

# Análisis de las características físicas y de equipamiento de vivienda en los hogares colombianos

Angela Marcela Vargas<sup>1</sup>, Jorgue Andrés Buitrago<sup>2</sup>

<sup>1-2</sup>Dpto. de Ingeniería y Ciencias Básicas

Universidad Central

Maestría en Analítica de Datos

Curso de Bases de Datos

Bogotá, Colombia

<sup>1</sup>angela.vargas@ucentral.edu.co, <sup>2</sup>jbuitragoc7@ucentral.edu.com

October 22, 2022

## Contents

<b>1</b>	<b>Introducción</b>	<b>3</b>
<b>2</b>	<b>Características del proyecto de investigación</b>	<b>3</b>
2.1	Título del proyecto de investigación . . . . .	3
2.2	Objetivo general . . . . .	3
2.2.1	Objetivos específicos . . . . .	4
2.3	Alcance . . . . .	4
2.4	Pregunta de investigación . . . . .	4
2.5	Hipótesis . . . . .	4
<b>3</b>	<b>Reflexiones sobre el origen de datos e información</b>	<b>5</b>
3.1	¿Cuál es el origen de los datos e información ? . . . . .	5
3.2	¿Cuáles son las consideraciones legales o éticas del uso de la información? . . . . .	5
3.3	¿Cuáles son los retos de la información y los datos que utilizará en la base de datos en términos de la calidad y la consolidación? . . . . .	5
3.4	¿Qué espera de la utilización de un sistema de Bases de Datos para su proyecto? . . . . .	5

<b>4</b>	<b>Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos)</b>	<b>7</b>
4.1	Características del SMBD (Sistema Manejador de Bases de Datos) para el proyecto . . . . .	7
4.2	Diagrama modelo de datos . . . . .	7
4.3	Imágenes de la Base de Datos . . . . .	8
4.4	Código SQL - lenguaje de definición de datos (DDL) . . . . .	10
4.5	Código SQL - Manipulación de datos (DML) . . . . .	10
4.6	Código SQL + Resultados: Vistas . . . . .	11
4.7	Código SQL + Resultados: Triggers ( <i>Primera entrega</i> ) . . . . .	11
4.8	Código SQL + Resultados: Funciones ( <i>Primera entrega</i> ) . . . . .	11
4.9	Código SQL + Resultados: procedimientos almacenados ( <i>Primera entrega</i> ) . . . . .	11
<b>5</b>	<b>Bases de Datos No-SQL (<i>Segunda entrega</i>)</b>	<b>12</b>
5.1	Diagrama Bases de Datos No-SQL ( <i>Segunda entrega</i> ) . . . . .	12
5.2	SMBD utilizado para la Base de Datos No-SQL ( <i>Segunda entrega</i> )	12
<b>6</b>	<b>Aplicación de ETL (Extract, Transform, Load) y Bodega de Datos (<i>Tercera entrega</i>)</b>	<b>13</b>
6.1	Ejemplo de aplicación de ETL y Bodega de Datos ( <i>Tercera entrega</i> )	13
<b>7</b>	<b>Lecciones aprendidas (<i>Tercera entrega</i>)</b>	<b>14</b>
<b>8</b>	<b>Bibliografía</b>	<b>15</b>

# 1 Introducción

La Encuesta Nacional de presupuesto de los Hogares – ENPH, es el instrumento que el país utiliza para estimar ingresos y gastos de los hogares, los cuales sirven para evaluar el nivel, estructura y las tendencias del bienestar económico de los hogares. Además los flujos de ingresos, los gastos en consumo y el ahorro-consumo futuro-, definen los límites del estilo de vida del hogar y su nivel de bienestar. (DANE, 2020). Este es el sentido principal de la encuesta, sin embargo, también se capturan “otras variables clasificatorias y de control, como las características generales, educativas y económicas de las personas del hogar.” (DANE, 2020). Los temas principales del uso de esta información son actualizar las ponderaciones del Índice de Precios al Consumidor (IPC), construir las cuentas de los hogares en el Sistema de Cuentas Nacionales (DANE, 2020) y la actualización de las líneas de pobreza con base en los cambios de las estructuras de gasto y los hábitos de consumo de los hogares (Fedesarrollo, 2020).

El presente documento se enfoca en el capítulo vivienda y hogares, que tiene como objetivo recolectar información relacionada características físicas y equipamiento de las viviendas, y también provee el campo de ingresos para poder hacer relacionamiento entre estos dos conceptos. Se pretende entonces, realizar un modelo de base de datos que permita analizar las características de vivienda según nivel de ingresos de los hogares encuestados en la ENPH.

