



**DEPARTAMENTO
DE COMPUTACION**

Facultad de Ciencias Exactas y Naturales - UBA

TP1

Relación PBI - Cantidad de sedes Argentinas en el exterior

February 19, 2024

Laboratorio de Datos

GRUPO 100

Integrante	LU	Correo electrónico
Chapana Puma, Joselin Miriam	1197/21	yoselin.chapana@gmail.com
Martinelli, Lorenzo	364/23	martinelli.lorenzo12@gmail.com
Padilla, Ramiro Martin	1636/21	ramiromdq123@gmail.com



Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

Ciudad Universitaria - (Pabellón I/Planta Baja)

Intendente Güiraldes 2610 - C1428EGA

Ciudad Autónoma de Buenos Aires - Rep. Argentina

Tel/Fax: (+54 +11) 4576-3300

<http://www.exactas.uba.ar>

1 Resumen

este es el resumen

2 Introducción

2.1 Objetivo y Fuente

El objetivo principal de este trabajo es encontrar una relación entre la cantidad de sedes de Argentina en un país y su PBI, para esto, trabajaremos con los siguientes datos,

- PBI per cápita de los países (1)
- Representaciones Argentinas en el exterior, donde tenemos, Datos básicos de las sedes, Datos completos de sedes y secciones (2)

(1) <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>

(2) <https://datos.gob.ar/dataset/exterior-representaciones-argentinas>

2.2 Procedimiento

Este trabajo tendrá varias etapas, comenzando con el planteo de un Diagrama de Entidad Relacional (DER) adecuado al objetivo de nuestro trabajo, continuando con la elaboración y normalización de un Modelo Relacional. Además, de llevar a cabo una limpieza de los datos utilizando ciertas métricas, en particular, utilizaremos el método GQM.

Una vez tengamos los datos limpios y el modelo normalizado, pasaremos al procesamiento y visualización de datos en Python utilizando librerías como Pandas, Inlinesql, Matplotlib entre otras con el fin de elaborar nuestras conclusiones, ¿Será mayor el PBI de aquellos países con sedes Argentinas? ¿Influirá la cantidad de sedes y/o secciones de estas?.

3 Procesamiento de Datos

3.1 Diagrama de Entidad Relacional

Una vez planteado nuestro objetivo, nos encargamos de ver que datos necesitabamos para alcanzarlo, y como estarian representados. Para esto, elaboramos el siguien diagrama de entidad relacional.

