

## Missing Pieces data dictionary

Last updated: 11/20/2017

Contact: [missingpieces@thetrace.org](mailto:missingpieces@thetrace.org)

Field	Definition	Data type
<b>RECORD ID</b>	A unique, sequential value that we assigned to each record in the master sheet. Begins at 1.	Number
<b>MATCH ID</b>	A value we assigned to identify match guns in the data. See the <b>Matches</b> section below for a full explanation.	Number
<b>STATE</b>	The state (plus Washington, D.C.) where the agency that provided the record is located. Standard two-character abbreviations.	String
<b>LOCATION</b>	The city, county, or other municipality of the agency that provided the record.	String
<b>AGENCY ORIGINAL</b>	The original name of the agency that provided the record.	String
<b>AGENCY</b>	<p>In most cases, the same as AGENCY ORIGINAL. In a few cases, where cities provided data broken down by department, AGENCY consolidates those departments into a single value, while the original values are preserved in AGENCY ORIGINAL. A list of all agencies included in the data is available <a href="#">here</a>.</p> <p>Example: <b>AGENCY ORIGINAL:</b> LAPD HOLLYWOOD <b>AGENCY:</b> LOS ANGELES PD</p>	String
<b>AGENCY STATUS</b>	The status of the record as reported by the agency.	String
<b>STATUS</b>	<p>Our grouped and standardized version of the record's status. Possible values:</p> <ul style="list-style-type: none"><li>● LOST: Indicates a firearm that was reported lost.</li><li>● LOST/STOLEN: Indicates a firearm that was reported lost or stolen. (Some agencies do not differentiate.)</li><li>● STOLEN: Indicates a firearm that was reported stolen.</li><li>● STOLEN (RECOVERED): Indicates a firearm that was reported stolen, where the record notes – either in the AGENCY STATUS column or elsewhere – that the firearm was later recovered.</li><li>● RECOVERED: Indicates a firearm that was seized or otherwise recovered by an agency.</li></ul>	String

- **RECOVERED (STOLEN):** Indicates a firearm that was seized or otherwise recovered by an agency, where the record notes that the firearm was originally stolen.

<b>SERIAL NUMBER</b>	A cleaned and formatted version of the serial number. Spaces and punctuation have been stripped, and non-serial number values such as “none,” “unknown,” and “obliterated” were set to blank. Values that we conclusively identified as non-serial number values, such as patent numbers and model numbers, were also set to blank.	String
<b>MAKE ORIGINAL</b>	The make of the firearm originally provided by the agency.	String
<b>MAKE</b>	The cleaned and standardized make of the gun. Variations, and typos, were standardized and clustered; <a href="#">NCIC field code abbreviations</a> were expanded to the full manufacturer name. In some cases, the make was inferred from the property description or the serial number. Values such as “unknown” were set to blank.  A make of “CONFIRMED SERIAL” means that we have manually verified that the record matches with another record with an identical serial number, but cannot identify the make.	String
<b>MODEL</b>	The model of the firearm. This column <b>has not been cleaned or standardized.</b>	String
<b>CALIBER</b>	The cleaned and standardized caliber of the firearm. All values in this column have been standardized according to the <a href="#">NCIC field codes for gauge/caliber</a> .	Number
<b>PROPERTY DESCRIPTION</b>	The property description provided by the agency.	String
<b>DATE</b>	The date of the record in DD/MM/YYYY format.	Date
<b>YEAR</b>	The year of the record in YYYY format.	Year
<b>OFFENSE ORIGINAL</b>	The primary offense associated with the record.	String
<b>OFFENSE</b>	Cleaned and standardized offenses. Primarily, this column replaces common penal code references with their associated offenses.  Example: <b>OFFENSE ORIGINAL:</b> 245(A)2 / 245A2 / 245 A 2 <b>OFFENSE:</b> ASSAULT WITH FIREARM	String

