

Proyecto de Bases de datos para el análisis de la calidad de vida de los habitantes de Bogotá D.C.

Cristian David Carvajal Parra¹, Alejandro Salazar Molina²

¹Facultad de ingeniería y ciencias básicas

Universidad Central

Maestría en Analítica de Datos

Curso de Bases de Datos

Bogotá, Colombia

²ccarvajalp2@ucentral.edu.co, ³esalazarm@ucentral.edu.co

April 21, 2023

Contents

1	Introducción (Max 250 Palabras) - (<i>Primera entrega</i>)	3
2	Características del proyecto de investigación que hace uso de Bases de Datos (Max 500 Palabras) - (<i>Primera entrega</i>)	6
2.1	Titulo del proyecto de investigación (Max 100 Palabras) - (<i>Primera entrega</i>)	6
2.2	Objetivo general (Max 100 Palabras) - (<i>Primera entrega</i>)	7
2.2.1	Objetivos especificos (Max 100 Palabras) - (<i>Primera entrega</i>)	7
2.3	Alcance (Max 200 Palabras) - (<i>Primera entrega</i>)	7
2.4	Pregunta de investigación (Max 100 Palabras) - (<i>Primera entrega</i>) .	7
2.5	Hipotesis (Max 100 Palabras) - (<i>Primera entrega</i>)	7
3	Reflexiones sobre el origen de datos e información (Max 400 Palabras) - (<i>Primera entrega</i>)	9
3.1	¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (<i>Primera entrega</i>)	9
3.2	¿Cuales son las consideraciones legales o eticas del uso de la información? (Max 100 Palabras) - (<i>Primera entrega</i>)	9
3.3	¿Cuales son los retos de la información y los datos que utilizara en la base de datos en terminos de la calidad y la consolidación? (Max 100 Palabras) - (<i>Primera entrega</i>)	10
3.4	¿Que espera de la utilización de un sistema de Bases de Datos para su proyecto? (Max 100 Palabras) - (<i>Primera entrega</i>)	10

4	Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos)(Primera entrega)	12
4.1	Características del SMBD (Sistema Manejador de Bases de Datos) para el proyecto (Primera entrega)	12
4.2	Diagrama modelo de datos (Primera entrega)	12
4.3	Imágenes de la Base de Datos (Primera entrega)	13
4.4	Código SQL - lenguaje de definición de datos (DDL) (Primera entrega)	16
4.5	Código SQL - Manipulación de datos (DML) (Primera entrega) . .	17
4.6	Código SQL + Resultados: Vistas (Primera entrega)	17
4.7	Código SQL + Resultados: Triggers (Primera entrega)	19
4.8	Código SQL + Resultados: Funciones (Primera entrega)	20
4.9	Código SQL + Resultados: procedimientos almacenados (Primera entrega)	21
5	Bases de Datos No-SQL (Segunda entrega)	21
5.1	Diagrama Bases de Datos No-SQL (Segunda entrega)	21
5.2	SMBD utilizado para la Base de Datos No-SQL (Segunda entrega)	21
6	Aplicación de ETL (Extract, Transform, Load) y Bodega de Datos (Tercera entrega)	22
6.1	Ejemplo de aplicación de ETL y Bodega de Datos (Tercera entrega)	22
7	Proximos pasos (Tercera entrega)	23
8	Lecciones aprendidas (Tercera entrega)	24
9	Bibliografía	25

1 Introducción (Max 250 Palabras) - (*Primera entrega*)

La calidad de vida es un tema recurrente entre la sociedad, haciendo alusión a su nivel de bienestar en sus vivencias diarias, de esta manera una persona puede ver influenciado este indicador por condiciones externas como el estado de las vías de su barrio, o factores internos como la salud mental. Este tópico ha sido estudiado por diferentes organizaciones a nivel mundial, llevando a múltiples discusiones y consensos para llegar a puntos en común que puedan compartirse desde los niveles locales, regionales e internacionales (Naciones Unidas, 1961)

En algunos casos, expertos han tratado de verificar cuáles son las variables que influyen en la calidad de vida de las personas, llegando a hallazgos importantes como la relevancia de variables internas y externas. Entre las variables internas se pueden encontrar la edad, el estado civil, los ingresos, el nivel educativo y de endeudamiento, entre otros; por su parte, entre los factores externos se pueden encontrar: la discriminación, la infraestructura, la urbanización, el uso de medicamentos, el acceso a internet y muchos más (Huerta-Riveros, Patricia, Yáñez-Alvarado, Miguel, Leyton-Pavez, Carolina, 2022).

Lo anterior también coincide con organizaciones internacionales como la ONU, la cual en 1961 aseguraba la importancia de tener en cuenta factores como la salud, el consumo de alimentos, la educación, el empleo, las condiciones de trabajo, la vivienda, la seguridad social, el vestido, el esparcimiento y las libertades humanas, como variables significativas para medir la calidad de vida (Naciones Unidas, 1961).

Sin embargo, dependiendo del campo del conocimiento desde el que se abarque el concepto de calidad de vida se pueden tener en cuenta diferentes aspectos para su definición. En economía, pueden ser relevantes medidas como el PIB, los ingresos de una nación, el consumo de una persona y su capacidad para adquirir bienes (Hurtado Briceño, Alberto Zerpa de Hurtado, Sadcidi, 2011). Por otra parte, en el caso de la salud son protagonistas las dimensiones como la movilidad, el dolor, la movilidad física, el nivel de esfuerzo y la preocupación frente al estado de salud, asimismo, en el campo de la psicología son tenidos en cuenta la repercusión del funcionamiento cognitivo, uso de la memoria y la capacidad de pensar con claridad (Ruiz, Miguel Pardo, Antonio, 2005).

Por otra parte, a nivel mundial se han desarrollado estudios para analizar la Calidad de Vida, los cuales nacen después de los años 60, momento en el que la ONU decide motivar a los países a analizar la calidad de vida bajo 12 variables. Con el paso del tiempo, empiezan a surgir nuevos ejercicios como el “The quality of american life: Perceptions, Evaluations and Satisfaction” realizado por el Institute for Social Research de la Universidad de Michigan, en el que se evaluó la percepción de la calidad de vida en un barrio, preguntándole

a más de 2.000 personas sobre sus relaciones personales, su trabajo, educación y nivel de satisfacción con la vida en general (Campbell A, Converse P, Rodgers E., 1976).

