

LatinR, Conferencia Latinoamericana sobre el Uso de R en Investigación + Desarrollo

De Excel a R: ¿cambio de una herramienta o apertura a nuevas posibilidades?

Carolina Soledad Fracchia¹ and Matías Lopez-Rosenfeld^{1,2}

¹ Unidad de Neurobiología Aplicada, Centro de Educación Médica e Investigaciones Clínicas, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

carolinafracchia@gmail.com, matiaslopez@gmail.com

² ICC-DC, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires

Resumen Keywords: R, repetición de análisis, metodología extensible, neurociencia

La psicología del desarrollo y la neurociencia cognitiva, han analizado la relación de variables socioambientales (e.g., pobreza) con el desempeño de niños y niñas en pruebas cognitivas durante programas de intervención. Los mismos, suelen involucran múltiples instancias de evaluación, que son comparadas posteriormente a través de diversos análisis estadísticos.

En este trabajo realizamos la comparación entre el análisis con planillas de cálculo en contraposición con R. Se documenta el proceso de migración y las perspectivas en un caso de estudio real. El esquema de trabajo obtenido (Figura 1) se focalizó en separar en módulos: la limpieza y estandarización de las bases de datos, la unificación de los datos, los distintos métodos de análisis, la combinación de análisis múltiples. Esto permitió incluir rutinas para chequear la integridad de los datos en todas las etapas (e.g., detección de registros duplicados), además de tener una arquitectura extensible para futuros análisis.

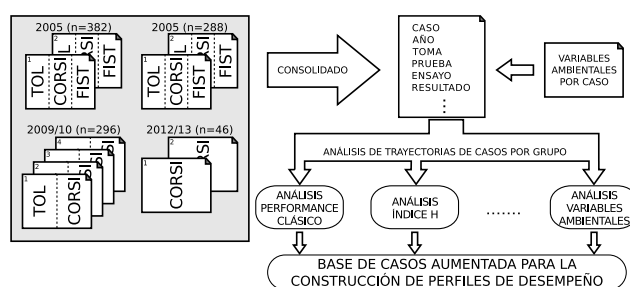


Figura 1. Esquema de trabajo

El pasaje a una metodología más flexible logró: (a) reducir errores humanos y tiempos de procesamiento; (b) incorporar nuevos datos con mayor facilidad; (c) extender los análisis con nuevos métodos en las posibles diversas direcciones; (d) reproducir el análisis con otros datos; y (e) empezar a generar una biblioteca de análisis plausible de ser compartida con la comunidad académica.