in-chicago?select=Chicago\_Crimes\_2012\_to\_2017.csv