

# Data Description – *Un Kilo de Ayuda*

Mauricio Fernández-Duque

October 2025

## 1 Purpose and Scope

This document is a brief introduction to the Mexican NGO [Un Kilo de Ayuda \(UKA\)](#) and a description of the data used in this teaching practice. The data shared with you is a privacy-preserving subset of UKA administrative and survey records prepared for instructional purposes. All personally identifiable information (PII) has been removed, and the files constitute a curated sub-sample suitable for classroom analysis.

### UKA and its program model

UKA serves children **from -9 to 60 months of age** (pregnancy to five years old) who live in rural and Indigenous communities with high or very high marginalization across Mexico, reaching ~23,000 children per year. UKA implements the *Modelo Integral para el Desarrollo Infantil Temprano* (MIDIT), organized around three mutually reinforcing axes:

1. **Neurological and Socio-emotional Development:** screening of neurodevelopment to identify lags and risks in gross/fine motor skills, language, socio-affective and cognitive domains; promotion of early stimulation via play-based workshops.
2. **Physical Development:** anthropometric assessment (weight, length/height) and nutritional status; capillary hemoglobin testing to detect, prevent, and treat iron-deficiency anemia; provision of preventive and therapeutic iron/micronutrient regimens.
3. **Community Development:** caregiver workshops on responsive parenting, malnutrition and anemia management, prevention of diarrheal and respiratory disease, early stimulation and play, emotional support; periodic delivery of nutrient-dense food packages to support household food security.

### Context: nutritional and developmental measurements

Under the Physical Development axis, UKA measures all children under five and computes WHO growth-standard indicators to evaluate nutritional status: *Weight-for-Age (W/A)*, *Weight-for-Length/Height (W/H)*, and *BMI*. These indicators classify normal nutrition, degrees of undernutrition, as well as overweight and obesity. Capillary hemoglobin is quantified

for children  $\geq 6$  months and  $< 5$  years to diagnose iron-deficiency anemia, with altitude adjustments based on the community's elevation. For child development, UKA applies the *Evaluación del Desarrollo Infantil (EDI)*, a screening tool designed and validated for Mexican populations that captures biological risk factors, warning signs, a brief neurological exam, and domain scores for gross/fine motor, language, social, and cognitive development.

**Measurement periodicity.** In routine operations, weight and length/height are collected approximately every three months, and capillary hemoglobin every six months. Children with moderate or severe anemia may receive a three-month follow-up measurement. During the COVID-19 pandemic (2020–2022), measurement frequency was temporarily reduced due to public-health restrictions.

## UKA Interventions

UKA provides a set of individualized interventions designed to promote children's health and early development.

Children with adjusted hemoglobin levels below 11 g/dL receive iron treatment in the form of iron syrup bottles, following clinical protocols for anemia management. These children are scheduled for follow-up hemoglobin testing to monitor progress and ensure recovery.

In the area of neurodevelopment, UKA uses the *Evaluación del Desarrollo Infantil (EDI)* to classify each child according to three levels: green (adequate development), yellow (at risk), and red (delay). Children classified as yellow receive counseling sessions for parents, aimed at stimulating the child's development in the domains where weaknesses are detected, such as fine or gross motor skills, language, or socio-affective areas. Children classified as red are referred to public health clinics for further evaluation and specialized attention.

Additionally, UKA delivers nutritional and multivitamin packages to families participating in the program and provides specialized therapeutic food to children who present malnutrition. Caregivers also attend educational workshops on topics such as early stimulation, play, nutrition, and responsive parenting, which seek to strengthen family practices that support early childhood development.

UKA periodically expands to new localities within the states where it operates, initiating the comprehensive set of interventions described above in newly incorporated communities. In 2023, UKA expanded its coverage to the state of Nuevo León, the most recent addition to its operational network. In 2024, the organization also implemented an emergency program in the state of Guerrero following a hurricane that caused severe damage in the region.

