SAmuel · Manual de usuario y actualización vigente

Guía vigente · 2 de noviembre de 2026

Esta edición conserva el manual base y añade el suplemento del 2 de noviembre de 2026: perfiles personalizados por funciones y procedimientos, asignación administrativa y recorrido de navegación de ocho pasos. Las páginas originales conservan sus fechas y contexto.

Orden de lectura: portada de actualización, manual base y suplemento desde la página 100. Para navegación y permisos vigentes consulta el suplemento y la guía web de la aplicación.



EPINex · SAmuel

Sistema de análisis médico,
epidemiológico y estadístico

MANUAL TÉCNICO DE USUARIO

SAmuel web · Versión 1 · 29 de octubre de 2026

EPINex - Marca y edición comercial

Propósito: guiar desde la preparación de datos hasta la interpretación, la comunicación científica y la reproducción en R y Python.

Contenido

Contenido 2

¿Por qué utilizar SAmuel? 7

1. Accesibilidad, orientación y aprendizaje . . . 8
2. Preparación, calidad y comprensión de los datos . . . 8
3. Rigor estadístico y epidemiológico . . . 9
4. Métodos avanzados integrados . . . 10
5. Interpretación, comunicación y reporte . . . 10
6. Gráficos, mapas y presentación científica . . . 11
7. Reproducibilidad, continuidad y programación . . . 11
8. Privacidad, seguridad e integridad . . . 12
9. Acceso, continuidad y mantenimiento . . . 12
10. Rendimiento, verificación y confianza proporcionada . . . 12

Un puente entre el menú y el código . . . 13

SAmuel y las plataformas estadísticas de tercera generación 14

Por qué el grupo de comparación directa es pequeño . . . 14
Qué diferencia el recorrido de SAmuel . . . 14
Comparación de características, facilitadores y módulos . . . 15
Matriz funcional resumida . . . 16
Lectura experta de la comparación . . . 16

1. Propósito, alcance y filosofía de uso 18

EPINex y cómo citar SAmuel . . . 18
Convenciones del manual . . . 18
Transición comercial a EPINex . . . 18

1. Acceso, licencias y activación 19

2. Inicio rápido 20

Guía de navegación del menú principal 21

Cómo moverse sin perder el contexto . . . 21
BÁSICO · ¿pa' que más? . . . 21
INTERMEDIO · El atrevido . . . 23
AVANZADO · Un semiDIOS . . . 28
PROGRAMADOR · The Goat . . . 33
ACCESOS INDEPENDIENTES · siempre visibles . . . 34

3. Bases demostrativas incluidas 36

3.1 Veinte recorridos de práctica . . . 36

4. Flujo analítico seguro y reproducible 38

4.1 Antes de abrir el programa . . . 38
4.2 Dentro de SAmuel . . . 38
4.3 Después del análisis . . . 38

5. Carga, exploración y diseño de muestras 39

5.1 Formatos y estructura . . . 39
5.2 Paso a paso . . . 39
5.3 Decimales independientes por columna . . . 39

6. Diccionario, naturaleza y nivel de medición 40

6.1 Habilitar o inhabilitar variables para análisis . . . 40
6.2 Guía visual de tipo y nivel . . . 41
6.3 Ejercicio guiado: edad . . . 41

7. Datos faltantes e imputación 42

7.1 Conceptos . . . 42
7.2 Uso . . . 42

8. Análisis univariado 43

8.1 Medidas disponibles . . . 43
8.2 Ejemplo: mediana de edad . . . 43
8.3 Media, DE, mediana y RIC . . . 44
8.4 Frecuencias cualitativas múltiples . . . 44
8.5 Tabla 1 científica: toda la base en una ejecución . . . 44
8.6 Contraste frente a un valor . . . 45

9. Asociación, comparación y correlación 46

9.1 Dos variables cualitativas . . . 46
9.2 Cuantitativa entre dos grupos . . . 46
9.3 Tres o más grupos . . . 46
9.4 Correlación . . . 47
9.5 Muestras relacionadas . . . 47
9.6 Cómo se construye la sugerencia . . . 47

10. RR, OR y razón de prevalencias en tablas 2×2 48

Ejemplo: RR crudo . . . 48

11. Modelos multivariados: explicación y ajuste 49

11.1 Modelos disponibles . . . 49
11.2 Modelos cuando las muestras son pareadas o dependientes . . . 49
11.3 Paso a paso: RR ajustado . . . 50
11.4 OR ajustado . . . 50
11.5 Confusión, mediación y colisión . . . 50
11.6 Diagnósticos mínimos . . . 50

