Laboratorio de Datos

Trabajo Práctico 01:

Escuelas y bibliotecas

23 de mayo de 2025

Lago de Batos

Integrante	LU	Correo electrónico
Barrios, Bruno	1369/24	brunolautarobarriosmarmier@gmail.com
Mur, Santiago	312/24	santi.mur2004@gmail.com
Operti, Bruno	422/24	opertibruno@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

➔ **Resumen:**

En este trabajo nos planteamos, a partir de datos reales oficiales obtenidos directamente de la web del INDEC, analizar una posible relación entre la cantidad de Establecimientos Educativos y Bibliotecas Públicas por cada provincia, a nivel nacional de la República Argentina. Con este fin, tendremos en cuenta para los distintos departamentos sociales dentro de cada provincia su distribución demográfica por rango etario, relacionándolo con la distribución tanto de establecimientos como de bibliotecas públicas con objetivo de llegar a una conclusión.

Contamos con tres bases de datos: el Padrón Oficial de Establecimientos Educativos, el Padrón de Bibliotecas Populares y los Datos de población por Departamento (datos del censo 2022). Inicialmente ideamos un DER, en base al cual diseñamos nuestro modelo relacional y la estructura de nuestro modelo de base de datos sobre el cual trabajaremos.

A partir del análisis realizado, no se identificó una relación directa entre la cantidad de establecimientos educativos y bibliotecas populares por provincia. Mientras que la cantidad de EE muestra una correlación con la población, la distribución de BP varía de forma más independiente. Esto sugiere que ambos tipos de instituciones responden a lógicas diferentes de planificación o necesidades territoriales.

➔ Introducción

Problemática a resolver:

Definir si, tras un análisis y manipulación de datos correspondiente, es posible encontrar o establecer una relación entre el número de Establecimientos Educativos y Bibliotecas Públicas en las provincias del país. Con este objetivo tendremos en cuenta la distribución o densidad tanto de la población como de los establecimientos nombrados anteriormente.

Objetivo general:

Llevar a cabo un posible diseño de base de datos, un procesamiento de nuestros datos originales para construir el modelo diseñado, y un posterior análisis de nuestro modelo de datos que nos ayude a determinar la existencia o no de la relación buscada.

Actividades a realizar:

Nuestra "hoja de ruta" para llevar a cabo nuestro objetivo consistió en:

- Descargar y hacer un primer análisis a nuestras fuentes de datos originales mencionadas anteriormente: analizar estructuras, información disponible y posibles conclusiones a partir de esta.
- Diseñar un DER (Diagrama de Entidad-Relación) que represente cómo va a verse nuestro modelo de datos, y cómo nuestras entidades (junto con sus atributos) van a relacionarse dentro de este.
- A partir del DER, diseñar un Modelo Relacional que lo represente correctamente.
- Realizar un análisis de calidad de nuestras bases de datos originales, utilizando el método GQM.
- Llevar a cabo una limpieza a nuestros datos originales, obteniendo así nuestras entidades del modelo relacional junto con sus atributos.
- Desarrollar el análisis de nuestros datos limpios a través de consultas SQL, y visualizaciones de la información obtenida. Evaluar la conclusión obtenida a través de nuestro desarrollo.

Acerca de este informe:

A continuación en nuestro documento, desarrollaremos acerca de:

- El procesamiento llevado a cabo sobre las bases de datos originales
- Las decisiones tomadas durante la ejecución de nuestra hoja de ruta
- El análisis sobre nuestro modelo de datos
- La conclusión obtenida a partir de este proceso.

➔ Procesamiento de datos

Formas normales de nuestras fuentes de datos

Tras analizar los datos originales, podemos ver que las 3 tablas se encontraban en la Primera Forma Normal (1FN), ya que todos sus atributos contienen valores atómicos sin repetición.

Establecimientos educativos: además de estar en 1FN, también cumple con la Segunda Forma Normal (2FN) al tomar como PK (*primary key*) al atributo "cueanexo". No existen DF (Dependencias Funcionales) parciales con respecto a la PK. Sin embargo, no cumple con la Tercera Forma Normal (3FN) debido a dependencias funcionales entre atributos no clave. Por ejemplo: el código postal determina el código de localidad, donde el código postal no es SK (super clave) del esquema ni atributo primo.

Bibliotecas Populares: también está en 1FN con atributos atómicos y tiene como PK el atributo "nro_conabip". Cumple con 2FN: no existen DFs parciales con respecto a esta clave simple. No obstante, no está en 3FN: existen DFs entre atributos no clave, por ejemplo: "id_provincia" → "provincia" donde "id_provincia" no es SK ni provincia atributo primo.

Calidad de Datos

- **Establecimientos Educativos**

Comenzamos con la base de datos de Establecimientos Educativos. En un primer vistazo, logramos identificar que: para el atributo "Mail", contamos con numerosos registros que contienen más de una dirección de mail para un único establecimiento, delimitados por caracteres como "," , ";" o "/". De este modo, estamos violando el principio de atomicidad de los datos, es decir, nuestros datos no están compuestos por una única unidad atómica e indivisible de información, sino por varias piezas de información.

