Neutrosophic Computing and Machine Learning, Vol. 45, 2026

University of New Mexico

Physically Complete, Semantically Indeterminate: Field-Level Evidence Coordinates for the Analytical Usability of Ecuador's Official Illicit-Firearms Records, January–July 2026

Físicamente completos, semánticamente indeterminados: Coordenadas de evidencia a nivel de campo para la utilidad analítica de los registros oficiales de armas de fuego ilícitas de Ecuador, enero-julio de 2026

Paola Alexandra Mora Pedroza [1], María José Chalán Castillo [1], Hernán Eduardo Solano León [1] and Ruth Elizabeth Soto Castillo [1,*]

[1] Master's Program in Data Analytics, Universidad Politécnica Salesiana, Quito, Ecuador; pmorap@est.ups.edu.ec; mchalanc@est.ups.edu.ec; hsolanol@est.ups.edu.ec; rsotoc@est.ups.edu.ec
* Correspondence: rsotoc@est.ups.edu.ec

Abstract. Official administrative records can be physically complete and still fail to carry usable information. We audited the open dataset on illicit firearms published by Ecuador's Ministry of the Interior for January–July 2026 (7,252 records, 27 fields) against its published data dictionary and proposed, for each record, separately grounded evidence coordinates for the proposition "the record is analytically usable": T from positive validity evidence (dictionary-consistent values in $K = 6$ attribute fields), I from explicit indeterminacy (sentinel values, values undefined in the dictionary, schema non-applicability) and F from contrary evidence in $M = 7$ cross-field consistency checks, with duplication risk D recorded separately and no unit-sum constraint. The file has no null cells, yet clase_arma holds the sentinel SE DESCONOCE in 1,797 records (24.78 %); the sentinel is itself a dictionary category, 1,702 of 1,704 REVOLVER records carry it, and the conditional entropy of clase_arma given the object type is 0.026 bits: the indeterminacy is schema-induced, not a recording failure. Contrary evidence was almost absent ($F > 0$ in 2 records). The national mean indeterminacy was $I = 0.048$ ($I > 0$ in 28.0 % of records) and fell to 0.009 when clase_arma was treated as inapplicable to revolvers. Across the 63 cantons with at least 30 records, mean I ranged from 0.017 to 0.096 and correlated at $r = 0.81$ with the value expected from each canton's object mix; quality-aware volume (Σ T) reproduced the raw canton ranking (Spearman $\rho = 0.999$). Exact repeated rows (1,972, in 714 configurations) clustered in same-day, same-subcircuit batches consistent with multi-weapon seizures and were not counted as falsity. A first version of this audit had labelled 34.7 % of records T, 48.0 % I and 17.3 % F with a mutually exclusive rule; the revised coordinates show that most of that indeterminacy and all of that falsity were artefacts of the partition. Results support quality-aware territorial review; they do not estimate crime rates, population risk, causal effects or police effectiveness.

Keywords: administrative records; data quality; semantic validity; sentinel values; structured missingness; neutrosophic evidence coordinates; indeterminacy; illicit firearms; Ecuador

Resumen. Los registros administrativos oficiales pueden estar físicamente completos y aun así no contener información utilizable. Auditamos el conjunto de datos abiertos sobre armas de fuego ilícitas publicado por el Ministerio del Interior de Ecuador para enero-julio de 2026 (7252 registros, 27 campos) contra su diccionario de datos publicado y propusimos, para cada registro, coordenadas de evidencia fundamentadas por separado para la proposición "el registro es analíticamente utilizable": T de evidencia de validez positiva (valores consistentes con el diccionario en campos de atributos), I de indeterminación explícita (valores centinela, valores no definidos en el diccionario, no aplicabilidad del esquema) y F de evidencia contraria en verificaciones de consistencia entre campos, con riesgo de duplicación D registrado por separado y sin restricción de suma de unidades. El archivo no tiene celdas nulas, sin embargo, clase_arma contiene el centinela SE DESCONOCE en 1797 registros (24,78 %); el centinela es en sí mismo una categoría de diccionario, 1702 de 1704 registros REVOLVER lo contienen, y la entropía condicional de clase_arma dado el tipo de objeto es 0,026 bits: la indeterminación es inducida por el esquema, no por un fallo de registro. La evidencia contraria fue casi inexistente (en 2 registros). La indeterminación media nacional fue (en el 28,0 % de los registros) y cayó a 0,009 cuando clase_arma se trató como no aplicable a los revólveres. En los 63 cantones con al menos 30 registros, la media I osciló entre 0,017 y 0,096 y se correlacionó en con el valor esperado de la mezcla de objetos de cada cantón; el volumen consciente de la calidad (Σ T) reprodujo la clasificación bruta del cantón (Spearman). Las filas repetidas exactas (1972, en 714 configuraciones) se agruparon en lotes del mismo día y del mismo subcircuito, consistentes con incautaciones de múltiples armas, y no se contabilizaron como falsedades. Una primera versión de esta auditoría había etiquetado el 34,7 % de los registros como verdaderos (V), el 48,0 % como falsos (I) y el 17,3 % como falsos (F) con una regla de exclusión mutua; las coordenadas revisadas muestran que la mayor parte de esa indeterminación y toda esa falsedad eran artefactos de la partición. Los resultados respaldan la revisión territorial con conciencia de la calidad; no estiman las tasas de criminalidad, el riesgo poblacional, los efectos causales ni la efectividad policial.

Palabras clave: registros administrativos; calidad de los datos; validez semántica; valores centinela; datos faltantes estructurados; coordenadas de evidencia neutrosófica; indeterminación; armas de fuego ilícitas; Ecuador

1 Introduction

Official administrative records are central to public-security planning, monitoring and accountability, but they reflect institutional recording processes as well as the phenomena being recorded. International guidance cautions that police and seizure statistics depend on operational capacity, reporting rules and data-production practices and should not be read as population crime rates [4, 12, 13]; peer-reviewed work on gun-market data makes the same point for seizure records, which measure enforcement activity as much as the underlying market [29], and measurement error in police-recorded categories is known to propagate into analyses built on them [30, 31]. In Ecuador the Ministry of the Interior publishes an open dataset on illicit firearms [1], and national research has documented the wider context of firearm trafficking and violence [3].

Data quality is multidimensional and must be evaluated relative to intended use. Established frameworks separate structural presence from accuracy, validity, consistency and interpretability [2, 5–11], and open-government-data assessments add traceability to published metadata [25, 26]. The distinction matters when fields are populated with institutional sentinel values such as SE DESCONOCE or SIN DATO: a database can be physically complete while part of its content remains semantically indeterminate. Database theory has long distinguished missing information that is applicable but unknown from information that is inapplicable to the record [18, 19], and the statistical literature distinguishes missingness mechanisms [20] and, more recently, structured missingness whose pattern is deterministic in other fields [21]. An audit that reports one completeness figure and one sentinel share cannot tell these situations apart.

Neutrosophic set theory represents truth, indeterminacy and falsity as separate components whose sum is not constrained to one [15, 17], and Neutrosophic Paraconsistent Logic (NPL) reads favourable evidence, contrary evidence and indeterminacy as separately grounded coordinates [16], in the tradition of four-valued and evidence logics in which a claim may be supported, opposed, both or neither [22, 23]. Neutrosophic relational models have used the same idea for databases with unknown attribute values [24]. The first version of this audit borrowed the T/I/F vocabulary as a mutually exclusive record partition in which F meant "additional exact repetition of a row". On reflection that construction has two defects: repetition is not contrary evidence about a record, and an exclusive partition forces every record into exactly one state, so that non-seizure administrative conditions were counted as indeterminacy and the three shares were made to sum to one by construction. The present version replaces the partition with field-level evidence coordinates that have explicit semantics, are not normalized, and keep duplication risk as a separate flag.

Four research questions are addressed. RQ1: How physically complete and how dictionary-consistent are the 7,252 records, and what kinds of non-information (physical missingness, semantic indeterminacy, non-applicability, dictionary drift, contrary evidence) does the file contain? RQ2: How is the principal sentinel distributed across object types, territories and months, and is its distribution explained by the recording schema? RQ3: Can record-level evidence coordinates T, I and F for analytical usability be grounded in verifiable field-level evidence without normalization, aggregated across the field → record → canton → province → month hierarchy with uncertainty, and used to compare raw with quality-aware territorial volumes? RQ4, the methodological question for this journal: what do such coordinates add over conventional data-quality indicators and over the exclusive T/I/F partition, and what did that partition get wrong?

The study makes four contributions: a reproducible dictionary-level audit of the official file, including the label variants between the 2025 dictionary and the 2026 data; the identification of the principal sentinel as schema-induced indeterminacy concentrated in one object type; an explicit definition of non-normalized evidence coordinates for record usability, with four elementary properties, an algorithm and a sensitivity analysis; and a quantified comparison with conventional indicators and with the exclusive partition, including the negative result that quality-aware volume does not change the territorial ranking. All results are record counts, shares and coordinates within the file; no population denominators, causal models or measures of police effectiveness are used.