12. Kaplan-Meier, Log-rank y Cox 51

12.1 Kaplan-Meier . . . 51
12.2 Log-rank . . . 51
12.3 Cox . . . 51

13. Desarrollo y evaluación de modelos predictivos 52

Ruta y reporte . . . 52

14. Cómo elegir la prueba o el modelo 53

15. Figuras científicas y editor de formato 54

15.1 Catálogo . . . 54
15.1.1 Barras de error . . . 54
15.1.2 Georreferenciación (motor Leaflet) . . . 54
15.1.3 Especiales Plotly . . . 55
15.2 Paso a paso: forest plot . . . 56
15.3 Menú completo de formato . . . 56
15.4 Calidad de publicación . . . 57

16. Interpretación, tablas e historial integral 58

16.1 Tres pestañas . . . 58
16.2 Redactar un resultado . . . 58

17. Reproducibilidad y extensión con R 60

Instalar paquetes . . . 60

18. Reproducibilidad y extensión con Python 61

19. Módulos estadísticos especializados 62

19.1 Secuencia común . . . 64
19.2 Mapa de cobertura funcional . . . 64
19.3 Guía técnica completa de los treinta módulos especializados . . . 64
19.4 Preparación de datos . . . 64
19.5 Tamaño de muestra . . . 65
19.6 Modelos longitudinales y mixtos . . . 65
19.7 Imputación múltiple y multinivel . . . 66
19.8 Validación predictiva . . . 67
19.9 Modelos flexibles y diagnósticos . . . 67
19.10 Comparaciones múltiples . . . 68
19.11 Tabla 1 y balance entre grupos . . . 68
19.12 Encuestas complejas . . . 69
19.13 Pruebas diagnósticas y DeLong . . . 69
19.14 DAG y plan causal . . . 70
19.15 Expediente reproducible del análisis . . . 70
19.16 Validación predictiva avanzada . . . 71
19.17 Supervivencia avanzada . . . 71
19.18 Auditoría de calidad de datos . . . 72
19.19 Inferencia causal aplicada . . . 72
19.20 Metaanálisis . . . 73
19.21 Asistente de publicación . . . 73
19.22 Análisis de sensibilidad . . . 74
19.23 Gobernanza de datos . . . 74
19.24 Epidemiología poblacional . . . 74
19.25 Inferencia bayesiana . . . 75
19.26 Mediación, moderación, CFA y SEM . . . 75
19.27 Metaanálisis avanzado . . . 76

19.28 Predicción clínica y explicabilidad . . . 76
19.29 Series temporales epidemiológicas . . . 77
19.30 Psicometría e instrumentos . . . 77
19.31 Arquitectura de extensiones . . . 77
19.32 Potencia por simulación . . . 78
19.33 Control de calidad y diseño experimental . . . 78

20. Guardar proyectos, copiar y exportar 79

21. Guía de interpretación de estadísticos 80

Significativo no equivale a importante . . . 80
No significativo no demuestra equivalencia . . . 80

22. Problemas frecuentes 81

23. Lista de verificación del análisis 82

24. Glosario esencial 83

25. Fichas técnicas de los procedimientos disponibles 84

25.1 Shapiro-Wilk y D'Agostino K^2 . . . 84
25.2 t y Wilcoxon de una muestra . . . 84
25.3 Chi-cuadrado de Pearson y prueba exacta de Fisher . . . 84
25.4 Student, Welch y Mann-Whitney . . . 84
25.5 ANOVA, ANOVA de Welch y Kruskal-Wallis . . . 85
25.6 Pearson, Spearman y Kendall . . . 85
25.7 t pareada, Wilcoxon pareado, McNemar y Friedman . . . 85
25.8 Regresión lineal . . . 85
25.9 Regresión logística binaria, multinomial y ordinal . . . 85
25.10 Poisson robusto, Poisson de conteos y binomial negativa . . . 86
25.11 Gamma con enlace log . . . 86
25.12 Kaplan-Meier, Log-rank y Cox . . . 86
25.13 PCA y matriz de correlación . . . 86
25.14 Bland-Altman . . . 86
25.15 Cómo usar una ficha en SAmuel . . . 86

26. Referencias bibliográficas y guías 87

27. Límites, privacidad y responsabilidad 89

28. Investigación integral en SAmuel web 90

Plantillas temáticas y programación reproducible . . . 90
Navegación y lectura del espacio de trabajo . . . 90
Cerrar la ventana y liberar recursos . . . 90
Portada, movimiento y orientación por tema . . . 90
Una ruta de investigación, no solo una calculadora . . . 91
Validar con otra población . . . 92
Tablas, figuras y mapas . . . 92
Guardar, continuar y revisar . . . 93

Programación autorizada y extensiones verificables 93
Cuenta, documentación y uso responsable 93

¿Por qué utilizar SAmuel?