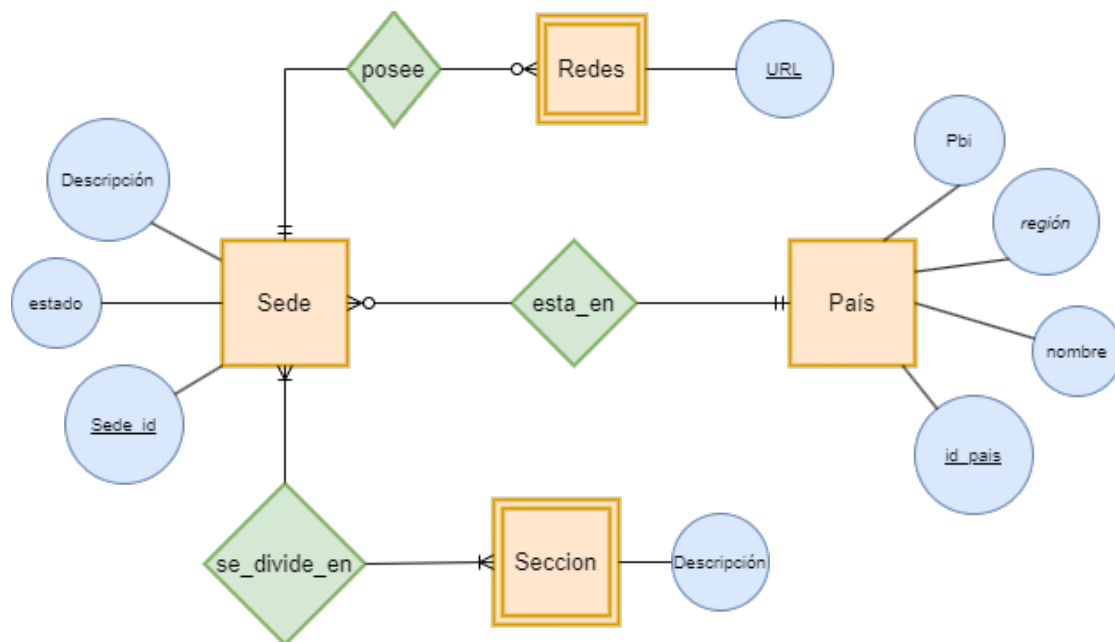


Figure 1: Diagrama de Entidad Relacional

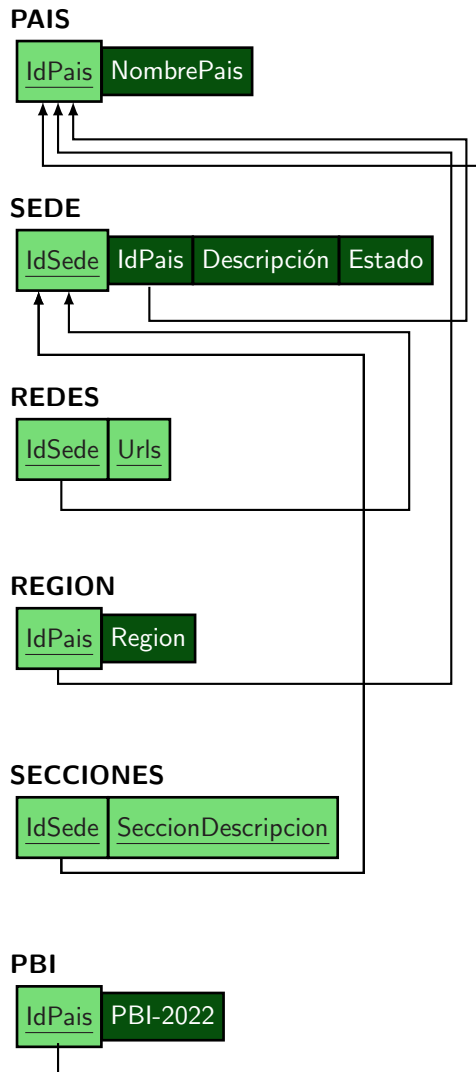
Como se puede ver en la Figure 1, consideramos que,

- Una sede esta en un país y solo en uno.
- Un País puede tener muchas o ninguna sede.
- Las secciones existen pues existen las sedes, entonces, lo consideramos una entidad debil.
- Una sección puede estar en una o muchas Sedes, por ejemplo, la Embajada en Brasil y en Chile tiene su sección Administración.
- Una sede tiene al menos una sección.
- Cuando hablamos de Redes, hablamos mas de un perfil en una red social, por lo tanto, una red puede pertenecer a una y solo una sede.

3.2 Modelo Relacional

Una vez que tenemos nuestro esquema gráfico, pasamos al planteo del modelo relacional. Notar que, las flechas representan las Foreign Keys y aquellos atributos subrayados representan las Primary Keys. En todas las relaciones, exceptuando a País, las Claves coinciden con las claves candidatas, esto se debe a que en País consideramos también a nombrePaís como una posible clave.

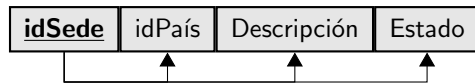
Otro detalle a tomar en cuenta, es que decidimos separar la relación País(IdPaís, región, nombre, Pbi) puesto que perdíamos la 2FN a raíz de que al ser nombre una clave candidata, región y Pbi dependían parcialmente de la clave primaria.



3.3 Normalización

Con los datos aún sin limpiar, nuestro modelo no está en 1FN, pues tiene columnas con varios datos en cada celda, por ejemplo, todas las redes sociales de una sede aparecen en una misma celda separadas por "/" . Sin embargo, podemos mostrar nuestro modelo respeta la 2FN Y 3FN. Para esto, veamos sus dependencias funcionales.

Sede



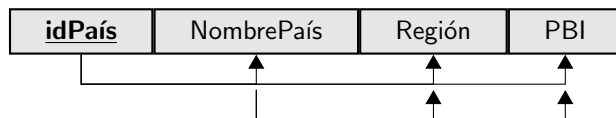
Veamos que la unica dependencia funcional que tenemos es,

$idSede \rightarrow \{idPaís, Descripción, Estado\}$

Puesto que con el id de un País no podemos inferir la descripción de una sede en particular ni su estado. Tampoco podemos inferir los demas atributos a partir de la descripción de una sede pues consideramos que podria haber descripciones repetidas o faltantes y no son suficiente para identificarlas. Por lo tanto, tenemos que todos los atributos no primos dependen de la Clave Primaria y está en 2FN, mas aún, está en 3FN. Notar además, que la DF es minimal puesto que posee un solo atributo en su lado izquierdo.

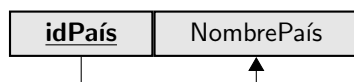
El detalle comentando en el modelo relacional en cuanto a la estructura y el porque de nuestra elección se puede ver en la siguiente tabla.

País



Es fácil notar que con la DF mas externa, estamos perdiendo la 2FN y por ende la 3FN, puesto que tenemos atributos no primos que dependen parcialmente de la clave primaria. Por lo tanto, nuestra solución fue dividir en tablas como se muestra a continuación,

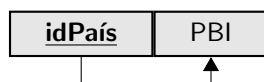
País



Región



PBI




De esta manera, estamos respetando la 2FN puesto que dependemos completamente de la PK y, en 3FN pues idPais es SK en todas las relaciones . Además, las DF detalladas son minimales puesto que dependen de un solo atributo.

Sección

<u>IdSede</u>	<u>Sección</u>
---------------	----------------

Redes

<u>IdSede</u>	<u>redURL</u>
---------------	---------------



Notemos que en Sección relaciones no podemos establecer dependendencias funcionales puesto que solo con el IdSede yo no puedo inferir una sección, y tampoco puedo realizar el camino inverso. Por otro lado, en Redes si puedo determinar la sede a partir de su url sin embargo, no puedo hacer el camino inverso pues tengo más de una url asociada a una unica sede.

Una vez hecha la aclaración, podemos ver que Sección esta en 3FN pues no posee DF. Y, por otro lado, Redes está también en 3FN pues IdSede es un atributo primo.

4 Decisiones tomadas

5 Análisis de datos

6 Conclusiones