## Data frames provided

This release includes three databases (more below):

1. **Anthropometry & Hemoglobin Panel.** This database contains repeated clinic visits for each child (one row per child/visit). It documents weight, growth and anemia monitoring over time.
2. **EDI Assessments.** This database stores developmental screening events (one row per EDI application). It summarizes outcomes across domains and flags potential risks.

3. **Beneficiaries Census (mother registry).** This database is a survey of participating caregivers. One row per mother/primary caregiver (single record).

**Included files (each is a 25% simple random sample of its source):**

- `panel_anon` — child-period nutrition panel,
- `edi_anon` — child developmental screening records,
- `censo_anon` — household census interviews.

### Consolidated Dataset via Fuzzy Matching

Due to historical data-management issues at UKA, **unique child identifiers are not consistently preserved across the three operational sources** (`panel_anon`, `edi_anon`, `censo_anon`). To address this, we constructed a single, consolidated dataset using a transparent, rule-based *fuzzy matching* routine that triangulates children across sources, enabling analyses that combine two or all three databases. In this consolidated file, note that only variables from `panel_anon` vary across periods for each child; variables originating in `edi_anon` and `censo_anon` appear as a single observation per child (although in the original EDI database some children had multiple applications). Variable labels are kept consistent with those of their source databases.

- `panelUKA_fuzzymerge` — consolidated database.

**Privacy protections (applied consistently).** Direct identifiers (names, contact details, string IDs) were removed. Geographic/program/person identifiers appear as numeric codes (`id_*`). Birthdates (where present) are truncated to YYYY-MM. Released files are subsamples of the original frames.

## 2 Shared Conventions

- **ID variables.** Integer codes starting at 1: `id_estado`, `id_municipio`, `id_localidad`, `id_grupo`, `id_cedit`; file-specific IDs (`id_madre`, `id_encuesta`, `id_edi`, `child_id`, `id_ninio`).
- **Dates.** Sensitive dates are truncated to YYYY-MM. Operational timestamps not considered PII remain unchanged where applicable.

## 3 `panel_anon`: Child-Period Nutrition Panel

### General Overview

- **Unit of observation:** One record per child per measurement round.
- **Structure:** Longitudinal; a child (`id_ninio`) appears in multiple periods (`periodo`).

## Identification and Structure

- **Child ID (numeric):** `id_ninio` (stable across rounds).
- **Period indicator:** `periodo` (e.g., P001–P013; official measurement rounds).
- **Mother linkage:** `id_madre` (siblings share the same value).
- **Geography/program codes:** `id_estado` (*state*), `id_municipio` (*municipality*), `id_localidad` (*locality*), `id_grupo` (*operational/community group*), `id_cedit` (*CEDIT*: UKA’s early childhood development center; an operational site where UKA delivers services. CEDITs are located across multiple Mexican states).
- **Birth months:** `fecha_nacimiento` (child) and `fnacmadre` (mother) as YYYY-MM.

## Main Variables and Glossary

**peso** Body weight at visit (kilograms).

**talla** Body length/height at visit (centimeters).

**pedz** Weight-for-age Z-score.

**tedz** Height/Length-for-age Z-score.

**p\_imc** *BMI-for-age Z-score*. Body mass index.

**petz** Weight-for-length/height Z-score.

**pednut, tednut, p\_imc2** Categorical nutritional status derived from the corresponding Z-scores (e.g., severe/moderate undernutrition, normal, risk of overweight/overweight). Labels are program-specific.

