

INFO133 - Base de Datos

Informe Parte B

Académico: Dr. Luis Veas {luis.veasc@inf.uach.cl}
Instituto de Informática, Universidad Austral de Chile
Alumnos: - Felipe Córdova {felipe.cordova@alumnos.uach.cl}
- Francisco Labrín {francisco.labrin@alumnos.uach.cl}
- Sebastián Montecinos {sebastian.montecinos02@alumnos.uach.cl}

Junio 14, 2023

Índice

1	Variables de Estudio	2
2	Diagrama Entidad-Relación	2
3	Modelo Relacional	3
4	Diccionario de Datos	4
5	Proceso de Descarga de Datos	5

1. Variables de Estudio

Se han seleccionado como variables de estudio: Salud, Educación, Seguridad y Entretenimiento.

Para una mejor visualización, las figuras mostradas en este informe se encuentran en el repositorio, específicamente en la ruta “Parte B/Imágenes”.

2. Diagrama Entidad-Relación

Realizando un estudio exhaustivo de los diversos enlaces proporcionados y explorados por nosotros, podemos construir el siguiente diagrama entidad-relación correspondiente al caso de estudio seleccionado [fig.1].

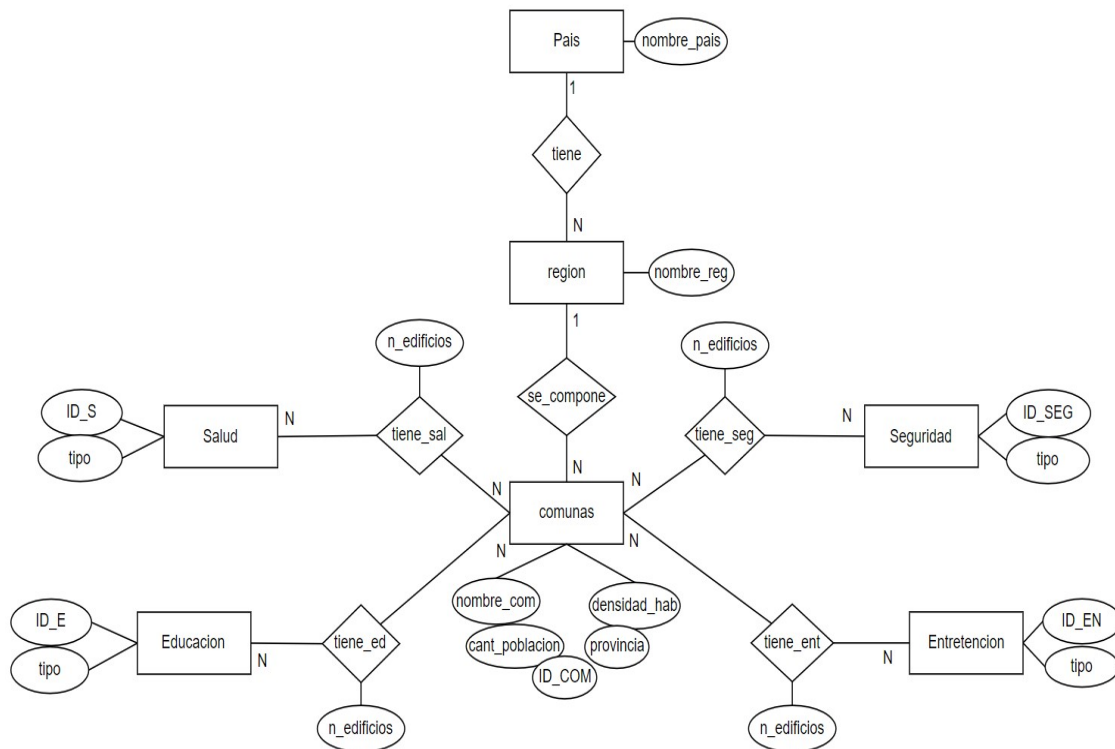


Figura 1: Diagrama Entidad-Relación

3. Modelo Relacional

Una vez completado el diagrama entidad-relación, procedemos a llevar a cabo la conversión correspondiente al modelo relacional, obteniendo la siguiente representación [fig.2].

Pais (**PK(nombre_pais)**)

region (**PK(nombre_reg)**, *FK(nombre_pais)*)

comuna (**PK(ID_COM)**, nombre_com, cant_poblacion, densidad_hab, provincia, *FK(nombre_reg)*)

salud (**PK(ID_S)**, tipo)

educacion (**PK(ID_E)**, tipo)

seguridad (**PK(ID_SEG)**, tipo)

entretencion (**PK(ID_EN)**, tipo)

tiene_sal (**PK(*FK(ID_S)*, *FK(ID_COM)*)**, n_edificios)

tiene_ed (**PK(*FK(ID_E)*, *FK(ID_COM)*)**, n_edificios)

tiene_seg (**PK(*FK(ID_SEG)*, *FK(ID_COM)*)**, n_edificios)

tiene_ent (**PK(*FK(ID_EN)*, *FK(ID_COM)*)**, n_edificios)

Figura 2: Modelo Relacional

4. Diccionario de Datos

A continuación, procedemos a completar el diccionario de datos [Figura 3], lo cual nos proporciona una comprensión más precisa de los datos que se obtendrán y que se insertarán posteriormente en la Base de Datos.