Los atributos de la calidad de esta base de datos afectados serían:

- **Disponibilidad**: la información del correo electrónico de los establecimientos se está en la base de datos, pero, su formato agrupado junto con uno o más mails dentro de la misma celda, hace que el dato no esté disponible para su utilización sin una transformación o limpieza previa. Es decir, el dato no está disponible para su utilización/manipulación inmediata.
- **Consistencia**: en una base de datos consistente, se espera tener para un mismo atributo una única forma uniforme y estructurada de representar sus datos correspondientes. En esta base de datos, contamos con varios formatos para representar los mails de los establecimientos, por lo que los datos no son consistentes.

Planteamos entonces el análisis GQM sobre el atributo "Mail" de la base de datos de Establecimientos Educativos.

Análisis GQM

- **Goal**: determinar en qué medida el atributo "Mail" de la base de datos original de establecimientos educativos cumple con todos los atributos de la calidad, especialmente su consistencia, disponibilidad y estructura por datos atómicos.
- **Question**: ¿Qué proporción de los registros del atributo "Mail" contienen más de una dirección de correo electrónico en la misma celda?

- **Metric** - definimos:

$Ntotal: \ cantidad\ de\ registros\ del\ atributo\ "Mail"$
$Nmult: \ cantidad\ de\ registros\ del\ atributo\ "Mail"\ con\ más\ de\ una\ dirección\ de\ mail.$
$\%mult = (Ntotal/Nmult) * 100: \ porcentaje\ de\ registros\ con\ más\ de\ una\ dirección\ de\ mail$

$Ntotal = 52874$
$Nmult = 14015$
$\%mult = (52874 / 14015) = 26.51\%$

Podemos ver de este modo que el 26.51% de los registros de Mail presentan estructura no atómica, por lo que impiden el uso instantáneo del dato para su manipulación y brindan inconsistencia sobre su formato. Podría tratarse de un problema del modelo de datos, el cual no permite detallar correctamente si un establecimiento cuenta con más de una dirección de correo electrónico.

Podríamos corregir este problema quedándose únicamente con la 1er dirección de mail correspondiente, o creando una tabla adicional que contenga el campo Mail para cada establecimiento, permitiendo más de una fila por establecimiento representando varias direcciones de mail.

- **Bibliotecas Populares:**

Para la base de datos de Bibliotecas Populares, podemos notar que no contamos con ningún dato correspondiente al atributo "subcategoria", es decir, para este atributo contamos únicamente con valores vacíos o Nulos.

Podemos decir que este fallo en la calidad afecta al atributo de **Relevancia** y de **Completitud** de la base de datos de Bibliotecas Populares. Esto se debe a que no brinda ningún tipo de información a la hora de realizar análisis o manipulación sobre los datos de esta base, y todos sus valores se encuentran vacíos.

Análisis GQM

- **Goal**: Evaluar completitud y relevancia de los registros del atributo "subcategoria" de nuestra tabla original de Bibliotecas Populares. Identificar si se evidencia un diseño inadecuado del modelo de datos.
- **Question**: ¿el atributo "subcategoria" contiene algún valor que lo justifique como atributo relevante? ¿En que proporcion?
- **Metric** - planteamos:

$Ntotal: \ cantidad\ de\ registros\ del\ atributo\ "subcategoria"$
$Nnull: \ cantidad\ de\ registros\ nulos\ del\ atributo\ "subcategoria"$
$\%null = (Ntotal/Nmult) * 100: \ porcentaje\ de\ registros\ nulos$
$\%null = 100\%$

Se trata de un problema del modelo de datos, teniendo un campo definido pero nunca utilizado. Por otro lado, es posible que se trate de un problema de instancia, si la columna debiera contener datos pero no se logró recolectarlos en el proceso de carga.

Este problema puede solucionarse eliminando el atributo de nuestra base de datos o, en caso de que el atributo sea teóricamente relevante, replantear el sistema de recolección de datos para que pueda representar correctamente este aspecto de nuestro dataset.

Diagrama Entidad-Relación (DER)

En esta sección se presenta el Diagrama Entidad Relación (DER) diseñado a partir de las tres fuentes de datos previamente procesadas: establecimientos educativos, bibliotecas

populares y padrón de población por nivel educativo. A partir del análisis y la limpieza de datos, se identificaron las entidades relevantes, sus atributos, y las relaciones necesarias para representar adecuadamente la información. El diseño busca normalizar los datos y reflejar de forma precisa las asociaciones entre departamentos, bibliotecas, establecimientos educativos, niveles educativos y población. A continuación, se describe brevemente cada entidad, sus claves primarias y relaciones, y se incluye el diagrama correspondiente.

Entidades:

- **Departamento**: Unidad geográfica y administrativa básica.
  Atributos: id, nombre, provincia
  Relación: Cada departamento puede tener cero o más bibliotecas, establecimientos y registros de población.

- **Biblioteca Popular**: Institución cultural sin fines de lucro.
  Atributos: id biblioteca, nombre, año de fundación, mail, id departamento
  Relación: Cada biblioteca está ubicada en un único departamento.

