

## Laboratorio de Datos

# Trabajo Práctico 1: Escuelas y Bibliotecas

**Fecha:** 21/5/2025

**Nombre del grupo:** Datos de Labo

**Nombre de los miembros:**

- **Denisse Britez**
- **Julieta Samosiuk**
- **Lautaro Alvarez Bertoya**



**DEPARTAMENTO  
DE COMPUTACION**

Facultad de Ciencias Exactas y Naturales - UBA

## 1. Resumen

---

El principal objetivo de este informe es analizar si existe una relación entre la cantidad de establecimientos educativos y la cantidad de bibliotecas populares en cada departamento de cada provincia de la República Argentina. Para ello, se trabajó con datos abiertos nacionales, que fueron inspeccionados, limpiados y transformados para su análisis.

El proceso incluyó la creación de un modelo de datos relacional, el análisis de formas normales, la evaluación de la calidad de los datos, la formulación de consultas SQL y la generación de visualizaciones que permitieron explorar patrones y relaciones entre las variables.

A partir del procesamiento y análisis de los datos, se obtuvieron resultados que permiten concluir en que no se observa una correlación clara entre ambas variables a nivel departamental. Mientras que la cantidad de establecimientos educativos presenta una asociación directa con la población del departamento, la distribución de bibliotecas populares evidencia una mayor dispersión, al no responder necesariamente a este mismo patrón.

Estos hallazgos sugieren que la distribución de ambos tipos de instituciones responden a factores distintos y que el acceso a establecimientos educativos y bibliotecas populares en el país no se desarrolla de manera equitativa ni complementaria.

## 2. Introducción

---

En este informe se aborda el análisis de datos abiertos correspondientes a los Establecimientos Educativos y las Bibliotecas Populares de la República Argentina. Este trabajo fue realizado con el objetivo principal de responder a la siguiente pregunta:

¿Existe una relación entre la cantidad de establecimientos educativos y la cantidad de bibliotecas populares en cada departamento de cada provincia del país?

En este contexto, se llevaron a cabo las siguientes actividades:

- Importación e inspección inicial de los datasets.
- Procesamiento de los datos importados, incluyendo la generación del Diagrama Entidad-Relación (DER), su mapeo al modelo relacional, el análisis de formas normales y la evaluación de la calidad de los datos.
- Formulación y ejecución de consultas en lenguaje SQL.
- Generación de visualizaciones que permiten describir y analizar patrones generales presentes en el conjunto de datos.
- Interpretación y análisis de los resultados obtenidos.

Como resumen del trabajo realizado, se logró resolver cada una de las etapas planteadas. A partir de la limpieza de los datasets, se formularon consultas y se desarrollaron visualizaciones que facilitaron el estudio de los datos. Todo este análisis permitió identificar posibles relaciones entre ambas variables, para llegar a conclusiones sobre su comportamiento a nivel departamental y provincial.

El informe continúa con una descripción detallada de cada una de las etapas mencionadas, incluyendo los resultados obtenidos y las decisiones tomadas a lo largo del proceso. Finalmente, se presentan las conclusiones generales del trabajo.

## 3. Procesamiento de Datos

---

### Limpieza de Datos

Para realizar las consultas y el modelado de las Bases de Datos, debemos primero hacer un análisis de las consignas del trabajo práctico, para luego poder proceder con la limpieza de los Datasets y trabajar de forma más cómoda y prolija.

Primero recorremos las columnas de cada tabla para analizar qué información nos aporta y qué información podemos descartar:

Arrancando por la base de **Padrón Población**, detectamos que la tabla tenía un formato muy incómodo con el cuál trabajar. Era una tabla llena de tablas que agrupaban habitantes según el departamento en el que se encontraban y, por cada departamento, había una tabla.

Decidimos organizarlo en una única tabla que contenga los identificadores de departamento y toda su información correspondiente. Luego, rescatamos las siguientes columnas:

- id\_departamento
- departamento
- edad
- habitantes

Para **Establecimientos Educativos**, además de analizar las columnas y limpiar los datos, tuvimos que hacer una agrupación de información sobre algunas columnas. Decidimos rescatar las siguientes columnas:

- cueanexo
- nombre
- código de localidad
- nivel inicial - jardín maternal (Común)
- nivel inicial - jardín de infantes (Común)
- primario (Común)
- secundario (Común)

En este caso, el Código de Localidad incluye el código de departamento. Como particularmente no vamos a trabajar con la localidad, decidimos rescatar solamente los dígitos que referencian al departamento (estos son los primeros 5 dígitos del código localidad).

Sobre los datos de los niveles de educación de cada establecimiento, decidimos unificarlos en una única columna, llamada Nivel educativo, que indique todos los niveles del establecimiento. Esto es porque, al encontrarse esta información en cuatro columnas con una importante cantidad de Nulls en cada una, es poco efectivo e incómodo de manejar.

Cómo quedaban casos en los que un establecimiento tenía varios niveles educativos (Valores multivariados) decidimos realizar una tabla que contenga la información de los niveles educativos, y otra que relacione los establecimientos con sus respectivos niveles. Notar que esto tiene sentido porque, al observar el DER, podemos confirmar que será una relación muchos-muchos.

La tabla de **Niveles Educativos** la conformamos por las columnas:

- id\_nivel\_educativo
- nombre
- desde\_edad
- hasta\_edad

El rango etario lo dejamos en columnas separadas para evitar valores multivariados. Y luego, la tabla de relación establecimiento-nivel-educativo tiene únicamente los valores de sus respectivos IDs.

