

Proyecto de Bases de datos

Diana Isabel Arévalo Durán

Universidad Central

Maestría en Analítica de Datos

Curso de Bases de Datos

Bogotá, Colombia

darevalod@ucentral.edu.co

October 21, 2022

Contents

1	Introducción	3
2	Características del proyecto	3
2.1	Titulo del proyecto	3
2.2	Objetivo general	3
2.2.1	Objetivos especificos	3
2.3	Alcance	4
2.4	Pregunta de investigación	4
2.5	Hipotesis	4
3	Reflexiones sobre el origen de datos e información	5
3.1	¿Cual es el origen de los datos e información?	5
3.2	¿Cuales son las consideraciones legales o eticas del uso de la información?	5
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?	6
3.4	¿Que espera de la utilización de un sistema de Bases de Datos para su proyecto?	6
4	Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos)	7
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)	11
4.5	Código SQL - Manipulación de datos (DML)	14

4.6	Código SQL + Resultados: Vistas	16
4.7	Código SQL + Resultados: Triggers	16
4.8	Código SQL + Resultados: Funciones	17
4.9	Código SQL + Resultados: procedimientos almacenados	17
5	Bases de Datos No-SQL (<i>Segunda entrega</i>)	18
5.1	Diagrama Bases de Datos No-SQL (<i>Segunda entrega</i>)	18
5.2	SMBD utilizado para la Base de Datos No-SQL (<i>Segunda entrega</i>)	18
6	Aplicación de ETL (Extract, Transform, Load) y Bodega de Datos (<i>Tercera entrega</i>)	19
6.1	Ejemplo de aplicación de ETL y Bodega de Datos (<i>Tercera entrega</i>)	19
7	Lecciones aprendidas (<i>Tercera entrega</i>)	20

1 Introducción

El presente trabajo se realiza con base en los datos recolectados en la gran encuesta de los hogares Colombianos 2020, con el objetivo de predecir la probabilidad de que los niños de Colombia, menores de 12 años al 2020, en el futuro accedan a educación superior o universitaria.

Para esto, inicialmente se realiza normalización de la base de datos obtenida del Archivo Nacional de Datos (ANDA) que permite la consulta por temática, operación estadística y año, de los diferentes microdatos anonimizados dispuestos al público en la página web del DANE, haciendo uso del sistema manejador de base de datos MySQL.

Posteriormente se aplican métodos de estadística exploratoria multidimensional, resumiendo el conjunto de variables en unas pocas nuevas variables, buscando agrupaciones en los datos y clasificando para identificar relaciones entre variables. Finalmente se contrastan los resultados contra la hipótesis planteada.

2 Características del proyecto

2.1 Título del proyecto

Predicción de la probabilidad que tienen los niños de los hogares colombianos menores de 12 años al 2020, de acceso a la educación superior, de acuerdo con condiciones de empleo, fuentes de ingreso y características generales de los hogares.

2.2 Objetivo general

Predecir la probabilidad de que los niños de Colombia, menores de 12 años al 2020, en el futuro accedan a educación superior o universitaria.

2.2.1 Objetivos específicos

- A través del uso de métodos de reducción de dimensionalidad generar visualizaciones de datos que permitan la fácil comprensión de los datos de la encuesta general de hogares colombianos.
- Por medio de la clasificación y regresión, analizar comportamientos para predecir la probabilidad de acceso a educación superior de los niños menores de 12 años.

2.3 Alcance

El alcance de este proyecto será la predicción probabilidad de que los niños de Colombia, menores de 12 años al 2020, en el futuro accedan a educación superior o universitaria.

2.4 Pregunta de investigación

Entre 0 y 12 años ¿Qué probabilidad tienen los niños de los hogares colombianos de la Gran encuesta integrada de hogares 2020, de acceder a la educación superior o universitaria de acuerdo con condiciones de empleo, fuentes de ingreso y características generales de los hogares?.

2.5 Hipotesis

La probabilidad de acceso a la educación superior o universitaria de los niños en Colombia está determinada por el nivel educativo de los padres.

3 Reflexiones sobre el origen de datos e información

La Gran encuesta integrada de hogares es una encuesta mediante la cual se solicita información sobre las condiciones de empleo de las personas, además de las características generales de la población como sexo, edad, estado civil y nivel educativo, se pregunta sobre sus fuentes de ingresos. Esta encuesta proporciona indicadores a nivel nacional del mercado laboral en Colombia que permiten conocer entre otros aspectos: la tasa de ocupación, la tasa de desocupación, la rama de actividad en que se desempeñan los colombianos y su remuneración, así como, el comportamiento del mercado laboral para jóvenes, mujeres y otros grupos poblacionales específicos.