- **Establecimiento Educativo**: Institución que brinda enseñanza en una o más modalidades.
  Atributos: id establecimiento, nombre, ámbito, modalidad común, modalidad especial, modalidad adultos, id departamento
  Relación: Cada establecimiento está ubicado en un único departamento. Puede estar vinculado a uno o más niveles educativos.

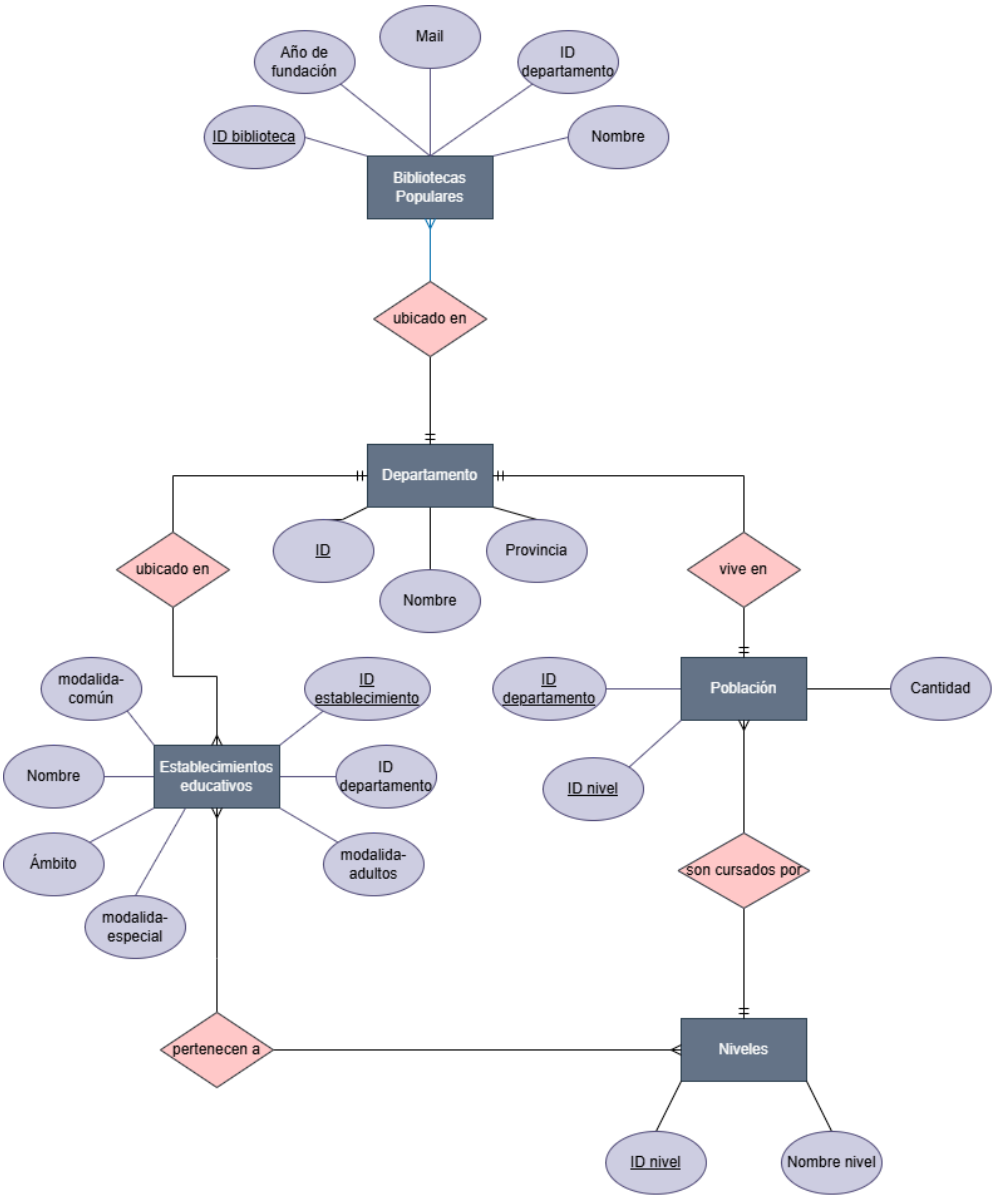
- **Nivel Educativo**: Representa un tipo de enseñanza (por ejemplo: inicial, primario, secundario).
  Atributos: id nivel, nombre nivel
  Relación: Cada nivel puede estar asociado a múltiples establecimientos educativos.

- **Población**: Datos demográficos agregados por departamento y nivel educativo.
  Atributos: id departamento, id nivel, cantidad
  Relación: Cada fila representa la cantidad de población correspondiente a un nivel específico dentro de un departamento.