2 Materials and Methods

2.1 Study design and data source

A quantitative, non-experimental, observational, retrospective and descriptive design was used, complemented by a data-quality audit. The source was the spreadsheet mdi_armasilicitas_pm_2026_enero_julio.xlsx, sheet 1. Armas Ilícitas, from the official open-data publication of the Ministry of the Interior [1]. The accompanying data dictionary, mdi_armasilicitas_dd_2025.xlsx, supplied field definitions and the institutionally valid category lists. The observation window was January through July 2026.

2.2 Dataset and unit of analysis

The complete worksheet was analysed: 7,252 rows and 27 variables. The unit of analysis is an administrative row concerning an illicit firearm. Because the file contains no unique event or weapon identifier, a row was not assumed to represent a unique incident or a unique weapon. Territorial analysis was restricted to province and canton. The province field contains 24 categories: 23 provinces and Mar Territorial (42 rows, canton code 0); Galápagos does not appear.

2.3 Dictionary audit and label variants

Every categorical field with a closed category list in the dictionary was compared with the values in the file after trimming, upper-casing and removing diacritics. Differences that are only plural or gender forms of the same label (PISTOLA/PISTOLAS, SEMIAUTOMÁTICA/SEMI AUTOMÁTICO, INCAUTADO/INCAUTADA) were recorded as label variants and treated as dictionary-consistent; values with no counterpart in the dictionary were recorded as outside the dictionary. Two fields defined in the dictionary (cantidad, calibre) are absent from the file, and one field (tipo_lugar) uses a 17-category taxonomy where the dictionary lists two; these are dictionary-version mismatches and were excluded from the coordinates.

2.4 Kinds of non-information

Five situations were distinguished before any coordinate was defined. Physical missingness: a null cell or empty string. Semantic indeterminacy: a populated cell whose value is an explicit indeterminacy marker (SE DESCONOCE, SIN DATO) or a value the published dictionary does not define. Non-applicability: a field whose category system cannot describe the object recorded, so that a marker is the schema's only admissible answer [18]; this cannot be read from a cell alone and was investigated empirically through the conditional distribution of the field given the object type. Contrary evidence: a populated value that contradicts the dictionary or another field of the same record. Duplication risk: exact identity with another row. Only the first four bear on the usability proposition; the fifth is reported separately because, without an identifier, identical rows may be one event recorded twice or several identical weapons recorded once each [27, 28].

2.5 Evidence coordinates for record usability

The proposition evaluated is Q_r: "record r is analytically usable for the territorial and temporal description of weapon characteristics". $K = 6$ attribute fields carry the positive and indeterminate evidence: nombre_objeto, clase_arma, fabricacion, calidad_de, tipo_delito and lugar. For each record r and field k, s_{rk} = 1 if the value is a determinate, dictionary-consistent category and i_{rk} = 1 if it is a sentinel or a value undefined in the dictionary (lugar), with $s_{rk} + i_{rk} \leq 1$ and both zero when the value contradicts the dictionary. $M = 7$ consistency checks carry the contrary evidence: nombre_objeto, clase_arma and calidad_de outside the dictionary after label variants; province code inconsistent with province name; canton code prefix inconsistent with province code; coordinates outside the national bounding box or date outside the window; zone code inconsistent with zone name. f_{rm} = 1 when check m fails. With equal weights, the coordinates are

$$T_r = \frac{1}{K}\sum_{k=1}^{K} s_{rk} \quad (1)$$

$$I_r = \frac{1}{K}\sum_{k=1}^{K} i_{rk} \quad (2)$$

$$F_r = \frac{1}{M}\sum_{m=1}^{M} f_{rm} \qquad (3)$$

and the duplication-risk flag is

$$D_r = \mathbb{1}[\,\exists\, r' \neq r : x_{r'} = x_r\,] \qquad (4)$$

where x_r is the full 27-field row. T_r and I_r are complementary at field level except where contrary evidence occurs, so $T_r + I_r \leq 1$ with equality for every record without a dictionary contradiction; F_r is grounded in different checks and is not $1 - T_r - I_r$. No statistical independence between coordinates is claimed. Equal weights are the default; no subjective field weights were introduced. Four properties follow from the definitions. Property 1 (non-complementarity of F): $T_r + I_r + F_r$ is not constrained to one and can exceed it; a record may be fully determinate ($T_r = 1$) and still fail a consistency check. Property 2 (duplication invariance): the coordinates of a record do not depend on whether another identical row exists; repetition affects D_r only, so a batch of k identical rows contributes k T_r to the quality-aware volume, that is, it counts weapons, not events. Property 3 (bounded non-applicability effect): removing one inapplicable field from K for a subset of records changes I_r by at most 1/K for those records and by zero for all others. Property 4 (aggregation linearity): group coordinates are means and quality-aware volumes are sums, so $QVol_g \leq n_g$ with equality only when every record in g has $T_r = 1$, and the raw and quality-aware rankings coincide whenever T_g is constant across groups.

2.6 Hierarchy, uncertainty and quality-aware volume

Record coordinates were aggregated to canton, province and month as means, with 95 % percentile bootstrap intervals from $B = 1{,}000$ resamples of the records in each group [32]:

$$T_g = \frac{1}{n_g}\sum_{r\in g} T_r,\quad I_g = \frac{1}{n_g}\sum_{r\in g} I_r,\quad F_g = \frac{1}{n_g}\sum_{r\in g} F_r \qquad (5)$$

Quality-aware volume, indeterminacy burden and contrary-evidence burden are the corresponding sums,

$$QVol_g = \sum_{r\in g} T_r,\quad IBurden_g = \sum_{r\in g} I_r,\quad FBurden_g = \sum_{r\in g} F_r \qquad (6)$$

and territorial units were ranked by raw count and by QVol. To test whether territorial differences in I reflect the object mix rather than recording practice, an indirectly standardized ratio was computed for each group, with π(o) the national sentinel share of object type o:

$$\rho_g = \frac{I_g}{\hat{E}_g},\quad \hat{E}_g = \frac{1}{K\,n_g}\sum_{r\in g} \pi(o_r) \qquad (7)$$

$\rho_g = 1$ means the group's indeterminacy is what its object mix implies; $\rho_g > 1$ means more indeterminacy than the mix implies.

2.7 Schema-induced indeterminacy

The dependence of clase_arma (C) on nombre_objeto (O) was measured with the conditional entropy

$$H(C \mid O) = -\sum_{o,c} p(o,c)\,\log_2 p(c \mid o) \qquad (8)$$

together with the share of records whose class is the modal class of their object type. A field that is almost a deterministic function of another field carries almost no information beyond it, and a sentinel concentrated in one object type is then a property of the schema rather than of the recording units. The non-applicability scenario removes clase_arma from K for object types with a sentinel share above 95 % and at least 20 records, and reports the resulting coordinates (T_{na}, I_{na}).

2.8 Statistical analysis

Categorical distributions are summarized with counts and within-dataset percentages. The first version tested the association between territory and the exclusive T/I/F state with Pearson chi-square; because the revised coordinates are record-level proportions and because the principal sentinel is determined by object type, that test was replaced by (i) a chi-square test of the binary sentinel indicator against object type, with expected cells reported; (ii) Kruskal–Wallis tests of I_r across provinces and months with ε^2 as effect size, repeated after excluding REVOLVER rows; and (iii) a permutation comparison (500 permutations of I_r across records) of the weighted squared deviation of provincial I from the object-mix expectation. Associations are read as recording composition, not causation.

2.9 Sensitivity analysis

Six scenarios were run: the primary definition; the non-applicability scenario; the catch-all categories OTROS, OTROS ARMAS CORTAS and OTROS ARMAS LARGAS counted as indeterminate; lugar values outside the dictionary counted as determinate; lugar excluded ($K = 5$); and tipo_delito and lugar excluded ($K = 4$). For each, the national means and the Spearman correlation of the canton QVol ranking with the primary and with the raw ranking are reported. Algorithm 1 summarizes the procedure.

Algorithm 1: Field-level evidence audit of record usability for an administrative dataset.

Step	Operation
1	Fix the proposition Q_r, the K attribute fields, the M consistency checks, the label-variant map and the sentinel list before inspecting results.
2	Audit each field against the dictionary: category counts, label variants, values outside the dictionary, sentinel rows (Table 1).
3	For each record compute s_{rk}, i_{rk} (Eqs. 1–2), f_{rm} (Eq. 3) and D_r (Eq. 4).
4	Measure schema dependence of any sentinel-heavy field on the object type (Eq. 8) and define the non-applicability scenario.
5	Aggregate to canton, province and month (Eq. 5) with bootstrap intervals; compute QVol, IBurden, FBurden (Eq. 6) and the standardized ratio ρ_g (Eq. 7).
6	Compare raw and quality-aware rankings; run the sensitivity scenarios; report negative results.
7	Report D separately with the batch structure of repeated configurations; never fold it into F.

2.10 Ethical considerations and reproducibility

The file contains administrative event and location attributes but no personal identifiers. Results are presented at province or canton level; no individual behaviour, population risk, causal effect or police effectiveness is inferred. Counts can be affected by operational intensity, institutional capacity, reporting practices and the probability of detection [4, 12, 13]. All computations are implemented in Python (pandas, SciPy, Matplotlib) in two scripts that read the original spreadsheets directly and export every table and figure; the record-level coordinates (record_quality_coordinates.csv), the group coordinates (group_quality_coordinates.csv) and the definitions (QUALITY_COORDINATE_DEFINITIONS.md) accompany the manuscript.

