



Universidad Austral de Chile

Conocimiento y Naturaleza

Información territorial basada en las comunas de Chile

INFO133 - Base de Datos

Integrantes:

- *Carmen Fuenzalida*
- *Andrés Torres*
- *Rodrigo Vergara*

Profesor: Luis Veas

Diagrama modelo entidad-relación

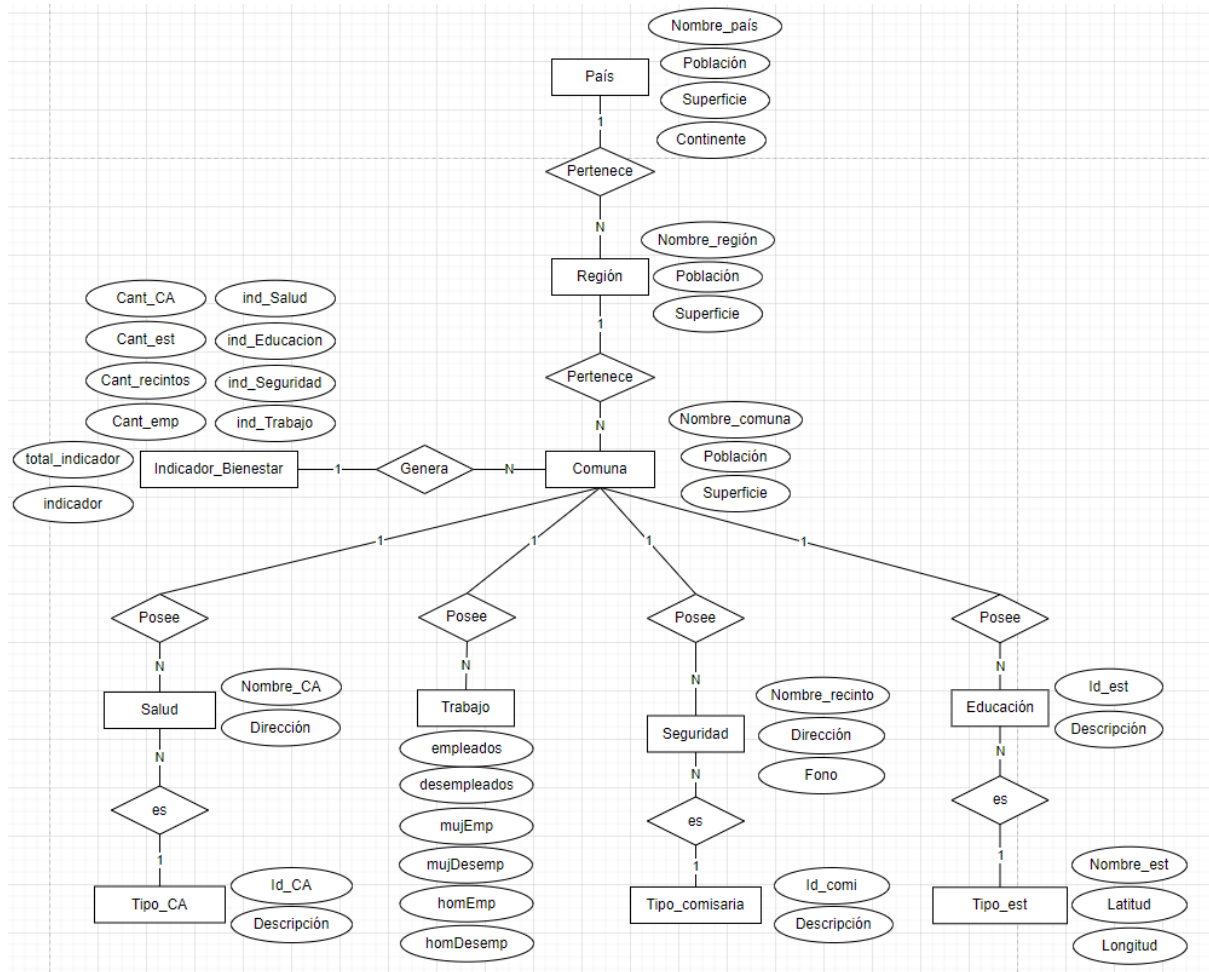


Diagrama modelo relacional

País(Id_pais, Nombre_país, Población, Superficie, Continente, Fk_Id_region)

Región(Id_región, Nombre_región, Población, Superficie, Fk_Id_país)

Comuna(Id comuna, Nombre comuna,Superficie, Población, Fk Id región)

Educación(Id_educación, Nombre_est, Latitud, Longitud, Fk_Id_est, Fk_Id_comuna)

Salud(Id_salud, Nombre_CA, Dirección, Fk_Id_CA, Fk_Id_comuna)

Trabajo(Id_trabajador, Empleados, Desempleados, mujEmp, homEmp, mujDesemp, homDesemp, Fk Id comuna)

Seguridad(Id_recinto, Nombre_recinto, Dirección, Fono, Fk_Id_comi, Fk_Id_comuna)

Tipo_est(Id_est, Descripción, Fk_Id_educación)

Tipo CA(Id CA, Descripción, Fk Id salud)

Tipo comisaría(Id comi, Descripción, Fk Id comisaría)

Indicador_Bienestar(Id_Indicador, Cant_CA, Cant_est, Cant_recintos, Cant_emp, ind_Salud, ind_Educacion, ind_Seguridad, ind_Trabajo, total_indicador, indicador, Id_Comuna)

Diccionario de datos

El diccionario de datos en archivo .xlsx se encuentra adjunto en la carpeta general.