3.1 ¿Cual es el origen de los datos e información?

Los datos corresponden a la Gran Encuesta Integrada de Hogares Colombianos 2020 realizada por el Departamento Administrativo Nacional de Estadística - DANE, a la población civil no institucional residente en hogares particulares. Esta población se estima con base en los censos de población, las estadísticas vitales y de migración.

Estos datos fueron obtenidos del Archivo Nacional de Datos (ANDA) que permite la consulta por temática, operación estadística y año, de los diferentes microdatos anonimizados dispuestos al público en la página web del DANE.

3.2 ¿Cuales son las consideraciones legales o eticas del uso de la información?

Toda la información recolectada para los censos y encuestas de los procesos estadísticos del DANE está protegida por la Ley 79 de 1993 o Ley de Reserva Estadística.

Según esta ley los datos suministrados al DANE a través de censos, encuestas u operaciones estadísticas "no podrán darse a conocer al público ni a las entidades u organismos oficiales, ni a las autoridades públicas, sino únicamente en resúmenes numéricos".

En consonancia con la Ley 1581 de 2012, las bases de datos y archivos regulados por la Ley 79 de 1993 están exceptuadas del régimen de protección de datos personales. Sin embargo, deben aplicarse los principios establecidos para su protección como son: el principio de legalidad en materia de tratamiento de datos, de finalidad, de libertad, de veracidad o calidad, de transparencia, de acceso y circulación restringida, de seguridad y de confidencialidad.

Por lo anterior, las bases anonimizadas y disponibles en el ANDA han surtido un proceso de modificación y transformación de los datos originales, con el

objetivo de garantizar la confidencialidad de las unidades de análisis, por lo cual es necesario tener en cuenta estas modificaciones para el uso de la información.

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?

Los datos obtenidos están representados con variables categóricas y el nombre de estas es dado con códigos, por lo que se debe realizar listas de homologación tanto para los nombres como para las categorías, con el objetivo de poder generar información sencilla y entendible.

Debido a que las bases publicadas por el DANE se encuentran anonimizadas y han surtido un proceso de modificación y transformación de los datos originales, la estructura de la base, entregada en el documento técnico de la DIAN, no corresponde con la publicada en la página, es por esto que se debe trabajar en la identificación de variables y homologaciones disponibles.

3.4 ¿Que espera de la utilización de un sistema de Bases de Datos para su proyecto?

Se hace uso del sistema de base de datos con el objetivo de organizar los datos obtenidos para poder acceder a búsquedas eficientes y consultas simplificadas que puedan ser reutilizadas.

Del sistema se espera extraer bases de datos unificadas con las variables más importantes de la operación estadística para realizar análisis de correspondencia.

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

El Sistema Manejador de Bases de Datos utilizado para el proyecto es MySQL y dentro de sus principales características están:

- Sistema de base de datos relacional de uso libre y gratuito.
- Soporta gran cantidad de datos, incluso con más de 50 millones de registros.
- Ejecución de transacciones y uso de claves foráneas.
- Conectividad segura.

