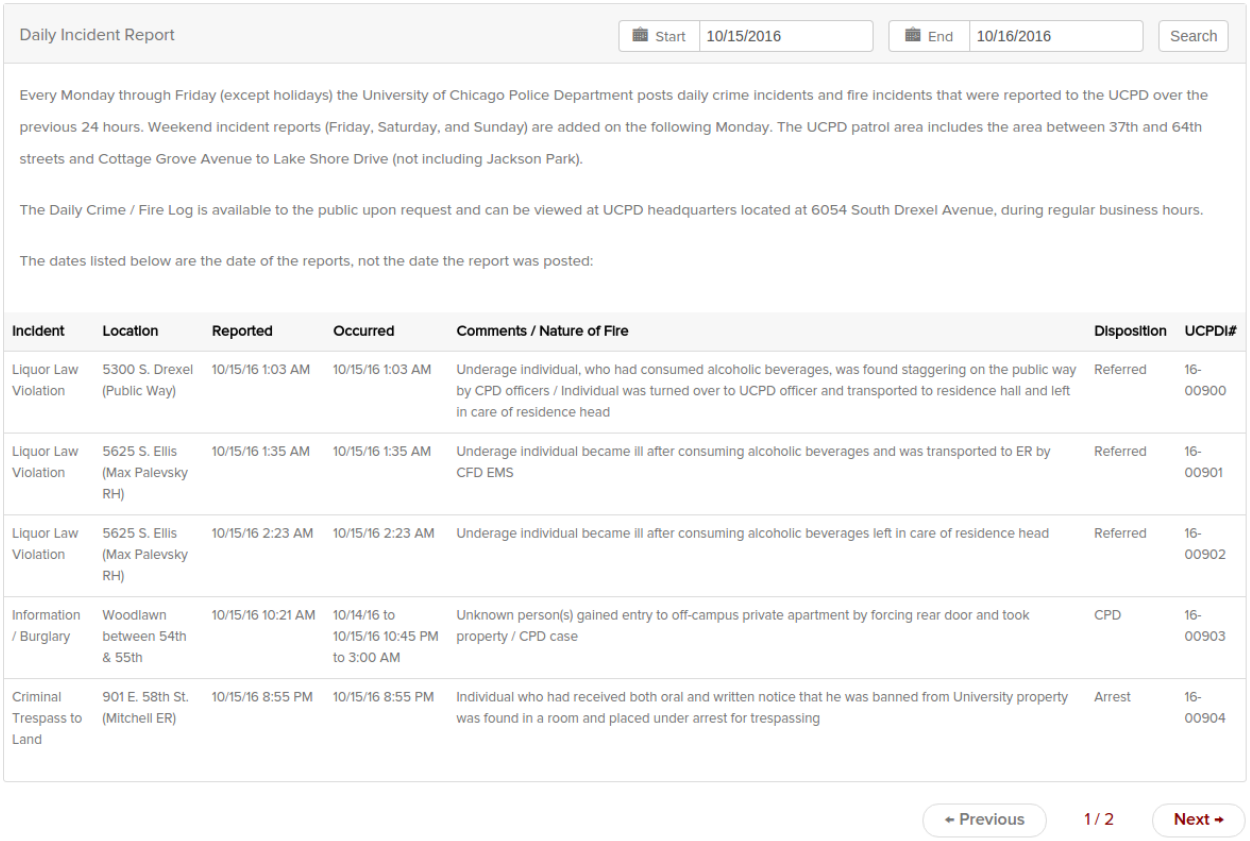
Intro to Spatial Data Science | Research Question

Julian McClellan

October 19, 2016

The University of Chicago Police (UCPD) maintains **a searchable database of incident reports** spanning July 1st, 2010 to the approximate present, subject to the update rules given in the top of Figure 1. The native format of this database is shown below. For analysis, the database will be scraped and put into CSV format, in addition to a number of other cleaning operations.