Relaciones:

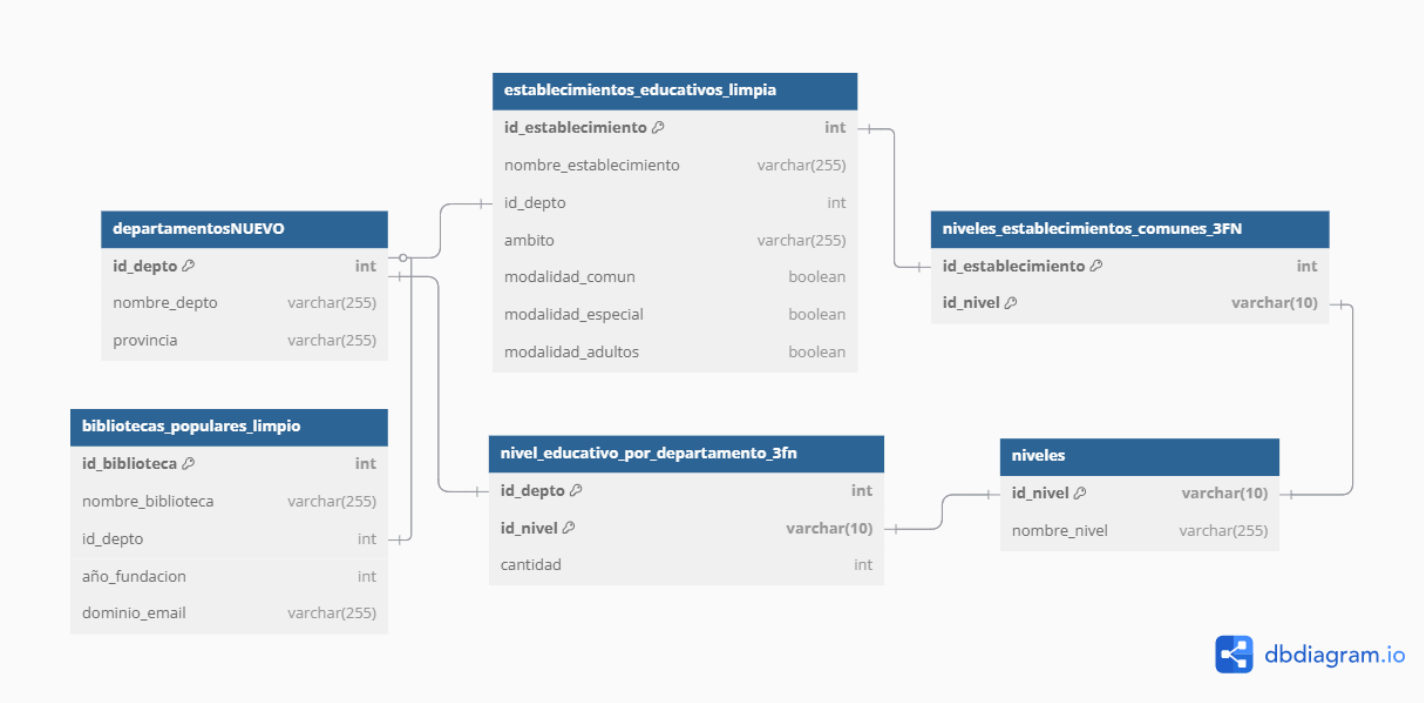
- **Ubicado en**: Cada biblioteca y establecimiento educativo se encuentra en un único departamento.
- **Vive en**: La población está segmentada por departamento y nivel educativo.
- **Son cursados por**: Cada nivel educativo tiene una cantidad de población asociada que lo cursa, en cada departamento.
- **Pertenecen a**: Los niveles educativos que ofrece un establecimiento se modelan mediante una relación entre ambas entidades (muchos a muchos).

A continuación, nuestro **Diagrama de Entidad-Relación**:



Modelo Relacional

Basándonos en el DER, construimos nuestro modelo relacional. En este, representamos las entidades y sus relaciones mediante tablas, de acuerdo al modelo de base de datos realizado. En cada relación contamos con *primary keys* (PK) y *foreign keys* (FK), vinculadas mutuamente. Sumado a esto tenemos *dependencias funcionales* (DF), las cuales describen las dependencias de los atributos con respecto a las claves. Al idear el modelo relacional, buscamos que todas las relaciones se encuentren en la **Tercera Forma Normal** (3FN) y cumplan con la **Forma Normal de Boyce-Codd** (BCNF). De este modo, no tendremos redundancias en la estructura.



Importación y limpieza de datos

Para generar los datasets o tablas asociadas al modelo relacional, desarrollamos distintos scripts de Python que se adapten a las distintas necesidades de limpieza de datos correspondientes a cada base de datos original. Procedimos de la siguiente forma:

- **Carga de datasets originales:** haciendo uso de la biblioteca Pandas, importamos los distintos archivos CSV y Excel correspondientes a las tablas de información originales (Establecimientos Educativos, Bibliotecas Populares, Población por departamento).
- **Limpieza y normalización:** nos aseguramos de no tener elementos duplicados, filas incompletas o datos no relevantes para nuestro análisis. Renombramos columnas para unificar nomenclaturas para PKs y FKs. Generamos las siguientes tablas, a partir de los distintos archivos fuente:

➢ ***bibliotecas_populares_limpio.csv :*** Representa cada biblioteca popular. A partir de la tabla de Bibliotecas Populares original, seleccionamos y renombramos los siguientes atributos relevantes para el análisis:
  - ➢ nro_conabip → id_biblioteca
  - ➢ nombre → nombre_biblioteca
  - ➢ id_departamento → id_depto
  - ➢ departamento → nombre_departamento
  - ➢ provincia → nombre_provincia
  - ➢ mail → dominio_email : a partir del atributo "mail", obtenemos el dominio de cada dirección email.
  - ➢ fecha_fundacion → año_fundacion : obtenemos únicamente el año de fundacion de cada biblioteca.

