

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 para las Variables	5
5.1	Comunas	5
5.2	Salud	5
5.3	Educación	5
5.4	Seguridad	6
5.5	Entretención	6

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].

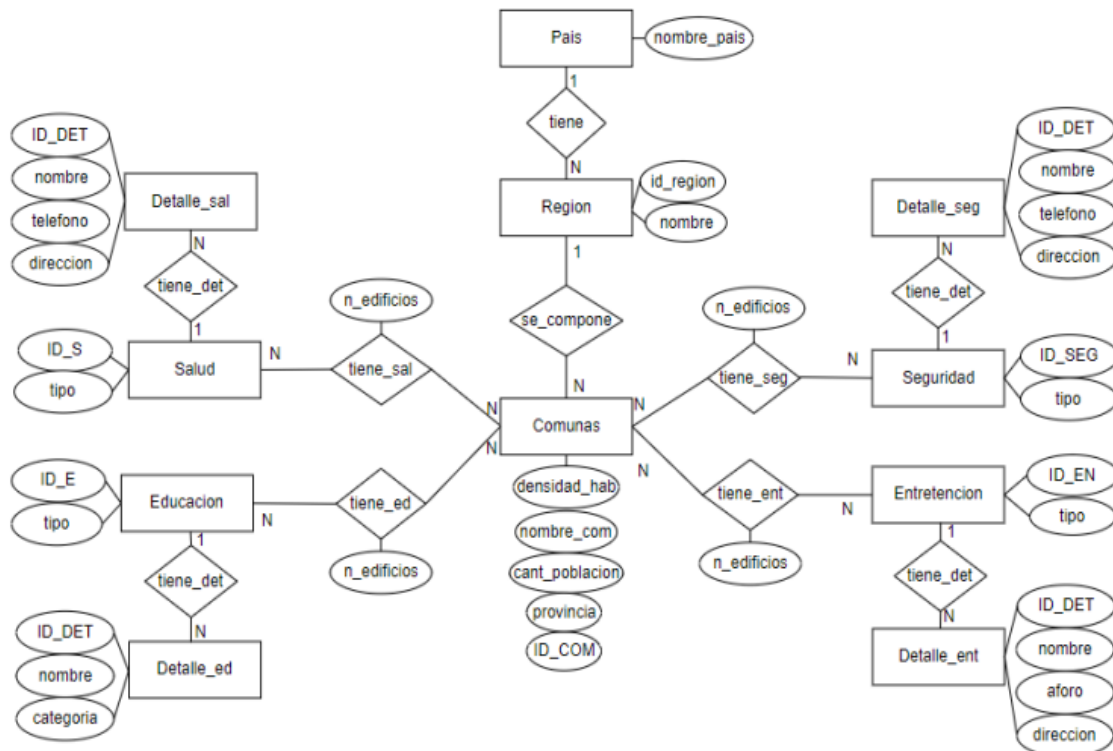


Figura 1: Diagrama Entidad-Relación

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(id_region), nombre, FK(nombre_pais) )  
Comuna ( PK(ID_COM), nombre_com, cant_poblacion, densidad_hab, provincia, FK(id_region) )  
Salud ( PK(ID_S), tipo )  
Detalle_sal( PK(ID_DET), nombre, telefono, direccion, FK(ID_S) )  
Educacion ( PK(ID_ED), tipo )  
Detalle_ed( PK(ID_DET), nombre, categoria, FK(ID_ED) )  
Seguridad ( PK(ID_SEG), tipo )  
Detalle_seg( PK(ID_DET), nombre, telefono, direccion, FK(ID_SEG) )  
Entretencion ( PK(ID_EN), tipo )  
Detalle_ent(PK(ID_DET), nombre, aforo, direccion, FK(ID_EN) )  
tiene_sal ( PK( FK(ID_S), FK(ID_COM) ) ), n_edificios )  
tiene_ed ( PK( FK(ID_ED), 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	PK	FK (referencia)	Descripción	Tipo de Dato	Largo	Valores aceptados	¿Obligatorio?	¿Acepta valores nulos?
Pais	nombre_pais	X		Nombre del país	VARCHAR	50	Chile, Peru, Bolivia, etc.	si	no
	id_region	X		Identificador único para cada región	VARCHAR	50	REG01, REG02, REG03, etc.	si	no
Region	nombre			Nombre de la región	VARCHAR	50	Los Rios, Los Lagos, etc.	si	no
	nombre_pais		X (Pais)					si	no
Comuna	ID_COM	X		Código de la comuna. El código 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
	nombre_reg		X (Region)					si	no
Salud	ID_S	X		Identificador único de Salud	VARCHAR	50	S01, S02, S03, etc.	si	no
	tipo			Tipo de sistema de salud	VARCHAR	50	Hospital, Clínica, Consultorio, etc.	si	no
detalle_sal	ID_DET	X		Identificador único de detalle de Salud	VARCHAR	50	DS01, DS02, DS03, etc.	si	no
	nombre			Nombre del recinto de Salud	VARCHAR	50	Clínica alemana, Hospital Base Valdivia, etc.	si	no
	telefono			Número de teléfono del recinto de Salud	VARCHAR	15	Números de teléfonos válidos en Chile.	no	si
	direccion			Dirección del recinto de Salud	VARCHAR	50	Direcciones reales de cierta comuna.	si	no
	ID_S		X (Salud)					si	no
Educacion	ID_ED	X		Identificador único de Educación	VARCHAR	50	E01, E02, E03, etc.	si	no