Para **Bibliotecas Populares**, decidimos rescatar las columnas:

- nro\_conabip
- id\_provincia
- id\_departamento
- nombre
- mail
- fecha\_fundacion

En esta tabla, los únicos valores que debimos limpiar fueron los de id\_Departamento y de mail. Para id\_Departamento solo tuvimos que eliminar los dos primeros caracteres. Esto es para poder separarlo del id\_Provincia, ya que existe una columna que tiene el valor del código de provincia. Unificamos valores variados para el caso null (" ", " ", "NULL", ".").

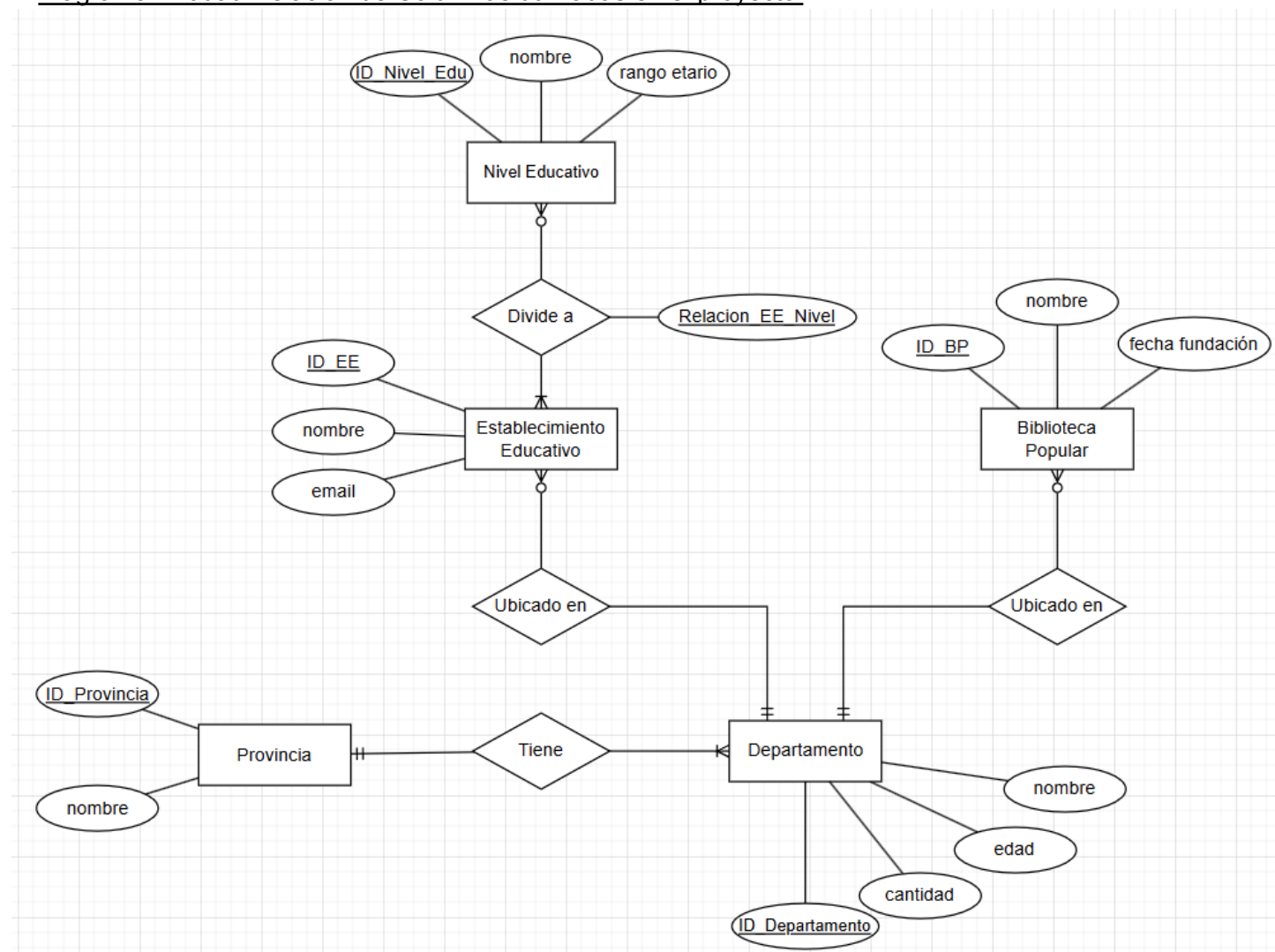
Luego, pasamos todos los mails a minúscula y borramos espacios entre caracteres. También, comprobamos que no existan casos de bibliotecas con más de un valor de mail. Particularmente, encontramos un único caso en donde el mismo valor del mail estaba repetido, así que eliminamos el duplicado.

Cabe destacar que decidimos renombrar varios atributos a lo largo de todas las tablas para obtener mayor declaratividad:

- Id\_EE: Cueanexo (Establecimientos Educativos)
- Id\_BP: nro\_conabip (Bibliotecas Públicas)
- Casos: Habitantes (Padrón población)

Además, decidimos separar id\_Departamento en dos columnas, porque este atributo también contenía a id\_Provincia. Los dos primeros dígitos del valor identificaban a la provincia (en los casos que inicializaban en 0, este también contaba como dígito) y el resto sí identificaba al departamento.

### 1. Diagrama Entidad Relación de Columnas utilizadas en el proyecto.



En el caso de la relación entre Nivel Educativo y Establecimiento Educativo, consideramos que un establecimiento puede tener muchos o no tener ningún Nivel. Esto es porque existen establecimientos que brindan únicamente Servicios Complementarios, y no son de ningún Nivel Educativo en particular.