➢ ***departamentosNUEVO.csv :*** Representa cada departamento. A partir de la tabla de Establecimientos Educativos, tomamos y renombramos los atributos relevantes:
  - ➢ Código de departamento → id_depto

- Departamento → nombre_depto
- Jurisdicción → nombre_provincia

- ***establecimientos_educativos_limpia.csv :*** Representa cada establecimiento educativo. A partir de la tabla de Establecimientos Educativos original, tomamos y renombramos los atributos:
  - Cueanexo → id_establecimiento
  - Nombre → nombre_establecimiento
  - Código de departamento → id_depto
  - Ámbito → ambito
  - Común → modalidad_común : si el establecimiento dicta la modalidad comun (1 o 0)
  - Especial → modalidad_especial : si el establecimiento dicta la modalidad especial (1 o 0)
  - Adultos → modalidad_adultos : si el establecimiento dicta la modalidad adultos (1 o 0)
- ***niveles.csv :*** Asocia el nombre de cada nivel educativo con su id. A partir de la tabla de Establecimientos Educativos original, generamos las siguientes asociaciones:
  - nombre_nivel = Jardin Maternal → id_nivel = jardin_maternal
  - nombre_nivel = Jardin de infantes → id_nivel = jardin_infante
  - nombre_nivel = Primario → id_nivel = primario
  - nombre_nivel = Secundario → id_nivel = secundario
  - nombre_nivel = SNU → id_nivel = snu
  - nombre_nivel = SNU - INET → id_nivel = snu_inet
  - nombre_nivel = SNU - Cursos → id_nivel = snu_cursos
- ***niveles_establecimientos_comunes_3FN.csv :*** Representa qué niveles educativos de modalidad común dicta cada establecimiento. Por cada nivel dictado por un establecimiento se cuenta con una fila. A partir de la tabla de Establecimientos Educativos original, tomamos los atributos relevantes y renombramos tal que:
  - Cueanexo → id_establecimiento
  - Nivel inicial - Jardìn maternal → jardin_maternal
  - Nivel inicial - Jardín de infantes → jardin_infante
  - Primario → primario
  - Secundario → secundario
  - Secundario - INET → secundario_inet
  - SNU → snu
  - SNU - INET → snu_inet
  - SNU - Cursos → snu_cursos
- ***nivel_educativo_por_departamento_3fn.csv :*** Representa la cantidad de personas que cursa cada nivel educativo en cada departamento. A partir de la tabla de Poblacion por departamento original, dividimos por rango etario y se lo asignamos al nivel correspondiente. Contamos con los atributos:
  - id_depto
  - id_nivel
  - cantidad → cantidad de personas con edad correspondiente a cada respectivo nivel

Decisiones tomadas

Durante el proceso de limpieza de datos y generación de entidades y relaciones, tomamos las decisiones de:

- ➔ Eliminar columnas vacías e irrelevantes
- ➔ Eliminar valores duplicados

- ➔ Para determinar cantidad de personas por nivel educativo, tomamos:
  - ◆ 0 a 5 años → Nivel inicial (Jardin)
  - ◆ 6 a 12 años → Nivel Primario
  - ◆ 13 a 18 años → Nivel Secundario
  - ◆ 19 a 51 años → Nivel Terciario
- ➔ Normalizamos los nombres correspondientes a atributos, tal que al utilizarlos como claves resulte más declarativo
- ➔ Para los establecimientos educativos, tomamos únicamente la primer dirección de mail en el campo mail, en caso de haberse ingresado mas de una direccion.

- ➔ **Análisis de Datos**

Consulta i: Cantidad de EE por nivel educativo y población por grupo etario por departamento

provincia	departamento	jardines	poblacion_jardin	primarias	poblacion_primaria	secundarios	poblacion_secundar
Buenos Aires	LA MATANZA	333	59108	335	225872	336	181212
Buenos Aires	LA PLATA	218	19886	201	77998	211	67326
Buenos Aires	LOMAS DE ZAMORA	168	19589	179	76967	191	65257
Buenos Aires	GENERAL PUEYRRE	177	15895	168	62565	173	57730
Buenos Aires	QUILMES	158	18414	145	70881	151	60085
Buenos Aires	MORENO	120	19407	138	76357	134	60359
Buenos Aires	ALMIRANTE BROWN	139	17241	137	67913	146	58227
Buenos Aires	MERLO	108	18543	120	71541	125	59718
Buenos Aires	LANUS	116	11148	117	44506	113	39855
Buenos Aires	PILAR	108	12657	112	49944	105	41528
Buenos Aires	TIGRE	120	12571	111	51961	107	45401
Buenos Aires	GENERAL SAN MAR	97	10861	98	44539	96	39634

Esta tabla se encuentra en la carpeta de consultasSQL con el nombre consulta_i. Este es un recorte de la misma.