Significativamente, conocer y cuestionar la calidad de vida se ha convertido en una necesidad mundial, por lo que entidades como la Organización para la Cooperación y el Desarrollo Económicos - OCDE han ejecutado 2 estudios How's life? con los que busca comparar la calidad de vida entre los países miembros. La organización ha ejecutado 2 reportes, uno en el 2011 y otro en el 2020, siendo este último relevante para Colombia. Llegando a hallazgos significativos como (OCDE, 2020):

- Casi el 40% de hogares de la OCDE son financieramente inseguros, lo que significa que pueden caer en pobreza si se abstienen de 3 meses de ingresos. Si bien el 12% de la población de la OCDE vive en una pobreza relativa de ingresos, la proporción de quienes informan tener dificultades para llegar a fin de mes en los países europeos de la OCDE es casi el doble, con un 21%.
- La mediana de la riqueza de los hogares disminuyó un 4, en promedio, desde alrededor de 2010, en aquellos países donde existen datos. Uno de cada cinco hogares de bajos ingresos gasta más del 40 de su ingreso disponible en costos de vivienda, dejando poco para otros elementos esenciales de la vida.
- Las experiencias de seguridad también contrastan fuertemente entre mujeres y hombres: por un lado, los hombres en los países de la OCDE tienen 4,5 veces más probabilidades de morir debido a una agresión, lo que refleja principalmente los altos valores observados en Colombia (donde los hombres tienen más de diez veces más probabilidades que mujeres a ser víctimas de homicidio) y México (donde la misma proporción es superior a ocho).
- Por otro lado, en promedio, ocho de cada diez hombres, pero solo seis de cada diez mujeres, informan que se sienten seguras cuando caminan solas por la noche, lo que posiblemente refleja el mayor riesgo de las mujeres de sufrir delitos de contacto y agresión sexual.
- La biodiversidad se ha perdido constantemente en muchos países de la OCDE (23) desde 2010. Las mayores disminuciones en el Índice de la Lista Roja para especies amenazadas generalmente han ocurrido en aquellos países con tasas de riesgo ya altas, incluidos Nueva Zelanda, México, Corea, Colombia, Chile, Reino Unido, Japón y Australia, además de Francia.

En el caso colombiano, entes gubernamentales como los Ministerios y otras entidades que buscan mejorar la calidad de vida de la población, necesitan contar con información veraz y actualizada de las condiciones de sus habitantes, siendo la fuente principal del análisis y toma de decisiones (DANE, 2021). Uno de los mecanismos principales para la obtención de esta información, es la aplicación de encuestas y entrevistas a la población objetivo, con las que se busca la participación de grupos poblacionales con diversas características o condiciones como tipos de territorios, razas, edades, etnias, principales ciudades e ideologías, entre otros (Cámara de Comercio de Medellín, 2021).

Teniendo en cuenta lo anterior, el Departamento Administrativo Nacional de Estadística – DANE, aplica las encuestas de calidad de vida anualmente “como una respuesta a la necesidad de caracterizar la población en los diferentes aspectos involucrados en el bienestar de los hogares”. Esta encuesta permite identificar en la población colombiana factores de pobreza desde dos enfoques: Necesidades Básicas Insatisfechas (NBI) y Línea de Pobreza, sin embargo, con el paso del tiempo se amplió el alcance de esta encuesta, incluyendo nuevas variables que permiten conocer más a profundidad la calidad de vida de los colombianos (DANE 2021).

En casos más específicos, grandes ciudades como Bogotá y Medellín, tienen condiciones que pueden afectar la calidad de vida como el tráfico, la inseguridad, la calidad del aire, y la falta de acceso a internet. En el 2021, se aplicó el Estudio de Calidad de Vida, desarrollado por la Pontificia Universidad Javeriana, la Cámara de Comercio de Bogotá, la Fundación Corona y el diario El Tiempo, en el que se analizaron 94 indicadores en factores en salud, empresa, pobreza, educación, ambiente, seguridad, movilidad y hábitat para conocer cómo están los bogotanos (Cámara de Comercio de Bogotá et al., 2021).

Este estudio dio a conocer grandes resultados como por ejemplo que en la capital colombiana mejoró el acceso a vivienda, basados en su calidad, en el 2020 el déficit denotaba un 12,7%, mientras que, en el 2021, la cifra se redujo llegando al 10,5%. Asimismo, el reporte informa que para el mismo año Bogotá mejoró económicamente llegando a un PIB del 9% y una reducción de la pobreza del 1,8%, a pesar estos valores positivos, el informe también señala que la ciudad es la más desigual del país (Cámara de Comercio de Bogotá et al., 2021).

Para el presente trabajo se usarán los datos de la Encuesta de Calidad de Vida del 2021, realizada por el DANE, empleando específicamente los registros obtenidos en la ciudad de Bogotá D.C. Con el fin de comparar los resultados obtenidos para conocer cuáles son las variables o factores que inciden en la calidad de vida de los habitantes de la capital colombiana, usando técnicas de machine learning que permitan este proceso. El machine learning es estratégico para este cometido porque su aplicabilidad permite el análisis de datos para la

toma de decisiones y para el estudio social de acontecimientos como la calidad de vida. Asimismo, este campo permite que los computadores aprendan por ellos mismos, cada vez teniendo un mejor desempeño, precisando sus habilidades para obtener resultados valiosos (Umaquina, Ana Suárez Zambrano, Luis Oña, Omar, 2018).

La aplicabilidad del machine learning es ideal para este proyecto porque permite analizar una cantidad de datos muy grande o pequeña, además, para su funcionamiento no se requiere una hardware de alto rendimiento. Asimismo, para su aplicabilidad divide las tareas en pequeñas actividades y luego combina los resultados para llegar a una conclusión. Por último, el tiempo de entrenamiento no es tan amplio, en comparación con otros métodos como el Deep Learning.

La calidad de vida es un tema recurrente entre la sociedad, haciendo alusión a su nivel de bienestar en sus vivencias diarias, de esta manera una persona puede ver influenciado este indicador por condiciones externas como el estado de las vías de su barrio, o factores internos como la salud mental.