4.2 Diagrama modelo de datos

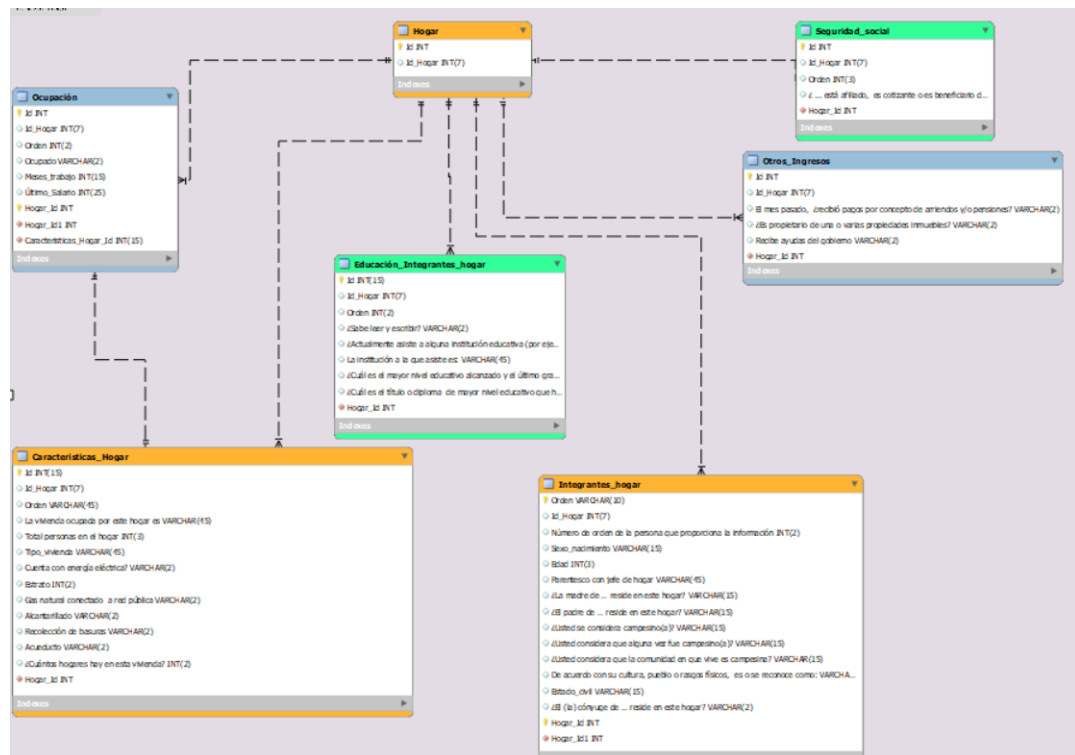


Figure 1: Diagrama Entidad Relación

4.3 Imágenes de la Base de Datos

Result Grid

Filter Rows:

Edits

Export/Import:

Wrap Cell Contents:

	Id	Id_Hogar	Orden	La vivienda ocupada por este hogar es	Total personas en el hogar	Tipo_vivienda	Cuenta con energía eléctrica?	Estrato	Gas natural conectado a red pública
▶	1	2120000	1	Posesion sin titulo	5	Casa	SI	0	NO
	2	2120000	2	Posesion sin titulo	5	Casa	SI	0	NO
	3	2120000	3	Posesion sin titulo	5	Casa	SI	0	NO
	4	2120000	4	Posesion sin titulo	5	Casa	SI	0	NO
	5	2120000	5	Posesion sin titulo	5	Casa	SI	0	NO
	6	2120001	1	Posesion sin titulo	1	Casa	SI	0	NO
	7	2120002	1	Posesion sin titulo	2	Casa	SI	0	NO
	8	2120002	2	Posesion sin titulo	2	Casa	SI	0	NO
	9	2120003	1	Posesion sin titulo	2	Casa	SI	0	NO
	10	2120003	2	Posesion sin titulo	2	Casa	SI	0	NO
	11	2120004	1	Posesion sin titulo	6	Casa	SI	0	NO

caracteristicas_hogar24

×

Apply

Revert

Result Grid

Form Editor

Field Types

Figure 2: Caracteristicas hogar

	Id	Id_Hogar	Orden	¿Sabe leer y escribe?	¿Actualmente asiste a alguna institución educativa?	La institución a la que asiste es:	¿mayor nivel educativo alcanzado?	¿título o diploma de su nivel?
1	2120000	1	SI	NO			Basica secundaria 6o 9o	
2	2120000	2	NO	NO			Ninguno	
3	2120000	3	SI	SI		Publica	Basica primaria 1o 5o	
4	2120000	4	NO	SI		Publica	Basica primaria 1o 5o	
5	2120000	5	NO	NO			Ninguno	
6	2120001	1	SI	NO			Basica primaria 1o 5o	
7	2120002	1	NO	NO			Ninguno	
8	2120002	2	SI	NO			Basica primaria 1o 5o	
9	2120003	1	NO	NO			Ninguno	
10	2120003	2	NO	NO			Ninguno	
11	2120004	1	SI	NO			Basica secundaria 6o 9o	

Figure 3: Educación integrantes hogar

	Id	Id_Hogar
1	212	
2	21200001	
3	21200002	
4	21200003	
5	21200004	
6	21200005	
7	21200006	
8	21200007	
9	21200008	
10	21200009	
11	2120001	
12	21200011	
13	21200012	