Podemos ver que: contamos con mayor cantidad de establecimientos de determinado nivel, mientras mayor cantidad de personas de su rango etario correspondiente se encuentren en su respectivo departamento.

Consulta ii: Bibliotecas Populares fundadas desde 1950 por departamento

provincia	departamento	cantidad_BP_funda
Buenos Aires	LA PLATA	15
Buenos Aires	LA MATANZA	15
Buenos Aires	MORENO	13
Buenos Aires	TIGRE	12
Buenos Aires	BAHIA BLANCA	12
Buenos Aires	GENERAL SAN MAR	10
Buenos Aires	TANDIL	10

Esta tabla se encuentra en la carpeta de consultasSQL con el nombre consulta_ii. Este es un recorte de la misma.

Notamos que la mayor parte de las Bibliotecas Públicas fundadas desde 1950 se encuentran en zonas de mayor densidad poblacional. Es decir, los departamentos mas poblados cuentan con mayor cantidad de bibliotecas publicas.

Consulta iii: Cantidad EEs, BPs y Población por departamento

provincia	departamento	cantEE	cantBP	poblacion_total
Buenos Aires	25 DE MAYO	120	1	24571
Buenos Aires	9 DE JULIO	115	2	36643
Buenos Aires	ADOLFO ALSINA	64	2	11713
Buenos Aires	ADOLFO GONZALES	47	2	8965
Buenos Aires	ALBERTI	43	2	8708
Buenos Aires	ALMIRANTE BROWN	504	11	440662
Buenos Aires	ARRECIFES	63	1	23322
Buenos Aires	AVELLANEDA	367	15	259086

Esta tabla se encuentra en la carpeta de consultasSQL con el nombre consulta_iii. Este es un recorte de la misma.

Notamos que, en la misma linea que el analisis anterior, los departamentos de provincias mas pobladas (Bs As, CABA, Cordoba) cuentan con mayor cantidad de establecimientos educativos y bibliotecas populares.

Consulta iv: Dominio más usado por BPs en cada departamento

Provincia	Departamento	Dominio_mas_frecu
Buenos Aires	ADOLFO ALSINA	hotmail
Buenos Aires	ADOLFO GONZALES	yahoo
Buenos Aires	ALBERTI	live
Buenos Aires	ALMIRANTE BROWN	yahoo
Buenos Aires	ARRECIFES	yahoo
Buenos Aires	AVELLANEDA	yahoo
Buenos Aires	AZUL	yahoo
Buenos Aires	BAHIA BLANCA	yahoo

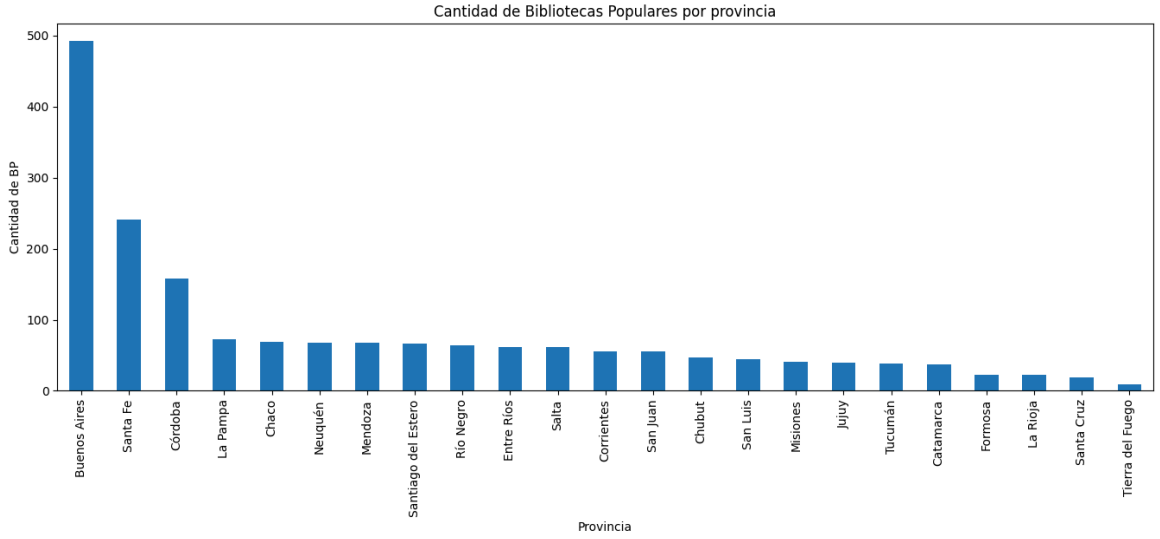
Esta tabla se encuentra en la carpeta de consultasSQL con el nombre consulta_iv. Este es un recorte de la misma.