	tipo			Tipo de sistema de Educación	VARCHAR	50	Colegio, Universidad, etc.	si	no
detalle_ed	ID_DET	X		Identificador único de detalle de Educación	VARCHAR	50		si	no
	nombre			Nombre del recinto	VARCHAR	15	UACH, Windsor, Salesiano, etc.	si	no
	categoria			Para ver si el recinto es público o privado	VARCHAR	50	Público, privado.	si	no
	ID_ED		X (Educacion)					si	no
Seguridad	ID_SEG	X		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
detalle_seg	ID_DET	X		Identificador único de detalle de Seguridad	VARCHAR	50	DSEG01, DSEG02, DSEG03, etc.	si	no
	nombre			Nombre del recinto	VARCHAR	50	Subcomisaría, Primera Comisaría, etc.	si	no
	telefono			Número de teléfono del recinto	VARCHAR	15	Números de teléfonos válidos en Chile.	si	no
	direccion			Dirección del recinto	VARCHAR	50	Direcciones reales de cierta comuna.	si	no
	ID_SEG		X (Seguridad)					si	no
Entretencion	ID_EN	X		Identificador único de Entretención	VARCHAR	50	EN01, EN02, EN03, etc.	si	no
	tipo			Tipo de Entretención	VARCHAR	50	Cine, Teatro, Estadio, etc.	si	no
detalle_ent	ID_DET	X		Identificador único de detalle de Entretención	VARCHAR	50	DENT01, DENT02, DENT03, etc.	si	no
	nombre			Nombre del recinto	VARCHAR	50	Cineplanet, Teatro municipal, etc.	si	no
	aforo			Capacidad de personas del recinto	INT		Números enteros positivos.	no	si
	direccion			Dirección del recinto	VARCHAR	50	Direcciones reales de cierta comuna.	si	no
	ID_EN		X (Entretencion)					si	no
tiene_sal	n_edificios			Cantidad de edificios	INT		Números enteros positivos (+0).	si	no
	ID_S	X	X (Salud)					si	no
tiene_ed	ID_COM	X	X (Comuna)					si	no
	n_edificios			Cantidad de edificios	INT		Números enteros positivos (+0).	si	no
tiene_seg	ID_ED	X	X (Educacion)					si	no
	ID_COM	X	X (Comuna)					si	no
tiene_seg	n_edificios			Cantidad de edificios	INT		Números enteros positivos (+0).	si	no
	ID_SEG	X	X (Seguridad)					si	no
tiene_ent	ID_COM	X	X (Comuna)					si	no
	n_edificios			Cantidad de edificios	INT		Números enteros positivos (+0).	si	no
tiene_ent	ID_EN	X	X (Entretencion)					si	no
	ID_COM	X	X (Comuna)					si	no

Figura 3: Diccionario de Datos

5. Proceso de Descarga de Datos para las Variables

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.