## 2 Características del proyecto de investigación

Con la información del capítulo de vivienda y hogares de la Encuesta Nacional de Presupuesto de los Hogares, se pretende dar respuesta a algunas preguntas de interés relacionadas con las características habitacionales de los hogares colombianos, así como también, enlazar estos conceptos a los niveles de ingreso reportados. Para tal fin, se elaborará un modelo de base de datos que permita segregar la información del cuadernillo, de modo tal que sean aplicables los conceptos aprendidos durante el curso. Es de resaltar el hecho de que la estructura de los datos es apropiada para lograr un modelo entidad relación que lleve al entendimiento de una manera más simple.

### 2.1 Título del proyecto de investigación

Análisis de las características físicas y de equipamiento de vivienda en los hogares colombianos.

### 2.2 Objetivo general

Analizar las características físicas y de equipamiento de vivienda en los hogares colombianos en función de otras variables clasificatorias de interés.

### **2.2.1 Objetivos específicos**

- Segregar los datos del cuadernillo de vivienda de la Encuesta Nacional de Presupuestos de los Hogares – ENPH.
- Identificar características físicas y de equipamiento de las viviendas con relación al nivel de ingresos, ciudad donde reside, número de personas que conforman el hogar, entre otras.

## **2.3 Alcance**

El presente documento tiene un alcance específico de caracterización de la información, es decir, se pretende con los datos disponibles generar análisis sobre variables clasificatorias de interés general para conocer a cerca de rasgos físicos de las viviendas de la población colombiana en función algunas otras de interés como nivel de ingresos, ciudad de residencia, número de personas que conforman el hogar, etc.

Esto en el marco de la Encuesta Nacional de presupuestos de los Hogares – ENPH, que fue realizada entre los años 2016 y 2017, cuya vigencia estadística es de 10 años, fue recolectada en 32 ciudades de departamento y 6 cabeceras adicionales. Esta caracterización es aplicable solo para el cuadernillo de vivienda y hogares disponible a nivel de microdatos de forma abierta en la página web del DANE.

## **2.4 Pregunta de investigación**

¿Qué características físicas y de equipamiento tienen las viviendas de los hogares colombianos en función de los ingresos reportados?

## **2.5 Hipótesis**

El estrato socioeconómico reportado en la Encuesta Nacional de Presupuestos de los Hogares – ENPH, no determina necesariamente el equipamiento y las condiciones físicas de la vivienda de los colombianos

### **3 Reflexiones sobre el origen de datos e información**

Los inicios de las mediciones de ingresos y gastos en los hogares colombianos se remontan al período de 1937 – 1940 en cabeza de la Contraloría General, con cobertura para las ciudades de Bogotá y Medellín, desde entonces y en intervalos de 10 años aproximadamente, se han realizado los diferentes esfuerzos para la actualización de esta información en materia de recolección y metodología.

Desde 1967, el Departamento Nacional de Estadística - DANE ha estado a cargo de la implementación de este proceso de investigación, mejorando las estructuras de captación de información y aplicando mejores prácticas sugeridas a nivel internacional por entidades como la OIT, EuroStat, CEPAL y OCDE, así como también, ampliando las coberturas para incluir la población rural, con el fin de entender dinámicas diferenciadas de las grandes ciudades. Entendiendo que dicha información es insumo para la toma de decisiones a nivel de política pública.

#### **3.1 ¿Cuál es el origen de los datos e información ?**

Encuesta Nacional de Presupuestos de los Hogares – ENPH 2016 – 2017, recolectada por el DANE y correspondiente al cuadernillo 1 cuyo tema principal es vivienda y hogares.

#### **3.2 ¿Cuáles son las consideraciones legales o éticas del uso de la información?**

Al ser una base de datos abierta al público para su consulta general, las consideraciones legales y éticas se enmarcan en los principios de los investigadores y sus fines académicos con el uso de la información.

#### **3.3 ¿Cuáles son los retos de la información y los datos que utilizará en la base de datos en términos de la calidad y la consolidación?**

Los principales retos de la información y los datos están en el entendimiento, la organización, partición y posterior uso. En primer lugar, el entendimiento de las estructuras del cuadernillo y su presentación en forma de base de datos, segundo la organización de la información relevante para cumplir con los objetivos del proyecto, tercero la partición de las bases de datos adecuadas para establecer un modelo entidad relación eficiente que permita la aplicabilidad de conceptos vistos durante el curso.