## 2. Modelo relacional

### **BibliotecaPopular(ID\_BP, nombre, fecha\_fundacion, ID\_Departamento)**

- PK: ID\_BP
- DF: ID\_BP → nombre, fecha\_fundacion, ID\_Departamento
- FK: Departamento(ID\_Departamento)

### **EstablecimientoEducativo(ID\_EE, nombre, email, ID\_Departamento)**

- PK: ID\_EE
- DF: ID\_EE → nombre, email, ID\_Departamento
- FK: Departamento(ID\_Departamento)

### **NivelEducativo(ID\_Nivel\_Edu, nombre, edad\_desde, edad\_hasta)**

- PK: ID\_Nivel\_Edu
- DF: ID\_Nivel\_Edu → nombre, edad\_desde, edad\_hasta
- FK: Inexistente

### **Provincia(ID\_Provincia, nombre)**

- PK: ID\_Provincia
- DF: ID\_Provincia → nombre
- FK: Inexistente

### **Departamento(ID\_Departamento, nombre, edad, cantidad, ID\_Provincia)**

- PK: ID\_Departamento
- DF: ID\_Departamento → nombre, edad, cantidad, ID\_Provincia
- FK: Provincia(ID\_Provincia)

### **NivelEdu\_EE(ID\_Nivel\_Edu, ID\_EE)**

- PK: (ID\_Nivel\_Edu, ID\_EE)
- DF: (ID\_Nivel\_Edu, ID\_EE) → ∅
- FKs: EstablecimientoEducativo(ID\_EE), NivelEducativo(ID\_Nivel\_Edu)

Notar que todas las relaciones están en 3FN al no haber dependencias transitivas. Además, todos los atributos no primos dependen únicamente de la clave primaria y la relación M:N está representada con la tabla intermedia DivideA.

## 3. Formas Normales

### 1. Biblioteca Popular

La tabla está en Primera Forma Normal (1FN) porque todos sus atributos contienen valores atómicos. Es decir, ningún campo tiene listas o múltiples valores en una misma celda. Como la tabla no especifica ninguna clave primaria explícita, vamos a suponer una para poder analizar si está en 2FN. Para esto, asumimos que la clave primaria es el atributo nro\_conabip, al ser único para cada biblioteca.

Luego, para que la tabla esté en Segunda Forma Normal (2FN), todos los atributos no clave deben depender completamente de la clave primaria. Notamos que esto no se cumple, ya que hay atributos no clave que no dependen de nro\_conabip, sino de otros atributos no clave. Por ejemplo: provincia, departamento y localidad dependen de id\_provincia, id\_departamento y id\_localidad respectivamente.

Además, al no estar en 2FN, podemos afirmar que la tabla tampoco estará en Tercera Forma Normal (3FN).

### 2. Establecimientos Educativos

La tabla no está en Primera Forma Normal (1FN) porque no todos sus atributos contienen valores atómicos. Por ejemplo: los atributos teléfono y mail contienen valores multivaluados, al tener más de un número de teléfono/dirección de mail para una misma celda. Esto viola el principio de atomicidad exigido por la 1FN.

Luego, al no estar en 1FN, sabemos que tampoco va a estar en Segunda Forma Normal (2FN) ni en Tercera Forma Normal (3FN).

#### 4. Calidad de Datos

##### **1. Establecimientos Educativos**

En esta tabla detectamos que en la columna CP, correspondiente a los códigos postales de los establecimientos educativos, existen registros con valores vacíos o nulos. Esto implica un problema en el atributo de calidad completitud, ya que no están presentes todos los valores que representan la realidad. Este tipo de problema es de instancia, porque estos datos no fueron almacenados con la precisión suficiente.

Luego, para determinar la gravedad de este problema, vamos a utilizar la técnica GQM.

- Goal/Objetivo: Verificar que la columna CP esté completa para todo establecimiento educativo.
- Question/Pregunta: ¿Cuál es la proporción de establecimientos educativos que tienen valores nulos en el atributo CP?
- Metric/Métrica: Cantidad total de códigos postales vacíos / Cantidad total de códigos postales →  $2678 \div 63662 \approx 0,042$

Para facilitar la interpretación, pasamos esta proporción a porcentaje:  $0,042 \times 100 = 4,2\%$ . Luego, podemos afirmar que hay un 4,2% de valores nulos registrados en la columna CP. Si bien no es un valor grande, puede comprometer la representación completa de la realidad.

Este problema pudo haberse generado por una carga manual incompleta, probablemente porque al momento de cargar la información no se contaba con el dato necesario para completar este campo. Esto podría resolverse completando los códigos postales faltantes con fuentes externas o, directamente, no trabajando con este atributo cuando el código postal no sea relevante.

##### **2. Bibliotecas Populares**

En esta tabla detectamos varias columnas que consideramos que no aportan información relevante para representar la realidad. Podemos clasificarlas en dos grandes grupos:

- Columnas completamente vacías: observacion, subcategoria, piso, web e informacion\_adicional. Estos atributos no contienen información en ninguna instancia. Por lo tanto, su inclusión no agrega valor a la tabla.
- Columnas con un único valor idéntico en todas las filas: categoria, fuente, anio\_actualizacion y tipo\_latitud\_longitud. Creemos que su contenido es relevante como información general, pero al tener el mismo valor para todas las filas, terminan no siendo relevantes como variables.

Esto implica un problema en el atributo de calidad relevancia, ya que no todos los atributos presentes aportan valor para representar la realidad. Este tipo de problema es de modelo de datos, porque el diseño del dataset incluye atributos innecesarios.

