

blockr universe: un ecosistema abierto para workflows sin código y análisis de datos en R

Autores: Nicolas Bennet, David Granjon, Christoph Sax, Angelica Becerra

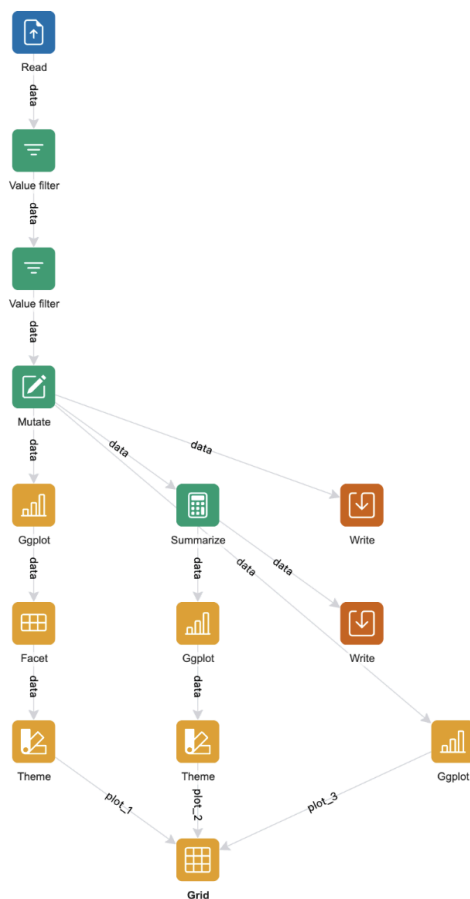
Palabras clave: no-code, frameworks expandibles, pipelines

Abstract

¿Se imaginan realizar análisis de datos complejos en R sin escribir una sola línea de código? Con blockr, esta visión se hace realidad. blockR es un framework que permite crear flujos de trabajo de análisis de datos mediante una interfaz gráfica intuitiva con solo apuntar y hacer clic. El ecosistema blockr puede ser utilizado por personas con poca o ninguna experiencia en programación, mientras que brinda a los desarrolladores una base que se puede extender para crear bloques y plugins. En esta charla, nos enfocaremos en el uso de la herramienta con el universo blockr y sus capacidades actuales para el análisis de datos, cerrando con una breve invitación a la comunidad latinoamericana de R a participar en su evolución.

¿Qué es blockr?

El concepto central de blockR son los "bloques": unidades fundamentales que realizan tareas específicas como cargar datos, filtrar, visualizar o modelar. Estos bloques se conectan mediante grafos acíclicos dirigidos (DAG por sus siglas en inglés) para crear flujos de trabajo completos y reproducibles, como se observa en la siguiente imagen.

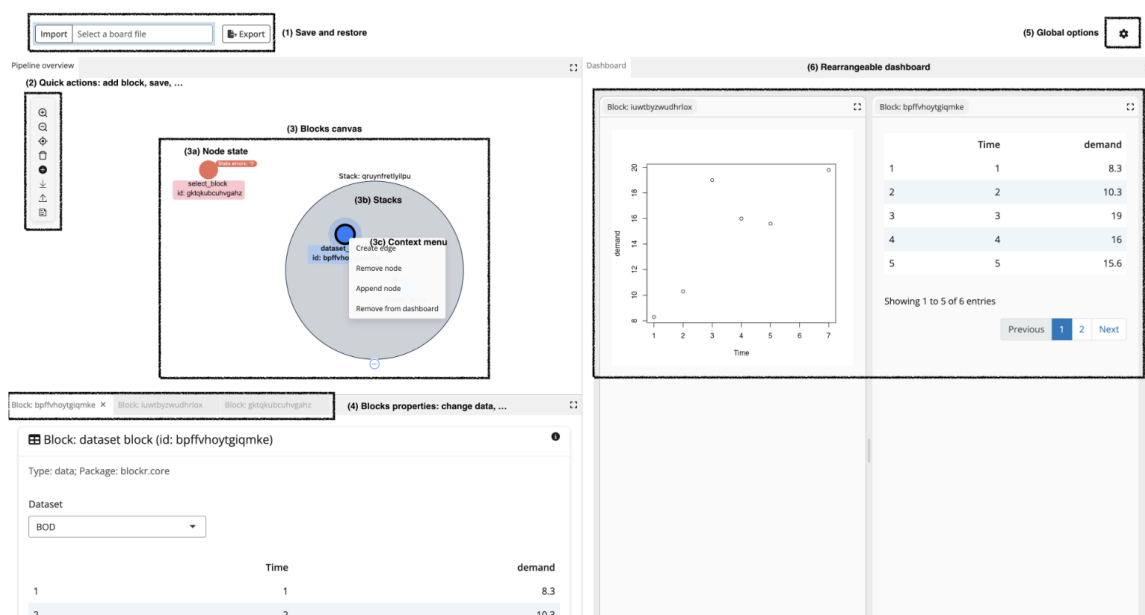


Gracias a esta estructura, blockr es un framework para análisis de datos que sirve a dos tipos de usuarios. Por un lado, permite que personas sin experiencia en programación creen análisis y workflows conectando bloques ya listos de forma sencilla. Y adicionalmente, ofrece a los desarrolladores las herramientas necesarias para construir bloques personalizados y expandir sus funcionalidades. En esta oportunidad nos centraremos en el uso práctico de blockr, mostrando cómo construir análisis mediante las extensiones ya existentes.

Capacidades actuales del ecosistema blockr

El paquete `blockr.core`, el cual es el núcleo del ecosistema, proporciona la infraestructura esencial y puede usarse de forma independiente o integrarse con otros paquetes. El ecosistema ya cuenta con extensiones especializadas como:

- Interfaz de usuario para la construcción visual de las pipelines.



- **blockr.dplyr**: Integración con dplyr para manipulación de datos
- **blockr.ggplot**: Bloques para generar visualizaciones de datos
- **blockr.ai**: Bloques que integran servicios de inteligencia artificial para potenciar el análisis de datos a través de interacciones con LLMs (Large Language Models)
- **blockr.io**: Manejo de múltiples formatos de entrada y salida de datos (CSV, XPT, etc.)

Algunas de las características claves de blockr son: la flexibilidad para reorganizar bloques dinámicamente, las herramientas para programadores para crear extensiones, la posibilidad de exportar el código para facilitar la reproducibilidad, y la interactividad y retroalimentación inmediata durante la construcción de pipelines.

Queremos expandir el blockr universe

Nuestros próximos pasos incluyen expandir bloques, mejorar herramientas para programadores, desarrollar tutoriales y construir una comunidad activa. La comunidad puede contribuir de diversas formas: desarrolladores creando extensiones blockR con sus paquetes, expertos diseñando bloques especializados para sus áreas (finanzas, bioinformática, etc.), y diseñadores mejorando la interfaz y accesibilidad.