<b>OFFENSE 2</b> <b>OFFENSE 3</b> <b>OFFENSE 4</b>	In cases where agencies provided multiple offenses associated with a record in separate columns, those values are listed in these additional offense columns, which have not been cleaned or standardized in any way.	String
<b>CASE NUMBER</b>	The reporting agency's internal case or record number.	String
<b>GUN TYPE</b>	The type of gun. All values in this column correspond with NCIC <a href="#">field codes for weapon type</a> .	String
<b>GUN DESC</b>	The description of the gun. All values in this column correspond with NCIC <a href="#">field codes for weapon description</a> .	String

---

### Caveats

Agencies have different systems for recording data. As a result, rows in the database do not uniformly correspond with individual incidents, or individual firearms. Some agencies supplied several rows of identical or nearly-identical data per incident. Others created one record per offense if a gun was used in multiple offenses. Not all agencies provided their internally used unique case identifiers in their data. On top of this, many records lack unique identifying information (makes and serial numbers) for the weapons themselves.

To avoid overcounting, we restricted our analysis to only the records that contained makes and serial numbers – records where we can conclusively identify individual guns.

While we have standardized the vast majority of values in the MAKE column, some makes that appear infrequently in the data have not been cleaned. In cases where the original make provided by the agency contained other information, like the gun's model or caliber, that information may not have been transferred to its appropriate field. Therefore, the MAKE ORIGINAL column may contain information about the firearm that is not present in any other field. The PROPERTY DESCRIPTION column may also contain information about a gun's make, model, or caliber.

Because the OFFENSE column contains offenses recorded by more than a thousand law enforcement agencies, we have standardized it very little. While we were able to clean some of the more common values, there are thousands more that are untouched.

Florida maintains gun theft data at the state level, but [state law](#) requires records related to stolen guns to be deleted in the event that those guns are recovered. As a result, the Florida data is highly incomplete and the scale of gun theft in the state cannot be known. The state advised us that they could provide the missing records going back two years; we have an outstanding request for those records and will add them once that information is obtained.

The data we were provided is of differing degrees of quality and completeness. Many records are missing important values, such as the make, serial number, or offense. We have included all records we received, even when missing information limits their usefulness.

---

## **Matches**

The primary goal of collecting this data was to identify what we have termed “matches” – guns that were stolen and then recovered. Matches were identified in one of two ways:

### *Make and serial number matches*

If multiple records share a common make and serial number, but different statuses, these are considered a match.

Records where either the make or serial number is blank are excluded. Records where the status is “LOST” are also excluded.

Example:

- Record IDs 54081 and 424626 indicate the theft of a Smith & Wesson, serial number DTZ1792, which was reported to the Sacramento, CA police department on 05/13/2010.
- Record ID 811920 indicates the recovery of the same firearm, Smith & Wesson DTZ1792, by the Chicago, IL police department on 09/15/2013.
- Because these three records correspond to the theft and recovery of the same firearm, they are all assigned the same match ID value of 1887.
- Without the record from Chicago, the two Sacramento records would not have been assigned a match ID. Although they both represent the same gun, they do not have different statuses, so there is no match.

### *Single-record matches*

If an individual record indicates the theft of a gun that was later recovered, or the recovery of a gun that was originally stolen, it is considered a match. Those two cases are identified in the status column as “STOLEN (RECOVERED)” and “RECOVERED (STOLEN)”, respectively.

Example:

- Record ID 189467 indicates a firearm, Smith & Wesson 169380, that was recovered by the St. Louis, MO police department on 01/03/2010.
- Although there are no other records pertaining to this gun in the data, because it has the status of “RECOVERED (STOLEN)” it is assigned the match ID 4190.

Many matches consist of multiple records, so the number of matches we have identified is not equal to the number of rows with a match ID – it is equal to the number of unique match IDs. In cases where a match contains a theft record and a recovery record, the OFFENSE value of the theft record typically corresponds to the crime in which the gun was stolen, and the OFFENSE value of the recovery record corresponds to the crime that the gun was recovered in connection with.