5.1. Comunas

Para la variable “Comunas” nos aseguramos de utilizar una fuente confiable proporcionado en el documento word con las instrucciones de este trabajo. A continuación, extrajimos información importante como el nombre de la comuna, nombre, cantidad de población, densidad habitacional, provincia y un código único territorial. La descripción detallada de los datos extraídos se encuentra disponible en el diccionario de datos proporcionado. Este proceso se repetirá para todas las comunas.

Enlace:

- $L_0 = \text{https://es.wikipedia.org/wiki/Anexo:Comunas_de_Chile}$

5.2. Salud

Para la variable “Salud” nos aseguramos de utilizar una fuente oficial del gobierno. Navegamos por el sitio para acceder a los datos relevantes. A continuación, extrajimos información importante como el tipo de establecimiento de salud, nombre, número de teléfono y la dirección. La descripción detallada de los datos extraídos se encuentra disponible en 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 de salud y para todas las comunas.

Enlace:

- $L_1 = \text{https://reportesdeis.minsal.cl/ListaEstablecimientoWebSite/}$

5.3. Educación

Para la variable “Educación” nos aseguramos de utilizar una fuente confiable del gobierno y otras fuentes validada por nuestro criterio. Navegamos por los sitios para acceder a los datos relevantes para nuestro estudio. A continuación, extrajimos información importante como el tipo de establecimiento de educación, nombre y categoría. La descripción detallada de los datos extraídos se encuentra disponible en el diccionario de datos proporcionado.

Enlaces:

- $L_2 = \text{https://www.bcn.cl/siit/estadisticasterritoriales/tema?id=52}$
- $L_3 = \text{https://www.altillo.com/universidades/universidades_chile.asp}$
- $L_4 = \text{https://es.wikipedia.org/wiki/Anexo:Universidades_de_Chile}$

Para nuestra Base de Datos, utilizaremos la información extraída de L_2 y L_4 , debido a su estructura y formato, lo cual facilitó la extracción de datos para su posterior almacenamiento en un archivo CSV. 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.

5.4. Seguridad

Para la variable “Seguridad” nos aseguramos de utilizar una fuente confiable otras fuentes validadas por nuestro criterio. Navegamos por los sitios para acceder a los datos relevantes para nuestro estudio. A continuación, extrajimos información importante como el tipo de institución, nombre del cuartel, teléfono y dirección. La descripción detallada de los datos extraídos se encuentra disponible en 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 seguridad y para todas las comunas.

- $L_5 = \text{https://es.wikipedia.org/wiki/Anexo:Comisarías_de_Chile}$
- $L_6 = \text{https://es.wikipedia.org/wiki/Anexo:Cuarteles_de_la_PDI}$

5.5. Entretención

Para la variable “Entretención” nos aseguramos de navegar por link confiables, como por ejemplo el de la ANFP. A continuación, extrajimos información importante como el tipo de entretención, nombre, aforo y dirección. La descripción detallada de los datos extraídos se encuentra disponible en el diccionario de datos proporcionado. Para cada comuna, registraremos la cantidad de edificios de entretención correspondiente a cada tipo. Este proceso se repetirá para cada tipo de establecimiento de salud y para todas las comunas.

- $L_7 = \text{https://es.wikipedia.org/wiki/Anexo:Estadios_de_futbol_de_Chile}$
- $L_8 = \text{https://es.wikipedia.org/wiki/Categoría:Teatros_de_Chile_por_región}$
- $L_9 = \text{https://es.wikipedia.org/wiki/Cinépolis_Chile}$
- $L_{10} = \text{https://www.anfp.cl/}$