#### **3.4 ¿Qué espera de la utilización de un sistema de Bases de Datos para su proyecto?**

Eficiencia para el consumo de información con el fin de generar la construcción del conocimiento en torno a los temas de la caracterización física y de equipamiento

de las viviendas de los hogares colombianos.

## 4 Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos)

### 4.1 Características del SMBD (Sistema Manejador de Bases de Datos) para el proyecto

Para el desarrollo del presente proyecto se utilizará la herramienta MySQL cuyas características destacables son:

- Disponibilidad en gran cantidad de plataformas y sistemas operativos.
- Soporta gran cantidad de datos y posee diferentes opciones para su almacenamiento.
- Transacciones y claves primarias y foráneas.
- Conectividad segura.
- Utiliza el lenguaje de acceso a datos SQL.
- No necesita requerimientos de software ni hardware de alto rendimiento.
- De código abierto.

### 4.2 Diagrama modelo de datos

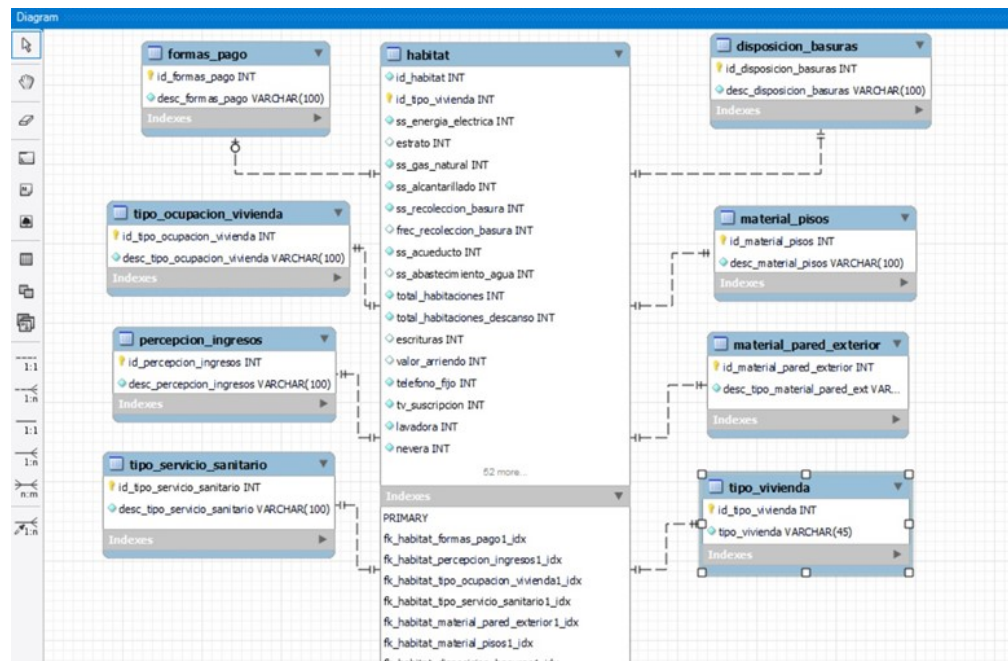


Figura 1. Modelo entidad relación

### 4.3 Imágenes de la Base de Datos

Model Overview

Tables (9 items)

Add Table

disposicion\_basuras

tpo\_ocupacion\_vivie...

tpo\_servicio\_sanitario

formas\_pago

tpo\_vivienda

habitat

material\_pared\_exterior

material\_pisos

percepcion\_ingresos

Views (0 items)

Add View

Routines (0 items)

habitat - Table

material\_pared\_exterior - Table

Table Name: habitat

Schema: mydb

Filter Rows

Edit

Export/Import

Wrap Cell Contents

Apply changes

id_habitat	id_tipo_vivienda	ss_energia_electric	estrato	ss_gas_natural	ss_alcantarillado	ss_recoleccion_bas	frec_recoleccion_bas	ss_acueducto	ss_abastecimiento	total_habitado
118329	ORIENTAL	RURAL DISPERSO	1	1	2	1	2	2	2	2
118331	ORIENTAL	BUCARAMANGA ...	1	5	2	1	1	2	1	1
118332	PACIFICA	TUMACO	1	2	2	1	1	2	2	1
118338	PACIFICA	POPAYAN	1	1	4	1	2	1	1	1
118344	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	2	1	1	1
118346	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	2	2	1	1
118354	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	2	1	1	1
118356	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	2	1	1	1
118362	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	3	2	1	1
118364	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	3	1	1	1
118372	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	3	2	1	1
118373	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	3	1	1	1
118380	CENTRAL	MEDELLIN Y A.M.	1	1	4	1	4	1	1	1
118381	ATLANTICA	BARRANQUILLA	1	1	2	1	1	2	2	1
118383	ATLANTICA	BARRANQUILLA	1	5	1	1	1	2	2	1