Luego, para determinar la gravedad de este problema, vamos a utilizar la técnica GQM.

- Goal/Objetivo: Verificar que todos los atributos del dataset sean relevantes.
- Question/Pregunta: ¿Cuál es la proporción de columnas de Bibliotecas Populares que no son relevantes?
- Metric/Métrica: Cantidad total de columnas no relevantes / Cantidad total de columnas del dataset →  $9 \div 25 = 0,36$

Para facilitar la interpretación, pasamos esta proporción a porcentaje:  $0,36 \times 100 = 36\%$ . Luego, podemos afirmar que el 36% de las columnas del Dataset son irrelevantes. Si bien es menos de la mitad, sigue siendo un número que podemos considerar. Esta proporción significativa de columnas no relevantes puede dificultar la limpieza de los datos o el trabajo con el dataset. Además, puede influir negativamente en la visualización o aumentar innecesariamente el tamaño del archivo.

Este problema pudo haberse generado por una exportación sin depuración desde las bases de datos completas o por falta de revisión en la definición del esquema de datos. Para el caso de las columnas vacías, este problema podría resolverse eliminándolas o, si se decide mantenerlas, completando sus campos para que no generen ambigüedad.

En el caso de las columnas con un valor único para todas las filas, proponemos extraer dicha información del cuerpo del dataset y trasladarla a la metadata del archivo. Notar que estos atributos pueden considerarse metadatos, ya que no describen a registros individuales, sino que aportan información general sobre el conjunto completo de datos.

Una forma concreta de hacer esto podría ser incluir estos valores en el nombre del archivo (por ejemplo, "bibliotecas\_populares\_CONABIP\_2023.csv"). Así, se preserva la información general, sin comprometer la relevancia del modelo de datos. Además, facilita la limpieza, interpretación y manipulación del dataset en etapas posteriores del análisis.

#### 4. Decisiones tomadas

Importación de Dataset de id Provincias. Obtenido de la página del INDEC para no tener que hacer una consulta desde el archivo de padrón. Esto es porque no sabemos si se encuentran todas las provincias con sus respectivos códigos.

Fuente: [https://www.indec.gob.ar/ftp/cuadros/menusuperior/clasificadores/anexo1\\_resol55\\_2019.pdf](https://www.indec.gob.ar/ftp/cuadros/menusuperior/clasificadores/anexo1_resol55_2019.pdf)

Rangos Etarios de Niveles Educativos. Decidimos clasificar los rangos etarios en:

- Jardín: de 0 a 5 años, incluido Maternal.
- Primario: de 6 a 12 años.
- Secundario: de 13 a 18 años.

Limpiar códigos de departamentos de CABA. Observamos que no coinciden los IDs de estos departamentos entre las tablas y, además, en la tabla de Bibliotecas Populares, no se explicitan las comunas de CABA (directamente se detalla "CABA"). Entonces, para no generar datos incoherentes intentando arreglar estos errores, decidimos "perder el detalle" de los datos de esos casos y agruparlos en un grupo mayor.

Unificación de archivos. Durante la realización de este informe, estuvimos trabajando sobre cinco archivos distintos: uno por cada Dataset, uno para las consultas SQL y uno para el análisis y la visualización de los datos. Es por esto que, al unificar estos archivos, puede que nos hayamos referido a las mismas variables con distintos nombres.

#### 5. Análisis de datos

##### Consultas SQL

##### **Consulta 1:**

Buscamos una tabla que, para cada departamento, muestre la provincia, el nombre del departamento, la cantidad de establecimientos educativos por nivel (Jardín, Primaria, Secundaria) y la población en el rango de edad correspondiente a cada nivel. Ordenamos alfabéticamente por provincia y, dentro de cada una, de mayor a menor según escuelas primarias.

Dividimos la consulta en tres partes:

- Subconsulta poblacion\_por\_nivel: agrupa la población por nivel educativo y departamento.
- Subconsulta establecimientos\_por\_nivel: cuenta los establecimientos por nivel educativo y departamento.
- Consulta principal: une los resultados de las subconsultas anteriores con la tabla de provincias para obtener el nombre legible de cada provincia, y organiza el resultado final ordenado alfabéticamente por provincia, y dentro de cada una, por cantidad descendente de escuelas primarias.

## **Consulta 2:**

Queremos saber cuántas bibliotecas populares fueron fundadas desde 1950 en cada departamento del país, mostrando también la provincia y el nombre del departamento. Ordenamos los resultados alfabéticamente por provincia y, dentro de cada una, de mayor a menor según cantidad de bibliotecas fundadas.

Para eso usamos:

- Tabla bibliotecas: contiene los datos principales, como la fecha de fundación y la ubicación.
- Tabla provincias: se une para obtener el nombre de la provincia.
- Tabla poblacion: se usa solo para recuperar el nombre del departamento.

Se filtran las bibliotecas según la fecha de fundación, se agrupan por provincia y departamento, y se cuenta cuántas cumplen la condición.

## **Consulta 3:**

Buscamos obtener, para cada departamento, la provincia, el nombre del departamento, la cantidad de bibliotecas populares, la cantidad de establecimientos educativos (modalidad común) y la población total. Incluimos todos los departamentos, incluso aquellos sin bibliotecas o establecimientos. El resultado se ordena por cantidad de establecimientos descendente, luego bibliotecas descendente, y finalmente por provincia y departamento en orden alfabético.

Dividimos la consulta en dos partes:

- Subconsulta poblacion\_total: calcula la población total por provincia y departamento, y obtiene el nombre del departamento.
- Consulta principal: une la población con las tablas de provincias, establecimientos y bibliotecas para contar la cantidad de EE y BP, respetando los casos sin coincidencias, y ordena el resultado.