Con este análisis, notamos que, en su mayoría, las BP cuentan con dominio de correo hotmail y yahoo, seguidos por Gmail. En otros casos puntuales, tenemos dominios propios de cada biblioteca.

➔ **Análisis de Visualizaciones**

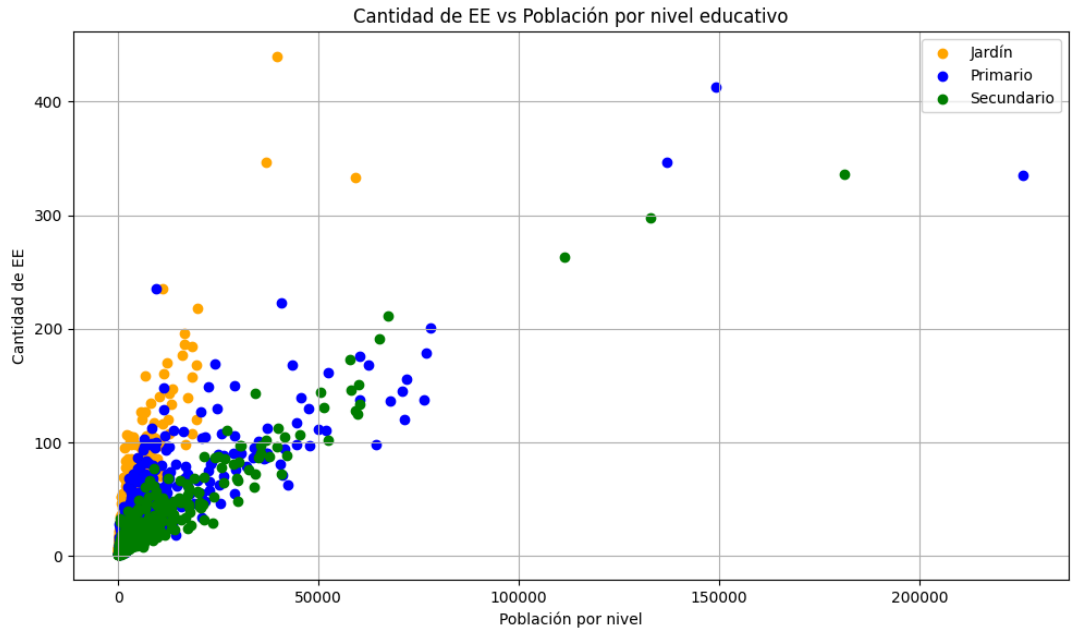
i) Cantidad de Bibliotecas Populares por provincia

El gráfico de barras muestra la cantidad total de Bibliotecas Populares por provincia. Se observa que **Buenos Aires** lidera ampliamente, seguida por **Córdoba** y **Santa Fe**. Esta distribución parece correlacionarse con los niveles de urbanización y concentración poblacional. En contrapartida, las provincias del norte y sur del país presentan una menor presencia de BP, lo cual podría atribuirse a factores geográficos, demográficos o económicos que dificultan la creación o el sostenimiento de estas instituciones.



ii) Cantidad de EE en función de la población por nivel educativo

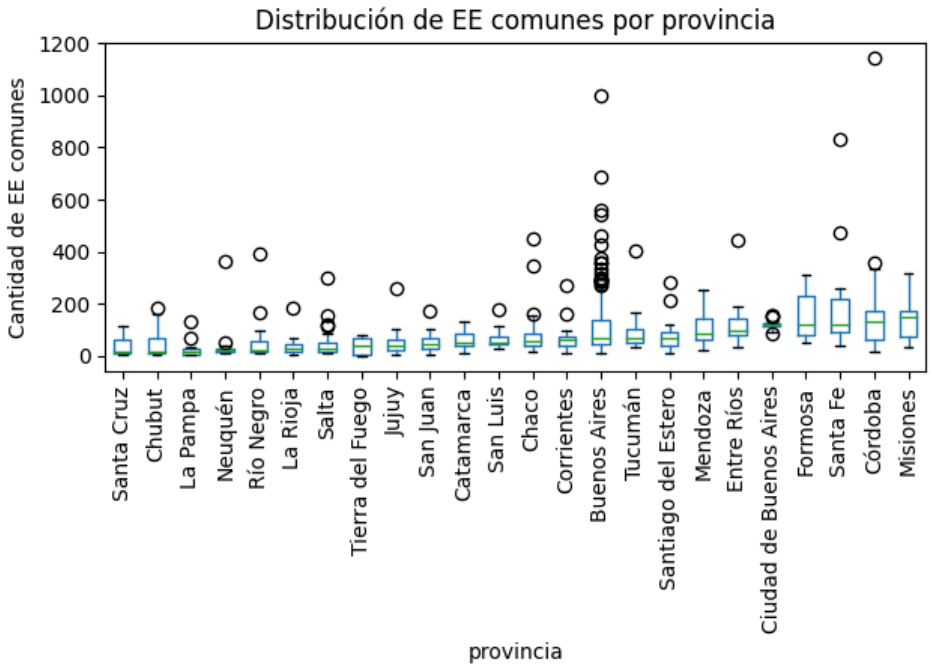
Este gráfico de dispersión representa la relación entre la cantidad de establecimientos educativos comunes por nivel (jardín de infantes, primario y secundario) y la población correspondiente a cada grupo etario. Se observa una **correlación esperada**: cuanto mayor es la población en edad correspondiente, mayor es la cantidad de EE. Sin embargo, se identifican ciertos departamentos con mayor cantidad de EE que lo esperado, lo cual podría deberse a criterios de distribución territorial que priorizan el acceso en zonas rurales o alejadas, independientemente de la densidad poblacional.