Entidad / Relación	Nombre Atributo	Descripción	Tipo de Dato	Largo	Valores aceptados	¿Obligatorio?	¿Acepta valores nulos?
Pais	nombre_pais	Nombre del país	VARCHAR	50	Chile, Peru, Bolivia, etc.	si	no
Region	nombre_reg	Nombre de la región	VARCHAR	50	Los Rios, Los Lagos, etc.	si	no
Comuna	ID_COM	Codigo de la comuna. El codigo de cada comuna es diferente	VARCHAR	50	15101, 15102, etc. Fuente: Wikipedia.	si	no
	nombre_com	Nombre de la comuna	VARCHAR	50	Valdivia, Lanco, Panguipulli, etc.	si	no
	cant_poblacion	Número de habitantes	INT		Números enteros positivos.	no	no
	densidad_hab	Habitantes por km²	FLOAT		Números flotantes positivos.	no	no
	provincia	Nombre de la provincia a la cual pertenece dicha comuna	VARCHAR	50	Valdivia, Ranco, Osorno, etc	si	no
Salud	ID_S	Identificador único de Salud	VARCHAR	50	S01, S02, S03, etc.	si	no
	tipo	Tipo de sistema de salud	VARCHAR	50	Hospital, Clinica, Farmacia, etc.	si	no
Educación	ID_E	Identificador único de Educación	VARCHAR	50	E01, E02, E03, etc.	si	no
	tipo	Tipo de sistema de Educacion	VARCHAR	50	Colegio, Universidad, etc.	si	no
Seguridad	ID_SEG	Identificador único de Seguridad	VARCHAR	50	SEG01, SEG02, SEG03, etc.	si	no
	tipo	Tipo de sistema de Seguridad	VARCHAR	50	Carabineros, PDI, Seguridad ciudadana, etc.	si	no
Entretencion	ID_EN	Identificador único de Entretención	VARCHAR	50	EN01, EN02, EN03, etc.	si	no
	tipo	Tipo de Entretencion	VARCHAR	50	Cine, Teatro, Estadio, etc.	si	no
tiene_sal tiene_ent tiene_seg tiene_ed	n_edificios	Cantidad de edificios	INT		Números enteros positivos (+0).	si	no

Figura 3: Diccionario de Datos

5. Proceso de Descarga de Datos

Se comienza por indagar en los enlaces proporcionados, buscando la información necesaria para el análisis. Posteriormente, se crea un archivo en formato CSV para almacenar los datos que serán extraídos. Considerando la diversidad de formatos en los que se encuentran disponibles los datos, se optó por hacer un proceso de extracción manual. Además, hemos agregado nuevos enlaces debido a las necesidades requeridas a partir del modelo obtenido anteriormente, para luego extraer la información en cada variable de estudio.

(a) Para la variable “Comunas” navegamos por [Comunas de Chile](#). Extraemos los datos para cada comuna según el diccionario de datos proporcionado. Este proceso se repetirá para todas las comunas

(b) Para la variable “Salud” navegamos por [Establecimientos de Salud](#). A continuación, clasificamos los establecimientos de salud según su tipo, siguiendo el diccionario de datos proporcionado. Para cada comuna, registraremos la cantidad de edificios de salud correspondiente a cada tipo. Este proceso se repetirá para cada tipo de establecimiento y para todas las comunas.

(c) Para la variable “Educación” navegamos por [Establecimientos de parvularia, Básica y Media](#) y [Establecimientos de Educación Superior](#). A continuación, clasificaremos los establecimientos educacionales según su tipo, siguiendo el diccionario de datos proporcionado. Para cada comuna, registraremos la cantidad de edificios de educación correspondiente a cada tipo. Este proceso se repetirá para cada tipo de establecimiento y para todas las comunas.

(d) Para la variable “Seguridad” navegamos por [Establecimientos de Carabineros](#) y [Establecimientos de PDI](#). A continuación, clasificamos los establecimientos de seguridad según su tipo, siguiendo el diccionario de datos proporcionado. Para cada comuna, registraremos la cantidad de edificios de seguridad correspondiente a cada tipo. Este proceso se repetirá para cada tipo de institución y para todas las comunas.

(e) Para la variable “Entretenimiento” navegamos por [Estadios de Chile](#), [Teatros en Chile](#) y [Cines de Chile](#). A continuación, clasificaremos los establecimientos de entretenimiento según su tipo, siguiendo el diccionario de datos proporcionado, registraremos la cantidad de edificios de entretenimiento correspondiente a cada tipo. Este proceso se repetirá para cada tipo de entretenimiento y para todas las comunas.