Grandes ciudades como Bogotá, tienen condiciones que pueden afectar la calidad de vida como el tráfico, la inseguridad, la calidad del aire, y la falta de acceso a internet. De acuerdo con el Estudio de Calidad de Vida del 2021, desarrollado por la Pontificia Universidad Javeriana, la Cámara de Comercio de Bogotá, la Fundación Corona y el diario El Tiempo, en el que se analizaron 94 indicadores en factores en salud, empresa, pobreza, educación, ambiente, seguridad, movilidad y hábitat para conocer cómo están los bogotanos.

Este estudio dio a conocer grandes resultados como por ejemplo que en la capital colombiana mejoró el acceso a vivienda, basados en su calidad, en el 2020 el déficit denotaba un 12,7%, mientras que, en el 2021, la cifra se redujo llegando al 10,5%. Asimismo, el reporte informa que para el mismo año Bogotá mejoró económicamente llegando a un PIB del 9% y una reducción de la pobreza del 1,8%, a pesar estos valores positivos, el informe también señala que la ciudad es la más desigual del país.

2 Características del proyecto de investigación que hace uso de Bases de Datos (Max 500 Palabras) *- (Primera entrega)*

2.1 Título del proyecto de investigación (Max 100 Palabras) - *(Primera entrega)*

Modelo de machine learning para el análisis de calidad de vida de los habitantes de Bogotá D.C.

2.2 Objetivo general (Max 100 Palabras) - (*Primera entrega*)

Caracterizar los factores que determinan el nivel de calidad de vida de los habitantes de Bogotá D.C.

2.2.1 Objetivos específicos (Max 100 Palabras) - (*Primera entrega*)

- Comparar las definiciones de calidad de vida que se han suministrado desde diferentes disciplinas como la economía, la salud, la psicología y la sociología, entre otras.
- Analizar los datos de calidad de vida obtenidos por el DANE en el año 2021 en Bogotá D.C, aplicando herramientas y procedimientos de estadística y machine learning.
- Identificar la incidencia de diversas características, situaciones o condiciones que permitan predecir la calidad de vida de los habitantes de Bogotá D.C.

2.3 Alcance (Max 200 Palabras) - (*Primera entrega*)

Si bien la "Encuesta nacional de calidad de vida" realizada por el DANE en el año 2021, proporciona información relevante y diversa de las condiciones de la población colombiana a nivel nacional, el presente trabajo tiene como enfoque principal los habitantes de Bogotá D.C, los cuales habitan una de las principales ciudades y la capital del país que a su vez, es elegida como destino habitacional por algunas personas que provienen de otras ciudades y países vecinos en búsqueda de mejores oportunidades y calidad de vida.

Así mismo, el modelo que se pretende desarrollar en el presente trabajo tiene como objetivo desarrollar y aplicar un modelo que permita predecir la calidad de vida de los habitantes de Bogotá D.C. y a su vez, categorizar las condiciones y factores que influyen significativamente en la misma.

2.4 Pregunta de investigación (Max 100 Palabras) - (*Primera entrega*)

¿Qué factores inciden en la calidad de vida de los habitantes de Bogotá D.C?

2.5 Hipotesis (Max 100 Palabras) - (*Primera entrega*)

Al abordar un estudio de calidad de vida en los habitantes de Bogotá D.C. permite plantear diferentes hipótesis que pueden ser analizadas en el desarrollo del presente trabajo, tales como:

- Las variables de ingreso y nivel educativo influyen la calidad de vida de los habitantes de Bogotá D.C.
- La calidad de vida de los habitantes de Bogotá D.C. mejora con el paso del tiempo.

- Los habitantes de Bogotá D.C. que tienen acceso al trabajo formal, tienen un alto nivel de calidad de vida.
- Los habitantes de Bogotá D.C. que presentan complicaciones de salud, tienen un bajo calidad de vida.

3 Reflexiones sobre el origen de datos e información

(Max 400 Palabras) - (*Primera entrega*)

A lo largo del tiempo, la calidad de vida de la población ha sido un tema de gran interés, el cuál ha tenido diversos aportes desde la perspectiva económica, política, social, psicológica, médica, entre otras; así mismo, se han implementado diferentes mecanismos en búsqueda de su constante mejoramiento.

En el caso Colombiano, los entes gubernamentales y la entidades que se encuentran dentro de su competencia, se encuentran constantemente en búsqueda del mejoramiento de la calidad de vida de su población, para lo que es necesario contar con información veraz y actualizada de las condiciones con las que cuentan sus habitantes, sirviendo de fuente principal del análisis y toma de decisiones en torno a los objetivos planteados.

Uno de los mecanismos principales para la obtención de esta información, es la elaboración y aplicación de encuestas a la población objetivo, en las cuales es importante asegurar su implementación en grupos poblacionales con diversas características o condiciones como tipos de territorios, razas, edades, etnias, principales ciudades, ideologías, entre otros.

3.1 ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (*Primera entrega*)

El Departamento Administrativo Nacional de Estadística – DANE, aplica las encuestas de calidad de vida anualmente “como una respuesta a la necesidad de caracterizar la población en los diferentes aspectos involucrados en el bienestar de los hogares”. Esta encuesta permite identificar en la población colombiana factores de pobreza desde dos enfoques Necesidades Básicas Insatisfechas (NBI) y Línea de Pobreza, sin embargo, con el paso del tiempo se amplió el alcance de esta encuesta, incluyendo nuevas variables que permiten conocer más a profundidad la calidad de vida de los colombianos.

De acuerdo con el DANE, el principal objetivo de la encuesta es “obtener información que permita analizar y realizar comparaciones de las condiciones socioeconómicas de los hogares que habitan en el país, las cuales posibiliten hacer seguimiento a las variables necesarias para el diseño e implementación de políticas públicas.” (DANE, 2021).

3.2 ¿Cuales son las consideraciones legales o éticas del uso de la información? (Max 100 Palabras) - (*Primera entrega*)

La información resultante de la “Encuesta Nacional de Calidad de Vida 2021” corresponde a datos abiertos que se encuentran publicados en la página web del DANE, por lo cual, su acceso es de libre consulta y su empleabilidad no puede incurrir en incumplimientos legales, partiendo del hecho que se encuentran

anonimizados con el principal objetivo de cumplir con la protección de datos personales contemplados en la ley de habeas data 1581 de 2012.

