



# **Informe de Integración de SGBD en la construcción de software.**

Integrantes:

-Miguel Ormeño.

-Diego Vidal.

Grupo: 19

Profesores: Matthieu Vernier, Luis Veas.

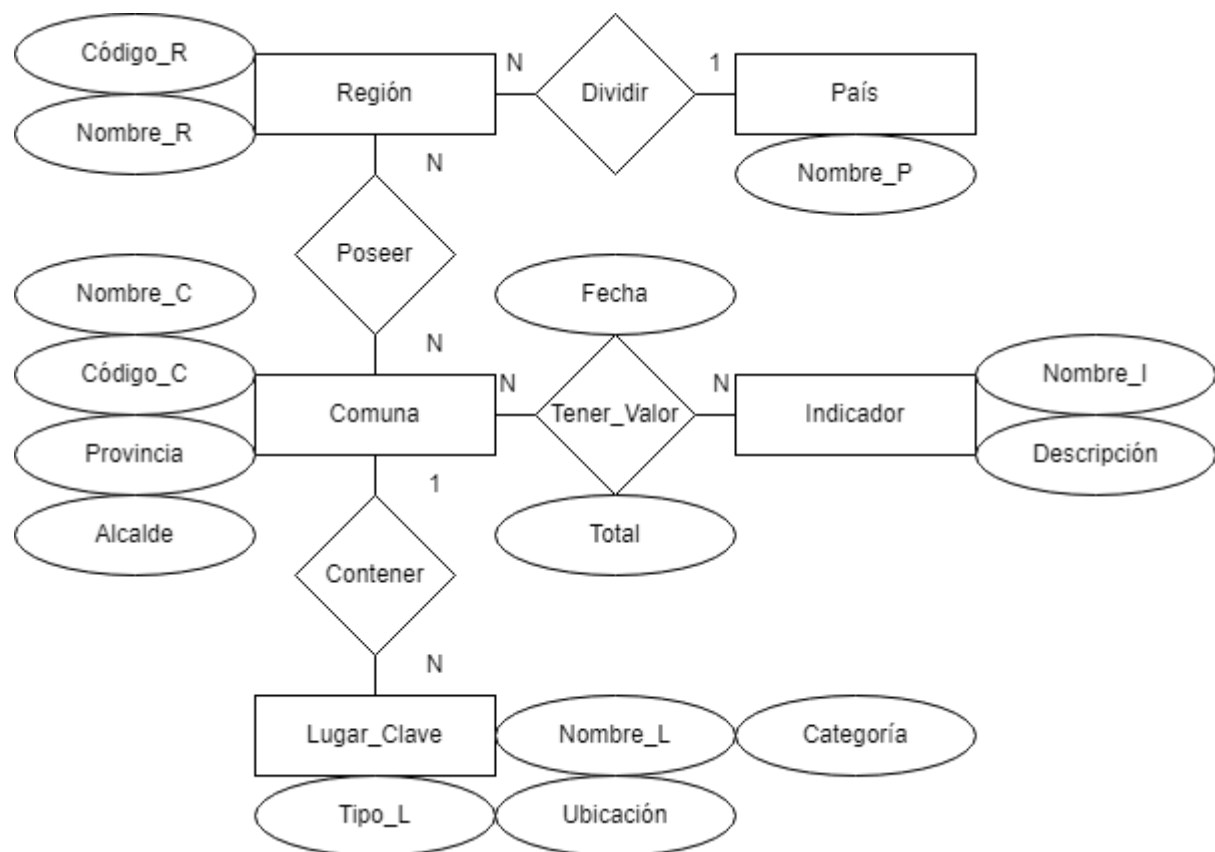
Asignatura: Base de Datos (INFO133).

Item	Descripción	puntuación
1	diagrama modelo entidad-relación	10
2	diagrama modelo relacional	10
3	diccionario de datos	10
4	script de base de datos	10
5	seleccionar las variables a estudiar (a lo menos 4)	5
6	documentar el proceso de descarga (puede ser uno o varios programas o un proceso manual)	15
7	programa de carga de datos	10
8	programa que entregue el resultado de las 14 sentencias sql	40
9	explicar su modelo de indicador de bienestar	10
10	creación/poblar/mostrar de indicador de bienestar	20

Repositorio de GitHub, conteniendo todos los archivos pertinentes de este trabajo:

<https://github.com/azurite-12/INFO133-equipo19>

1.



Un país de cierto nombre se divide en varias comunas, cada una con nombre y numeración. Estas regiones poseen comunas con nombre, código, provincia y un alcalde. Las comunas pueden tener valores ya medidos como superficie y personas trabajando en cierta fecha, los cuales se describen; pero también existen datos como la cantidad de antenas o estadios que no vienen medidas, estos los llamaremos lugares clave. Un lugar clave tiene nombre y ubicación, además de tipo de lugar y categoría (por ejemplo, tipo: farmacia y categoría: salud).

