A Data Story on Drug-Related Deaths:

An Exploration of Drug Deaths, Opioid Prescription Rates, and Supervised Learning to predict Heroin vs non-Heroin related deaths on Open Datasets

Sackler Institute
Predictive Analytics Unit

Isabel Metzger[1], Yindalon Aphinyanaphongs[1], Kasthuri Kannan[2]
[1]New York University Langone Health Predictive Analytics Unit, [2]New York University School of Medicine

INTRODUCTION

In this work, we explore aspects of the opioid epidemic using the Connecticut Accidental Drug-Related Deaths (CTAD) database, Center for Disease Control (CDC) databases, and other databases. We use a combination of machine learning and descriptive statistics applied to these databases to demonstrate a series of analyses regarding the current epidemic.

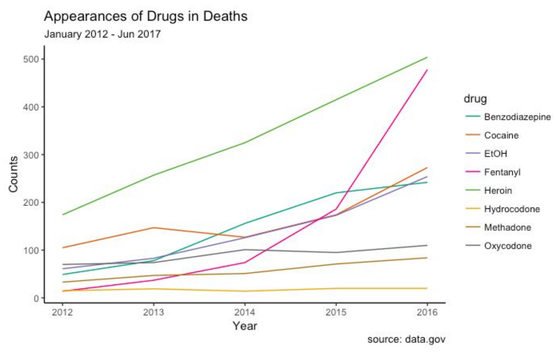
Connecticut Accidental Drug-Related Deaths Database

The Connecticut 'Accidental Drug-Related Deaths Jan 2012-Jun2017' consisted of over 3,500 unique observations and included data such as Case Numbers, Date of accidental drug death (from January 2012 to June 2017), Sex, Race, Age, Residence City, Death City, Injury Location, Death Location, County, Lat / Lon coordinates, Cause of Injury (Officer Notes), Heroin in Y/N format, Fentanyl Y/N format, Cocaine Y/N format, and other drugs provided in the Y/N format.

Data are derived from an investigation by the Office of the Chief Medical Examiner which includes the toxicity report, death certificate, as well as a scene investigation.

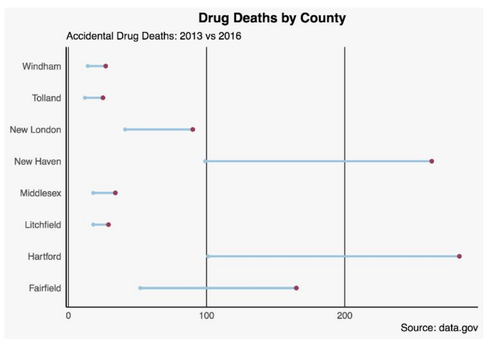
We explored this data, and make several conclusions.

(1) Heroin death are on the rise.

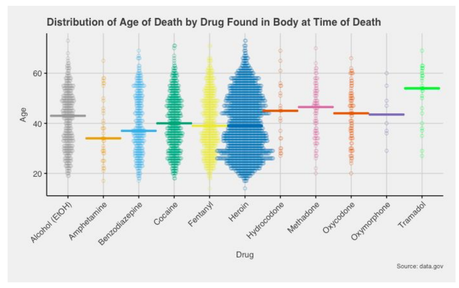


There were 104 appearances of heroin in drug deaths in 2012 and 504 in 2016. **Fentanyl jumped from 105 in 2012 to 478 in 2016.** 49 Benzodiazepine in 2012, 242 in 2016. Oxymorphone (*not included in this plot*) appeared in only 9 deaths/toxicology reports in 2016. Oxymorphone was the only drug which decreased.

(2) Specific counties have higher death rates.

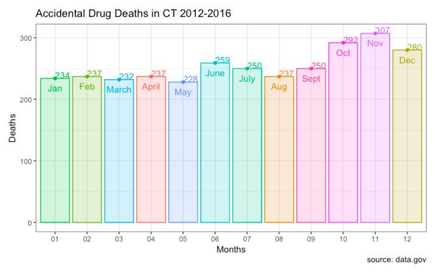


Hartford County and New Haven had the most accidental drug deaths in total. These counties also had the highest heroin and fentanyl related deaths. Heroin appeared in 50 percent of all New Haven drug-related deaths and 57% Hartford drug-related deaths. Hartford Hospital was listed as the most common place of death in Hartford.



