

Experiencia del INE Uruguay en el uso de R para el cálculo del IPC

Abstract

El Índice de Precios del Consumo (IPC) es un indicador del nivel general de precios de los bienes y servicios que los hogares residentes consumen. La variación que experimenta el IPC a lo largo del tiempo es una medida de la inflación en la economía. El Instituto Nacional de Estadística (INE) es el encargado de publicar mensualmente su estimación, así como de introducir mejoras metodológicas para su cálculo y llevar adelante cambios de base a partir de ediciones renovadas de la Encuesta de Gastos e Ingresos de los Hogares (ENGIH). En 2022 el INE presentó la última actualización del índice.

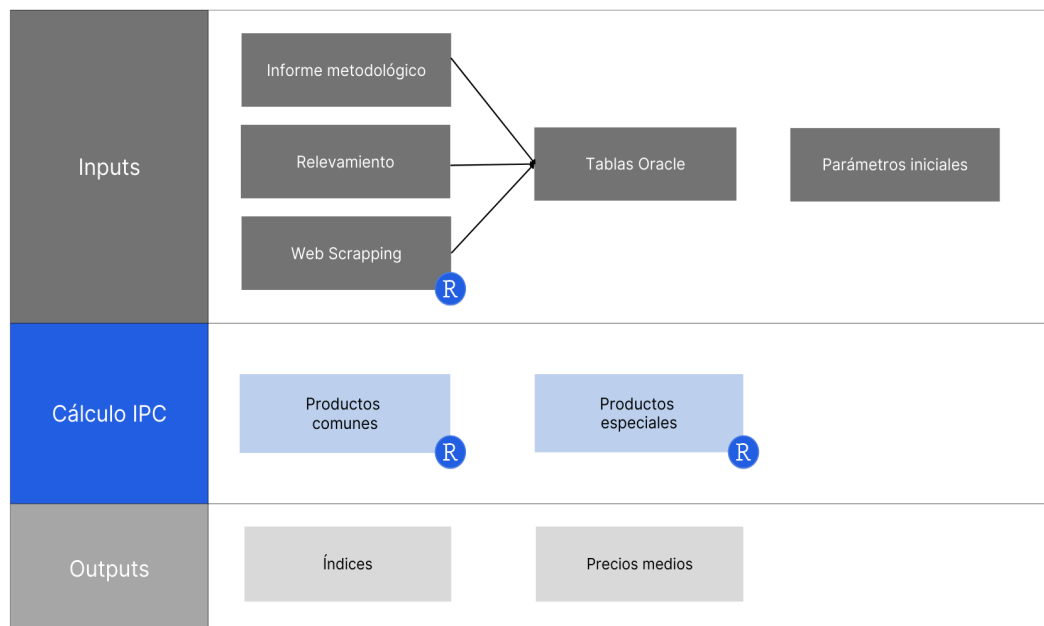
Como parte de un proceso institucional de migración progresiva desde softwares con licencia, la programación del nuevo cálculo del IPC se desarrolló en R. En este proceso, se utilizó principalmente el ambiente *tidyverse*, junto con otras librerías auxiliares como *RJDBC* y *rJava* para la conexión con la base de datos Oracle, donde se almacena la información relevante. Además, con el objetivo de modernizar y mejorar el cálculo, se usaron las librerías *rvest* y *RSelenium* para obtener datos de pasajes de avión a través de *Web Scraping*. Asimismo, se están desarrollando procedimientos para ampliar las fuentes de datos mediante el uso de *APIs*.

Para desarrollar la programación y definir la estructura de los datos de entrada y salida, se formaron dos grupos de trabajo. Cada grupo elaboró un código independiente, con el objetivo de aumentar la robustez mediante un control cruzado de los resultados. Ambos equipos partieron del documento metodológico y la base de datos que contenía la información recolectada en el campo. Durante el proceso, fue fundamental mantener una interacción constante y una retroalimentación continua entre el grupo de programación y todas las áreas involucradas, incluyendo el equipo de relevamiento, informática y metodología.

La migración hacia R ha facilitado la participación, el acceso y la comprensión del cálculo por parte de un número mayor de funcionarios/as, dado que es un lenguaje cada vez más utilizado dentro del INE. Esta migración también ha permitido reducir los costos asociados a licencias y a la contratación de desarrolladores externos.

Palabras clave: IPC, Estadísticas económicas, Uruguay, Tidyverse, Web Scraping

Figura 1. Esquema del proceso



- **Informe metodológico:** documento que contiene la metodología del cálculo realizado por un consultor externo
- **Relevamiento:** precios y descripción de productos relevados de manera presencial, a través de formularios web e información enviada por grandes superficies.
- **Web Scrapping:** datos relevados automáticamente desde la web, por medio de un código R.
- **Tablas Oracle:** los datos anteriores se registran en tablas Oracle que serán insumo directo para el cálculo.
- **Parámetros iniciales:** refieren al año, mes y semana de cálculo que se especifican en la Web donde se ejecuta el cálculo y son relevados por R.
- **Productos comunes:** engloban la gran mayoría de productos considerados en la canasta, se caracterizan por tener un procedimiento de cálculo único.
- **Productos especiales:** a diferencia de los comunes tienen distintas casuísticas relacionadas tanto a la fuente como al cálculo.
- **Índices:** comprende los índices a nivel de producto, así como en los distintos niveles de agregación. Es el output principal del cálculo.
- **Precios medios:** promedio aritmética de los precios agrupados según variedad y tipo de empresa.

Referencias

INE (2022) Metodología Índice de Precios del Consumo. [Link](#)

John Harrison (2022). RSelenium: R Bindings for 'Selenium WebDriver'. R package version 1.7.9.

Simon Urbanek (2021). rJava: Low-Level R to Java Interface. R package version 1.0-6.

Simon Urbanek (2022). RJDBC: Provides Access to Databases Through the JDBC Interface. R package version 0.2-10.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Golemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). “Welcome to the tidyverse.” *_Journal of Open Source Software_*, *4*(43), 1686. doi: 10.21105/joss.01686