Figura 2. Base de datos - Habitat



disposicion_basuras - Table	
<div> <div>Table Name: <input type="text" value="disposicion_basuras"/></div> <div>Schema: <b>mydb</b></div> </div> <div> <div>Filter Rows:</div> <div>Edit: </div> <div>Export/Import: </div> <div>Wrap Cell Content: </div> <div>Apply changes:  </div> </div>	
id_disposicion	desc_disposicion_basuras
1	Por recolección pública o privada
2	La tiran a un río, quebrada, caño o la...
3	La tiran a un patio, lote, zanja o baldío
4	La queman o entierran
5	La eliminan de otra forma

material_pared_exterior - Table	
<div> <div>Table Name: <input type="text" value="material_pared_exterior"/></div> <div>Schema: <b>mydb</b></div> </div> <div> <div>Filter Rows:</div> <div>Edit: </div> <div>Export/Import: </div> <div>Wrap Cell Content: </div> <div>Apply changes:  </div> </div>	
id_material_pai	desc_tipo_material_pared_ext
1	Ladrillo, bloque, material prefabricad...
2	Madera pulida
3	Adobe o tapia pisada
4	Bahareque
5	Madera burda, tabla, tablón

material_pisos - Table	
<div> <div>Table Name: <input type="text" value="material_pisos"/></div> <div>Schema: <b>mydb</b></div> </div> <div> <div>Filter Rows:</div> <div>Edit: </div> <div>Export/Import: </div> <div>Wrap Cell Content: </div> <div>Apply changes:  </div> </div>	
id_material_pis	desc_material_pisos
1	Tierra, arena
2	Cemento, grava
3	Madera burda, tabla, tablón, otro vegetal
4	Baldosin, ladrillo, vinisol, otros materiales sintéticos
5	Mármol
6	Madera pulida
7	Alfombra o tapete de pared a pared