**hb** Hemoglobin concentration (g/dL).

**hb\_hbaj** Altitude-adjusted hemoglobin (g/dL). Adjustment uses `altitud`.

**hb\_anemicat** Categorical anemia status (e.g., none/mild/moderate/severe).

**hb\_anemiadicot** Binary anemia flag (1 = anemic, 0 = not anemic).

**altitud** Elevation of the child’s location (meters above sea level), used for Hb adjustment.

**\_ini, \_fin** First/last available measurement in the panel for the child (same units as the base variable).

**nummed\_peso, num\_med\_hb** Count of available measurements for weight or Hb for the child within the panel.

**dias\_desde\_anterior\_hb** Days since the previous Hb measurement.

## 4 edi\_anon: Child Developmental Screening

### General Overview

- **What is the EDI?** The *Evaluación del Desarrollo Infantil* (EDI) is a standardized screening tool (not a diagnostic test) designed to detect early signs of developmental delay in children from 1 month to 5 years 11 months. It combines caregiver report and direct observation of age-appropriate items.
- **Objective:** Rapid identification of children who may need referral and closer follow-up in domains of gross motor, fine motor, language, social/personal, and cognition, plus quick neurological checks and risk factors.
- **Administration:** Trained personnel apply the age-appropriate item set; some items are caregiver questions, others require observation.

### 4.1 Global scoring: color classification

After administering the items, the evaluator integrates results across five axes to assign one of three outcomes:

**Green (Normal development)** Expected milestones across all domains, with no alarms and a normal neurological screen.

**Yellow (Developmental lag)** Does not meet all current-group milestones but has achieved those of the previous group; may have risk factors or alert signs.

**Red (Risk of developmental delay)** Fails milestones (often even those of the previous group) and/or has alarm signs or an abnormal neurological screen.

The EDI is **qualitative**: the final result is a *color category*, not a numeric score, based on the group-specific tables.

### Identification and Structure

- **Record ID:** `id_edi` (one screening event).
- **Child ID (within EDI database):** `child_id` (numeric).
- **Mother linkage:** `id_madre`.
- **Geography/program codes:** `id_estado`, `id_municipio`, `id_localidad`, `id_grupo`, `id_cedit`.
- **Birth month:** `FechaNacimiento` stored as YYYY-MM; other operational dates remain unchanged.
- **Age grouping at application:** `GrupoEDI` (current) and `GrupoEDIANterior` (previous group label if tracked). Use `GrupoEDI` for comparability across records.

## Scoring and Interpretation (Color Outcome)

The EDI integrates domain results with neurological checks, alarm/alert signs, and risk factors to assign a **color**:

**Green (Normal)** Age-appropriate skills, no alarms, normal quick neurological screen.

**Yellow (Developmental lag)** Not all current-group milestones met, but previous group achieved; monitoring/early support advised.

**Red (Risk of delay)** Clear shortfalls (often including previous group) and/or alarm signs/abnormal neuro screen; referral recommended.

In `edi_anon`, this global classification is summarized in **Resultado** (Green=1, Yellow=2, Red=3).

## Main Variables and Glossary

(All categorical “result” fields below use the **same ordered numeric encoding** stated above: green = 1, yellow = 2, red = 3.)

**ResultadoMotrizGruoso** Gross motor domain outcome (e.g., posture, locomotion).

**ResultadoMotrizFino** Fine motor domain outcome (e.g., grasping, manipulation).

**ResultadoLenguaje** Language/communication domain outcome.

**ResultadoSocial** Social/personal domain outcome.

**ResultadoConocimiento** Cognitive/knowledge domain outcome.

**ResultadoSeñalesAlarma** Summary of alarm/alert signs; typically coded with the same 1–3 scheme. (If your extract stores counts/flags instead, treat as integers.)

**ResultadoFRB** Biological risk factors summary (e.g., prematurity, neonatal complications, very low birthweight); typically coded 1–3.

**Resultado** Global EDI classification (the color outcome for the screening event).

**EDAD\_MESES\_APLICACION** Child age in months at screening (continuous).

**PESO\_MAS\_CERCANO, talla\_MAS\_CERCANO, hb\_MAS\_CERCANO** Nearest observed weight/height/Hb around the screening (context covariates; not part of the EDI score).