3 Results

3.1 Physical completeness and dictionary consistency

The worksheet contained no null cells and no empty strings: physical completeness was 100 % for all 27 variables. Table 1 gives the dictionary audit. After label variants, nombre_objeto had 4 rows outside the dictionary (LANZA GRANADA, 2; SE DESCONOCE, 2), clase_arma, calidad_de, fabricacion and subtipo_objeto_robado had none, and lugar had 293 (4.04 %), mostly VIVIENDA/ALOJAMIENTO (194) and institutional sites absent from the 2025 list. The principal sentinel, SE DESCONOCE in clase_arma, occurred in 1,797 records (24.78 %) and is itself one of the four categories the dictionary defines for that field; SIN DATO in tipo_delito occurred in 4 records and SE DESCONOCE in nombre_objeto in 2. The dictionary also defines cantidad and calibre, which the 2026 file does not contain, and describes tipo_lugar with two categories where the file uses seventeen. Completeness, dictionary consistency and semantic determinacy are therefore three different properties of the same file.

Table 1: Dictionary audit of the categorical fields (values compared after trimming, upper-casing and removing diacritics; label variants = plural/gender forms treated as consistent).

Field	Dictionary categories	Distinct in file	Label variants	Rows outside dictionary	Sentinel is a dictionary category	Sentinel rows	Use in coordinates
subtipo_objeto_robado	2	2	1	0	no	0	check only
nombre_objeto	13	15	7	4	no	2	K field
clase_arma	4	4	2	0	yes	1797	K field
calidad_de	8	8	2	0	no	0	K field

fabricacion	2	2	0	0	no	0	K field
tipo_delito	22	27	0	25	no	4	K field
lugar	64	83	0	293	no	0	I (undefined value)
tipo_lugar	2	17	0	7252	no	0	excluded: dictionary version
codigo_zona	10	10	0	0	no	0	check only

Note: tipo_delito has no closed list in the dictionary; its 25 rows outside the list are category-name differences and were not counted as contrary evidence. cantidad and calibre are defined in the dictionary but absent from the file.

3.2 Territorial and monthly distribution

Guayas contained 2,551 records (35.18 %), followed by Los Ríos with 1,056 (14.56 %) and Manabí with 739 (10.19 %); together 4,346 records (59.93 %). Guayaquil was the leading canton with 1,365 records (18.82 %). Monthly totals ranged from 843 (June) to 1,194 (January); Guayas ranked first in every month with a within-month share between 26.28 % (April) and 47.82 % (January). Figure 1 shows the monthly totals with the quality-aware volume, and Figure 2 the recorded coordinates of all rows with provincial bubbles. These percentages describe the composition of the file, not population rates.

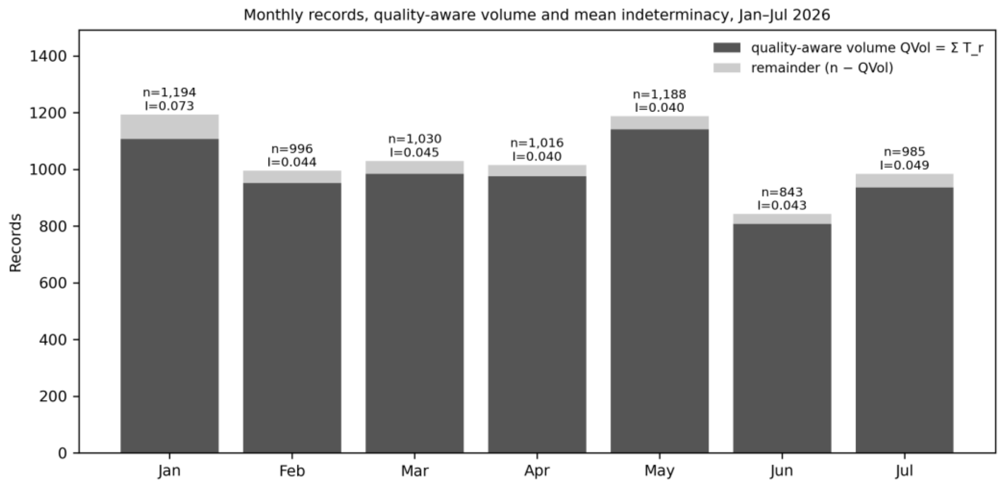

Figure 1: Monthly records, January–July 2026, split into quality-aware volume ($\Sigma\, T_r$, dark) and remainder (light), with mean record-level indeterminacy per month. The January excess of indeterminacy is due to one batch of 230 records on 14 January (Section 3.5).

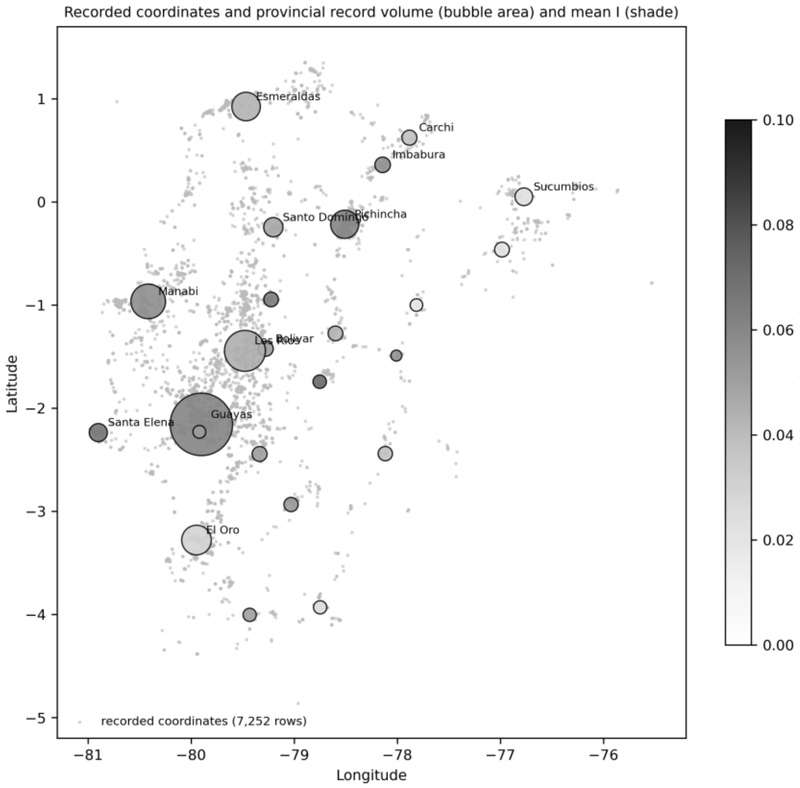

Figure 2: Recorded coordinates of the 7,252 rows (grey points; the national silhouette is drawn by the data themselves, no external geometry is used) and provincial bubbles at the median recorded coordinate, with area proportional to record count and shade proportional to the provincial mean indeterminacy I. Not population-normalized.

3.3 Weapon and administrative characteristics

The fabrication field classified 3,786 records as ARTESANAL (52.21 %) and 3,466 as IMPORTADA (47.79 %); the majority of craft-produced weapons is consistent with international evidence on improvised and craft-produced small arms [14]. The most frequent object types were PISTOLA (2,599; 35.84 %), REVOLVER (1,704; 23.50 %), ESCOPETA (1,299; 17.91 %) and CARTUCHERA (571; 7.87 %). The leading administrative conditions were APREHENDIDO (2,510; 34.61 %) and INCAUTADO (1,835; 25.30 %), followed by ABANDONADO (1,010), RETENIDO (930), ENTREGADO (494), ENCONTRADO (380), DECOMISO (81) and RECUPERADO (12). All eight conditions are determinate dictionary categories; the first version of this audit had treated the six conditions other than APREHENDIDO and INCAUTADO as indeterminate, which conflated "not a seizure" with "unknown".

3.4 Schema-induced indeterminacy in clase_arma

Table 2 gives the distribution of the principal sentinel by object type. The class field is almost a deterministic function of the object type: the modal class of each object type accounts for 99.39 % of records and the conditional entropy H(clase_arma | nombre_objeto) is 0.026 bits against an unconditional entropy of 1.832 bits. Pistols are SEMIAUTOMÁTICA (99.96 %), shotguns, carbines and cartucheras are TIRO A TIRO, rifles, sub-machine guns and machine guns are AUTOMÁTICA, and revolvers are SE DESCONOCE (1,702 of 1,704; 99.88 %), as are the catch-all categories OTROS ARMAS CORTAS (20 of 20) and OTROS ARMAS LARGAS (13 of 14). The three admissible classes (automatic, semi-automatic, single-shot) describe the firing mechanism of self-loading and single-shot weapons; a revolver is neither, so the schema's only admissible answer for it is the sentinel. The sentinel is thus inapplicability in the sense of database theory [18] and structured missingness in the statistical sense [21]: it is determined by another field, adds no information about the recording unit, and cannot be reduced by better recording without changing the category system. The chi-square test of the binary sentinel indicator against object type is accordingly extreme ($\chi^2(14) = 7{,}092$, Cramér's $V = 0.99$, 7 expected cells below five).

Table 2: Distribution of the principal sentinel (clase_arma = SE DESCONOCE) and of the modal class by object type (object types with at least 10 records).

