



Universidad de
SanAndrés

Nivel educativo y reincidencia delictiva:

Un análisis del sistema penitenciario de la provincia de Buenos Aires

Ciencias de Datos

7 de diciembre 2024

Profesor: Ignacio Spiousas

Zahira Chaia

Martina Lucas

Introducción

Según el Censo Nacional de 2023, en Argentina había 111.967 personas privadas de su libertad, sin embargo, el sistema penitenciario permanece profundamente marginado del interés y conocimiento de la sociedad en general. La mayoría de las personas desconocen por completo su funcionamiento, lo que refuerza la invisibilidad de éste sector de la sociedad y dificulta la mejora del sistema en pos de la readaptación social.

Actualmente, en el sistema penitenciario argentino se establece la obligatoriedad de brindar a los presos educación formal y no formal (Ley Penitenciaria Nacional, N.º 412). Con este marco legal, es importante evaluar la relación de la educación con la delincuencia, considerando que el acceso limitado a instituciones educativas en ciertas áreas parecería estar relacionado con la prevalencia de la misma y con la reincidencia de los internos (Millán-Valenzuela et al., 2019).

Diversos estudios como el de Biswal & Mishra, 2020 o Schepers, 2017 mencionan el rol de factores externos como el contexto socioeconómico o el nivel educativo en relación con la tendencia a delinquir. Con esta base, se pone en evidencia la relevancia del contexto de la persona, tanto en la probabilidad de delinquir como en la dificultad de reinserirse en la sociedad, ligada a la repetición de acciones delictivas. Teniendo esto en cuenta, es muy importante evaluar qué aspectos del sistema pueden mejorarse para proponer intervenciones efectivas en evitar la reincidencia.

Desde esta base, en este informe nos planteamos responder a la pregunta: *¿El nivel educativo alcanzado por los individuos es un factor predictor de la reincidencia delictiva en los centros de detención de la provincia de Buenos Aires?*

Llevar a cabo esta propuesta representa un aporte valioso a la literatura existente debido a la escasez de estudios locales que analicen el impacto de la educación en la reincidencia delictiva en Argentina. Aunque investigaciones internacionales vinculan la educación con una menor propensión a reincidir, los estudios en contextos latinoamericanos son limitados, lo que deja un vacío importante en la comprensión de esta relación en sistemas penitenciarios como el argentino. Además, el estudio tiene implicancias prácticas, ya que puede ofrecer lineamientos para ajustar estrategias de reinserción social basadas en la educación, favoreciendo la empleabilidad y la reintegración de los internos en la sociedad.

Revisión bibliográfica

En primer lugar, es importante definir a qué nos referimos con “reincidencia”, definida por el CELIV (2022) como “las personas que cometieron delitos punibles con penas de privación de libertad, que hayan estado privados de libertad anteriormente, siempre y cuando no haya pasado desde su cumplimiento, más tiempo que la duración de la condena impuesta.”

En cuanto a la relación entre educación y reincidencia, estudios previos realizados en otros países establecen importantes hallazgos. Por ejemplo, Groot y van den Brink (2007) analizan la relación entre el nivel educativo y la delincuencia en Estados Unidos, encontrando una relación causal entre el aumento de los años de instrucción académica y una reducción en la probabilidad de cometer ciertos delitos,

(excluyendo evasión de impuestos). Además, destacan que factores como la edad, el género y las creencias religiosas también influyen en dicha probabilidad.

En concordancia con esto, el estudio de Esperian (2010) realizado también en Estados Unidos aborda el impacto de los programas educativos dentro de la prisión en la reincidencia. Los resultados respaldan la hipótesis de que los programas educativos dentro de la prisión reducen significativamente la reincidencia, con un énfasis particular en la influencia de completar una educación superior. Además, esta investigación resalta que los programas educativos también aumentan las posibilidades de empleo tras la liberación.