## 2. Modelo relacional.

País (PK\_nombreP);

Región (PK\_codigoR, FK\_nombreP, nombreR);

Comuna (PK\_codigoC, nombreC, provincia, alcalde);

Lugar\_Clave (PK\_ubicacion, FK\_codigoC, nombreL, tipoL, categoria);

Indicador (PK\_nombreI, descripcion);

Poseer (PK(FK\_codigoR, FK\_codigoC));

Tener\_Valor (PK(FK\_codigoC, FK\_nombreI), fecha, total);

## 3.

	Nombre	Descripción	Tipo de Datos	Tamaño Máx.	Ejemplo	¿Obligatorio?
País	Nombre_P	Nombre del País	VarChar	30	Chile	Sí
Región	Código_R	Número de la Región	TinyInt	127	9	Sí
	Nombre_R	Nombre de la Región	VarChar	50	Región de Los Ríos	Sí
Comuna	Código_C	Código de la Comuna	SmallInt	32767	15101	Sí
	Nombre_C	Nombre de la Comuna	VarChar	30	La Serena	Sí
	Provincia	Nombre de la Provincia a la cual Pertenecer la Comuna	VarChar	30	Parinacota	Sí
	Alcalde	Nombre del Alcalde de la Comuna	VarChar	100	Franco Jiménez	Sí
Lugar_Clave	Nombre_L	Nombre del Lugar	VarChar	100	Hospital Regional	Sí
	Categoría	Propósito del Lugar	VarChar	30	Entretenimiento	Sí
	Tipo	Clase de Lugar	VarChar	20	Farmacia	Sí
	Ubicación	Dirección del Lugar	VarChar	100	Arroyo Carmesí #999	Sí
Indicador	Nombre_I	Nombre del Indicador	VarChar	20	IDH, Calidad_Vida, Seguridad, Superficie, Habitantes, Transporte o Laboral	Sí
	Descripción	Descripción del Indicador	VarChar	100	Índice de Calidad de Vida Urbana de la Comuna	Sí
Tener_Valor	Valor	Valor de lo Indicado	Float	3,40E+47	387164,102	Sí
	Fecha	Año en que se obtuvo el dato	SmallInt	32767	2021	No

## 4. Ver carpeta sql del GitHub

## 5. Nuestras variables que decidimos estudiar son las siguientes:

- IDH.
- Laboral.
- Seguridad.
- Habitantes.

#### 6. Documentación del proceso de obtención de datos:

- Para el IDH, extraemos información de [Anexo:Comunas de Chile - Wikipedia, la enciclopedia libre](#) del 2005, para eso creamos un script en python usando los principios básicos del *web scraping*, creando un documento csv en el proceso que posteriormente usaremos en nuestro análisis.
- Para seguridad, vamos a [Plataforma de datos comunales](#) y descargamos un archivo .xlsx de la sección de Ranking Comunal, luego usamos el módulo pandas de python para transformar el archivo a .csv, extraer la información de las columnas que nos interesan y analizarla.
- Para las variables de tipo laboral y habitantes, ingresamos al sitio <https://www.observatorionacional.cl/panorama-laboral/datos-comunales> y se descargaron manualmente por cada región los archivos .csv con toda la información, luego con un script de python y pandas se pueden extraer los datos pertinentes para usarse en el proyecto en cuestión.

#### 8. Sentencias SQL a utilizar:

- 1- Región más segura.
- 2- Región menos segura.
- 3- Región con mayor IDH.
- 4- Región con menor IDH.
- 5- lista de personas con edad de trabajar en la comuna de Valdivia, ordenados por región, comuna, ambos ordenados en orden ascendente.
- 6- Comuna con más desempleo.
- 7- Lista de regiones con más desempleo, mostrando el nombre de la comuna y la cantidad de desempleados.
- 8- Comuna más poblada.
- 9- Comuna con mayor igualdad de género en el ámbito laboral.
- 10- Comuna con menor igualdad de género en el ámbito laboral.
- 11- Comuna con mayor IDH en cada región.
- 12- Lista de comunas ordenadas según su IDH en forma descendente.
- 13- Comuna menos segura por región.
- 14- Lista de comunas según índice de seguridad, indicando la región a la que pertenece.

9. Nuestro modelo de indicador de bienestar consiste en tomar ciertos factores que demostrablemente aumentan la calidad de vida de las personas. El índice de desarrollo humano, IDH, mide los niveles de desarrollo social y económico de una comuna dada. Mientras que los índices de seguridad reportan lo expuestas (o no expuestas) que las personas están al peligro y al crimen, como a mayor tasa hay más crimen, se usa con proporción inversa. Por otra parte, las métricas laborales miden la accesibilidad que las personas viviendo en dichas comunas tienen a trabajos estables y abundantes, lo que las provee de un consistente sueldo para llevar la vida diaria sin problemas.

Aquí la representación matemática de nuestro modelo ya descrito:

$$\text{Bienestar} = \text{IDH} * \left( \frac{\text{Personas Ocupadas}}{\text{Población en Edad de Trabajar}} + \frac{100}{\text{Tasa de Crímenes}} \right)$$