Object type	Records	SE DESCONOCE	Sentinel share (%)	Modal class	Modal share (%)
PISTOLA	2,599	1	0.04	SEMIAUTOMATICA	99.96
REVOLVER	1,704	1,702	99.88	SE DESCONOCE	99.88
ESCOPETA	1,299	2	0.15	TIRO A TIRO	99.85
CARTUCHERA	571	2	0.35	TIRO A TIRO	99.65
CARABINA	456	1	0.22	TIRO A TIRO	99.78
FUSIL	288	0	0.00	AUTOMATICA	100.00
SUBAMETRALLADORA	119	0	0.00	AUTOMATICA	100.00
OTROS	87	52	59.77	SE DESCONOCE	59.77
RIFLE	50	0	0.00	AUTOMATICA	100.00
REPETIDORA	27	0	0.00	AUTOMATICA	100.00
OTROS ARMAS CORTAS	20	20	100.00	SE DESCONOCE	100.00
OTROS ARMAS LARGAS	14	13	92.86	SE DESCONOCE	92.86
AMETRALLADORA	14	0	0.00	AUTOMATICA	100.00

Note: H(clase_arma | nombre_objeto) = 0.026 bits; modal-class accuracy 99.39 %. Object types with sentinel share above 95 % and $n \geq 20$ *(REVOLVER, OTROS ARMAS CORTAS) define the non-applicability scenario.*

3.5 Exact repeated rows as batches

Exact comparison across all 27 variables identified 714 repeated configurations involving 1,972 rows (27.19 %), with 1,258 additional repetitions after the first occurrence of each configuration. Most configurations are pairs (524) or triples (110); the largest hold 56 and 53 rows (Appendix Figure A1). Because identical rows share the date, the subcircuit and the coordinates by definition, the batch structure is informative: the largest same-day, same-subcircuit batch comprised 194 repeated rows on 2026-01-14 in subcircuit 09D05C06S04 (Guayaquil), and 33.4 % of all repeated rows belong to same-day batches of at least five rows. Such batches are consistent with multi-weapon seizures recorded one row per weapon, which the dictionary's (absent) cantidad field would have made explicit; they are not evidence that the rows are false. On 14 January the file holds 230 rows, 203 of them repeated, which raises the January mean indeterminacy to 0.073 against 0.049 for the rest of the month. Duplication risk D is therefore reported alongside, not inside, the coordinates.

Table 3: Exact repeated-row audit (duplication risk D, reported separately from the evidence coordinates).

Indicator	Value	Reading
Repeated configurations	714	distinct 27-field rows occurring more than once
Rows in repeated configurations	1,972 (27.19 %)	D_r = 1
Additional exact repetitions	1,258 (17.35 %)	rows beyond the first occurrence
Distinct rows after exact deduplication	5,994	
Configuration sizes	2: 524; 3: 110; 4: 36; 5: 16; 6: 7; 7: 7; 8: 2; 10: 1; 11: 3; 12: 1; 13: 1; 20: 2; 21: 1; 25: 1; 53: 1; 56: 1	rows per configuration: count
Largest same-day same-subcircuit batches	2026-01-14 / 09D05C06S04: 194; 2026-02-04 / 09D05C04S01: 35; 2026-04-10 / 23D02C03S01: 20; 2026-02-17 / 08D02C03S03: 16	date / subcircuit code: rows
Repeated rows in same-day batches of ≥ 5	33.4 %	consistent with multi-weapon events

Note: without a unique event or weapon identifier, exact repetitions are not confirmed duplicated events.

3.6 Evidence coordinates: national values and comparison with the exclusive partition

Table 4 gives the national coordinates. The mean favourable evidence was $T = 0.952$ and the mean indeterminacy $I = 0.048$; 5,223 records (72.0 %) had T_r = 1, 2,029 (28.0 %) had I_r > 0, and 85.3 % of the latter owed their indeterminacy to the clase_arma sentinel alone. Contrary evidence was almost absent: F_r > 0 in 2 records (the two LANZA GRANADA rows, outside the dictionary), and every cross-field check (province code and name, canton prefix, zone code and name, coordinates and dates) passed for all 7,252 rows. $T_r + I_r$ = 1 for all but those two records, and $T+I+F$ was not constrained. Under the non-applicability scenario, T_{na} = 0.991 and I_{na} = 0.009, with I_{na} > 0 in 5.1 % of records.

The exclusive partition of the first version had assigned 34.67 % of records to T, 47.99 % to I and 17.35 % to F. Every one of those three figures is reproduced by the accompanying script, and each is now attributable to a construction choice rather than to the data: the F share is the additional-repetition rate (D), which is not contrary evidence; the I share combines the sentinel rows with the 2,907 records whose administrative condition is determinate but is not a seizure; and the T share is what remains after both subtractions. Conventional indicators (completeness 100 %; validity rate of clase_arma 75.2 %; 1,801 rows with any sentinel; additional-duplicate rate 17.35 %) are also in Table 4. What the coordinates add over them is stated in Section 3.9.

Table 4: National evidence coordinates, duplication risk and conventional indicators, compared with the exclusive T/I/F partition of the first version.

Indicator	Primary coordinates ($K=6, M=7$)	Non-applicability scenario	Exclusive partition (first version)
T (favourable evidence, mean)	0.952	0.991	—
I (indeterminacy, mean)	0.048	0.009	0.480 (share of rows)
F (contrary evidence, mean)	0.0000	0.0000	0.173 (= repetition rate)
Records with $I>0$	2,029 (28.0 %)	371 (5.1 %)	3,480
Records with $F>0$	2	2	1,258
Records with $T=1$	5,223 (72.0 %)	6,881	2,514
D (rows in repeated configurations / additional)	1,972 / 1,258	same	folded into F
Completeness (null cells)	100 % (0)	—	—
Validity rate of clase_arma (non-sentinel)	75.22 %	—	—
Rows with any sentinel	1,801	—	—

3.7 Territorial and monthly coordinates

Table 5 gives the provincial coordinates with bootstrap intervals. Provincial mean I ranged from 0.018 (Napo) to 0.066 (Chimborazo); F was zero in every province. The standardized ratio ρ_a lay between 0.97 and 1.39, so no province had less than the indeterminacy its object mix implies and Loja, Guayas and Azuay had the most excess. Across the 63 cantons with at least 30 records (Figure 3), mean I ranged from 0.017 to 0.096, correlated at $r=0.78$ with the canton's revolver share and at $r=0.81$ with the value expected from its object mix, and the largest excess ratios were Guayaquil (1.51), Esmeraldas (1.44) and Sucre (2.76). The Kruskal–Wallis test of I_r across provinces was significant (H = 132.8, $p<0.001$) with a negligible effect size (ε^2 = 0.015), which persisted after excluding revolver rows (ε^2 = 0.018); by month the effect was also negligible (ε^2 = 0.017). In the permutation comparison the observed weighted deviation of provincial I from the object-mix expectation was smaller than in every one of 500 random reassignments of I across records: the provincial pattern is what the object composition implies, not a territorial recording effect. Figure 2 and Table 5 must therefore be read as maps of where revolvers are recorded, not of where recording is worse.

Table 5: Provincial evidence coordinates (means over records), duplication-free; bootstrap intervals from $B=1$,000 resamples.

Territorial category	n	T	I (95 % bootstrap)	F	Revolver share (%)	Expected I from object mix	ρ (observed / expected)	QVol	Rank raw / QVol
Guayas	2,551	0.943	0.057 (0.053–0.060)	0.000	24.7	0.044	1.30	2,406.3	1 / 1
Los Rios	1,056	0.958	0.042 (0.038–0.047)	0.000	23.7	0.041	1.02	1,011.3	2 / 2
Manabi	739	0.947	0.053 (0.047–0.059)	0.000	26.9	0.046	1.14	700.0	3 / 3
El Oro	522	0.973	0.027 (0.022–0.033)	0.000	13.4	0.025	1.08	507.7	4 / 4
Esmeraldas	479	0.959	0.041 (0.034–0.047)	0.000	19.4	0.035	1.18	459.5	5 / 5
Pichincha	464	0.944	0.056 (0.049–0.064)	0.000	29.7	0.052	1.08	437.8	6 / 6
Santo Domingo de los Tsáchilas	182	0.956	0.044 (0.033–0.055)	0.000	24.2	0.041	1.06	174.0	7 / 7
Santa Elena	152	0.939	0.061 (0.048–0.075)	0.000	30.9	0.053	1.15	142.7	8 / 8

Sucumbios	142	0.979	0.021 (0.012–0.031)	0.000	9.2	0.018	1.20	139.0	9 / 9
Imbabura	94	0.949	0.051 (0.037–0.067)	0.000	28.7	0.048	1.07	89.2	10 / 10
Bolivar	86	0.953	0.047 (0.031–0.064)	0.000	27.9	0.047	1.00	82.0	11 / 11
Carchi	85	0.965	0.035 (0.022–0.049)	0.000	18.8	0.032	1.12	82.0	12 / 12
Canar	82	0.953	0.047 (0.030–0.063)	0.000	26.8	0.045	1.04	78.2	13 / 13
Tungurahua	79	0.960	0.040 (0.025–0.057)	0.000	24.1	0.041	0.97	75.8	14 / 15
Orellana	79	0.981	0.019 (0.008–0.032)	0.000	11.4	0.019	0.99	77.5	14 / 14
Cotopaxi	76	0.941	0.059 (0.042–0.077)	0.000	27.6	0.053	1.11	71.5	16 / 16
Morona Santiago	70	0.964	0.036 (0.019–0.052)	0.000	18.6	0.034	1.05	67.5	17 / 17
Azuay	68	0.951	0.049 (0.032–0.069)	0.000	20.6	0.039	1.26	64.7	18 / 18
Chimborazo	53	0.934	0.066 (0.041–0.088)	0.000	35.8	0.063	1.05	49.5	19 / 19
Loja	50	0.953	0.047 (0.027–0.070)	0.000	20.0	0.033	1.39	47.7	20 / 21
Zamora Chinchipe	50	0.977	0.023 (0.010–0.040)	0.000	12.0	0.023	1.00	48.8	20 / 20
Mar Territorial	42	0.952	0.048 (0.024–0.071)	0.000	26.2	0.044	1.09	40.0	22 / 22
Napo	38	0.982	0.018 (0.004–0.035)	0.000	10.5	0.018	0.99	37.3	23 / 23
Pastaza	13	0.949	0.051 (0.013–0.090)	0.000	30.8	0.051	1.00	12.3	24 / 24

Note: ρ from Eq. (7); Mar Territorial is a source category, not a province.