Figure 4: Hogar

Result Grid					
Filter Rows:					
Edit: Export/Import: Wrap Cell Content:					
	Id	Id_Hogar	Orden	El mes pasado, recibí pagos por arriendos o pensiones	¿Es propietario de una o varias propiedades inmuebles?
1	2120000	1	NO		NO
2	2120000	2	NO		NO
3	2120000	3			NO
4	2120000	4			NO
5	2120000	5			NO
6	2120001	1	NO		NO
7	2120002	1	NO		NO
8	2120002	2	SI		NO
9	2120003	1	NO		NO
10	2120003	2	NO		NO
11	2120004	1	NO		SI
12	2120004	2			NO

Figure 5: Otros ingresos

Result Grid						
Filter Rows:						
Edit: Export/Import: Wrap Cell Content:						
	Id	Id_Hogar	Orden	Ocupado	Ocupacion	Meses_trabajo
9	2120003	1	SI	Asalariado	240	0
17	2120005	1	SI	Asalariado	8	900000
32	2120008	1	SI	Asalariado	10	908526
38	2120009	2	SI	Asalariado	2	600000
45	2120010	3	SI	Asalariado	24	0
79	2120020	2	SI	Asalariado	24	908526
86	2120021	3	SI	Asalariado	3	0
100	2120024	2	SI	Asalariado	24	0
122	2120037	3	SI	Asalariado	8	0
125	2120039	1	SI	Asalariado	492	2400000
133	2120042	4	SI	Asalariado	72	1200000
137	2120043	3	SI	Asalariado	12	908526
141	2120044	2	SI	Asalariado	13	910000

Figure 6: Ocupación

Result Grid				
Filter Rows:				
Edit: Export/Import: Wrap Cell Content:				
	Id	Id_Hogar	Orden	¿está afiliado entidad de seguridad social
1	2120000	1	SI	Subsidiado? (EPS-S)
2	2120000	2	SI	Subsidiado? (EPS-S)
3	2120000	3	SI	Subsidiado? (EPS-S)
4	2120000	4	SI	Subsidiado? (EPS-S)
5	2120000	5	SI	Subsidiado? (EPS-S)
6	2120001	1	SI	Subsidiado? (EPS-S)
7	2120002	1	SI	Subsidiado? (EPS-S)
8	2120002	2	SI	Subsidiado? (EPS-S)
9	2120003	1	SI	Subsidiado? (EPS-S)
10	2120003	2	SI	Subsidiado? (EPS-S)
11	2120004	1	SI	Subsidiado? (EPS-S)
12	2120004	2	SI	Subsidiado? (EPS-S)

Figure 7: Seguridad Social

Result Grid									
Filter Rows:									
Edit: Export/Import: Wrap Cell Content: <input type="checkbox"/>									
	Id	Orden	Id_Hogar	Número de orden de la persona	Sexo_nacimiento	Edad	Parentesco con jefe de hogar	¿La madre de ... reside en este hogar?	¿El padre de ... reside en este hogar?
▶	3	3	2	2120000	Masculino	7	Hjoea. hjaastroa	SI	SI
	4	4	2	2120000	Masculino	6	Hjoea. hjaastroa	SI	SI
	5	5	2	2120000	Masculino	3	Hjoea. hjaastroa	SI	SI
	12	2	2	2120004	Femenino	12	Hjoea. hjaastroa	SI	NO
	13	3	1	2120004	Femenino	9	Hjoea. hjaastroa	SI	NO
	14	4	1	2120004	Femenino	6	Hjoea. hjaastroa	SI	NO
	15	5	1	2120004	Masculino	5	Hjoea. hjaastroa	SI	NO
	16	6	1	2120004	Masculino	1	Hjoea. hjaastroa	SI	NO
	19	3	2	2120005	Femenino	1	Hjoea. hjaastroa	SI	SI
	20	4	2	2120005	Masculino	0	Hjoea. hjaastroa	SI	SI
	23	3	2	2120006	Femenino	7	Hjoea. hjaastroa	SI	SI