The median age of a person who died with Amphetamine was the youngest at 34 years old while the median age of a person who died with Tramadol in their system was much higher, at 54 years old. The median age of death for heroin and fentanyl were both 39. The youngest person in the dataset died of **'Heroin and Fentanyl Intoxication'** at **age 14** at Hartford Hospital.

Most deaths occurred during the fall months of October and November.

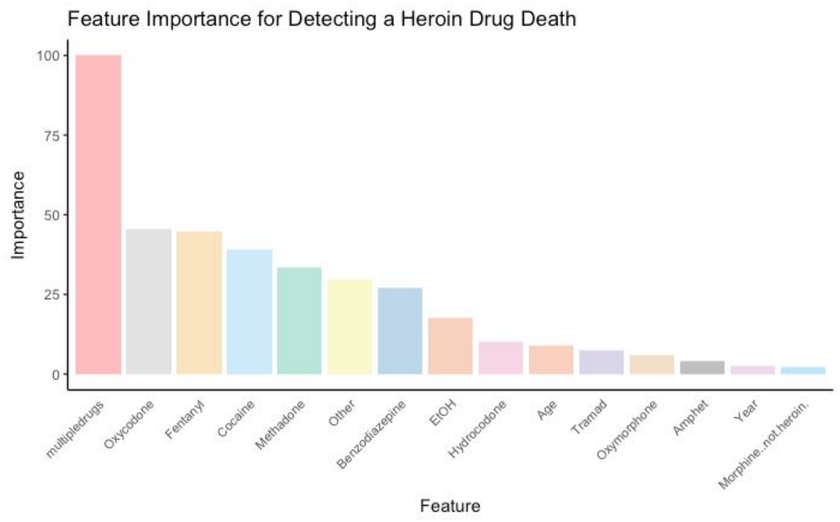


Exploring Correlative Features with Explanatory Machine Learning.

Because heroin remains the highest drug-related cause of death, an explanatory model was built to differentiate between a Heroin Death (deaths which reported "Y (Yes)" for heroin in the toxicology report and "N (No)" for heroin.

For the explanatory features,, we used the original features in the dataset but also added constructed features from other government open data resources. Additional features included the county of death's number of naloxone prescribing pharmacies and the number of opioid prescribers. This resulted in 237 features.

We applied a stochastic gradient boosting model. The model is highly explanatory with an AUC of 0.99.



Variable importance features identified by this model are shown above. Multiple drugs tend to be used by heroin users and other explanatory features are shown.

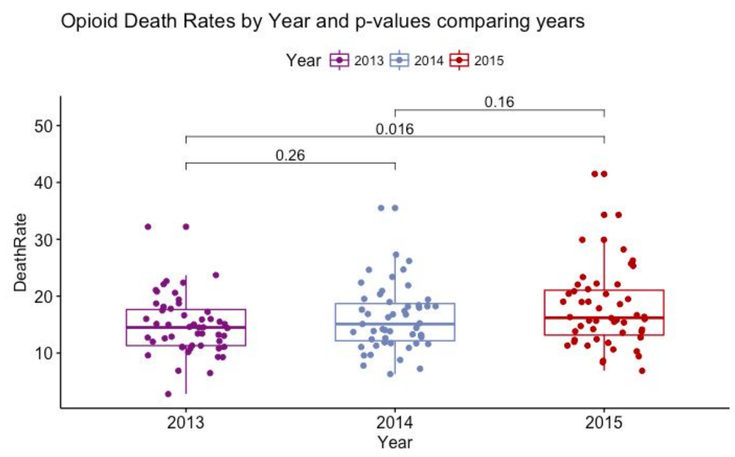
FUTURE

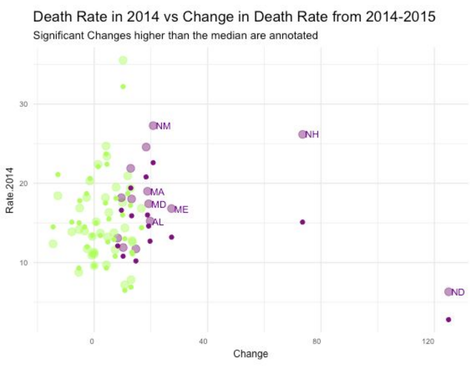
Open data was essential to this project. This datastory reveals how we used open data sources and various models to describe and quantify our data.