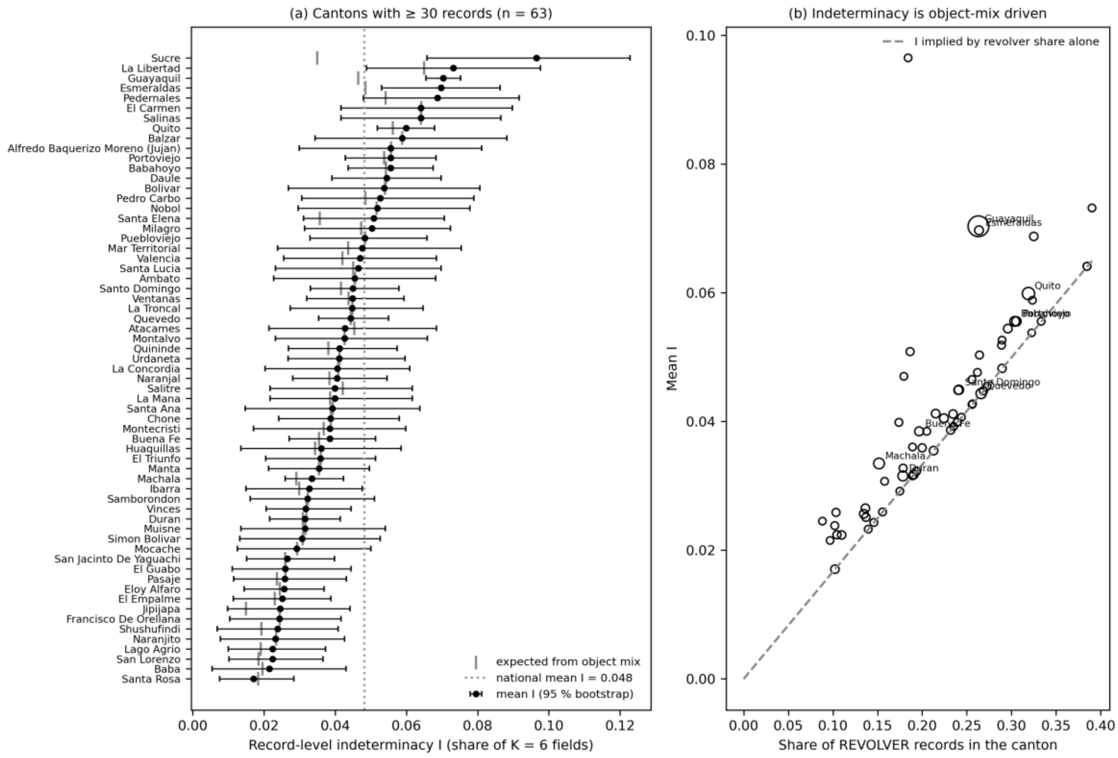

Figure 3: Canton-level indeterminacy. (a) Mean record-level I with 95 % bootstrap intervals for the 63 cantons with at least 30 records, with the value expected from each canton's object mix (grey ticks) and the national mean (dotted). (b) Mean I against the canton's share of REVOLVER records; the dashed line is the indeterminacy implied by the revolver share alone; marker area proportional to record count.

3.8 Raw versus quality-aware ranking

Because T_r is 1 for 72 % of records and at least 1 − 1/K for all but 69 of the rest, quality-aware volume is nearly proportional to raw count. The Spearman correlation between the raw and the QVol canton rankings was $\rho = 0.999$ (1.000 in the non-applicability scenario); among the 20 largest cantons no rank moved by more than one position (Figure 4, Table 6). The first version had reported that restricting the ranking to T records moved Machala from third to second, Quito from second to third, Santa Rosa from eighteenth to ninth and El Guabo from fortieth to a tie for fifteenth. Those movements are reproduced by the exclusive rule and are caused by it: they track the local share of non-seizure conditions and of revolvers, neither of which is a quality defect. The negative result is the finding: in this file, a quality-aware volume does not reorder the territory, because indeterminacy is low, contrary evidence is absent and both are spread in proportion to the object mix.

Table 6: The twelve largest cantons: evidence coordinates, standardized indeterminacy ratio and rankings.

Canton	n	T	I (95 % bootstrap)	Revolver share (%)	ρ	QVol	Rank raw	Rank QVol	Rank QVol (non-appl.)
Guayaquil	1,365	0.929	0.070 (0.065–0.075)	26.3	1.51	1,268.7	1	1	1
Quito	373	0.940	0.060 (0.052–0.068)	31.9	1.07	350.7	2	2	2
Machala	244	0.967	0.033 (0.026–0.042)	15.2	1.15	235.8	3	3	3
Quevedo	203	0.956	0.044 (0.035–0.055)	26.6	1.00	194.0	4	4	4
Duran	185	0.968	0.032 (0.022–0.041)	17.8	1.02	179.2	5	5	5
Babahoyo	168	0.944	0.056 (0.044–0.067)	30.4	1.02	158.7	6	6	6
Portoviejo	144	0.944	0.056 (0.043–0.068)	30.6	1.03	136.0	7	7	7
Santo Domingo	141	0.955	0.045 (0.033–0.058)	24.1	1.08	134.7	8	8	8
Buena Fe	117	0.962	0.038 (0.027–0.051)	19.7	1.08	112.5	9	9	9
Esmeraldas	110	0.930	0.070 (0.053–0.086)	26.4	1.44	102.3	10	11	10
Naranjal	107	0.960	0.040 (0.028–0.055)	22.4	1.05	102.7	11	10	11
Vinces	105	0.968	0.032 (0.021–0.044)	19.0	1.00	101.7	12	12	12

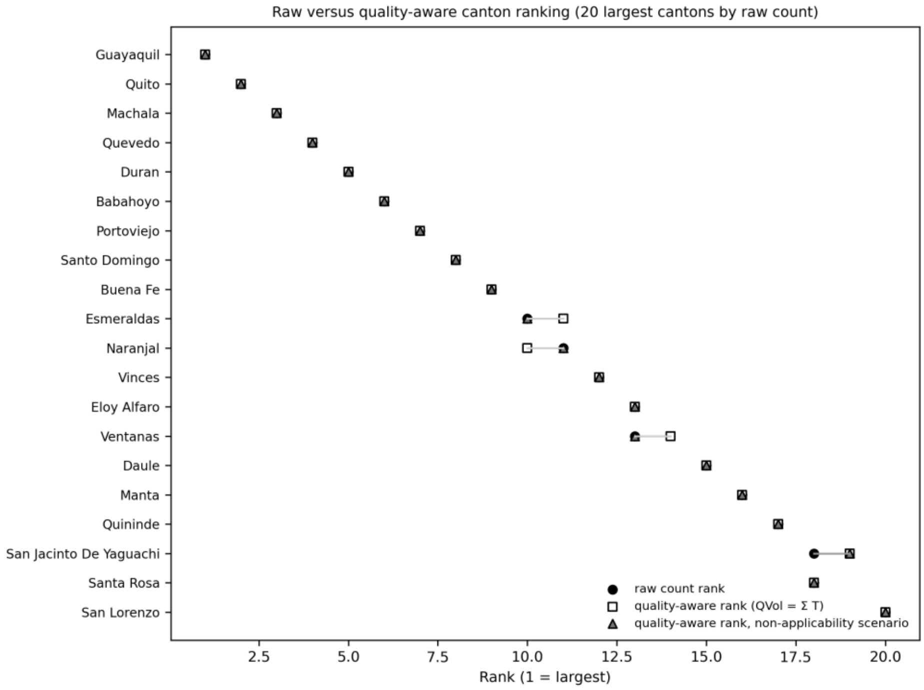

Figure 4: Raw count rank (circles), quality-aware rank by QVol = Σ T_r (squares) and quality-aware rank in the non-applicability scenario (triangles) for the 20 largest cantons. Rankings are descriptive and do not measure population risk or institutional performance.