integrantes_hogar23 x

Apply

Revert

Result Grid

Form Editor

Field Types

Figure 8: Integrantes hogar

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

```
21 -----
22 -- Table `mydb`.`Hogar`
23 -----
24 ● CREATE TABLE IF NOT EXISTS `Encuesta_hogares`.`Hogar` (
25     `Id` INT NOT NULL AUTO_INCREMENT,
26     `Id_Hogar` INT(7) NULL,
27     PRIMARY KEY (`Id`))
28 ENGINE = InnoDB;
29
30
31 -----
32 -- Table `mydb`.`Integrantes_hogar`
33 -----
34 ● CREATE TABLE IF NOT EXISTS `Encuesta_hogares`.`Integrantes_hogar` (
35     `Id` INT NOT NULL AUTO_INCREMENT,
36     `Orden` VARCHAR(10) NULL,
37     `Id_Hogar` INT(7) NULL,
38     `Número de orden de la persona que proporciona la información` INT(2) NULL,
39     `Sexo_nacimiento` VARCHAR(15) NULL,
40     `Edad` INT(3) NULL,
41     `Parentesco con jefe de hogar` VARCHAR(45) NULL,
42     `¿La madre de ... reside en este hogar?` VARCHAR(15) NULL,
43     `¿El padre de ... reside en este hogar?` VARCHAR(15) NULL,
44     `¿Usted se considera campesino(a)?` VARCHAR(15) NULL,
45     `¿Usted considera que alguna vez fue campesino(a)?` VARCHAR(15) NULL,
46     `¿Usted considera que la comunidad en que vive es campesina?` VARCHAR(15) NULL,
47     `De acuerdo con su cultura se reconoce como:` VARCHAR(45) NULL,
48     `Estado_civil` VARCHAR(15) NULL,
49     `¿El (la) cónyuge de ... reside en este hogar?` VARCHAR(2) NULL,
50     `Hogar_Id` INT NOT NULL,
51     `Hogar_Id1` INT NOT NULL,
52     PRIMARY KEY (`Id`, `Hogar_Id`),
53     CONSTRAINT `fk_Integrantes_hogar_Hogar1`
54     FOREIGN KEY (`Hogar_Id1`)
55     REFERENCES `mydb`.`Hogar` (`Id`)
56     ON DELETE NO ACTION
57     ON UPDATE NO ACTION)
58 ENGINE = InnoDB;
```

```

63 -- Table `Encuesta_hogares`.`Caracteristicas_Hogar`
64 -----
65 ● CREATE TABLE IF NOT EXISTS `Encuesta_hogares`.`Caracteristicas_Hogar` (
66     `Id` INT(15) NOT NULL AUTO_INCREMENT,
67     `Id_Hogar` INT(7) NULL,
68     `Orden` VARCHAR(45) NULL,
69     `La vivienda ocupada por este hogar es` VARCHAR(45) NULL,
70     `Total personas en el hogar` INT(3) NULL,
71     `Tipo_vivienda` VARCHAR(45) NULL,
72     `Cuenta con energía eléctrica?` VARCHAR(2) NULL,
73     `Estrato` INT(2) NULL,
74     `Gas natural conectado a red pública` VARCHAR(2) NULL,
75     `Alcantarillado` VARCHAR(2) NULL,
76     `Recolección de basuras` VARCHAR(2) NULL,
77     `Acueducto` VARCHAR(2) NULL,
78     `¿Cuántos hogares hay en esta vivienda?` INT(2) NULL,
79     `Hogar_Id` INT NOT NULL,
80     PRIMARY KEY (`Id`),
81     CONSTRAINT `fk_Caracteristicas_Hogar_Hogar1`
82         FOREIGN KEY (`Hogar_Id`)
83         REFERENCES `mydb`.`Hogar` (`Id`)
84         ON DELETE NO ACTION
85         ON UPDATE NO ACTION)
86     ENGINE = InnoDB;
91 -- Table `Encuesta_hogares`.`Seguridad_social`
92 -----
93 ● CREATE TABLE IF NOT EXISTS `Encuesta_hogares`.`Seguridad_social` (
94     `Id` INT NOT NULL AUTO_INCREMENT,
95     `Id_Hogar` INT(7) NULL,
96     `Orden` INT(3) NULL,
97     `¿ está afiliado entidad de seguridad social` VARCHAR(45) NULL,
98     `A qué regimen` VARCHAR(45) NULL,
99     `Hogar_Id` INT NOT NULL,
100     PRIMARY KEY (`Id`),
101     CONSTRAINT `fk_Seguridad_social_Hogar1`
102         FOREIGN KEY (`Hogar_Id`)
103         REFERENCES `mydb`.`Hogar` (`Id`)
104         ON DELETE NO ACTION
105         ON UPDATE NO ACTION)
106     ENGINE = InnoDB;