Así mismo, las características de la información orienta su uso explícitamente para la exploración y análisis de las condiciones de la calidad de vida de los colombianos, así como también a proporcionar el entorno real para la toma de decisiones e implementaciones de políticas públicas en beneficio de la sociedad.

3.3 ¿Cuales son los retos de la información y los datos que utilizara en la base de datos en terminos de la calidad y la consolidación? (Max 100 Palabras) - (Primera entrega)

La encuesta está conformada por once capítulos que han sido aplicados de manera continua durante los últimos años. Además, existen capítulos especiales que se aplican periódicamente para satisfacer las necesidades específicas de los usuarios, los cuales son:

- Cada capítulo cuenta con una base de datos en formato .csv agrupada por un número de identificación del hogar encuestado.
- Cada hogar encuestado cuenta con un número de identificación de la persona encuestada.
- Se encuestaron 88.723 hogares a nivel nacional.
- Se encuestaron 257.589 habitantes a nivel nacional.
- El promedio de habitantes por hogar a nivel nacional es de 2,9.
- Se encuestaron 2.777 hogares en Bogotá.
- Se encuestaron 7.692 habitantes en Bogotá.
- El promedio de habitantes por hogar en Bogotá es de 2,7.
- La mayoría de las variables de la base de datos son categóricas.
- La base de datos cuenta con variables numéricas y categóricas.

3.4 ¿Que espera de la utilización de un sistema de Bases de Datos para su proyecto? (Max 100 Palabras) - (Primera entrega)

Los principales propósitos del uso de bases de datos para la implementación del proyecto son:

- Contar con un repositorio en donde se encuentre la información almacenada y disponible para cuando se requiera procesar. Permitir realizar las consultas necesarias a la información cuando se requiera.

- Contar con la información organizada y clasificada de la manera en que se requiera para el análisis de los datos, construcción e implementación de los modelos.
- Hacer uso de una herramienta que permita la conexión con diferentes procesadores de código empleado para la construcción e implementación de los modelos.
- Determinar las relaciones entre los datos y las variables.
- Realizar operaciones entre las variables con el fin de reducir columnas para mejorar el procesamiento de los datos.

4 Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos) *(Primera entrega)*

4.1 Características del SMBD (Sistema Manejador de Bases de Datos) para el proyecto *(Primera entrega)*

Para este proyecto se decide emplear el Sistema Manejador de Bases de Datos MySQL porque es una herramienta de gestión de bases de datos muy completa y versátil que permite diseñar, modelar, generar y mantener bases de datos MySQL de una manera fácil y rápida.

Además, proporciona una interfaz gráfica de usuario (GUI) intuitiva que facilita el trabajo con bases de datos. Cuenta con múltiples funcionalidades que permiten manipular y visualizar los datos de forma eficiente, como la edición de tablas, la creación de consultas y la ejecución de scripts.

Es una herramienta multiplataforma, lo que significa que se utilizar en diferentes sistemas operativos, como Windows, macOS y Linux, lo que la hace muy accesible para diferentes tipos de usuarios.

Se destaca por ser gratuita y de código abierto, lo que significa que no se necesita invertir grandes cantidades de dinero para utilizarla y además se tiene la libertad de modificar y personalizar la herramienta según las necesidades del administrador.

Por último, cuenta con una amplia comunidad de usuarios que proporciona soporte técnico, documentación, tutoriales y recursos adicionales que pueden ayudar a resolver problemas.

4.2 Diagrama modelo de datos *(Primera entrega)*

Para este ejercicio se emplearon 3 tablas de la encuesta de Calidad de Vida del Dane del 2021. A continuación, se presenta el modelo de base de datos, en las imágenes se evidencia la relación de cada columna por la llave primaria DIRECTORIO, siendo la característica que enlace toda la bases de datos,

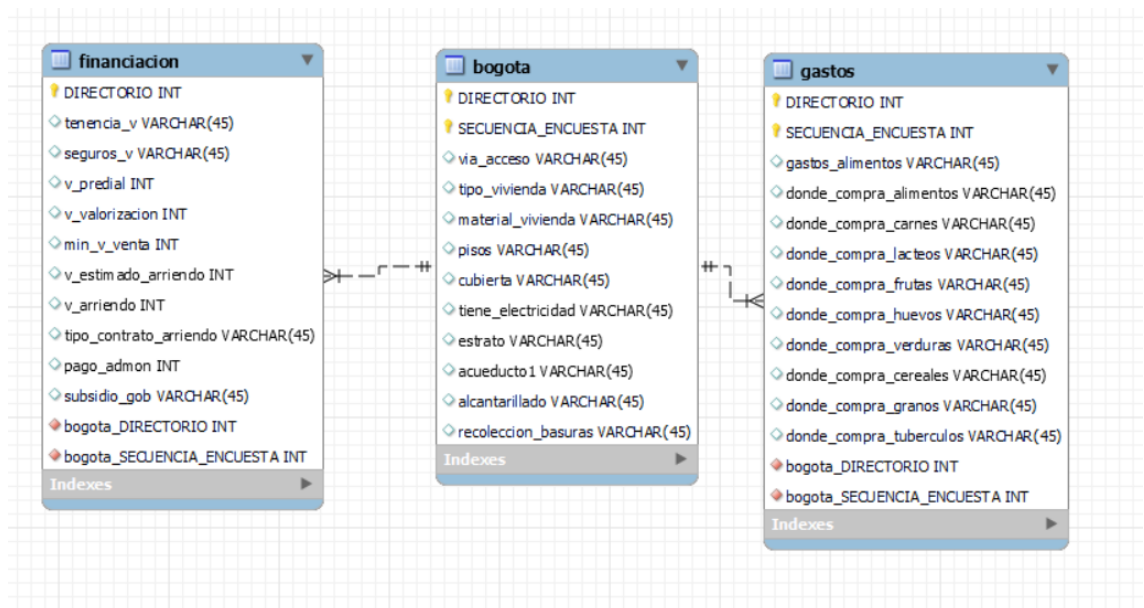


Figura 1

4.3 Imágenes de la Base de Datos (*Primera entrega*)

Se relacionan las visualizaciones de la base de datos, en primer lugar, se creó un esquema llamado calidad en el que se incluyeron las tablas: gastos, Bogotá y financiación.