3.9 Sensitivity and what the coordinates add

Table 7 gives the sensitivity scenarios. The national mean I moved between 0.009 and 0.062 and the share of records with $I > 0$ between 5.1 % and 28.1 %; the canton QVol ranking correlated at $\rho \geq 0.999$ with the primary and at $\rho \geq 0.999$ with the raw ranking in every scenario. The coordinates are thus stable in what they say and in what they do not change.

Relative to the conventional indicators of Table 4, the coordinates add four things that can be pointed to. First, a separation that the indicators lack: completeness (100 %), determinacy (I) and contradiction (F) are three numbers with three sources, and the file scores perfectly on the first, moderately on the second and perfectly on the third; a single validity rate would have merged the last two. Second, a level of aggregation: the record-level I_r can be averaged along any hierarchy with an interval, standardized by object mix and summed into a volume, whereas a field-level sentinel share cannot be attached to a canton without recomputation. Third, a diagnosis: measuring I per field made it visible that 85 % of indeterminate records owe their state to one field, and Eq. (8) then showed that field to be schema-determined. Fourth, a correction: the exclusive partition, which used the same vocabulary, produced an I share ten times larger and an F share that was in fact D; the revised construction is what makes that error legible. What the coordinates do not add, in this file, is a different ranking.

Table 7: Sensitivity of the coordinates to the field set and to the treatment of catch-all and undefined values.

Scenario	K	mean T	mean I	records with $I > 0$	ρ canton rank vs primary	ρ canton rank vs raw
primary ($K = 6$)	6	0.952	0.048	28.0 %	1.000	0.999
non-applicability: clase_arma dropped for objects with sentinel share > 95 %	6	0.991	0.009	5.1 %	0.999	1.000
catch-all OTROS* counted as indeterminate	6	0.949	0.051	28.1 %	1.000	0.999
lugar outside dictionary counted as determinate	6	0.959	0.041	24.8 %	1.000	0.999
lugar excluded ($K = 5$)	5	0.950	0.050	24.8 %	1.000	0.999
tipo_delito and lugar excluded ($K = 4$)	4	0.938	0.062	24.8 %	1.000	0.999

4 Discussion

4.1 Principal finding

The official file is physically complete and, by the evidence available in the file and its dictionary, almost free of contradiction; its semantic problem is one field whose sentinel is the schema's only admissible value for a quarter of the weapons recorded. Stated in the coordinates: $T = 0.952, I = 0.048$, F ≈ 0 at record level; I falls to 0.009 if the field is treated as inapplicable to revolvers. The territorial variation of I is the variation of the revolver share, and the quality-aware volume ranks cantons as the raw count does.

4.2 Why the exclusive partition misled

The first version's 34.7 / 48.0 / 17.4 partition was reproducible and wrong in its semantics. Three choices produced it: repetition was named F, so a property of pairs of rows was attributed to single rows as falsity; six determinate administrative conditions were named I, so "not a seizure" became "unknown"; and the states were made exclusive with precedence F > I > T, so the shares summed to one by construction and any change in one moved the others. None of these is a property of the data. A neutrosophic representation is only as good as the evidential semantics of its coordinates, and an exclusive partition with a forced unit sum manufactures indeterminacy and falsity from ordinary categories. This is the methodological lesson the paper offers to the journal: T, I and F must each be traced to a verifiable kind of evidence, and where a kind of evidence is absent the coordinate should be near zero rather than filled by a proxy.

4.3 Four kinds of non-information

Physical missingness (none here), semantic indeterminacy (the sentinels and the undefined lugar values), non-applicability (the revolver case) and dictionary drift (tipo_lugar, the missing cantidad and calibre, the plural and gender variants) are distinct, and the audit only became informative once they were separated. The distinction between applicable-but-unknown and inapplicable information is older than most data-quality frameworks [18] and is now studied as structured missingness [21]; sentinel values are the form it takes in published administrative files. Contrary evidence is a fifth kind and was nearly empty, which is itself a finding: the file's internal consistency (codes, names, prefixes, zones, coordinates, dates) is complete.

4.4 Schema-induced indeterminacy and the revolver

With H(clase_arma | nombre_objeto) = 0.026 bits, clase_arma is redundant with nombre_objeto for every object type and uninformative for revolvers. Two consequences follow for data owners. The field either needs a category that applies to revolving-cylinder weapons or should be documented as inapplicable to them, in which case SE DESCONOCE should be replaced by an explicit NOT APPLICABLE code so that users do not read it as unknown. And the first version's ranking of provinces by sentinel share (Chimborazo 37.7 %, Santa Elena 32.9 %, Pichincha 31.7 %) should not be read as a ranking of recording performance: standardized by object mix, no province falls below what its mix implies and the excess ratios are modest (1.0–1.4). The pattern is a metadata issue, as the first version already suspected, and the present analysis shows why.

4.5 Duplication is not falsity

Exact repeated rows are a duplication risk, not contrary evidence about any row. The batch structure, with a third of repeated rows in same-day, same-subcircuit groups of five or more and one group of 194 rows on a single day, is what multi-weapon seizures look like when a file records one row per weapon and omits the quantity field its dictionary defines. Without an identifier the alternative explanation (double entry) cannot be excluded for pairs, and record-linkage methods [27, 28] cannot be applied where all fields are identical. Property 2 keeps the coordinates independent of this ambiguity and makes the choice explicit: QVol counts rows, and therefore weapons if the batch reading is right.

4.6 What the coordinates add and what they do not

Section 3.9 lists the four additions: separation of sources, aggregation with uncertainty, field-level diagnosis and the correction of the earlier partition. The negative result is equally important and should not be softened: quality-aware volume does not reorder the cantons ($\rho = 0.999$), because the only substantial indeterminacy is schema-induced and proportional to the object mix. A reader who wants a dashboard indicator gets one, T_a, and is told that in this file it changes nothing. Neutrosophic quality indicators must be judged by what they reveal, not by whether they produce a different answer.

4.7 Relation to Neutrosophic Paraconsistent Logic and neutrosophic data models

The coordinates borrow from NPL the separation of favourable evidence, contrary evidence and indeterminacy without a unit-sum constraint [16] and from neutrosophic set theory the independence of components in the sense that their sum is unconstrained [17]; they do not use NPL's evidence-synthesis procedure, and I here is evidential (a marker in a cell, a value outside a dictionary, an inapplicable field), not ontological. Neutrosophic relational models represent unknown attribute values inside the database [24]; the present audit measures them from outside, on a published file, and adds the non-applicability reading. Two design rules from the comparison with the first version generalize: an evidence coordinate must not be a relabelled property of a different object (repetition is a property of pairs), and normalizing or partitioning coordinates changes what they say, as normalization changes rank orders in neutrosophic triples generally [33].

4.8 Evidence-aware dashboarding

The results support a quality-aware dashboard for periodic territorial review in which every count is shown with its T_a and I_a, every percentage with its denominator, the revolver share beside the sentinel share, the raw and QVol rankings side by side, and the repeated-row batches listed rather than deleted. It supports situational review and data-quality monitoring of a static file; it does not support automated deployment, policing recommendations or claims about institutional performance, for which unit-level metadata, capacity data and a comparison design would be required.

4.9 Limitations

Only counts, shares and coordinates within the file were analysed; no population or exposure denominator exists, so nothing here is a crime rate or a risk estimate. The design is observational and retrospective. The coordinates depend on the choice of K fields, on equal weights and on the label-variant map; the sensitivity analysis covers the first two, and the map is published. Contrary evidence is limited to checks that the file and dictionary make possible; cross-source validation against police or judicial records was not available. The non-applicability reading of the revolver sentinel is an inference from the conditional distribution, not a statement of the data owner. Coordinates were not audited for positional accuracy beyond the national bounding box, and 18 coordinate pairs are used by ten or more rows, which suggests facility-level geocoding for some records. Coverage is seven months of one year. The dictionary is a 2025 document applied to a 2026 file.

4.10 Future research

Four directions follow. First, validation of the field semantics with the data owner: whether clase_arma is inapplicable to revolvers, whether cantidad was dropped, and whether same-day batches are single operations. Second, linkage to institutional metadata (recording unit, operation type, original report format) to test whether the residual excess ratios ρ_a above one are unit effects. Third, extension of the coordinates to positional evidence (geocoding precision, facility placeholders) and to multi-year files, where dictionary drift can be tracked as a time series of I. Fourth, on the neutrosophic side, a comparison of these field-grounded coordinates with neutrosophic relational representations of the same file [24] and with NPL evidence synthesis across sources [16] once a second source exists; any such extension should keep the rule applied here, that a coordinate whose value can be produced by a construction choice rather than by evidence is not a coordinate.

5 Conclusions

Empirically, the official illicit-firearms file for January–July 2026 is physically complete, internally consistent and semantically indeterminate in one field: clase_arma carries SE DESCONOCE in 24.78 % of records because the schema has no class for revolvers, and the field is otherwise a function of the object type. Territorial differences in indeterminacy are differences in the revolver share; quality-aware volume ranks cantons as raw counts do; exact repeated rows form same-day batches consistent with multi-weapon seizures.