```

```

111 -- Table `Encuesta_hogares`.`Educación_Integrantes_hogar`
112 -----
113 • CREATE TABLE IF NOT EXISTS `Encuesta_hogares`.`Educación_Integrantes_hogar` (
114     `Id` INT(15) NOT NULL AUTO_INCREMENT,
115     `Id_Hogar` INT(7) NULL,
116     `Orden` INT(2) NULL,
117     `¿Sabe leer y escribir?` VARCHAR(2) NULL,
118     `¿Actualmente asiste a alguna institución educativa?` VARCHAR(2) NULL,
119     `La institución a la que asiste es:` VARCHAR(45) NULL,
120     `¿mayor nivel educativo alcanzado?` VARCHAR(45) NULL,
121     `¿título o diploma de mayor nivel educativo?` VARCHAR(45) NULL,
122     `Hogar_Id` INT NOT NULL,
123     PRIMARY KEY (`Id`),
124     CONSTRAINT `fk_Educación_Integrantes_hogar_Hogar1`
125     FOREIGN KEY (`Hogar_Id`)
126     REFERENCES `mydb`.`Hogar` (`Id`)
127     ON DELETE NO ACTION
128     ON UPDATE NO ACTION)
129     ENGINE = InnoDB;
130
131 -- Table `Encuesta_hogares`.`Ocupación`
132 -----
133 • CREATE TABLE IF NOT EXISTS `mydb`.`Ocupación` (
134     `Id` INT NOT NULL,
135     `Id_Hogar` INT(7) NULL,
136     `Orden` INT(2) NULL,
137     `Ocupado` VARCHAR(2) NULL,
138     `Meses_trabajo` INT(15) NULL,
139     `Último_Salario` INT(25) NULL,
140     `Hogar_Id` INT NOT NULL,
141     `Hogar_Id1` INT NOT NULL,
142     `Características_Hogar_Id` INT(15) NOT NULL,
143     PRIMARY KEY (`Id`, `Hogar_Id`),
144     CONSTRAINT `fk_Ocupación_Hogar`
145     FOREIGN KEY (`Hogar_Id1`)
146     REFERENCES `mydb`.`Hogar` (`Id`)
147     ON DELETE NO ACTION
148     ON UPDATE NO ACTION,
149     CONSTRAINT `fk_Ocupación_Características_Hogar1`
150     FOREIGN KEY (`Características_Hogar_Id`)
151     REFERENCES `mydb`.`Características_Hogar` (`Id`)
152     ON DELETE NO ACTION
153     ON UPDATE NO ACTION)
154     ENGINE = InnoDB;

```

```

161 -- Table `Encuesta_hogares`.`Otros_Ingresos`
162 -----
163 CREATE TABLE IF NOT EXISTS `Encuesta_hogares`.`Otros_Ingresos` (
164   `Id` INT NOT NULL,
165   `Id_Hogar` INT(7) NULL,
166   `¿recibió pagos por concepto de arriendos, pensiones?` VARCHAR(2) NULL,
167   `¿Es propietario de una o varias propiedades inmuebles?` VARCHAR(2) NULL,
168   `Recibe ayudas del gobierno` VARCHAR(2) NULL,
169   `Hogar_Id` INT NOT NULL,
170   PRIMARY KEY (`Id`),
171   CONSTRAINT `fk_Otros_Ingresos_Hogar1`
172     FOREIGN KEY (`Hogar_Id`)
173     REFERENCES `mydb`.`Hogar` (`Id`)
174     ON DELETE NO ACTION
175     ON UPDATE NO ACTION)
176 ENGINE = InnoDB;
177
178
179 SET SQL_MODE=@OLD_SQL_MODE;
180 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;
181 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;

```

4.5 Código SQL - Manipulación de datos (DML)

```

SELECT hogar.Id_Hogar, integrantes_hogar.Orden, integrantes_hogar.Sexo_nacimiento, integrantes_hogar.Edad,
educación_integrantes_hogar.¿Sabe leer y escribir? FROM encuesta_hogares.hogar, encuesta_hogares.integrantes_hogar,
encuesta_hogares.educación_integrantes_hogar
where integrantes_hogar.Id_Hogar=hogar.Id_Hogar and hogar.Id_Hogar=educación_integrantes_hogar.Id_Hogar
and educación_integrantes_hogar.¿Sabe leer y escribir?='SI';