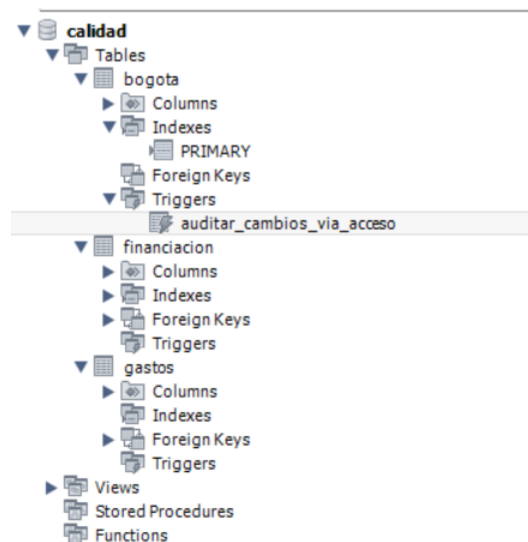


Figura 2

Local instance MySQL80 x

File Edit View Query Database Server Tools Scripting Help

Navigator: Filter objects

SCHEMAS

- calidad
 - Tables
 - bogota
 - financiacion
 - gastos
 - Views
 - Stored Procedures
 - Functions
- sys

Query 1 bogota x financiacion gastos bogota financiacion gastos

1 • `SELECT * FROM calidad.bogota;`

Result Grid

	DIRECTORIO	SECUENCIA	via_acceso	tipo_vivienda	material_vivienda	pisos	cubierta	tiene_electricidad	estrato
▶	7567439	1	1	1	1	4	1	1	3
	7567440	1	4	1	1	4	1	1	3
	7567441	1	2	2	1	4	1	1	6
	7567442	1	4	2	1	4	1	1	6
	7569259	1	4	1	1	4	1	1	3
	7569260	1	4	2	1	2	1	1	6
	7569261	1	1	2	1	4	1	1	3
	7569262	1	4	2	1	4	1	1	3
	7574885	1	4	1	1	4	1	1	3
	7574886	1	4	1	1	4	1	1	4
	7574887	1	2	2	1	2	1	1	4
	7574888	1	4	2	1	4	1	1	3
	7574889	1	4	1	1	4	1	1	4
	7574890	1	4	1	1	4	1	1	4
	7574891	1	4	1	1	4	1	1	3

Administration Schemas

Information: Table: gastos

bogota 1 x

Apply Revert

Figura 3

SCHEMAS

Filter objects

- calidad
 - Tables
 - bogota
 - financiacion
 - gastos
 - Views
 - Stored Procedures
 - Functions
- sys

Administration Schemas Information

Table: **financiacion**

1 • `SELECT * FROM calidad.financiacion;`

Limit to 1000 rows

Result Grid

	DIRECTORIO	tenencia_v	seguros_v	v_predial	v_valorizacion	min_v_venta	v_estimado_arriendo	v_arriendo	tip
▶	7567439	3	0	0	0	0	0	600000	2
	7567440	1	2	1923000	98	500000000	800000	0	0
	7567441	1	1	9500000	0	120000000	1500000	0	0
	7567442	3	0	0	0	0	0	3000000	1
	7569259	3	0	0	0	0	0	400000	1
	7569260	3	0	0	0	0	0	2800000	2
	7569261	3	0	0	0	0	0	1100000	1
	7569262	1	2	4000000	99	800000000	1200000	0	0
	7574885	1	2	99	98	900000000	1000000	0	0
	7574886	2	1	1800000	18000000	450000000	2000000	0	0
	7574887	3	0	0	0	0	0	1600000	2
	7574888	4	0	0	0	0	2000000	0	0
	7574889	1	2	98	1800000	620000000	2000000	0	0
	7574890	3	0	0	0	0	0	2300000	2
	7574891	1	2	1500000	1500000	300000000	300000	0	0

financiacion 1 x

Apply Revert

Figura 4

SCHEMAS

Filter objects

- calidad
 - Tables
 - bogota
 - financiacion
 - gastos
 - Views
 - Stored Procedures
 - Functions
- sys

Administration Schemas Information

Table: **financiacion**

Query 1 bogota financiacion **gastos**

1 • `SELECT * FROM calidad.gastos;`

Limit to 1000 rows

Result Grid

	DIRECTORIO	SECUENCIA_ENCUESTA	gastos_alimentos	donde_compra_alimentos	donde_compra_carne	donde_compra
▶	7567439	1	1	0	6	2
	7567440	1	1	0	2	2
	7567441	1	2	0	1	1
	7567442	1	1	0	1	1
	7569259	1	1	0	2	2
	7569260	1	1	0	1	1
	7569261	1	1	0	2	2
	7569262	1	1	0	3	3
	7574885	1	1	0	1	1
	7574886	1	1	0	1	1
	7574887	1	1	0	8	2
	7574888	1	1	0	2	2
	7574889	1	1	0	2	6
	7574890	1	1	0	8	1
	7574891	1	1	0	3	3