---

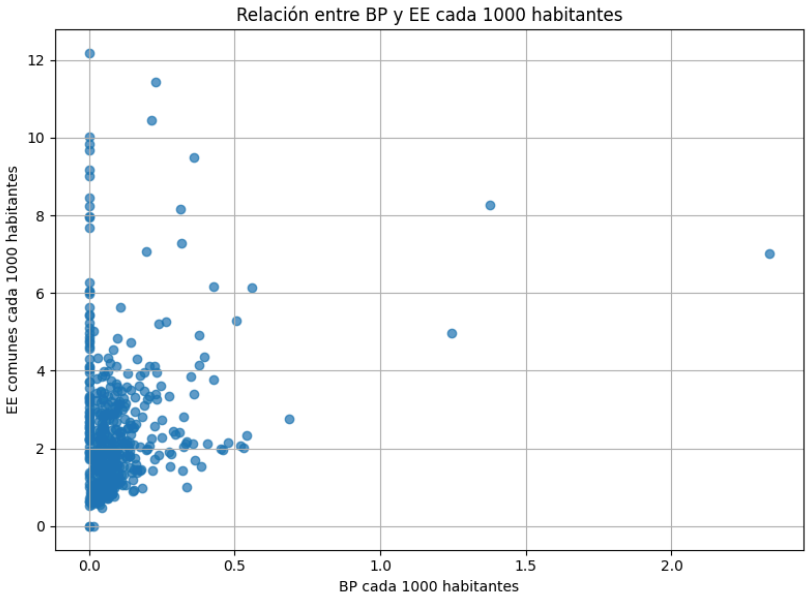
iii) Distribución de EE comunes por departamento en cada provincia

El boxplot muestra la **variabilidad de EE comunes entre departamentos** dentro de cada provincia. El ordenamiento por mediana permite observar cuáles provincias tienen mayor "centralidad" en su distribución. **Buenos Aires**, por ejemplo, presenta una dispersión muy amplia, lo cual refleja su gran heterogeneidad territorial. Por otro lado, provincias como **La Pampa** o **Formosa** muestran una distribución más homogénea, con menor diferencia entre departamentos. Esto puede estar relacionado con políticas educativas más uniformes o con una estructura territorial menos fragmentada.



iv) Relación entre BP y EE por cada 1000 habitantes

Este gráfico explora la relación entre la cantidad de Bibliotecas Populares y Establecimientos Educativos comunes **normalizada por población**. Se observa que **departamentos con baja población tienden a tener más instituciones por cada 1000 habitantes**, lo cual sugiere un esfuerzo por garantizar el acceso equitativo a servicios culturales y educativos. En departamentos densamente poblados, la proporción tiende a disminuir, aunque no necesariamente indique menor cobertura absoluta. Este análisis permite evaluar no solo la cantidad, sino también la **intensidad de acceso institucional en relación con la población**.



Conclusiones

Nosotros nos propusimos explorar si existe una relación entre la cantidad de establecimientos educativos y la cantidad de bibliotecas populares en las distintas provincias del país. A partir del análisis de las tablas generadas mediante consultas SQL y los gráficos construidos sobre ellas, pudimos obtener algunas conclusiones relevantes.

En primer lugar, observamos que la cantidad de establecimientos educativos por provincia guarda una relación positiva con la cantidad de habitantes, como se logra evidenciar en el gráfico "Cantidad EE en función de la población". En cambio, la cantidad de bibliotecas populares no presenta una relación directa con la población provincial, tal como se observa en el gráfico "Cantidad de BP Provincia". Además, al analizar la relación entre EE y BP cada mil habitantes, notamos que no existe una correspondencia lineal entre ambas variables. Incluso, en algunos casos, los departamentos con mayor cantidad de EE por mil habitantes no cuentan con ninguna biblioteca popular. Este dato resulta especialmente llamativo, ya que revela que existen departamentos sin presencia de BP, mientras que los EE están presentes en al menos uno por departamento.

En conclusión, no se encontró una relación clara entre la cantidad de establecimientos educativos y la de bibliotecas populares a nivel departamental. Ambos indicadores parecen variar de forma independiente. En todo caso, podría afirmarse que la cantidad de EE está más influenciada por la estructura administrativa y urbana de cada provincia, como se aprecia en el gráfico "Cantidad de EE por Departamento por Provincia".