```

	Id_Hogar	Orden	Sexo_nacimiento	Edad	¿Sabe leer y escribir?
▶	2120004	2	Femenino	12	SI
	2120004	2	Femenino	12	SI
	2120004	2	Femenino	12	SI
	2120004	3	Femenino	9	SI
	2120004	3	Femenino	9	SI
	2120004	3	Femenino	9	SI
	2120004	4	Femenino	6	SI
	2120004	4	Femenino	6	SI
	2120004	4	Femenino	6	SI
	2120004	5	Masculino	5	SI
	2120004	5	Masculino	5	SI

Figure 9: Sentencia 1

```

20 • select Id_Hogar, sum(Último_Salario) from encuesta_hogares.ocupación
21 group by Id_Hogar ;
22
23
24
25

```

Id_Hogar	sum(Último_Salario)
2120055	900000
2120057	908526
2120058	600000
2120059	1300000
2120060	910000
2120064	910000
2120065	1020000

Result 30 x

Read Only

Figure 10: Sentencia 2

```

23 • select `La vivienda ocupada por este hogar es`, sum(`Total personas en el hogar`) from encuesta_hogares.caracteristicas_hogar
24 group by `La vivienda ocupada por este hogar es`;
25
26
27

```

La vivienda ocupada por este hogar es	sum(`Total personas en el hogar`)
En usufructo	24
Posecion sin titulo	458

Result Grid

Read Only

Figure 11: Sentencia 3

4.6 Código SQL + Resultados: Vistas

```
7 • CREATE VIEW ESCRITURA AS
8 SELECT count(*) FROM encuesta_hogares.hogar, encuesta_hogares.integrantes_hogar,
9 encuesta_hogares.educación_integrantes_hogar
10 where integrantes_hogar.Id_Hogar=hogar.Id_Hogar and hogar.Id_Hogar=educación_integrantes_hogar.Id_Hogar
11 and educación_integrantes_hogar.`¿Sabe leer y escribir?`='NO' and integrantes_hogar.Edad>5;
12
13 • SELECT * FROM ESCRITURA;
14
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [↗](#)

count(*)
18

Result Grid

Figure 12: Vista 1

4.7 Código SQL + Resultados: Triggers