## **Consulta 4:**

Buscamos identificar, para cada departamento, la provincia, el nombre del departamento y cuál es el dominio de correo electrónico más utilizado por las bibliotecas populares.

Dividimos la consulta en tres partes:

- Subconsulta dominios\_contados: extrae el dominio del mail de cada biblioteca y cuenta cuántas veces aparece en cada provincia y departamento, descartando mails inválidos o nulos.
- Subconsulta máximos: calcula, para cada provincia y departamento, la máxima cantidad de uso de un dominio.
- Consulta principal: une los resultados para devolver el dominio o dominios más frecuentes por departamento, junto con la provincia y el nombre del departamento, ordenados alfabéticamente.

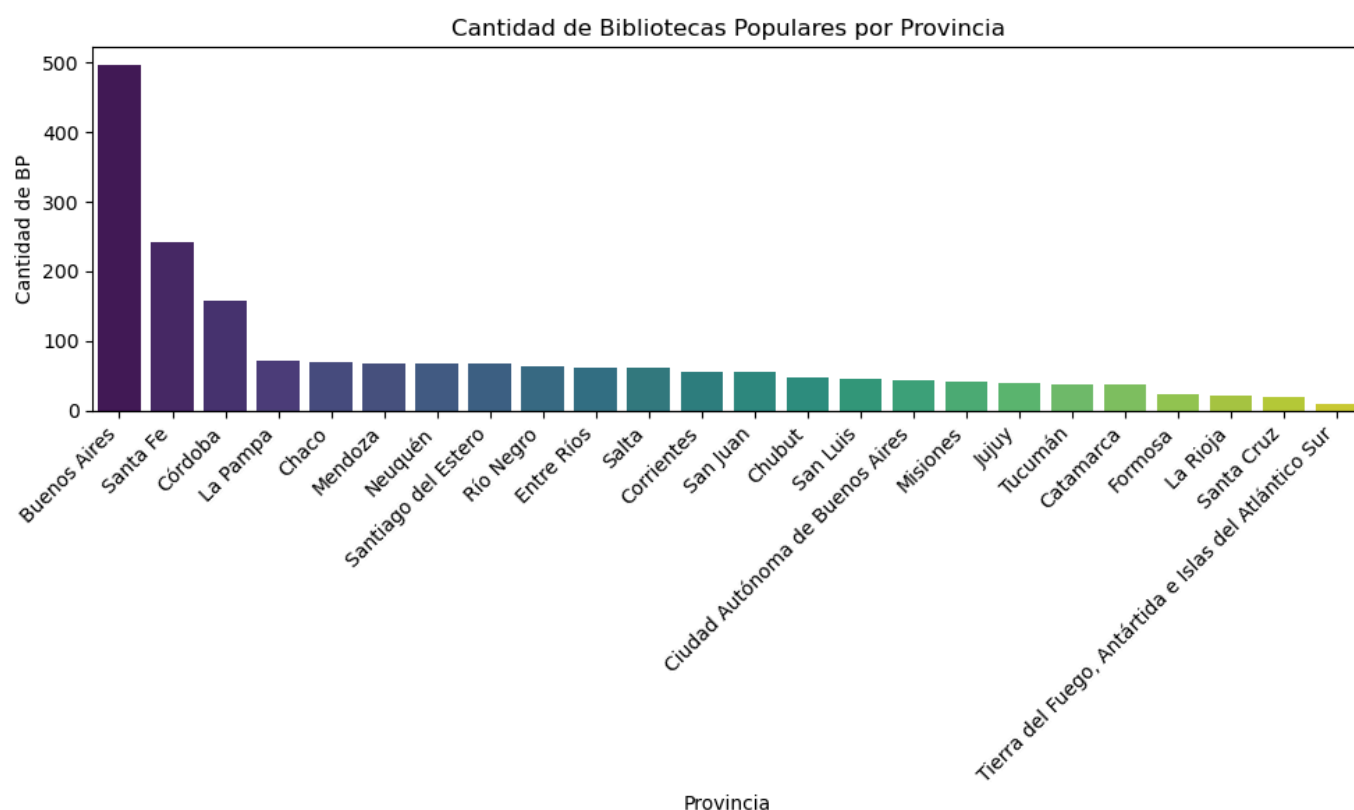


### Gráfico 1: Cantidad de bibliotecas populares por provincia, ordenadas de mayor a menor.

Para este gráfico necesitamos dos tablas:

- bibliotecas.
- provincias.

En este caso, consideramos más adecuado trabajar con un **gráfico de barras** porque muestra comparaciones categóricas con claridad: cada barra representa una provincia y su longitud indica la cantidad de BP. Ordenar las barras de mayor a menor resalta visualmente cuáles provincias tienen más bibliotecas, y el formato es fácil de leer e interpretar, incluso con muchas provincias. Además, es más apropiado que un gráfico de torta cuando hay muchas categorías o cuando las diferencias entre valores son sutiles.



**Conclusión:** Buenos Aires, Santa Fe y Córdoba lideran con diferencia en la cantidad de bibliotecas populares, lo que podría reflejar una combinación de factores como tamaño poblacional, políticas de promoción cultural y recursos disponibles. Otras provincias presentan cantidades notablemente menores, lo que pone en evidencia una distribución desigual de este tipo de instituciones en el país.