Methodologically, field-level evidence coordinates with explicit sources, no unit sum and duplication kept apart reproduce every figure of the earlier exclusive partition and show that its 48 % indeterminacy and 17 % falsity were produced by the partition, not by the data. Neutrosophic quality assessment of administrative records should therefore begin by naming the kind of evidence behind each coordinate; where the evidence is absent, the honest value is near zero, and where the evidence is schema-induced, the remedy is metadata, not recording. Territorial rankings should be accompanied by these coordinates and never read as population crime rates, causal effects or police effectiveness.

Declarations

Author Contributions: Conceptualization, methodology, validation, interpretation, writing and final approval were performed collectively by the listed authors.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. The study used an open administrative dataset without personal identifiers and involved no intervention or contact with human participants.

Informed Consent Statement: Not applicable.

Data Availability: The analysed dataset is available through Ecuador's official open-data portal: https://www.datosabiertos.gob.ec/dataset/armas-ilicitas. The interactive dashboard reported by the authors is available at https://radiant-medovik-4c97ff.netlify.app/. The reproducibility package accompanying this manuscript contains the source spreadsheets, the two analysis scripts, the record- and group-level coordinate files, the coordinate definitions, the derived tables and the figures.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration of Generative AI Use: Generative artificial intelligence tools (Antigravity CLI with Gemini, ChatGPT/Codex, and Claude) were used to assist with code review, the second-round reanalysis, language editing and document formatting. All numerical results were computed from the original Ministry of the Interior spreadsheets by the accompanying scripts and verified by the authors, who retain full responsibility for methodological decisions, interpretation, source verification and the final content.

Appendix A. Canton-level coordinates and repeated-configuration sizes

Table A1: Evidence coordinates for the 63 cantons with at least 30 records.

Canton	n	T	I	I 95 % bootstrap	F	Rows in repeated conf. (%)	Revolver share (%)	ρ	Rank raw	Rank QVol
Guayaquil	1,365	0.929	0.070	0.065–0.075	0.000	33.5	26.3	1.51	1	1
Quito	373	0.940	0.060	0.052–0.068	0.000	9.7	31.9	1.07	2	2
Machala	244	0.967	0.033	0.026–0.042	0.000	27.0	15.2	1.15	3	3
Quevedo	203	0.956	0.044	0.035–0.055	0.000	20.2	26.6	1.00	4	4
Duran	185	0.968	0.032	0.022–0.041	0.000	12.4	17.8	1.02	5	5
Babahoyo	168	0.944	0.056	0.044–0.067	0.000	23.2	30.4	1.02	6	6
Portoviejo	144	0.944	0.056	0.043–0.068	0.000	31.9	30.6	1.03	7	7
Santo Domingo	141	0.955	0.045	0.033–0.058	0.000	28.4	24.1	1.08	8	8
Buena Fe	117	0.962	0.038	0.027–0.051	0.000	18.8	19.7	1.08	9	9
Esmeraldas	110	0.930	0.070	0.053–0.086	0.000	14.5	26.4	1.44	10	11
Naranjal	107	0.960	0.040	0.028–0.055	0.000	38.3	22.4	1.05	11	10
Vinces	105	0.968	0.032	0.021–0.044	0.000	26.7	19.0	1.00	12	12
Eloy Alfaro	104	0.974	0.026	0.014–0.037	0.000	40.4	13.5	1.05	13	13
Ventanas	104	0.955	0.045	0.032–0.059	0.000	17.3	24.0	1.03	13	14
Daule	98	0.946	0.054	0.039–0.070	0.000	26.5	29.6	1.00	15	15
Manta	94	0.965	0.035	0.021–0.050	0.000	24.5	21.3	1.00	16	16
Quininde	93	0.959	0.041	0.027–0.057	0.000	30.1	21.5	1.08	17	17

San Jacinto De Yaguachi	88	0.973	0.027	0.015–0.040	0.000	55.7	13.6	1.02	18	19
Santa Rosa	88	0.983	0.017	0.008–0.028	0.000	35.2	10.2	0.93	18	18
San Lorenzo	82	0.978	0.022	0.010–0.037	0.000	28.0	11.0	1.21	20	20
Urdaneta	81	0.959	0.041	0.027–0.060	0.000	30.9	23.5	1.00	21	21
Pedernales	80	0.931	0.069	0.048–0.092	0.000	32.5	32.5	1.27	22	22
Puebloviejo	76	0.952	0.048	0.033–0.066	0.000	18.4	28.9	1.00	23	23
El Empalme	73	0.975	0.025	0.011–0.039	0.000	30.1	13.7	1.09	24	24
Chone	69	0.961	0.039	0.024–0.058	0.000	30.4	23.2	1.00	25	25
La Troncal	67	0.955	0.045	0.027–0.065	0.000	50.7	26.9	1.00	26	27
Lago Agrio	67	0.978	0.022	0.010–0.037	0.000	31.3	10.4	1.17	26	26
El Triunfo	65	0.964	0.036	0.021–0.051	0.000	24.6	20.0	1.00	28	28
Samborondon	62	0.968	0.032	0.016–0.051	0.000	30.6	19.4	1.00	29	29
Santa Elena	59	0.949	0.051	0.031–0.071	0.000	20.3	18.6	1.42	30	31
Pasaje	58	0.974	0.026	0.011–0.043	0.000	8.6	10.3	1.09	31	30
Ibarra	56	0.967	0.033	0.015–0.048	0.000	19.6	17.9	1.09	32	32
Milagro	53	0.950	0.050	0.031–0.072	0.000	18.9	26.4	1.06	33	33
El Carmen	52	0.936	0.064	0.042–0.090	0.000	28.8	38.5	1.00	34	34
Salinas	52	0.936	0.064	0.042–0.087	0.000	25.0	38.5	1.00	34	34
Shushufindi	49	0.976	0.024	0.007–0.041	0.000	34.7	10.2	1.23	36	36
Francisco De Orellana	48	0.976	0.024	0.010–0.042	0.000	8.3	14.6	0.99	37	37
Salitre	46	0.960	0.040	0.022–0.062	0.000	34.8	23.9	0.95	38	38
La Mana	46	0.960	0.040	0.022–0.062	0.000	23.9	17.4	1.03	38	38
El Guabo	45	0.974	0.026	0.011–0.044	0.000	35.6	15.6	0.99	40	40
Nobol	45	0.948	0.052	0.030–0.078	0.000	33.3	28.9	1.00	40	41
Ambato	44	0.955	0.045	0.023–0.068	0.000	15.9	27.3	1.00	42	42
Montalvo	43	0.957	0.043	0.023–0.066	0.000	46.5	25.6	1.00	43	44
Naranjito	43	0.977	0.023	0.008–0.043	0.000	39.5	14.0	0.99	43	43

Santa Lucia	43	0.953	0.047	0.023–0.070	0.000	32.6	25.6	1.03	43	45
Mar Territorial	42	0.952	0.048	0.024–0.075	0.000	19.0	26.2	1.09	46	46
La Concordia	41	0.959	0.041	0.020–0.061	0.000	22.0	24.4	1.00	47	47
La Libertad	41	0.927	0.073	0.049–0.098	0.000	19.5	39.0	1.13	47	49
Mocache	40	0.971	0.029	0.013–0.050	0.000	25.0	17.5	1.00	49	48
Alfredo Baquerizo Moreno (Jujan)	39	0.944	0.056	0.030–0.081	0.000	46.2	33.3	1.00	50	54
Atacames	39	0.957	0.043	0.021–0.068	0.000	10.3	25.6	0.94	50	51
Valencia	39	0.953	0.047	0.026–0.068	0.000	15.4	17.9	1.12	50	52
Montecristi	39	0.962	0.038	0.017–0.060	0.000	20.5	20.5	1.04	50	50
Pedro Carbo	38	0.947	0.053	0.031–0.079	0.000	28.9	28.9	1.09	54	55
Simon Bolivar	38	0.969	0.031	0.013–0.053	0.000	36.8	15.8	0.99	54	53
Sucre	38	0.904	0.096	0.066–0.123	0.000	47.4	18.4	2.76	54	58
Huaquillas	37	0.964	0.036	0.014–0.059	0.000	8.1	18.9	1.05	57	57
Muisne	37	0.968	0.032	0.014–0.054	0.000	35.1	18.9	1.00	57	56
Balzar	34	0.941	0.059	0.034–0.088	0.000	29.4	32.4	1.00	59	61
Jipijapa	34	0.975	0.025	0.010–0.044	0.000	38.2	8.8	1.65	59	59
Santa Ana	34	0.961	0.039	0.015–0.064	0.000	29.4	23.5	1.00	59	60
Baba	31	0.978	0.022	0.005–0.043	0.000	54.8	9.7	1.10	62	62
Bolivar	31	0.946	0.054	0.027–0.081	0.000	35.5	32.3	1.00	62	63

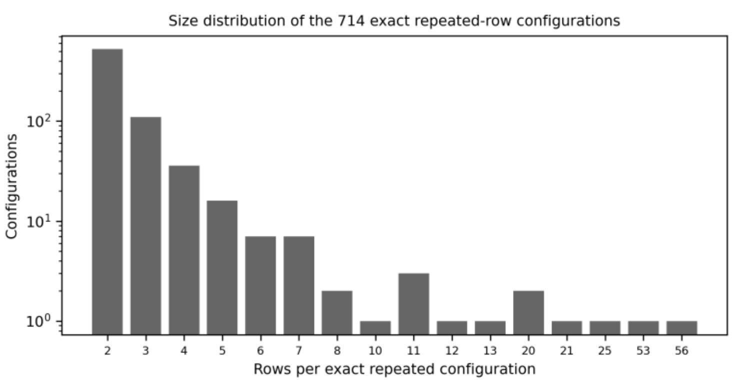

Figure A1: Number of exact repeated-row configurations by size (rows per configuration; logarithmic scale).