```
1 • CREATE trigger Ingreso_nuevo_hogar3 BEFORE insert on hogar FOR EACH ROW INSERT INTO nuevos_hogares (Id, Id_Hogar, Fecha_actualizacion)
2 values (new.Id, new.Id_Hogar, NOW());
18 • INSERT INTO encuesta_hogares.hogar (Id, Id_Hogar) values (25000, 3456783);
15 • select * FROM encuesta_hogares.nuevos_hogares;
16
```

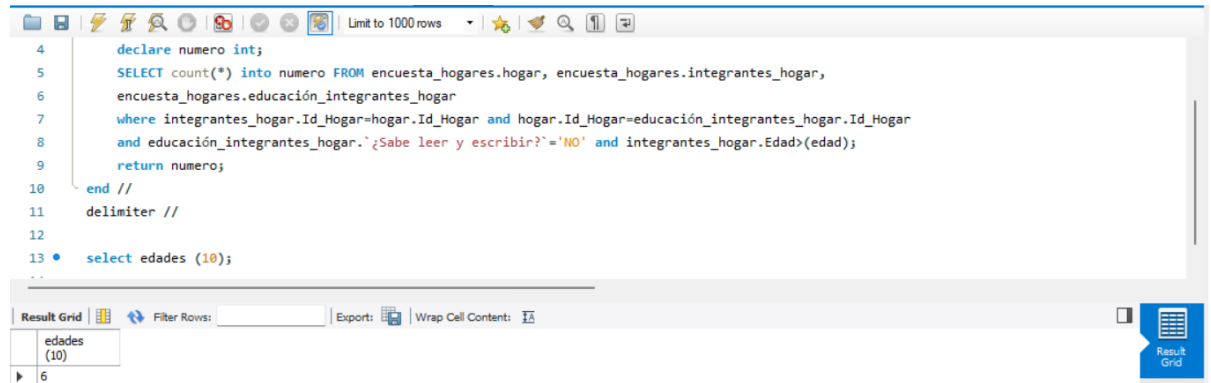
Result Grid | Filter Rows: | Edit: | Export/Imports: | Wrap Cell Content: [↗](#)

Id	Id_Hogar	Fecha_actualizacion
25000	3456783	2022-10-21

Result Grid

Figure 13: Trigger 1

4.8 Código SQL + Resultados: Funciones



The screenshot shows a SQL IDE with a code editor and a result grid. The code editor contains the following SQL code:

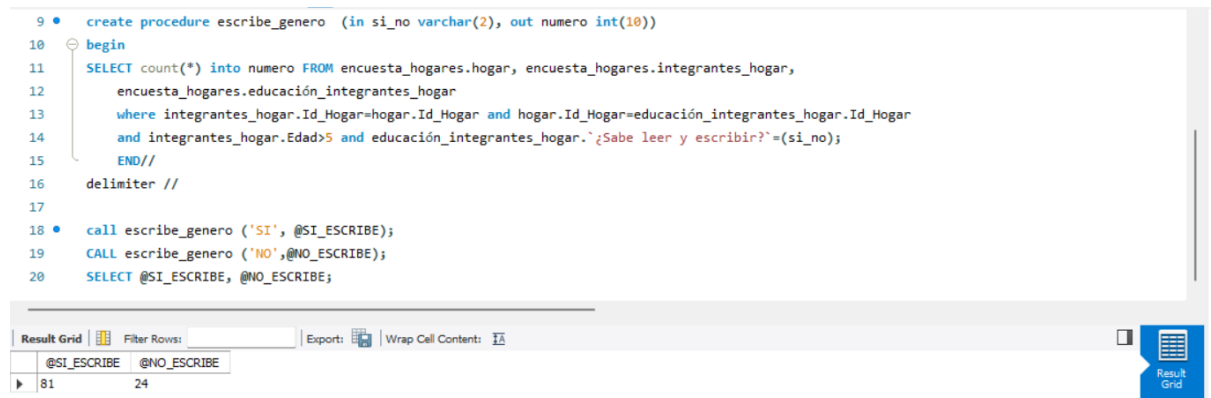
```
4 declare numero int;
5 SELECT count(*) into numero FROM encuesta_hogares.hogar, encuesta_hogares.integrantes_hogar,
6 encuesta_hogares.educación_integrantes_hogar
7 where integrantes_hogar.Id_Hogar=hogar.Id_Hogar and hogar.Id_Hogar=educación_integrantes_hogar.Id_Hogar
8 and educación_integrantes_hogar.`¿Sabe leer y escribir?`='NO' and integrantes_hogar.Edad>(edad);
9 return numero;
10 end //
11 delimiter //
12
13 • select edades (10);
14
```

The result grid shows the output of the function call:

edades (10)
6

Figure 14: Función 1

4.9 Código SQL + Resultados: procedimientos almacenados



The screenshot shows a SQL IDE with a code editor and a result grid. The code editor contains the following SQL code:

```
9 • create procedure escribe_genero (in si_no varchar(2), out numero int(10))
10 begin
11 SELECT count(*) into numero FROM encuesta_hogares.hogar, encuesta_hogares.integrantes_hogar,
12 encuesta_hogares.educación_integrantes_hogar
13 where integrantes_hogar.Id_Hogar=hogar.Id_Hogar and hogar.Id_Hogar=educación_integrantes_hogar.Id_Hogar
14 and integrantes_hogar.Edad>5 and educación_integrantes_hogar.`¿Sabe leer y escribir?`=(si_no);
15 END//
16 delimiter //
17
18 • call escribe_genero ('SI', @SI_ESCRIBE);
19 CALL escribe_genero ('NO',@NO_ESCRIBE);
20 SELECT @SI_ESCRIBE, @NO_ESCRIBE;
```

The result grid shows the output of the stored procedure call:

@SI_ESCRIBE	@NO_ESCRIBE
81	24

Figure 15: Procedimiento 1

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*)

References

LAZARO (2019) Cómo crear y usar una función almacenada MySQL Obtenido de <https://www.neoguias.com/funciones-almacenadas-mysql/>: :text=Las

SANCHEZ 2022 Triggers, procedimientos y funciones en MySQL Obtenido de <https://josejuansanchez.org/bd/unidad-12-teoria/index.htmlfunciones>

DANE (2021) Gran Encuesta Integrada de Hogares Obtenido de http://microdatos.dane.gov.co/index.php/catalog/MICRODATOS/about_collection/23/1

DANE Consideraciones para el uso de la información Obtenido de <https://microdatos.dane.gov.co/home>