CIEE Análisis

Renato Vargas

## Introducción

En este documento se detalla los pasos utilizados para el análisis.

## Hallazgos

El primer paso consiste en cargar las librerías necesarias para el análisis.

library(RSQLite)  
library(DBI)

Nos conectamos a la base de datos a través del paquete DBI y obtenemos objetos de datos usando SQL. Idealmente, la base de datos cuenta con la documentación adecuada, pero en una primera instancia se explora la base de datos para saber con qué tablas y variables cuenta. Utilizamos la función dbConnect() habilitada por el paquete DBI para abrir una conexión con la base de datos SQLite del Sistema de Cuentas Nacionales. Una vez concluido el análisis, se utiliza dbDisconnect() para cerrar la conexión como puede verse al final de esta memoria de análisis.

con <- dbConnect(RSQLite::SQLite(), "C:/Users/renato/GitHub/scn\_scae\_gt/datos/scn.db")  
dbListTables(con)

[1] "areas\_columnas" "areas\_filas" "columnas"   
 [4] "cuadros" "energia" "filas"   
 [7] "naeg" "npg" "ntg2"   
[10] "oferta\_utilizacion"

La mayoría de tablas en la base datos se utilizan para explicar los valores de los campos de las tablas principales oferta\_utilizacion y balances\_energeticos. Por esa razón, es importante saber con qué campos cuentan estas dos tablas.

dbListFields(con, "oferta\_utilizacion")

[1] "anio" "id\_precios" "id\_cuadro" "id\_fila" "id\_columna"  
[6] "valor" "id\_unidad"

Puesto que esta base de datos contiene los cuadros de oferta y utilización de los años 2013 a 2020 y esta cuenta con cientos de miles de registros, se trató de dejar la tabla principal con la menor cantidad de columnas y dentro de estas, con la menor cantidad de valores, con el objeto de reducir el tamaño. Se puede inspeccionar las primeras filas de la tabla oferta\_utilizacion. Hay varias formas de hacer esto, tanto en R, como en SQL. Para mantener la consistencia con las consultas de SQL, aquí se utilizará este lenguaje.

dbGetQuery(con ,"SELECT \* FROM oferta\_utilizacion LIMIT 5")

anio id\_precios id\_cuadro id\_fila id\_columna valor id\_unidad  
1 2013 1 1 f001 oc001 5.346939 1  
2 2013 1 1 f002 oc001 4734.946097 1  
3 2013 1 1 f003 oc001 101.132353 1  
4 2013 1 1 f004 oc001 0.000000 1  
5 2013 1 1 f005 oc001 0.000000 1

dbGetQuery(con, "  
SELECT  
 anio,  
 cuadro,  
 sum(valor)  
FROM   
 oferta\_utilizacion  
JOIN  
 cuadros  
ON  
 oferta\_utilizacion.id\_cuadro  
 =  
 cuadros.id\_cuadro  
GROUP BY   
 anio, oferta\_utilizacion.id\_cuadro  
ORDER BY  
 anio, oferta\_utilizacion.id\_cuadro  
")

anio cuadro sum(valor)  
1 2013 Oferta 863742.7  
2 2013 Utilización 863742.7  
3 2014 Oferta 914170.8  
4 2014 Utilización 914170.8  
5 2015 Oferta 937376.9  
6 2015 Utilización 937376.9  
7 2016 Oferta 957463.2  
8 2016 Utilización 957463.2  
9 2017 Oferta 1001273.8  
10 2017 Utilización 1001273.8  
11 2018 Oferta 1064436.4  
12 2018 Utilización 1064436.4  
13 2019 Oferta 1129280.7  
14 2019 Utilización 1129280.7  
15 2020 Oferta 1103069.6  
16 2020 Utilización 1103069.6

Finalmente, es necesario cerrar la conexión que se abrió al inicio de este apartado, usando la función dbDisconnect().

dbDisconnect(con)