



Proyecto 1

Libro de códigos

Data Science

Sección 20

Mónica Alejandra Salvatierra Chacón - 22249

Derek Fabián Arreaga Kelson - 22537

Paula Rebeca Barillas Alvarez – 22764

CÓDIGO

- Identificador único del establecimiento educativo
- Formato: XX-XX-XXXX-XX
 - Primer Bloque: Código del departamento
 - Segundo Bloque: Código del municipio dentro del departamento
 - Tercer Bloque: Identificador interno del centro educativo
 - Cuarto Bloque: Identificador de nivel o modalidad diferente (como diversificado, jornada específica o plan educativo autorizado)

DISTRITO

- Código del distrito escolar al que pertenece el establecimiento
- Formato: XX-XXX
 - Primer Bloque: Código del departamento
 - Segundo Bloque (últimos dígitos): Identifican zonas operativas/educativas regionales

DEPARTAMENTO

- Nombre del departamento de Guatemala donde se ubica el establecimiento

MUNICIPIO

- Municipio donde se encuentra el establecimiento

ESTABLECIMIENTO

- Nombre oficial completo del centro educativo

DIRECCION

- Ubicación del establecimiento educativo

TELEFONO

- Número telefónico de contacto del establecimiento

SUPERVISOR

- Nombre completo del supervisor distrital asignado al establecimiento

DIRECTOR

- Nombre completo del director o responsable del centro educativo

NIVEL

- Máximo nivel educativo ofrecido por el establecimiento

SECTOR

- Sector al que pertenece el establecimiento según su administración
 - PRIVADO: Instituciones privadas
 - OFICIAL: Instituciones públicas del gobierno central
 - MUNICIPAL: Instituciones administradas por municipalidades
 - COOPERATIVA: Instituciones administradas por cooperativas

AREA

- Clasificación geográfica según ubicación del establecimiento
 - URBANA: Centros ubicados en áreas urbanas
 - RURAL: Centros ubicados en áreas rurales
 - SIN ESPECIFICAR: Sin clasificación definida

STATUS

- Estado de funcionamiento del centro educativo

MODALIDAD

- Modalidad lingüística de enseñanza
 - MONOLINGUE: Enseñanza en español únicamente
 - BILINGUE: Enseñanza en español e idioma maya

JORNADA

- Horario de funcionamiento del establecimiento
 - MATUTINA: Horario matutino
 - VESPERTINA: Horario vespertino/tarde
 - NOCTURNA: Horario nocturno
 - DOBLE: Dos jornadas (mañana y tarde)
 - INTERMEDIA: Horario intermedio
 - SIN JORNADA: Sin jornada específica

PLAN

- Plan o modalidad educativa implementada
 - DIARIO(REGULAR): Plan educativo regular de lunes a viernes
 - FIN DE SEMANA: Clases los fines de semana
 - A DISTANCIA: Educación completamente a distancia
 - SEMIPRESENCIAL: Combinación de presencial y a distancia
 - SEMIPRESENCIAL (FIN DE SEMANA): Semipresencial los fines de semana
 - SEMIPRESENCIAL (UN DÍA A LA SEMANA): Un día presencial por semana
 - SEMIPRESENCIAL (DOS DÍAS A LA SEMANA): Dos días presenciales por semana
 - VIRTUAL A DISTANCIA: Educación virtual
 - SABATINO: Clases los sábados
 - DOMINICAL: Clases los domingos
 - INTERCALADO: Horario intercalado
 - MIXTO: Combinación de modalidades

DEPARTAMENTAL

- Nombre del departamento o subdivisión regional administrativa

SUPERVISORES_SIMILARES.CSV

- Registra nombres de supervisores con alta similitud pero escritura diferente
- Columnas:
 - **SUPERVISOR_BASE**: Nombre considerado correcto
 - **SUPERVISOR_SIMILAR**: Variante encontrada con errores tipográficos
 - **SIMILITUD**: Porcentaje de similitud calculado (0-100)

DIRECTORES_SIMILARES.CSV

- Registra nombres de directores con alta similitud pero escritura diferente
- Columnas:
 - **DIRECTOR_BASE**: Nombre considerado correcto
 - **DIRECTOR_SIMILAR**: Variante encontrada con errores tipográficos
 - **SIMILITUD**: Porcentaje de similitud calculado (0-100)

ESTABLECIMIENTOS_SIMILARES.CSV

- Registra nombres de establecimientos con variaciones menores
- Columnas:
 - **Nombre_Base:** Nombre estándar del establecimiento
 - **Nombre_Similar:** Variante encontrada (con tildes, abreviaturas, etc.)

ERRORES_COMBINACIONES.CSV

- Registra combinaciones de datos que presentan inconsistencias, en donde los registros donde la combinación departamento-municipio no coincide con códigos oficiales
- Columna:
 - **ES_VALIDO:** Indica si la combinación es válida según datos oficiales