Seleccionar las variables a estudiar

Para nuestra base de datos hemos decidido estudiar las siguientes variables:

- *Educación*
- *Salud*
- *Trabajo*
- *Seguridad*

Documentar el proceso de descarga (puede ser uno o varios programas o un proceso manual)

Para nuestro proceso de descarga de datos tuvimos acceso a diferentes páginas web confiables de donde obtuvimos la información necesaria para nuestra base de datos. Creamos tablas en Excel donde organizamos estos datos que fueron extraídos manualmente.

Para la tabla Trabajo obtuvimos los datos manualmente del observatorio laboral del Servicio Nacional de Capacitación y Empleo (SENCE) que ocupa cifras directamente del último censo realizado por el Instituto Nacional de Estadísticas (INE). Para poder obtener los datos de esta tabla llenamos un archivo .csv con las columnas id_region, Mujeres empleadas, Mujeres desempleadas, Hombres empleados, Hombres desempleados, Total empleados y Total desempleados, todos los datos extraídos de la fuente mencionada anteriormente extraídos de <https://www.observatorionacional.cl/>.

Para la creación de la tabla Seguridad tuvimos ciertos problemas ya que no existía otra página confiable más que la página oficial de Carabineros de Chile: <https://www.carabineros.cl/detalleUnidad.php>. Para estos datos no existía un listado específico sino que se debía seleccionar: región, comuna, y hacer click uno a uno sobre los diferentes puntos que habían en el mapa para obtener la información de cada una de las diferentes unidades de carabineros. Todo esto con un proceso manual, de donde logramos extraer 711 datos de unidades incluyendo nombre del recinto, dirección, fono y el tipo de recinto.

Para la tabla Salud obtuvimos los datos de la página del <https://deis.minsal.cl/> (DEIS) subdivisión del Ministerio de Salud (MINSAL) en este la extracción de los datos fue un poco más simple ya que la tabla de esta página se podía exportar directamente a excel y lo que hicimos fue cambiar el formato a csv y hacer un barrido de los caracteres especiales que podrían entrar en conflicto con la base de datos. De aquí pudimos extraer información sobre el nombre del centro de salud, dirección y si se trata de un centro público o privado.

Para la tabla Educación se descargó una base de datos directamente desde la página <https://datosabiertos.mineduc.cl> del ministerio de educación , para poder obtener los datos

de esta tabla se tuvo que filtrar los datos que no necesitamos de la base de datos y crear el archivo .csv para posteriormente subirlo a la base de datos con el script. Incluimos el nombre del establecimiento y su ubicación con dos parámetros: latitud y longitud.

Posteriormente convertimos los archivos que .xlsx (excel) a .csv (comma separated value) para poder ser manipulados en Python. En el archivo enviado están adjuntos los archivos de Excel y los CSV de cada tabla, con el fin de mantener todo ordenado y claro. Por otra parte el script de creación de la base de datos y de las tablas está en archivo .sql el cual fue creado con algunos problemas al inicio pero finalmente se logró la correcta incorporación al código.

Indicador de bienestar

Para cada una de las variables buscamos la cantidad mínima y máxima mediante consultas sql a nuestra base de datos. En la siguiente tabla se reflejan los intervalos establecidos: malo, medio y bueno. Siendo malo un bajo nivel de bienestar y bueno un alto nivel de bienestar.

En el proceso de creación del indicador de bienestar en nuestra base de datos, decidimos asignar números a los intervalos con el fin de simplificar y agilizar el manejo de los datos.

Malo = 1, Medio = 2, Bueno = 3

Variables	Malo (1)	Medio (2)	Bueno (3)
Trabajo	207 - 89.413	89.414 - 178.619	178.620 - 267.827
Seguridad	0 - 5	6 - 11	12 - 16
Salud	0 - 26	27 - 53	54 - 80
Educación	1 - 104	105 - 209	210 - 313

Preguntas SQL:

El siguiente listado de preguntas fue el modelo base para la creación del programa .py que muestra las consultas SQL en el terminal. Como fueron nuestra guía para elaborar el programa decidimos dejarlas adjuntas en el informe.

Las consultas finales se encuentran disponibles dentro del programa principal por lo que al ejecutarlo se verán reflejadas las 'últimas versiones' de las consultas.

1. ¿Cuántas comunas pertenecen a cada región?
2. ¿Cuál es la región con la mayor población?
3. ¿Cuáles son las comunas cuya población es mayor a 100,000 habitantes?
4. ¿Cuál es el promedio de superficie de las comunas en cada región?
5. ¿Cuál es la población total en cada región?
6. ¿Cuál es la región con la mayor superficie total de comunas?
7. ¿Cuál es el establecimiento de educación con la menor longitud de coordenadas (latitud + longitud)?

8. ¿Cuál es la región con la menor población total en sus comunas?
9. ¿Cuál es el promedio de población en el sector de trabajo en todas las comunas?
10. Ordena las regiones de Chile de mayor a menor población
11. ¿Cuál es la comuna con el mayor número de desempleados?
12. Lista todos los establecimientos educacionales de la comuna de *
13. ¿Cuál es la comuna con menor cantidad de Centros de Salud?
14. ¿Cuál es la cantidad de recintos de tipo comisaría por región?*

(Sea * una variable a elección de quien realiza la consulta)