En Argentina, Natalia Ojeda (2024) llevó a cabo una investigación etnográfica enfocada en evaluar si la gestión penitenciaria promueve efectivamente la readaptación social de las personas privadas de la libertad. Aunque Ojeda adopta un enfoque crítico hacia el sistema penitenciario actual, sus argumentos se alinean con la Ley Penitenciaria Nacional N.º 412, que establece la obligatoriedad de brindar educación formal y no formal a los internos con el objetivo de fomentar su reinserción social.

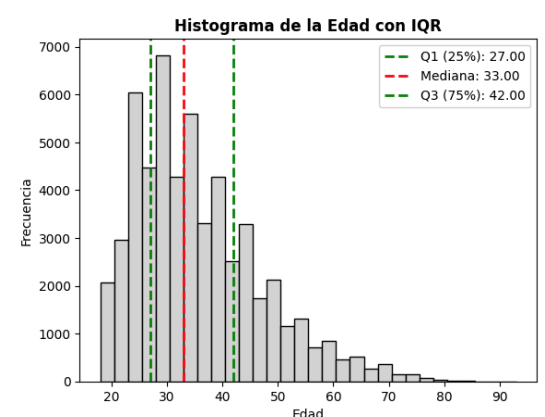
A partir de esta revisión bibliográfica resulta pertinente evaluar si en Argentina parece haber una relación entre el nivel educativo que alcanza una persona y la probabilidad de reincidencia. Se espera que alcanzar un mayor nivel educativo se relacione con una menor probabilidad predicha de reincidir. En este sentido se busca destacar la importancia de implementar planes educativos en los centros penitenciarios, ya que estos no solo pueden favorecer la resocialización de los internos, sino también generar beneficios para la sociedad en general

3. Base de datos

La información disponible proviene del Sistema Nacional de Estadísticas sobre Ejecución de la Pena (SNEEP), que lleva a cabo un censo anual en los establecimientos carcelarios. El SNEEP compila la estadística penitenciaria oficial del país y muestra la evolución y las características de la población privada de libertad en unidades penitenciarias. Particularmente extrajimos una base de datos que contiene datos únicamente de los penales de Buenos Aires en el año 2023.

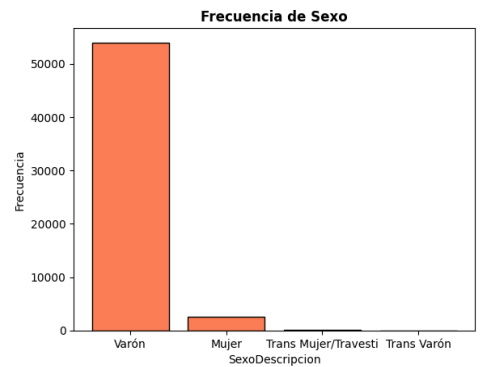
Comenzamos evaluando la base de datos de forma general. Ésta cuenta con 56736 observaciones y 84 columnas, 47 continuas y 37 categóricas. Se identificaron 44 registros duplicados. Además, se observó que el total de datos faltantes en todo el Data Frame es 651.851, lo que representa una cantidad significativa. Igualmente, detectamos que la gran mayoría de datos faltantes estaban concentrados en ciertas columnas específicas, como “MujerTieneHijosIntramuro” (54185); éstas son variables que aplican a unos pocos internos y no se relacionan con nuestro análisis. En nuestras variables de interés no se encontraron datos faltantes.

En cuanto a la distribución de edad de las personas privadas de libertad, la edad promedio de la población es 35 años. Se calcularon los cuartiles de la distribución. El primer cuartil (Q1), que indica el punto en el cual el 25% de los datos son menores, se encuentra en 27 años, mientras que el tercer cuartil (Q3), que indica el punto en el cual el 75% de los datos son menores, está en 42 años.

