Read me notes and caveats	"Codebook" in development
This Google sheet is managed by D. Brian Burghart of Fatal Encounters Dot Org, fatalencounters.org To download, go to the Google spreadsheet, link below, go under File>Download as> and pick your format. We recommend comma-seperated values.	Column A, Fatal Encounters' Unique ID: Generally speaking, the UID works like other UIDs work; new IDs are added to new incidents without consideration of the date of the incident or the date of its inclusion into the dataset. However, twice since 2012, we've had to rebuild UIDs because of problems with the Google Spreadsheet. This in no way infers a problem with the sheet, more likely operator error, but in one instance, formulas crept in, and in the other, UIDs, which are generated manually, developed duplicate values. Also, if a duplicate record is discovered in the data, the duplicate is replaced with a non-duplicate, in order to keep the UIDs sequential, and to enable visual verification that the UIDs are working as intended. Column A is manually replicated in Column AA as a backup to enable replacement of Column A, if problems with the UIDs ever arise again. Fatal Encounters recommends researchers note download dates. The current sheet always includes the most up-to-date data and accurate data.
https://docs.google.com/spreadsheets/d/1dKmaV_JiWcG8XBoRgP8b4e9Eopkpgt7FL7nyspvzAsE/edit#gid=0	Column B, Subject's name: Names contain all information Fatal Encounters has been able to collect, including nicknames if available. Often additional information for names comes through obituaries or social media. In case of "Names withheld by police," the names are sometimes also voluntarily withheld by news media--especially in the cases of suicide--but police fail to publish the names through public disclosures. In the cases of "aka," sometimes this indicates errors or variations of names reported in news media, and sometimes it indicates aliases used by the decedent. In instances of transgender individuals, it's a regretable necessity because news media and police often refuse to identify transgender people by their chosen names and gender.
Fatal Encounters documents non-police deaths that occur when police are present or are precipitated by police action or presence. Officer deaths are included when caused by another officer, including friendly fire incidents, and criminal actions—like domestic violence—and suicides that occur when other officers are present. Officer vehicle-related deaths are included when they are caused by another officer. Homicides of officers by felons or deaths in the regular course of duties are not generally documented in the database.	Column C, Subject's age: Ages are generally reported in news media, official documents and obituaries. Ages frequently change with updates in articles, with early reporting being the least accurate, neither police nor news media using "time between date" calculators, instead subtracting birth year from current year. In cases where police and media reports and obituaries conflict, and there is no birth date available, Fatal Encounters generally goes with the age stated in the obituary.
We believe we include all the available records for all 50 states and DC back to 2000, but there are several data points that we think are too poorly reported in the news media to result in accurate results for analysis: disposition and mental state. Our racial data (Column D) is the best that exists, but it's pretty spotty and gets worse prior to 2013. Beginning in 2020 we added two columns that regard imputed race. We generally do weekly updates on Tuesdays (although for practical reasons, sometimes that's extended later into the week), so be aware we're usually a few days behind. Government data also suggests that police chase deaths are often not reported in news media, so our data almost certainly understates those totals.	Column D, Subject's gender: Male, Female, Transgender, or empty cell
This data is available for anyone to use for whatever purpose they choose. The only requirement for use is if users spot an errors, please report it to d.brian@fatalencounters.org. We don't require attribution, but your users probably will. We're FatalEncounters.org. If you have questions, just email d.brian@fatalencounters.org.	Column E, Subject's race: In Column E, race is usually reported based on visual evidence or official reports. Visual evidence includes images in news stories, obituaries, or body camera or other surveilance videos. Sometimes race is disclosed in a news article as an identifying characteristic of who was killed, particularly when police are withholding a name. Occasionally, FE researchers will contact a family to determine the race. Sometimes race is determined by language or a photo on a tombstone.
Some words of caution. The vast majority of these records come from media sources and police records. That means only one version of the story—the police story—is generally told in the descriptions. Rarely do news media seek out family members and friends to balance or contradict police narratives. While we verify our data against media reports, sometimes the information presented is so wildly inaccurate that we instead include accurate but conflicting information. Also, be aware that this document is "living." We repair any errors or eliminate any duplicates as we discover them. Errors can be reported through the fatalencounters.org website. We also sometimes discover new incidents in older time periods. Researchers who use the data should note when they download the spreadsheet so they can note changes in newer documents. We don't own many of the photos on the spreadsheet or displayed on the website, and they are not always explicitly authorized by the copyright owner. While we're using the images under "Fair Use" doctrine, for-profit news outlets should probably obtain permission from the copyright owner. That being said, taxpayer-funded images, like mugshots, aren't owned by the police agency that took them or any news agency that happens to display them.	Column F, Subject's race with imputations: A combined column with both non-imputed race codes (Column E) and imputed race codes.
This project owes a great debt to the many volunteers and paid researchers who got us this far. Carla DeCeros' volunteer work was amazing. She did the initial work on more than 20 percent of this data. Walt Lockley also did incredible work, completing research on several key states. Christopher Cox did the bulk of the work on Texas. I can't begin to mention all the people who contributed, but those three moved the data collection ahead by years. We've also had two funders who came forward when we desperately needed help. Jan Schaffer, executive director of J-Lab: The Institute for Interactive Journalism enabled us to move from a primarily crowdsourced, volunteer project to a more professional effort with a $12,000 grant. Our largest individual funder has been Jeff Moe, and his contributions have meant the difference between moving forward and shuttering the project.	Column G, Imputation probability: This column shows whether a race was imputed or not imputed and the probability of accuracy in the imputed races in Column F. While the numbers have been relatively small, we have found errors in imputed races.