gastos 1 x

Apply Revert

Figura 5

4.4 Código SQL - lenguaje de definición de datos (DDL) (Primera entrega)

Se comparte el código con el que se crearon las tablas, la asignación de las llaves primarias, el tipo de datos que irá por columna. Los datos se importaron directamente sobre cada tabla desde un archivo CSV.

```
use calidad;
create table Bogota (DIRECTORIO INT PRIMARY KEY,
SECUENCIA INT NOT NULL,
via_acceso INT NOT NULL,
tipo_vivienda INT NOT NULL,
material_vivienda INT NOT NULL,
pisos INT NOT NULL,
cubierta INT NOT NULL,
tiene_electricidad INT NOT NULL,
estrato INT NOT NULL,
acueducto1 INT NOT NULL,
alcantarillado INT NOT NULL,
recoleccion_basuras INT NOT NULL);
```

```
create table gastos (
DIRECTORIO INT PRIMARY KEY,
SECUENCIA_ENCUESTA INT NOT NULL,
gastos_alimentos INT NOT NULL,
donde_compra_alimentos INT NOT NULL,
donde_compra_carnes INT NOT NULL,
donde_compra_lacteos INT NOT NULL,
donde_compra_frutas INT NOT NULL,
donde_compra_huevos INT NOT NULL,
donde_compra_verduras INT NOT NULL,
donde_compra_cereales INT NOT NULL,
donde_compra_granos INT NOT NULL,
donde_compra_tuberculos INT NOT NULL );
```

```
create table financiacion (
DIRECTORIO INT PRIMARY KEY,
tenencia_v INT NOT NULL,
seguros_v INT NOT NULL,
v_predial INT NOT NULL,
v_valorizacion INT NOT NULL,
min_v_venta INT NOT NULL,
v_estimado_arriendo INT NOT NULL,
v_arriendo INT NOT NULL,
tipo_contrato_arriendo INT NOT NULL,
pago_admon INT NOT NULL,
subsidio_gob INT NOT NULL)
```


4.5 Código SQL - Manipulación de datos (DML) (*Primera entrega*)

Se hace uso de los comandos SELECT, para indicar las columnas a las que se añadirá un registro a través de la función INSERT TO.

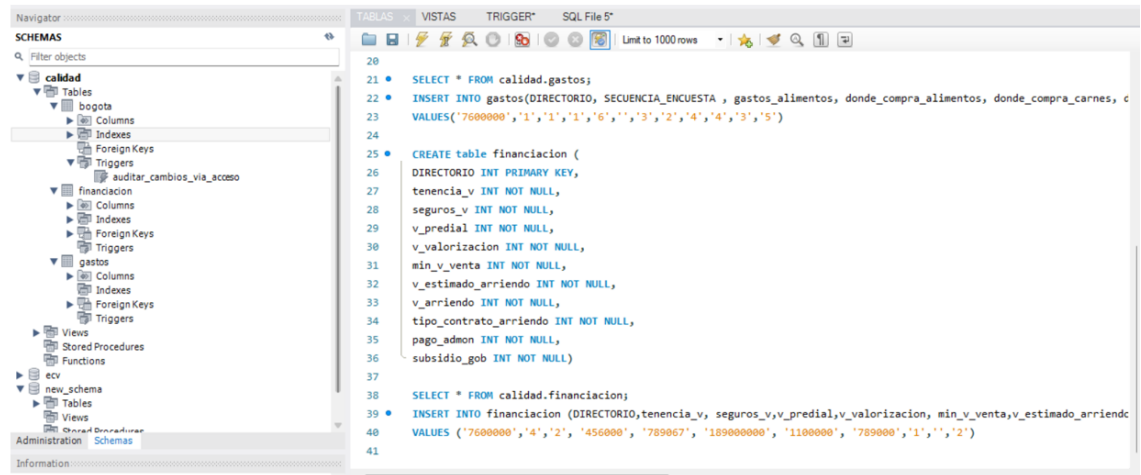


Figura 6

4.6 Código SQL + Resultados: Vistas (*Primera entrega*)

Se generaron las siguientes vistas, las cuales actúan como filtros que facilitan el acceso a la información.

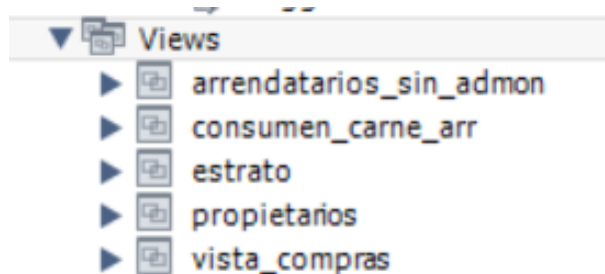


Figura 7: se evidencian 5 vistas.