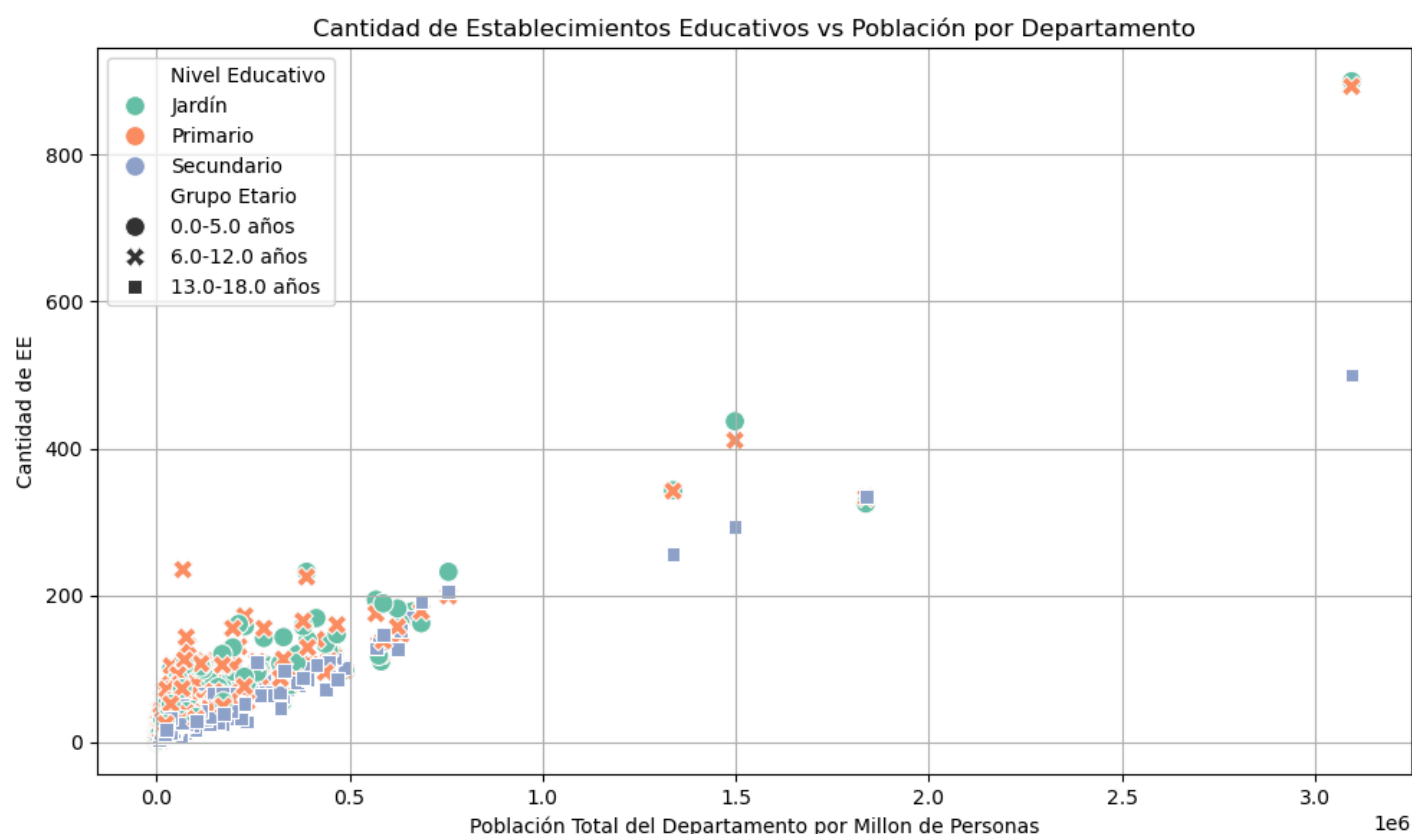
### Gráfico 2: Cantidad de establecimientos educativos por departamento en función de la población, diferenciando por nivel educativo y grupo etario.

Para este gráfico necesitamos cuatro tablas:

- establecimientos.
- poblacion.
- relación\_EE\_nivel.
- niveles.

En este caso, consideramos más adecuado trabajar con un **gráfico de dispersión (scatter plot)** porque permite relacionar dos variables numéricas (población y cantidad de EE) y, al mismo tiempo, incorporar información adicional mediante el color y la forma de los puntos, diferenciando tanto el nivel educativo como el

grupo etario. Este tipo de gráfico es útil para detectar patrones como si existe proporcionalidad entre EE y población o si ciertos niveles educativos están más presentes en determinados departamentos.