Our racial imputations were added on January 17, 2020, and will be updated periodically. The imputations were calculated by Joseph Gibbons, PhD Associate Professor Department of Sociology Associate Director Center for Human Dynamics in our Mobile Age (HDMA) San Diego State University For most entries without a reported race/ethnicity, race/ethnicity was imputed with estimates derived with Bayesian Improved Surname Geocoding (BISG). This method uses both surnames and demographics of place of residence of victims to estimate their race/ethnicity (Elliott et al. 2008). This method has been demonstrated in other fields, including mortgage research, to be a useful method to determine race/ethnicity (Consumer Financial Protection Bureau 2014). Our estimates were validated by singling out encounters where one's race was determined and comparing the estimated values to the reported values. As reported below, we found the correlation of estimates to reported values to be moderate to strong and in all cases statistically significant. In the dataset (Columns F and G), we report the probabilities for each imputed race/ethnicity. Correlations Between Imputed Race/Ethnicity and Reported Race/Ethnicity White, 0.61; Black, 0.64; Hispanic, 0.82; Asian, 0.68; Native American, 0.46 Citations Consumer Financial Protection Bureau. 2014. “Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity.” Washington, D.C.: Consumer Financial Protection Bureau. Elliott, MarcN, Allen Fremont, PeterA Morrison, Philip Pantoja, and Nicole Lurie. 2008. “A New Method for Estimating Race/Ethnicity and Associated Disparities Where Administrative Records Lack Self-Reported Race/Ethnicity.” Health Services Research 43 (5 Pt 1): 1722–36. https://doi.org/10.1111/j.1475-6773.2008.00854.x.	Column H, URL of image of deceased: We use images to help in establishing race. We don't own any of these images and have removed them when asked by the owners of the images (twice). We began keeping images on our own servers because image links go rotten quickly, and we don't have an automated method for updating rotten links.
The racial imputations work was funded by USC and the Eunice Kennedy Shriver National Institute of Child Health and Human Development, NICHD R01HD093382	Column I, Date of injury resulting in death (month/day/year): Seems self-explanatory.

Read me notes and caveats	"Codebook" in development
Also beginning in January 2020, we have begun adding "exclusions" to the "Dispositions" column with the intention of backfilling when time is available. The idea is to provide a means for our researchers to easily remove incidents that are not typically considered intentional use of force. The convention will have results following a backslash, including /off-duty or /collaterol death, officer present and others. Again, we suggest that outside researchers do not use the column for analysis because its contents are in development.	
	Column J, Location of injury (address): The best street address of the location where the injury causing death happened that we can find using public documents, news reports, and videos. All of our geocoding uses Google Maps. Sometimes this doesn't allow for a street address, for example, when a death happens on a freeway, highway or in the middle of a national forest. We use the best location we can acquire.
	Column K, Location of death (city): The best city location where the injury causing death happened that we can find using public documents, news reports, and Google Maps. All of our geocoding uses Google Maps, so occasionally, the city or township or other sub-county designation mentioned in the article or document does not match the geolocation provided by Google Maps.
	Column L, Location of death (state): The best state location where the injury causing death happened that we can find using Google Maps.
	Column M, Location of death (zip code): The best zip code location where the injury causing death happened that we can find using Google Maps. We concatenate street address, city, and state and search Google Maps to arrive at the Zip code.
	Column N, Location of death (county): We generally use the counties provided by the National Association of Counties using Zip codes. Frequently, we must manually find the correct county because cities and Zip codes can cross county lines.
	Column O, Full Address: Google Maps concatenates Street address, city, state, Zip, county or burrough to process in the Google Maps API. We use a Google Sheets add-on called Awesome Table to develop Latitude and Longitude.
	Column P, Latitude: From Google Maps API. After the initial completion of Lat/Longs in December 2017, locations for which there is not a street address are hand coded. During the initial batch geocoding, only addresses that produced an error were hand-coded.
	Column Q, Longitude: From Google Maps API. After the initial completion of Lat/Longs in December 2017, locations for which there is not a street address are hand coded. During the initial batch geocoding, only addresses that produced an error were hand-coded.
	Column R, Agency(ies) involved in death:
	Columns below this line are moving around as they're developed. Please compare the headings to the spreadsheet to see what is describing what.
	Column S, ORI7s: In development, to be added at a later date
	Column T, Cause of death:
	Column U, Armed or Unarmed: In development, to be added at a later date
	Column V, Alleged weapon: In development, to be added at a later date
	Column W, Physical movement: In development, to be added at a later date
	Column X, Fleeing/Not fleeing: In development, to be added at a later date
	Column Y, A brief description of the circumstances surrounding the death:
	Column Z, Dispositions/Exclusions INTERNAL USE, NOT FOR ANALYSIS:
	Currently Column V. Column AA, Intentional Use of Force (Developing) Codes will be Yes; No; Intentional use of deadly force; Suicide, Vehicle/Pursuit. At the onset of development, our plan is to further define Vehicle/Pursuit. Yes will indicate an intentional use of force with or to the vehicle, for example, when an officer intentionally runs down an alleged criminal or uses a PIT maneuver or spike strips; Vehicle will indicate an accident, for example, an officer hits a pedestrian or another vehicle when not engaged in a pursuit; Pursuit will indicate whether the officer was or had been in pursuit.
	Column AB, Link to news article or photo of official document:
	Column AC, Were police aware of symptoms of mental illness before interaction? INTERNAL USE, NOT FOR ANALYSIS:
	Column AD, Video:
	Column AE, Date&Description:
	Column AF, Unique ID formula INTERNAL USE, NOT FOR ANALYSIS:
	Column AF, Unique identifier (redundant):
	Column AH, Date (Year):