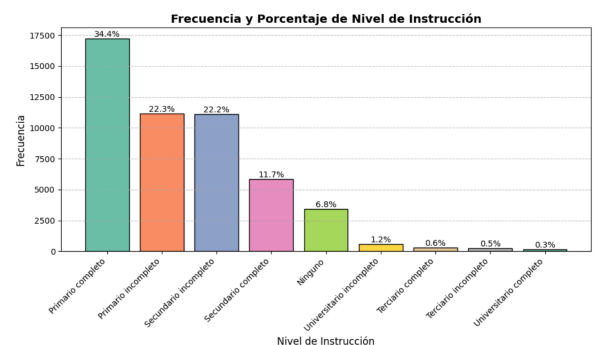


El género se presenta en la base como una variable categórica. La mayoría de la población penitenciaria en su mayoría de sexo masculino, con 53978

(aproximadamente el 95% de la población), 1,827 mujeres (3.73%) y 110 personas transgénero (0.22%), distribuidos entre trans mujeres y trans varones. En cuanto a la nacionalidad, Argentina es el país de origen de la gran mayoría de los internos, con 46,444 registros (95%), seguida por Paraguay, con 1,237 (2.5%) y Perú, con 350 registros (0.7%). La base de datos también incluye una pequeña representación de personas provenientes de otros países.



Analizamos la variable categórica referida al nivel de instrucción de la población penitenciaria. Se observa que la mayoría de los internos tienen una formación primaria, con 17,313 registros de primario completo (34%) y 11,319 de primario incompleto (22%). En conjunto, hay un 58% de los internos que no sobrepasaron el nivel primario. Luego, hay 10,926 registros de secundaria incompleta (22%) y 5,435 completa (11%). Además, se encuentra un pequeño porcentaje con nivel terciario y universitario incompleto o completo: 262 registros de terciario completo (0.62%) y 279 de universitario incompleto (0.66%), mientras que los que no tienen ningún nivel de instrucción ascienden a 3,816 registros (6.84%).

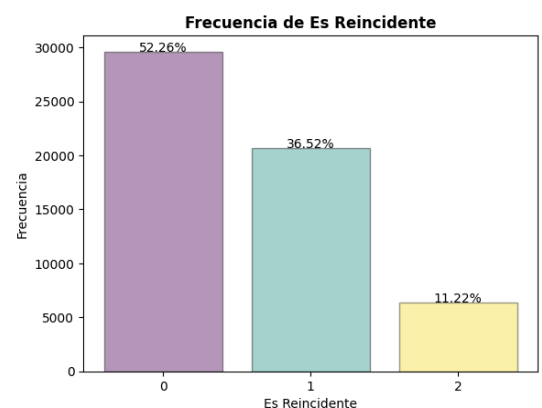


Por último, nuestra variable de outcome “EsReincidente”, presente en la base de datos como una variable categórica que va del 0 al 3, donde 0 representa los internos sin información, 1 lo primario/a (no reincidente), 2 los reiterante y 3 los reincidentes. La categoría “reiterante” es un término criminológico referente a casos particulares de personas que fueron procesadas y privadas de su libertad previamente pero no fueron condenadas formalmente (CELIV, 2022). En nuestro

análisis los tomaremos como residentes, ya que la distinción entre reiterantes y reincidentes no siempre se realiza y es principalmente de interés legal. De esta forma, transformamos la variable en una de tipo binario, donde 0 corresponde a los internos primarios y 1 a los reiterantes o reincidentes.

Dicho esto, al analizar la variable se encontraron 29616 registros iguales a 0 (52%), 20694 personas identificadas como “primarias” (36,5%) y 6358 reincidentes (%11).