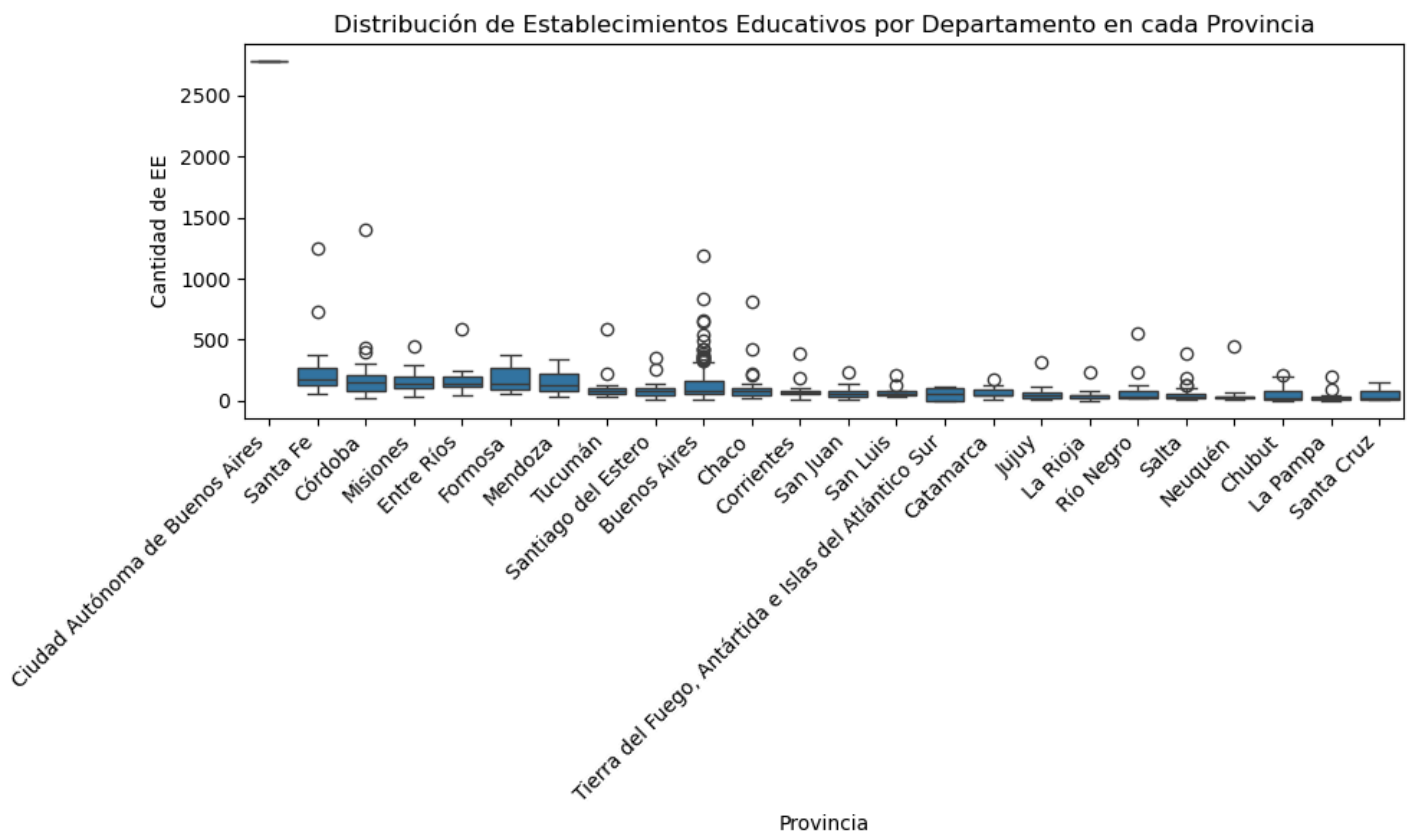
**Conclusión:** existe una correlación clara entre la población del departamento y la cantidad de EE, especialmente en los niveles inicial y primario. Sin embargo, se observa una menor presencia relativa de EE secundarios en muchos departamentos, lo que podría indicar una menor cobertura educativa para los adolescentes en ciertas regiones. Esta información es valiosa para planificar políticas públicas orientadas a equilibrar la oferta educativa en todo el país.

### Gráfico 3: Distribución de establecimientos educativos por departamento en cada provincia.

Para este gráfico necesitamos dos tablas:

- establecimientos.
- provincias.

En este caso, utilizamos un **boxplot** por provincia, donde cada “caja” representa la distribución de EE entre sus departamentos. Es una elección adecuada porque permite visualizar la mediana, el IQR y los valores atípicos de cada provincia. Además, se ordenaron los boxplots según la mediana de EE por departamento, lo que facilita la comparación relativa entre provincias. Esta representación permite detectar fácilmente si una provincia tiene departamentos muy desbalanceados o si la distribución es más homogénea.



**Conclusión:** provincias como CABA (Luego aclaramos este caso), Santa Fe y Córdoba muestran medianas más altas en cantidad de EE por departamento, aunque también presentan dispersión. Buenos Aires, si bien tiene algunos departamentos con muchos EE, muestra una mediana más baja, reflejando una alta concentración de EE en pocos departamentos.

En contraste, provincias como Santa Cruz, La Pampa o Chubut exhiben una distribución más acotada y con menor cantidad de EE por departamento, lo cual podría estar relacionado con la densidad poblacional y la extensión geográfica. Notar que, como realizamos una unificación de todas las comunas de CABA, este caso particular se ve representado en la visualización como una única línea.

