

# 1. Características de la base de datos de la Encuesta sobre Condiciones Sociales de la Ciudad de México 2024

## 1.1. Objetivo

La encuesta tiene como objetivo generar información estadística sobre el bienestar de las personas que habitan en la Ciudad de México, con énfasis en el acceso a derechos como la vivienda, la alimentación, la educación, la salud, el empleo y la movilidad, así como al acceso y percepción de programas sociales.

## 1.2. Conformación de la base de datos

La base de datos de la Encuesta sobre Condiciones Sociales de la Ciudad de México 2024, está conformada por tres tablas de datos provenientes de dos cuestionarios que conforman la encuesta y se dividen por su nivel de desagregación.

## Documentos

Las bases de datos se conforman por tres tablas en un formato nativo de Stata para almacenar conjuntos de datos (dta), un descriptor de datos en formato de archivo para hojas de cálculo de Microsoft Excel (xlsx) y un descriptor de la base de datos en Formato de Documento Portátil (pdf).

## Conjunto de datos

- **Individual:** Datos provenientes de la Encuesta sobre Condiciones Sociales de la Ciudad de México 2024 del cuestionario individual.
- **Hogar:** Datos provenientes de la Encuesta sobre Condiciones Sociales de la Ciudad de México 2024 del cuestionario hogares.
- **Individual\_hogar:** Datos provenientes de la Encuesta sobre Condiciones Sociales de la Ciudad de México 2024 del cuestionario hogares de la sección de características sociodemográficas (reportadas por el jefe del hogar).

## 1.3. Unidad de análisis

La unidad de análisis son las personas y los hogares, para explotar la base de datos se debe considerar:

La relación entre la tabla **Individual** e **Individual hogar** es uno a uno mediante las variables llave *FV*, *FH* y *FI*, mientras que para la tabla Hogar son *FV* y *FH*, cuando se quiere analizar las características de los residentes del hogar.

- *FV*: Folio de la vivienda
- *FH*: Folio del hogar
- *FI*: Folio individual

Esta información es únicamente de referencia y por cuestiones de confidencialidad, protección de datos y diseño muestral no es posible identificar ningún nivel geográfico diferente a la Ciudad de México.

## 1.4. Factor de expansión

Una característica de la encuesta es que busca representar a toda la población de la Ciudad de México, dado que es una muestra cada observación puede representar a personas u hogares, según sea el caso que presentan condiciones similares.

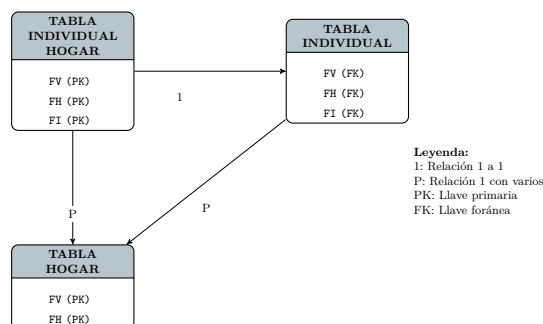
La encuesta cuenta con dos factores de expansión, *FACTOR\_POBLACION* que se debe usar únicamente cuando se quiere analizar a toda la población, mientras que *FACTOR\_HOGAR* sirve para analizar el total de hogares.

## 1.5. Relación de tablas

La relación de tablas es de uno a uno entre las tablas **Individual** e **Individual\_hogar**, mientras que para la tabla **Hogar** a las tablas **Individual** e **Individual\_hogar** es de uno a muchos.

## 1.6. Diagrama de relación

Figura 1: Diagrama de relación



Fuente: elaboración del EVALÚA.

## 2. Descripción de los datos

### 2.1. Descripción de las tablas

La base de datos se conforma de tres tablas, la primera **Individual** que provienen del cuestionario con el mismo nombre, mientras que las otras dos tablas se conformaron con la información recolectada en el cuestionario hogar, está dividida en dos tablas, porque la tabla **Hogar** hace referencia a la información de los hogares en muestra y la tabla **Individual hogar** es la información que el jefe del hogar proporcionó por cada uno de los integrantes del hogar.

- **Individual\_hogar**: Base de datos con 3,244 observaciones y 49 variables (Contiene el total de la población). [Llave primaria *FV*, *FH* y *FI*]
- **Individual**: Base de datos con 2,883 observaciones y 328 variables (Contiene a la población de 13 años o más). [Llave foránea *FV*, *FH* y *FI*]
- **Hogar**: Base de datos con 1,200 observaciones y 331 variables (Contiene a todos los hogares). [Llave foránea *FV* y *FH*]