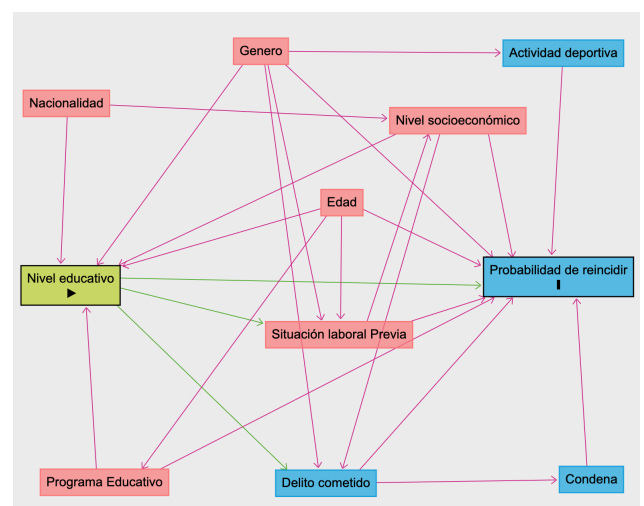
De cara a nuestro análisis, es importante resaltar que no vamos a tener en cuenta los registros iguales a 0 para entrenar el modelo, sinó que tomaremos a variable como una dummy, asignando 0 a la categoría “primario/a” y 1 a la categoría “reincidente”.



Metodología

Para abordar la pregunta de investigación, proponemos utilizar un modelo de regresión logística, con el objetivo de medir cómo cada una de las variables se relaciona con la probabilidad de que ocurra la variable respuesta. Este modelo es adecuado para variables dependientes binarias, como en este caso, donde el resultado sería 1 si la persona es reincidente y 0 en caso contrario.

El primer paso de nuestro análisis fue plantear el siguiente *Directed Acyclic Graph* donde plasmamos las variables relevantes para la relación estudiada a partir de la revisión bibliográfica y la observación de nuestra base de datos.



Comenzamos incluyendo variables de tipo demográfico como (edad, género y nivel socioeconómico), ya que estudios previos como el de Groot y van den Brink (2007) y el informe CELIV (2022) destacan el rol de éste tipo de indicadores en la reincidencia y su relación con el nivel educativo alcanzado por una persona (García de Fanelli, 2014). Por otro lado, la bibliografía respalda la inclusión de variables mediadoras como el delito cometido, la realización de actividad deportiva y la situación laboral previa (Groot y van den Brink, 2007; CELIV, 2022).

Comenzamos utilizando R Studio para identificar los backdoor paths, es decir, los caminos no directos que conectan la variable predictora y el outcome. Luego de analizar las posibles variables que podíamos controlar identificamos que un conjunto de ajuste efectivo es “Género”, “Edad” y “Nacionalidad”.

A continuación se encuentra nuestro modelo teórico de regresión logística:

$$\text{logit}(Y) = \beta_0 + \beta_1 . \text{Nivel Educativo} + \beta_2 . \text{Género} + \beta_3 . \text{Edad} + \beta_4 . \text{Nacionalidad} + \varepsilon$$

En primer lugar, antes de desarrollar el modelo se realizará una transformación de las variables mediante one-hot encoding, técnica utilizada para convertir variables categóricas en un formato numérico, de manera que el modelo pueda procesarlas adecuadamente. Además, determinaremos un valor $\alpha = 0.05$, ampliamente aceptado como un estándar para controlar el riesgo de cometer errores de Tipo I y una potencia $(1 - \beta)$ de 0.95 es deseable.

Para evaluar y validar el modelo de regresión logística, se implementaría la validación cruzada (K-Fold Cross-Validation). Ésta técnica divide los datos en múltiples subgrupos que se usan para entrenar y evaluar el modelo y es muy útil para evitar el *overfitting* y garantizar la generalización del modelo. Por otro lado,

aplicaríamos medidas de desempeño como la matriz de confusión para identificar verdaderos y falsos positivos y la curva ROC/AUC para analizar la capacidad del modelo de distinguir entre clases. Además, se podrían calcular métricas como la precisión, el *recall* y el F1-score para evaluar el equilibrio entre los errores del modelo.