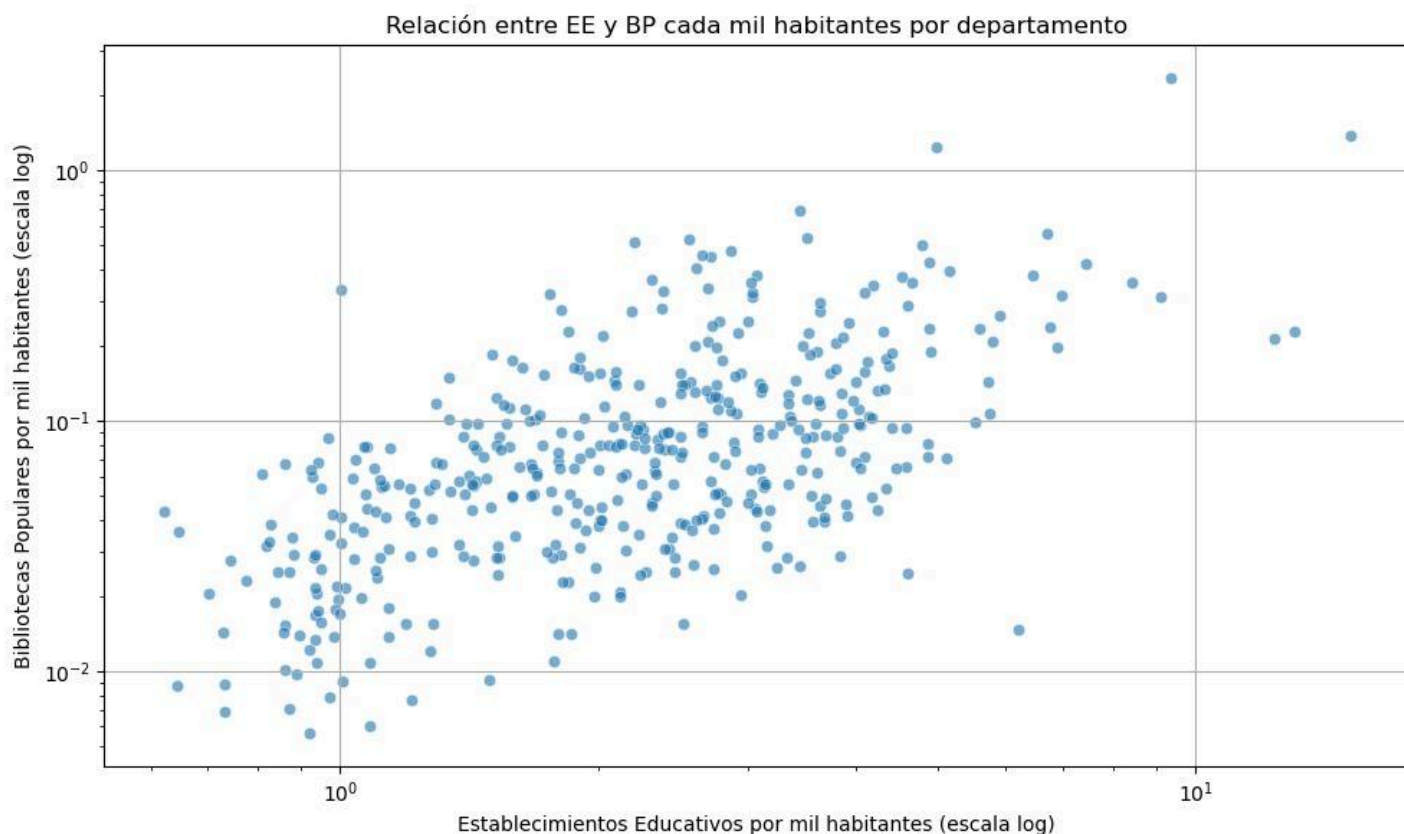
**Gráfico 4: Relación entre la cantidad de BP cada mil habitantes y de EE cada mil habitantes por departamento.**

El objetivo es analizar cómo se relacionan la cantidad de bibliotecas populares (BP) y establecimientos educativos (EE) por cada mil habitantes en los distintos departamentos del país.

Para este gráfico necesitamos tres tablas:

- establecimientos.
- bibliotecas.
- padron\_poblacion\_unificado.

En este caso, consideramos más adecuado trabajar con un **gráfico de dispersión**, porque permite visualizar la relación entre dos variables cuantitativas: EE y BP por mil habitantes. Cada punto del gráfico representa un departamento. La posición horizontal indica la cantidad de EE por mil habitantes, y la vertical, la cantidad de BP por mil habitantes. Esta visualización es útil para detectar tendencias, agrupamientos o valores atípicos (outliers) entre los departamentos.



**Conclusión:** si bien se observa una débil tendencia creciente, donde los departamentos con más escuelas también tienden a tener más bibliotecas, la alta dispersión de los datos indica que no existe una relación fuerte ni uniforme entre ambas variables. La mayoría de los departamentos presenta bajas tasas relativas de ambos servicios, concentrándose en la parte inferior izquierda del gráfico. La escala logarítmica permite identificar con mayor claridad las diferencias relativas entre departamentos, destacando aquellos con proporciones excepcionalmente altas o bajas.

## 6. Conclusiones

Luego de realizar este trabajo, podemos llegar a varias conclusiones en relación a nuestro objetivo principal: determinar si existe una relación entre la cantidad de bibliotecas populares y la cantidad de establecimientos educativos en Argentina.

En primer lugar, concluimos que no hay una relación significativa entre ambos indicadores. Esto se evidencia en la existencia de departamentos con una alta cantidad de establecimientos educativos pero pocas bibliotecas populares. Son escasos los casos donde ambos valores son elevados, ya que la mayoría de los departamentos se agrupan en un rango bajo o medio para ambas variables. Esto indica que la presencia de bibliotecas populares y establecimientos educativos no se desarrolla de manera paralela ni uniforme en el país.

Por otra parte, también observamos que la distribución de bibliotecas populares en Argentina es desigual. La Provincia de Buenos Aires concentra la mayor cantidad, seguida por Santa Fe y Córdoba. Esto es esperable, al ser las tres provincias más pobladas del país. En contraste, la Patagonia y el norte tienen una presencia significativamente menor de bibliotecas. Esto sugiere que el acceso a estas varía considerablemente entre provincias.

En cuanto a los establecimientos educativos, identificamos una tendencia positiva entre la población y la cantidad de establecimientos: a mayor población del departamento, mayor cantidad de establecimientos educativos. Particularmente, observamos que la cantidad de establecimientos educativos secundarios es siempre menor. Todo esto refuerza la idea de que el acceso educativo está más desarrollado y organizado que el acceso a bibliotecas populares.