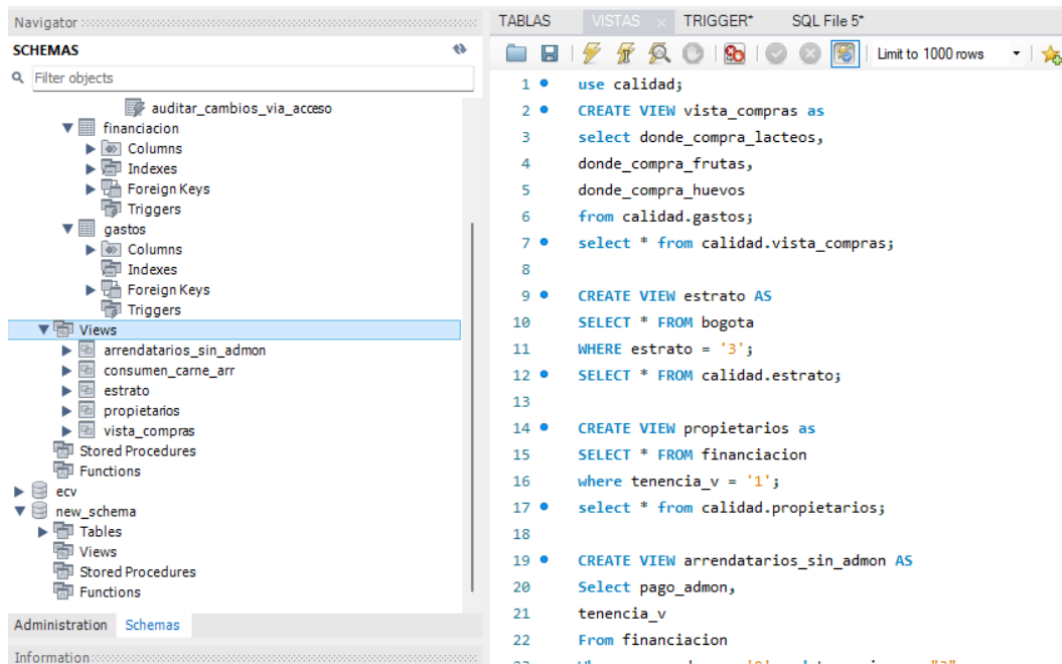


Figura 8: Creación vista

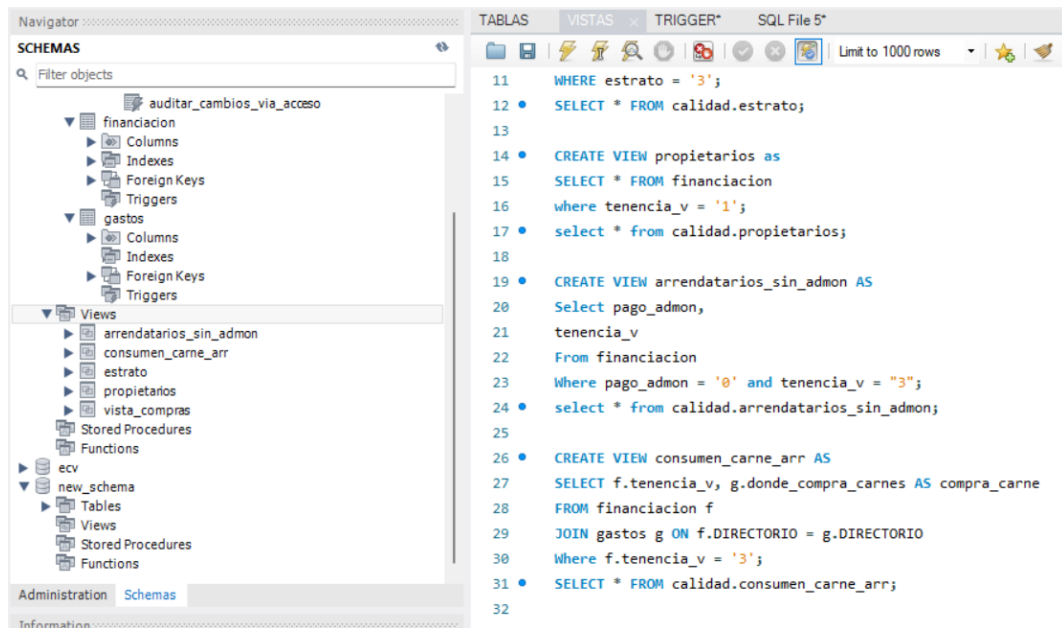


Figura 9: Creación vista

1 • SELECT * FROM calidad.arrendatarios_sin_admon;

Result Grid

pago_admon	tenencia_v
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3
0	3

1 • SELECT * FROM calidad.consumen_carne_arr;

Result Grid

tenencia_v	compra_carne
3	6
3	1
3	2
3	1
3	2
3	8
3	8
3	1
3	2
3	1
3	2
3	2
3	8
3	2
3	1
3	1

1 • SELECT * FROM calidad.vista_compras;

Result Grid

donde_compra_lacteos	donde_compra_frutas	donde_compra_huevos
2	2	2
2	2	6
1	1	1
1	5	1
2	2	2
1	1	1
2	2	2
3	3	3
1	1	1
1	1	1
2	2	8

1 • SELECT * FROM calidad.estrato;

Result Grid

DIRECTORIO	SECUENCIA	via_acceso	tipo_vivienda	material_vivienda	pisos	cubierta	tiene_electricidad	estrato	acueducto1	alcantarillado	recoleccion_basuras
7567439	1	1	1	1	4	1	1	3	1	1	1
7567440	1	4	1	1	4	1	1	3	1	1	1
7569259	1	4	1	1	4	1	1	3	1	1	1
7569261	1	1	2	1	4	1	1	3	1	1	1
7569262	1	4	2	1	4	1	1	3	1	1	1
7574885	1	4	1	1	4	1	1	3	1	1	1
7574888	1	4	2	1	4	1	1	3	1	1	1
7574891	1	4	1	1	4	1	1	3	1	1	1
7576534	1	2	2	1	5	1	1	3	1	1	1
7576537	1	4	2	1	4	1	1	3	1	1	1
7576538	1	4	1	1	4	1	1	3	1	1	1

Figura 10: Conjunto de vistas

4.7 Código SQL + Resultados: Triggers (Primera entrega)

Diseño de trigger para auditar los datos que ingresan a la columna vía acceso, los datos que se ingresen se almacenarán en una tabla aparte, lo que permitirá analizar cómo se ha hecho la modificación de la variable.

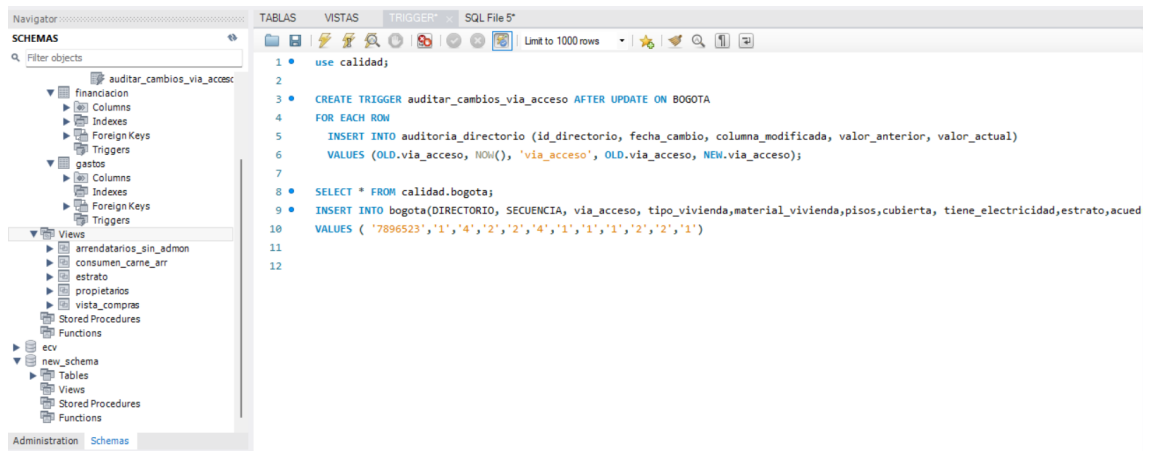


Figura 11: Trigger

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

La función suma gastos registros permite hacer una suma de 4 variables de la tabla financiación, lo que permite conocer cuánto se ha gastado en pago del impuesto predial, arriendo, administración y valorización.