Conclusiones y limitaciones

Este estudio busca analizar la relación entre el nivel educativo y la reincidencia delictiva en los centros penitenciarios de la provincia de Buenos Aires. A partir de la revisión bibliográfica, se plantea la hipótesis de que un mayor nivel educativo alcanzado podría relacionarse con una menor probabilidad predicha de reincidir. El modelo de regresión logística propuesto permitiría evaluar esta relación controlando por variables como género, nacionalidad y edad.

Asimismo, se espera identificar diferencias relacionadas a las variables sociodemográficas. Por ejemplo, las mujeres privadas de libertad podrían mostrar menores tasas de reincidencia comparadas con los hombres, alineándose con estudios previos que destacan cómo género y edad pueden influir en la propensión a reincidir.

El proyecto presenta ciertas limitaciones metodológicas y contextuales a considerar. En primer lugar, la muestra presenta cierta homogeneidad, al utilizar datos únicamente de penales de la provincia de Buenos Aires. Éste factor genera que los resultados puedan no ser generalizables a otras provincias o al país en su totalidad, especialmente considerando las diferencias regionales en el sistema penitenciario argentino.

Otra limitación a considerar es el desbalance de clases en las variables “reincidente”, “sexo”, ya que se observa un porcentaje relativamente bajo de reincidentes frente a primarios y de mujeres frente a hombres. Esto podría ser un problema para el ajuste de nuestro modelo, ya que la regresión logística puede ser menos efectiva con una distribución desequilibrada. Si este fuera el caso podrían considerar métodos de transformación de las variables como SMOTE (Synthetic Minority Over-sampling Technique). Aunque ésta técnica podría ser útil para objetivo es aumentar artificialmente el número de observaciones en las categorías minoritarias, ésta técnica presenta ciertas limitaciones como el riesgo de generar ruido.

A pesar de estas limitaciones, el proyecto tiene el potencial de generar información valiosa sobre la relación entre la educación y la reincidencia en Argentina, proporcionando una base para futuras intervenciones. Esperamos que estos resultados contribuyan a subrayar la importancia de la educación como una prioridad para la sociedad Argentina. Es importante comprender el rol de la educación como factor determinante para las personas, que tiene la potencialidad no solo de impulsar el bienestar general de la sociedad sino también de reinventar las posibilidades de ciertos grupos de la sociedad en situación de vulnerabilidad. Los resultados podrían contribuir a la formulación de políticas más efectivas de reinserción social que prioricen la educación como una herramienta clave en la reducción de la reincidencia delictiva.

Referencias

- Ley Penitenciaria Nacional, N.º 412 (1958). Recuperado de <http://www.saij.gob.ar/412-nacional-ley-penitenciaria-nacional-Ins0001686-1958-01-14/123456789-0abc-defg-g68-61000scanyel>
- Millán-Valenzuela, Henio, & Pérez-Archundia, Eduardo. (2019). Educación, pobreza y delincuencia: ¿nexos de la violencia en México?. *Convergencia*, 26(80), 01. <https://doi.org/10.29101/crcs.v26i80.10872>
- Groot, W., & van den Brink, H. M. (2007). The effects of education on crime. *Applied Economics*, 42(3), 279–289. <https://doi.org/10.1080/00036840701604412>
- GARCÍA DE FANELLI, ANA. (2014). INCLUSIÓN SOCIAL en la Educación superior argentina: Indicadores y políticas en torno al acceso y a la graduación. *Páginas de Educación*, 7(2), 124-151. Recuperado en 16 de abril de 2024, de http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S1688-74682014000200007&lng=es&tlng=es.
- Esperian, J. H. (2010). The effect of prison education programs on recidivism. *Journal of Correctional Education (1974-)*, 61(4), 316–334. Recuperado de <http://www.jstor.org/stable/23282764>
- Centro de Estudios Latinoamericanos sobre Inseguridad y Violencia (CELIV). (2022). *Reincidencia en Argentina*. Universidad Nacional de Tres de Febrero. Recuperado de <http://www.celiv.untref.edu>.