For the next release of the data, we hope to amplify the data in new ways, and utilize the Latitude and Longitude coordinates of "death location" to more closely understand the location data.

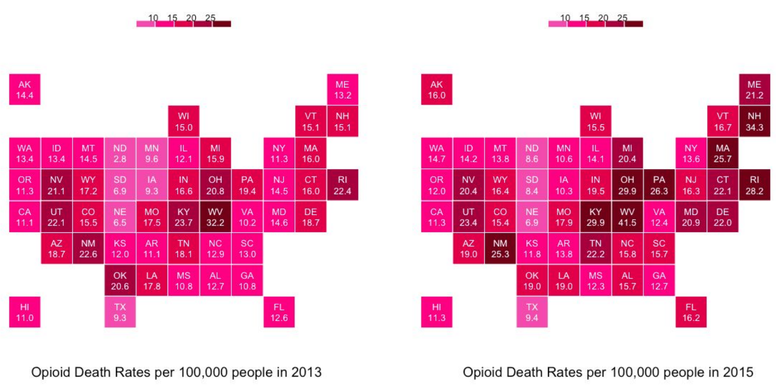
Center for Disease Control database

Nationally, opioid death rates have risen **significantly.**





When death rates are adjusted for age and population, West Virginia has the highest rates of opioid overdose deaths not only in the South, but for all states. However, other states have seen more drastic changes, for example, New Hampshire has nearly doubled, with a rate of 35.5 in 2015.



REFERENCES

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook. Computer and Information Technology: Computer Programmers. Jan 2018.

Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook. Office and Administrative Support: Customer Service Representatives. Jan 2018.

John A. Ramey (2014). noncensus: U.S. Census Regional and Demographic Data. R package version 0.1. https://CRAN.R-project.org/package=noncensus

CDC. Wide-ranging online data for epidemiologic research (WONDER). Atlanta, GA: CDC, National Center for Health Statistics; 2016. Available at http://wonder.cdc.gov.

Centers for Disease Control and Prevention. CDC Health Advisory: Increases in Fentanyl Drug Confiscations and Fentanyl-related Overdose Fatalities. HAN Health Advisory. October 26, 2015. http://emergency.cdc.gov/han/han00384.asp

Rudd RA, Seth P, David F, Scholl L. Increases in Drug and Opioid-Involved Overdose Deaths — United States, 2010–2015. MMWR Morb Mortal Wkly Rep. ePub: 16 December 2016. DOI: http://dx.doi.org/10.15585/mmwr.mm6550e1.

CMS.gov. CMS 2016 Star Ratings Fact Sheet. Centers for Medicare and Medicaid Services. Available from: https://aishealth.com/sites/all/files/cms-2016-star-ratings-fact-sheet.pdf

CMS.gov. CMS Fast Facts. Centers for Medicare and Medicaid Services. February 2018 . Available from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Fast-Facts/index.html

"Accidental Drug Related Deaths 2012-June 2017." *Accidental Drug Related Deaths 2012-June 2017 - Data.gov*, Publisher Data.ct.gov, updated 3 Feb. 2018, catalog.data.gov/dataset/accidental-drug-related-deaths-january-2012-sept-2015

"Pharmacies Offering Narcan, Evzio and Other Brands of Naloxone." Pharmacies Offering Narcan, Evzio and Other Brands of Naloxone - Data.gov, Publisher Data.ct.gov, updated 28 Feb. 2018, data.gov/dataset/pharmacies-offering-narcan-evzio-and-other-brands-of-naloxone

Please email any questions to isabel.metzger@nyumc.org

Special thanks to my advisor Yindalon Aphinyanaphongs, MD, PhD and the Predictive Analytics Unit at NYU Langone Health!