The Association Between Felonies in NYC and Weather and Temporal Conditions

Brian Weinstein

Columbia University
STAT W4201: Advanced Data Analysis

May 2, 2016

Abstract

Background: The New York City Police Department recently released incident-level felony data to the New York City Open Data portal. The dataset includes timestamped information for all felonies committed in NYC.

We first examine the association between the daily number of felonies committed in NYC in 2015 and temperature, presence of precipitation, day of week, federal and New York holidays, and school days. Second, we examine the association between large increases in temperature ($> 8°$F from the previous day) and increases in the number of felonies.

Methods and Results: We initially test for a difference between the number of felonies on warmer and cooler days ($\geq 51.98°$F and $< 51.98°$F, respectively — the NYC 2015 median), finding overwhelming evidence of a difference, with there being 44 more felonies, on average, on warmer days than on cooler day (95% CI 38 to 51 felonies; two sided p-value < 0.000001 from a two-sample t-test).

After accounting for presence of precipitation, holidays, school days, and day of week, the data provides overwhelming evidence that for every $1°$F increase in temperature there are, on average, 1.4 additional felonies per day (95% CI 1.3 to 1.6 felonies; two-sided p-value < 0.000001 for a test that the linear regression coefficient is 0).

We next find that, after accounting for day of week, there is little evidence to suggest that increases in felonies from the previous day are associated with large increases in temperature (two-sided p-value 0.1001 for a test that the linear regression coefficient is 0).

Conclusions: There is a clear association between warmer temperatures and an increased number of felonies. There is no evidence that large increases in temperature from the previous day are associated with increases in the number of felonies.

1 Introduction

1.1 Questions of Interest

After many years of pressure, the New York City Police Department (NYPD) recently released incident-level felony data to the NYC Open Data portal, as part of their initiative to improve their accessibility, transparency, and accountability. Prior to this release, felony data had only been provided in an aggregated format (by week and police precinct), and was done so only in PDF and Excel files on a weekly and quarterly basis.

In this paper, we use the newly-released data to examine the association between the daily number of felonies committed in New York City (NYC) and: day of week, outside air temperature, precipitation, federal and New York (NY) holidays, and public school days. Although there is no causal connection, we measure the association between the number of felonies and these variables, and for a given set of these conditions, estimate how many felonies the NYPD can reasonably expect.

We first examine the association between felonies and temperature, and then measure the association between felonies and the additional variables outlined above, after taking temperature into account. Second, we examine if large ($> 8°$F) increases in temperature, as compared the previous day, can explain spikes in the number of felonies.

1.2 the data set

1.2.1 Data sources

- NYPD 7 Major Felony Incidents: data.cityofnewyork.us/d/hyij-8hr7 The incident-level felonies included in the dataset are burglary, felony assault, grand larceny, grand larceny of motor vehicle, murder and non-negligent manslaughter, rape, and robbery.
- NYPD Motor Vehicle Collisions: data.cityofnewyork.us/d/h9gi-nx95
- National Centers for Environmental Information (weather conditions and temperature data): ncdc.noaa.gov/cdo-web/search
- New York State Holidays: cs.ny.gov/attendance_leave/2015_legal_holidays.cfm
- Federal Holidays: opm.gov/policy-data-oversight/snow-dismissal-procedures/federal-holidays/#url=2015
- School Attendance: github.com/ajschumacher/NYCattends/tree/master/xml (The NYC Department of Education also publishes this data to the NYC Open Data portal, but historical data is only retained there for the current school year. Instead Ill scrape the daily XML snapshots stored in the linked Github repository.)