Figura 3. Base de datos - Otras

#### 4.4 Código SQL - lenguaje de definición de datos (DDL)

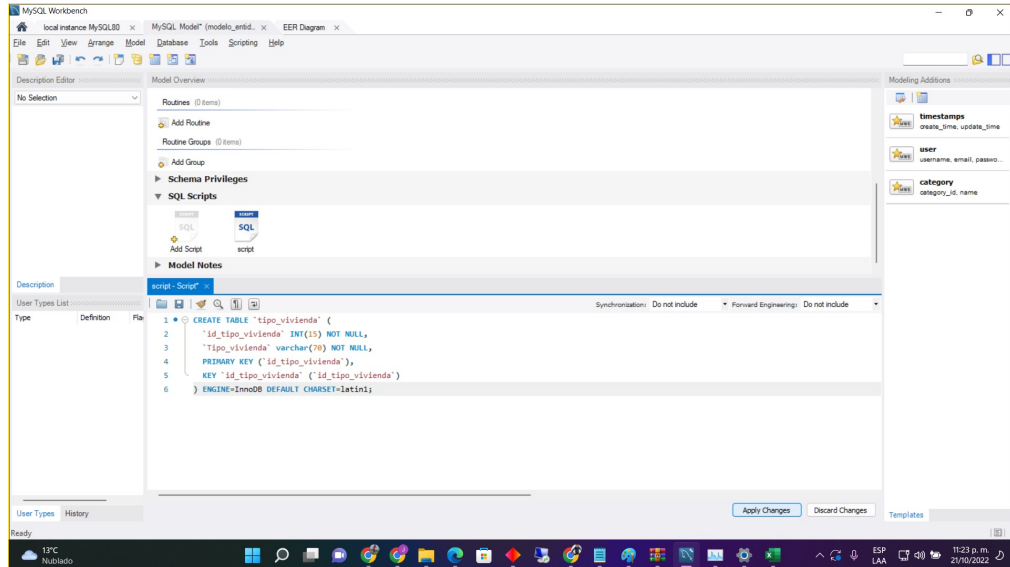


Figura 4. Base de datos - Creación

#### 4.5 Código SQL - Manipulación de datos (DML)

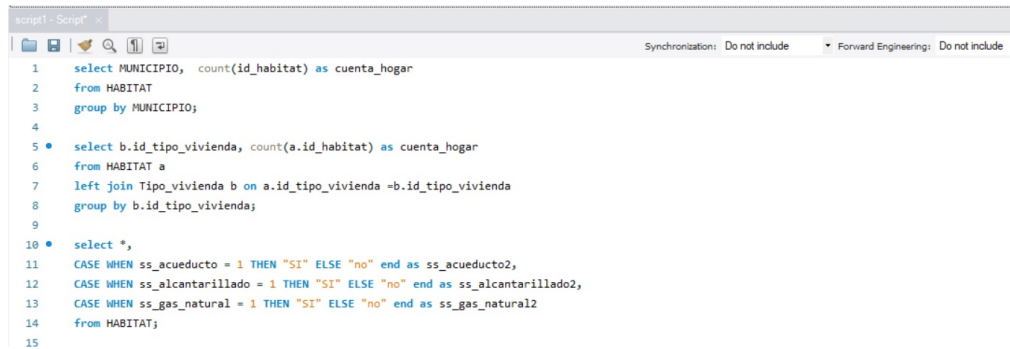
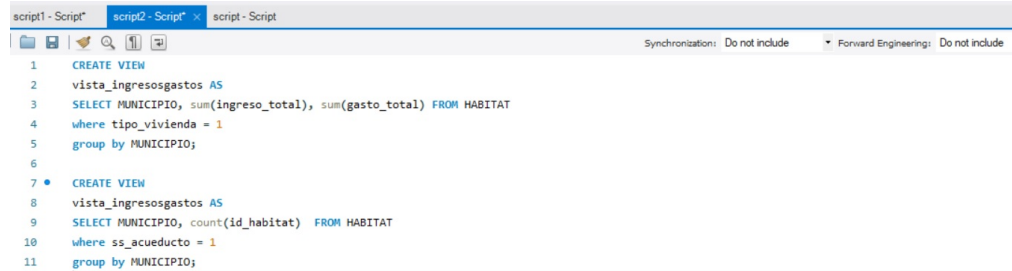


Figura 5. Base de datos - Manipulación

## 4.6 Código SQL + Resultados: Vistas



```
1 CREATE VIEW
2 vista_ingresosgastos AS
3 SELECT MUNICIPIO, sum(ingreso_total), sum(gasto_total) FROM HABITAT
4 where tipo_vivienda = 1
5 group by MUNICIPIO;
6
7 • CREATE VIEW
8 vista_ingresosgastos AS
9 SELECT MUNICIPIO, count(id_habitat) FROM HABITAT
10 where ss_acueducto = 1
11 group by MUNICIPIO;
```

Figura 6. Base de datos - Vistas

## 4.7 Código SQL + Resultados: Triggers (*Primera entrega*)

## 4.8 Código SQL + Resultados: Funciones (*Primera entrega*)

## 4.9 Código SQL + Resultados: procedimientos almacenados (*Primera entrega*)

## 5 Bases de Datos No-SQL (*Segunda entrega*)

### 5.1 Diagrama Bases de Datos No-SQL (*Segunda entrega*)

### 5.2 SMBD utilizado para la Base de Datos No-SQL (*Segunda entrega*)

## 6 Aplicación de ETL (Extract, Transform, Load) y Bodega de Datos *(Tercera entrega)*

### 6.1 Ejemplo de aplicación de ETL y Bodega de Datos *(Tercera entrega)*

## 7 Lecciones aprendidas (*Tercera entrega*)

## 8 Bibliografía

1. DANE. (2020). COLOMBIA - Encuesta Nacional de Presupuestos de los Hogares - ENPH - Julio 2016 - Julio 2017. Obtenido de <http://microdatos.dane.gov.co/index.php>
2. DANE. (2020). Metodología General Encuesta Nacional de Presupuestos de Los Hogares – ENPH.
3. Fedesarrollo. (2020). Tendencia Económica. Obtenido de <https://www.repository.fedesarrollo.org.co/handle/11445/4050>
4. International Household Survey Network - IHSN. (2019). Encuesta Nacional de Presupuestos de los Hogares 2016-2017. Obtenido de <https://catalog.ihsn.org/index.php/catalog/7949>