

# STA 6923 Homework 1, Fall 2021

Due date: Tuesday, August 31st by 11:59 pm

In this problem, first read some introduction here and then you answer all the questions below.

Crime is an international concern, but it is documented and handled in very different ways in different countries. In the United States, violent crimes and property crimes are recorded by the Federal Bureau of Investigation (FBI). Additionally, each city documents crime, and some cities release data regarding crime rates. The city of Chicago, Illinois releases [crime data from 2001 onward online](#).

Chicago is the third most populous city in the United States, with a population of over 2.7 million people. The city of Chicago is shown in the map below, with the state of Illinois highlighted in red.



Please download the file [ChicagoCrime.csv](#) for this problem and import it into your Rstudio. In this data set, only “theft” data was extracted from the database and also not all the columns were included. If you want to see the full data set, you can download it in [here](#). Here is a list of descriptions of the variables:

- ID: a unique identifier for each observation.
- Date: the date the crime occurred.
- Description: describe what kind of theft crime occurred.
- LocationDescription: the location where the crime occurred.

- Arrest: whether or not an arrest was made for the crime (TRUE if an arrest was made, and FALSE if an arrest was not made).
- Domestic: whether or not the crime was a domestic crime, meaning that it was committed against a family member (TRUE if it was domestic, and FALSE if it was not domestic).
- CommunityArea: the community area in which the crime occurred. Since the 1920s, Chicago has been divided into what are called “community areas”, of which there are now 77. The community areas were devised in an attempt to create socially homogeneous regions.
- Year: the year in which the crime occurred.
- Latitude: the latitude of the location at which the crime occurred.
- Longitude: the longitude of the location at which the crime occurred.

Read the dataset into R, using the `read.csv` function, and call the data frame “CrimeData”. Remember to navigate to the directory on your computer containing the file `ChicagoCrime.csv` first. It may take a few minutes to read in the data, since it is pretty large.