## 5 censo\_anon: Household Socioeconomic Census

### General Overview

- **Unit of observation:** One completed household interview.
- **Scope:** Demographics, dwelling characteristics, services, expenditures, and diet indicators.

## Identification and Structure

- **Survey ID:** `id_encuesta` (numeric; original survey ID not included).
- **Geography/program codes:** `id_cedit`, `id_municipio`, `id_grupo`.
- **Interview date:** `Fecha`.

### 5.1 How to read `censo_anon` (Spanish → English)

Most columns are survey questions in Spanish. Keep `Fecha` (interview date), and the program/geography IDs: `id_encuesta`, `id_cedit`, `id_municipio`, `id_grupo`. Everything else are respondent answers. Below we translate common blocks.

#### Household roster & dwelling

*“¿Cuántos años tiene?”* Respondent age (years).

*“¿Cuántas personas en total viven su hogar?”* Total household size (count).

*“¿Cuántos cuartos tiene su hogar para dormir?”* Number of bedrooms used for sleeping (count).

*“¿Cómo es la propiedad de su vivienda?”* Tenure (e.g., owner, renter, borrowed).

*“¿De qué material es la mayor parte del piso/techo/paredes?”* Main floor/roof/wall material (categorical). **Interpretation:** proxies for housing quality; you can build an index.

*“Para la disposición de excretas...”* Sanitation (toilet/latrine type).

#### Digital access & services

*“¿El celular del que dispone es smartphone?”* Smartphone availability (Yes/No).

*“¿En su comunidad tiene señal de celular?”* Cell signal in community (Yes/No).

*“¿En su hogar disponen de conexión a Internet?”* Internet at home (Yes/No).

#### Culture and language

*“De acuerdo con su cultura, ¿Usted se considera indígena?”* Self-identifies as Indigenous (Yes/No).

*“¿Usted habla alguna lengua indígena?”* Speaks an Indigenous language (Yes/No).

*“¿Qué lengua indígena habla?”* Name of the language (text; may include misspellings/variants).

## Expenditures and income proxies

*“¿Cuánto es el ingreso aproximado del hogar a la semana?”* Weekly household income (categorical ranges).

*“Del ingreso familiar ¿Cuánto gasta a la semana en alimentación/salud/educ...?”* Weekly pesos on food/health/education/services (numeric; may have zeros).

*Total\_calorias\_semana, Total\_hierro\_semana* Constructed proxy indices from diet items.

## Water, storage, and food safety

*“¿En su vivienda tienen agua?”* Water sources (multi-response). **Note:** can contain multiple options in a single cell.

*“¿Cómo almacena el agua...?”* Water storage (multi-response).

*“¿Qué tratamiento le da al agua...?”* Water treatment (single or multi-response).

*“¿Cuántos días a la semana llega el agua...?”* Water frequency (days/week; may be numeric or text).

*“¿Dónde almacena los alimentos?”* Food storage (multi-response).

## Child care, spaces, and health access

*“¿Lleva a sus hijos a... primera infancia?”* Uses ECCE/childcare (Yes/No).

*“¿Lleva a sus hijos para jugar en espacios públicos...?”* Uses public play spaces (Yes/No or list).

*“En su localidad ¿Qué servicio de salud usa con mayor frecuencia?”* Usual health service (multi-response allowed).

*“¿Cuántos días a la semana está disponible...?”* Availability (days/week).

*“¿Usa actualmente el servicio de vacunación infantil?”* Uses vaccination services (Yes/No).

*“¿Cuánto tiempo tardaría en llegar a un hospital...?”* Emergency travel time (minutes; may be categorical).

## Diet items (binary/frequency)

Items like “Res o cerdo,” “Pollo,” “Leche,” “Frijol,” “Frutas,” “Verduras,” “Refrescos,” “Jugos azucarados,” “Frituras o pastelitos,” etc. appear as numeric frequencies of items consumed in a typical week.