References

[1] Ministry of the Interior of Ecuador. Estadísticas de Armas Ilícitas (enero–julio de 2026); Dirección de Estadística y Economía de la Seguridad: Quito, Ecuador, 2026. Available online: https://www.datosabiertos.gob.ec/dataset/armas-ilicitas (accessed on 22 August 2026).

[2] Batini, C.; Scannapieco, M. Data and Information Quality: Dimensions, Principles and Techniques; Springer: Cham, Switzerland, 2016. https://doi.org/10.1007/978-3-319-24106-7

[3] Observatorio Ecuatoriano de Crimen Organizado; Pan American Development Foundation. ¿El paraíso perdido? Tráfico de armas de fuego y violencia en Ecuador; OECO/PADF: Quito, Ecuador, 2024. Available online: https://oeco.padf.org/el-paraiso-perdido-trafico-de-armas-de-fuego-y-violencia-en-ecuador/ (accessed on 22 August 2026).

[4] United Nations Office on Drugs and Crime. Global Study on Firearms Trafficking 2020; United Nations: Vienna, Austria, 2020. Available online: https://www.unodc.org/documents/data-and-analysis/Firearms/2020_REPORT_Global_Study_on_Firearms_Trafficking_2020_web.pdf (accessed on 22 August 2026).

[5] United Nations Statistics Division. United Nations National Quality Assurance Frameworks Manual for Official Statistics; United Nations: New York, NY, USA, 2019. Available online: https://unstats.un.org/unsd/methodology/dataquality/un-nqaf-manual/ (accessed on 22 August 2026).

[6] United Nations Statistics Division. Module for Quality Assurance When Using Administrative and Other Data Sources to Produce Official Statistics; United Nations: New York, NY, USA, 2025. Available online: https://unstats.un.org/UNSDWebsite/statcom/session_56/documents/BG-3r-Module_for_QA_when_using_AOS-2025-02-03-E.pdf (accessed on 22 August 2026).

[7] Wang, R.Y.; Strong, D.M. Beyond accuracy: what data quality means to data consumers. Journal of Management Information Systems 1996, 12(4), 5–33. https://doi.org/10.1080/07421222.1996.11518099

[8] Pipino, L.L.; Lee, Y.W.; Wang, R.Y. Data quality assessment. Communications of the ACM 2002, 45(4), 211–218. https://doi.org/10.1145/505248.506010

[9] Batini, C.; Cappiello, C.; Francalanci, C.; Maurino, A. Methodologies for data quality assessment and improvement. ACM Computing Surveys 2009, 41(3), Article 16. https://doi.org/10.1145/1541880.1541883

[10] Eurostat. European Statistical System Quality Assurance Framework, Version 2.0; Publications Office of the European Union: Luxembourg, 2019. Available online: https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf (accessed on 22 August 2026).

[11] Daas, P.; Ossen, S.; Vis-Visschers, R.; Arends-Tóth, J. Checklist for the Quality Evaluation of Administrative Data Sources; Statistics Netherlands / Eurostat: Luxembourg, 2009. Available online: https://ec.europa.eu/eurostat/documents/64157/4374310/45-Checklist-quality-evaluation-administrative-data-sources-2009.pdf (accessed on 22 August 2026).

[12] United Nations Office on Drugs and Crime. International Statistics on Crime and Justice; United Nations: Vienna, Austria, 2010. Available online: https://www.unodc.org/documents/data-and-analysis/Crime-statistics/International_Statistics_on_Crime_and_Justice.pdf (accessed on 22 August 2026).

[13] United Nations Office on Drugs and Crime. Criminal Justice Assessment Toolkit: Criminal Justice Information; United Nations: New York, NY, USA, 2006. Available online: https://www.unodc.org/documents/justice-and-prison-reform/cjat_eng/1_Criminal_Justice_Information.pdf (accessed on 22 August 2026).

[14] Hays, G.; Jenzen-Jones, N.R. Beyond State Control: Improvised and Craft-produced Small Arms and Light Weapons; Small Arms Survey / Armament Research Services: Geneva, Switzerland, 2018. Available online: https://www.smallarmssurvey.org/resource/beyond-state-control-improvised-and-craft-produced-small-arms-and-light-weapons (accessed on 22 August 2026).

[15] Smarandache, F. Introduction to Neutrosophic Statistics; Sitech & Education Publishing: Craiova, Romania, 2014. arXiv:1406.2000. https://arxiv.org/abs/1406.2000

[16] Leyva-Vázquez, M.; Smarandache, F. Neutrosophic Paraconsistent Logic: evidence degrees, ontological indeterminacy, and scientific evidence synthesis. Neutrosophic Computing and Machine Learning 2026, 43, 211–221. https://doi.org/10.5281/zenodo.20150483

[17] Smarandache, F. Degree of dependence and independence of the (sub)components of fuzzy set and neutrosophic set. Neutrosophic Sets and Systems 2016, 11, 95–97. https://doi.org/10.5281/zenodo.571359

[18] Codd, E.F. Missing information (applicable and inapplicable) in relational databases. ACM SIGMOD Record 1986, 15(4), 53–78. https://doi.org/10.1145/16301.16303

[19] Libkin, L. SQL's three-valued logic and certain answers. ACM Transactions on Database Systems 2016, 41(1), 1–28. https://doi.org/10.1145/2877206

[20] Rubin, D.B. Inference and missing data. Biometrika 1976, 63(3), 581–592. https://doi.org/10.1093/biomet/63.3.581

[21] Mitra, R.; McGough, S.F.; Chakraborti, T.; Holmes, C.; Copping, R.; Hagenbuch, N.; Biedermann, S.; Noonan, J.; Lehmann, B.; Shenvi, A.; et al. Learning from data with structured missingness. Nature Machine Intelligence 2023, 5(1), 13–23. https://doi.org/10.1038/s42256-022-00596-z

[22] Belnap, N.D. A useful four-valued logic. In Modern Uses of Multiple-Valued Logic; Dunn, J.M., Epstein, G., Eds.; D. Reidel: Dordrecht, 1977; pp. 5–37. https://doi.org/10.1007/978-94-010-1161-7_2

[23] Carnielli, W.; Rodrigues, A. An epistemic approach to paraconsistency: a logic of evidence and truth. Synthese 2019, 196(9), 3789–3813. https://doi.org/10.1007/s11229-017-1621-7
[24] Arora, M.; Biswas, R.; Pandey, U.S. Neutrosophic relational database decomposition. International Journal of Advanced Computer Science and Applications 2011, 2(8), 121–125. https://doi.org/10.14569/ijacsa.2011.020822
[25] Sadiq, S.; Indulska, M. Open data: quality over quantity. International Journal of Information Management 2017, 37(3), 150–154. https://doi.org/10.1016/j.ijinfomgt.2017.01.003
[26] Vetrò, A.; Canova, L.; Torchiano, M.; Minotas, C.O.; Iemma, R.; Morando, F. Open data quality measurement framework: definition and application to open government data. Government Information Quarterly 2016, 33(2), 325–337. https://doi.org/10.1016/j.giq.2016.02.001
[27] Fellegi, I.P.; Sunter, A.B. A theory for record linkage. Journal of the American Statistical Association 1969, 64(328), 1183–1210. https://doi.org/10.1080/01621459.1969.10501049
[28] Elmagarmid, A.K.; Ipeirotis, P.G.; Verykios, V.S. Duplicate record detection: a survey. IEEE Transactions on Knowledge and Data Engineering 2007, 19(1), 1–16. https://doi.org/10.1109/TKDE.2007.250581
[29] Braga, A.A.; Wintemute, G.J.; Pierce, G.L.; Cook, P.J.; Ridgeway, G. Interpreting the empirical evidence on illegal gun market dynamics. Journal of Urban Health 2012, 89(5), 779–793. https://doi.org/10.1007/s11524-012-9681-y
[30] Pina-Sánchez, J.; Buil-Gil, D.; Brunton-Smith, I.; Cernat, A. The impact of measurement error in regression models using police recorded crime rates. Journal of Quantitative Criminology 2023, 39(4), 975–1002. https://doi.org/10.1007/s10940-022-09557-6
[31] Simpson, R.; Orosco, C. Re-assessing measurement error in police calls for service: classifications of events by dispatchers and officers. PLOS ONE 2021, 16(12), e0260365. https://doi.org/10.1371/journal.pone.0260365
[32] DiCiccio, T.J.; Efron, B. Bootstrap confidence intervals. Statistical Science 1996, 11(3), 189–228. https://doi.org/10.1214/ss/1032280214
[33] Smarandache, F. The rank order of the neutrosophic triplets (T, I, F) may change after normalization. International Journal of Research in Industrial Engineering 2024, 13(4), 427–435. https://doi.org/10.22105/riej.2024.487239.1489