Daily Incident Report

Start 10/15/2016 End 10/16/2016 Search

Every Monday through Friday (except holidays) the University of Chicago Police Department posts daily crime incidents and fire incidents that were reported to the UCPD over the previous 24 hours. Weekend incident reports (Friday, Saturday, and Sunday) are added on the following Monday. The UCPD patrol area includes the area between 37th and 64th streets and Cottage Grove Avenue to Lake Shore Drive (not including Jackson Park).

The Daily Crime / Fire Log is available to the public upon request and can be viewed at UCPD headquarters located at 6054 South Drexel Avenue, during regular business hours.

The dates listed below are the date of the reports, not the date the report was posted:

Incident	Location	Reported	Occurred	Comments / Nature of Fire	Disposition	UCPD#
Liquor Law Violation	5300 S. Drexel (Public Way)	10/15/16 1:03 AM	10/15/16 1:03 AM	Underage individual, who had consumed alcoholic beverages, was found staggering on the public way by CPD officers / Individual was turned over to UCPD officer and transported to residence hall and left in care of residence head	Referred	16-00900
Liquor Law Violation	5625 S. Ellis (Max Palevsky RH)	10/15/16 1:35 AM	10/15/16 1:35 AM	Underage individual became ill after consuming alcoholic beverages and was transported to ER by CFD EMS	Referred	16-00901
Liquor Law Violation	5625 S. Ellis (Max Palevsky RH)	10/15/16 2:23 AM	10/15/16 2:23 AM	Underage individual became ill after consuming alcoholic beverages left in care of residence head	Referred	16-00902
Information / Burglary	Woodlawn between 54th & 55th	10/15/16 10:21 AM	10/14/16 to 10/15/16 10:45 PM to 3:00 AM	Unknown person(s) gained entry to off-campus private apartment by forcing rear door and took property / CPD case	CPD	16-00903
Criminal Trespass to Land	901 E. 58th St. (Mitchell ER)	10/15/16 8:55 PM	10/15/16 8:55 PM	Individual who had received both oral and written notice that he was banned from University property was found in a room and placed under arrest for trespassing	Arrest	16-00904

← Previous 1 / 2 Next →

Figure 1:

I will be looking to identify clusters and/or spatial heterogeneity in regards to the "Disposition" and/or "Incident" factors of the reports. For example, in Figure 1, one can see that the "Disposition" of the 4th case is "CPD" (Chicago Police Department), meaning that the CPD, rather than the UCPD is handling the case. With the ~8000 reports contained in the database, one might be able to determine if there is a spatial pattern affecting whether or not a case is handed over to the CPD. Depending on the data itself (which has been scraped but not fully cleaned yet), one might be able to ask this question for a variety of "Incident" attributes and/or other "Dispositions". Figure 1 shows "Referred" and "Arrest" as other possible "Dispositions", but a preliminary pass of the scraped data has shown that there is also "Open" and "Closed" (Cases), and there might be interesting spatial patterns that correspond with these dispositions.

Additionally, the Winter of 2015-2016 occurred during an "El Niño" year, which, in Chicago and the Midwest United States, is characterized by milder winters. It could prove an interesting exercise to analyze spatial patterns of "El Niño" winters against "La Niña" (~normal) winters.

Looking at the "Location" of these incidents, there are sometimes precise addresses, and other times approximate cross streets given. Thankfully, the Google Maps API can convert all of these into GPS coordinates. The precision of those GPS coordinates can vary somewhat. Looking at the 4th case (from the top) in Figure 1, the Location is given as "Woodlawn between 54th and 55th", hardly a precise address like the rest of the entries, and this is quite common in the dataset. This case in particular poses a unique problem as there are two 54th streets: 54th Street, and 54th place. In this case, the Google Maps API does function, but not with as much precision as precise addresses. Steps may need to be taken to ensure better coordinate accuracy for these types of "Locations". It may be possible to simply hand code the GPS coordinates for any "Location" containing the word "between".