{ARcenso}: primeros pasos desarrollando un paquete en comunidad con rOpenSci

**Palabras clave:**  R, paquete, censo, Estadísticas de población, Argentina, comunidad.

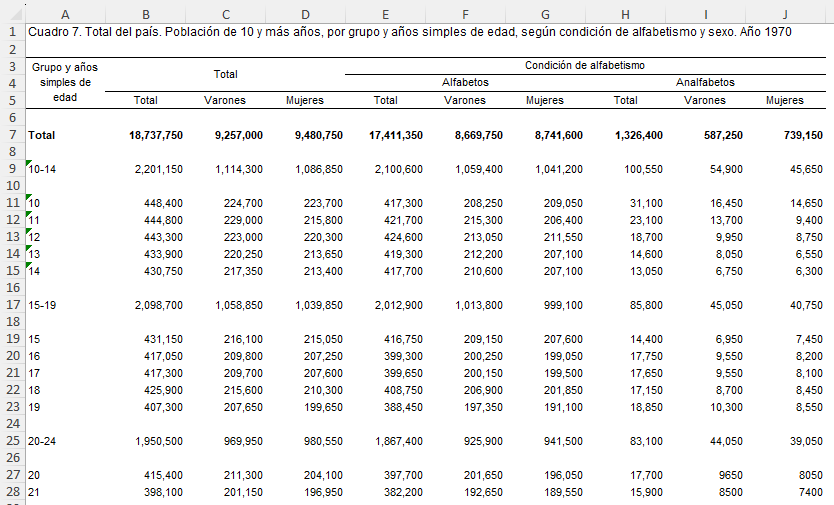
En el marco del programa de Campeones y Campeonas de rOpenSci[[1]](#footnote-0) y con el objetivo de identificar, reconocer y recompensar a los miembros apasionados de la comunidad que ayudan a la comunidad a crecer y mejorar, nace el proyecto de crear el paquete {ARcenso}.



{ARcenso} es un paquete en desarrollo que permite acceder a los datos oficiales de los censos nacionales de población en Argentina provenientes del Instituto Nacional de Estadísticas y Censos - INDEC desde 1970 a 2022.

Actualmente los resultados históricos de los censos nacionales de población están disponibles únicamente a través de archivos en formato excel y en libros físicos, no pensados para ser tratados o procesados como base de datos (Figura 1). Además a través de los censos la presentación de los datos no está homogeneizada dificultando la comparación a través del tiempo de la información disponible.

Figura 1. Información disponible en la web de INDEC.

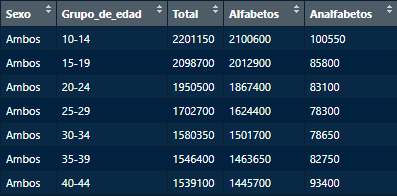


En este sentido contar con un paquete de información censal permitirá al sector público y privado, a los ciudadanos y a otros actores de la sociedad acceder a información actual e histórica sobre la población, los hogares y las viviendas de Argentina de una manera más accesible amplificando el ecosistema de la ciencia abierta.

El proyecto {ARcenso} consiste en generar un paquete que permita disponer de datos censales homogeneizados y listos para su uso. Lo que implica no sólo la conversión de los tabulados publicados, sino también aportar herramientas que ayuden a la visualización de los datos y el armado de funciones para simplificar la comparabilidad y utilización conjunta de la información de múltiples años censales.

Disponer de un paquete con estas características requiere de un extenso proceso de revisión, consolidación y conversión de tabulados por cada indicador, por cada referencia geográfica y por cada año censal. Acompañado de un marco metodológico y conceptual para obtener tablas en un formato “tidy” (Figura 2) preparado para ser utilizados de forma directa.

Figura 2. Mismo tabulado en formato “tidy”



En esta presentación nos proponemos compartir la experiencia de armar un paquete desde cero acompañados por rOpenSci y toda su comunidad, con las capacitaciones, la mentoría 1:1 y la sesiones de coworking que delimitaron los puntos de partida y estrategias para el armado del proyecto, la toma de decisiones, los desafíos y aprendizajes a la hora de construir una herramienta que le permita a la comunidad de usuarios de R, a los usuarios de estadísticas públicas y a la ciudadanía un mayor acercamiento a la producción pública de datos.

1. [¡Convocatoria abierta para la presentación de candidaturas al Programa de Campeonas y Campeones de rOpenSci!](https://ropensci.org/es/blog/2023/07/14/launch-champions-program-es/) [↑](#footnote-ref-0)