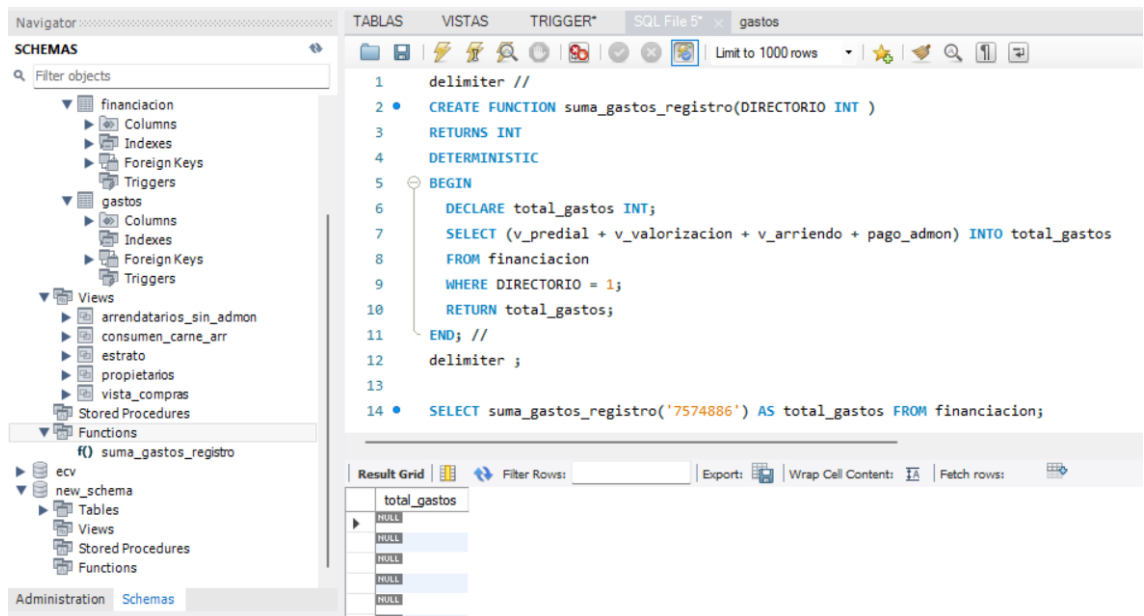


Figura 12: Funciones

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

Este procedimiento permitirá añadirle un sufijo a los datos duplicados para dejar en una variable registros únicos.

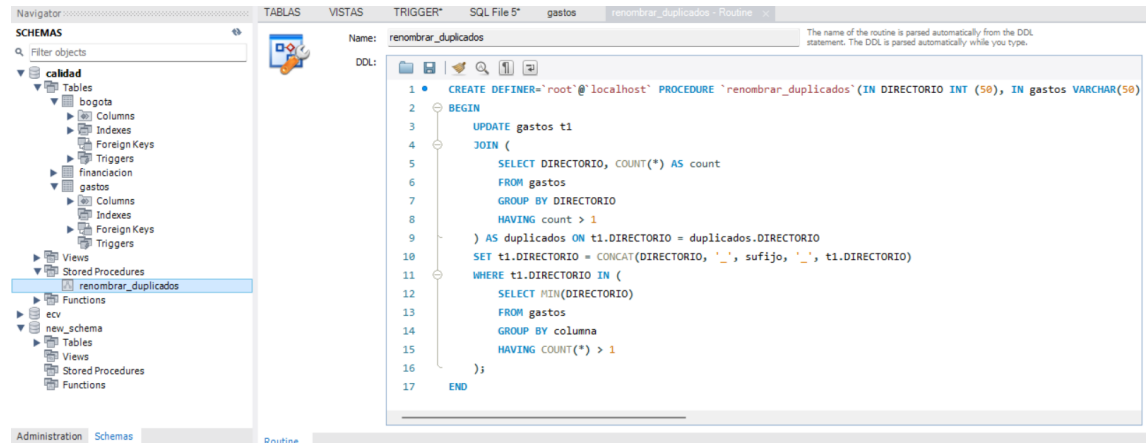


Figura 13: Procedimientos almacenados

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 Proximos pasos (*Tercera entrega*)

8 Lecciones aprendidas (*Tercera entrega*)

9 Bibliografía

References

- [1] DANE., *Encuesta Nacional de Calidad de Vida 2018 [dataset]*, <https://microdatos.dane.gov.co/index.php> (2021).
- [2] Bogotá Cómo Vamos., *Informe de calidad de vida 2021*, <https://bogotacomovamos.org/> (2021).
- [3] Cámara de comercio de Bogotá, Fundación Corona, Pontificia Universidad Javeriana, El Tiempo., *Informe de calidad de vida en Bogotá, 2021: Una mirada al derecho a la ciudad*, <https://s3.documentcloud.org/documents/22123000/icv-2021-ii.pdf> (2022).
- [4] United Nations. *Definición y medición internacional del nivel de vida. Department of Economic and Social Affairs, United Nations.*, https://unstats.un.org/unsd/publication/seriee/seriee_cn3270_evs.pdf (1961).
- [5] Huerta-Riveros, Patricia, Yáñez-Alvarado, Miguel, Leyton-Pavez, Carolina., *Factores que impactan en la calidad de vida y su incidencia en las directrices sanitarias. Hacia la Promoción de la Salud*, <https://doi.org/10.17151/hpsal.2022.27.2.10> (2022).
- [6] Campbell A, Converse P, Rodgers E., *The Quality of american life: Perceptions, Evaluations and Satisfaction* (2022).
- [7] Hurtado Briceño, Alberto Zerpa de Hurtado, Sadcidi., *Perspectivas teóricas del estudio de la calidad de vida en economía*, (2011).
- [8] Ruiz, Miguel Pardo, Antonio., *Calidad de vida relacionada con la salud: Definición y utilización en la práctica médica*, (2005)
- [9] OECD., *How's Life? 2020: Measuring Well-being*, <https://doi.org/10.1787/9870c393-en> (2020)
- [10] Umaquinga, Ana Suárez Zambrano, Luis Oña, Omar., *Machine learning: Importance, advances, techniques and applications*, (2018)