SAmuel integra en una sola aplicación diseño, preparación de datos, análisis epidemiológico, causalidad, predicción, interpretación, gráficos y reporte reproducible sin exigir programación. La revisión 41 conserva los motores estadísticos y la activación local estable, y prepara una segunda etapa centralizada sin habilitarla hasta completar su piloto de seguridad.

Comprender y decidir

La interfaz guía desde la naturaleza y el nivel de medición hasta una prueba compatible, explica qué falta para ejecutar y conserva visibles el diseño, el evento, la referencia y los supuestos.

Analizar con rigor

Integra estadística clásica y avanzada con controles que evitan publicar resultados inválidos por codificación binaria, convergencia, fuga de información o estructuras de dependencia mal declaradas.

Interpretar y comunicar

Cada salida procura indicar qué se hizo, qué variables se utilizaron, cuáles son los estadísticos relevantes y cómo expresarlos brevemente en lenguaje estadístico y epidemiológico.

Conservar y reproducir

El proyecto `.samuel`, el historial, el guardado automático, la reejecución, los editores R/Python y las exportaciones permiten continuar, revisar y enseñar el análisis.

Proteger información

El procesamiento principal es local; los temporales, proyectos y exportaciones incorporan controles de integridad, minimización de datos y protección frente a contenidos inseguros.

Trabajar y continuar

Acceda con su cuenta, conserve proyectos y retome configuraciones. La recuperación por cuenta y los respaldos complementan su copia propia; no ejecute archivos de origen desconocido.

1. Accesibilidad, orientación y aprendizaje

Beneficio	Qué hace SAmuel	Valor para el usuario
Uso guiado sin programar	Selecciona variables, pruebas, modelos y gráficos mediante menús temáticos, buscadores, autocompletado, asistentes y ayudas contextuales.	Permite comenzar desde una pregunta aplicada y avanzar sin memorizar comandos.
Cuatro niveles de trabajo	Organiza el contenido en Básico, Intermedio, Avanzado y Programador, con submenús por temas y un solo módulo visible a la vez.	Evita listas extensas y permite encontrar procedimientos por intención, no por nombres aislados.
Univariado en tres decisiones	Separa tipo de variable, grupo de análisis y medida estadística; admite una sola variable, selección por tipo y todas las variables compatibles.	Reduce errores de selección y acelera análisis cualitativos, cuantitativos, de supervivencia y resúmenes completos.
Estado de configuración visible	Indica si falta tipo, grupo, medida, variable, tiempo, evento o comparación y confirma cuándo el análisis está listo.	Convierte un botón aparentemente inactivo en una instrucción concreta para continuar.
Interfaz adaptable	Incluye paneles minimizables, desplazamiento, maximización de gráficos y formato, navegación por teclado y selectores que muestran nombre, naturaleza y nivel.	Facilita trabajar en pantallas pequeñas, ampliadas o con escalado alto sin perder controles.
Aprendizaje progresivo	Relaciona el análisis guiado con scripts R y Python, resultados, explicación y bibliografía.	El usuario puede aprender la lógica estadística y luego reutilizar el código de forma independiente.

2. Preparación, calidad y comprensión de los datos

Beneficio	Qué hace SAmuel	Valor para el usuario
Importación amplia	Abre CSV, Excel, TSV, Stata, SPSS, JSON y SAS XPT; permite exportar la base nuevamente a Excel.	Reduce conversiones manuales y conserva un flujo de trabajo familiar.
Reconocimiento inicial	Describe filas, columnas, faltantes, tipos sugeridos, categorías y patrones de observaciones independientes, pareadas o repetidas.	Ayuda a detectar problemas estructurales antes de calcular estadísticas.
Diccionario modificable	Permite revisar etiqueta, naturaleza, nivel, papel analítico, moneda y separador de miles; protege inicialmente identificadores explícitos sin eliminarlos.	Evita usar claves administrativas como variables analíticas y mejora la presentación de cantidades monetarias.
Auditoría y gobernanza	Examina faltantes, duplicados, rangos, fechas, categorías, riesgo de reidentificación, consentimiento y restricciones de uso.	Hace visibles problemas de calidad y gobierno sin corregir silenciosamente la base.
Datos faltantes e imputación	Describe el patrón de faltantes y ofrece preparación, imputación múltiple y opciones multinivel, con advertencias sobre MCAR, MAR y MNAR.	Evita convertir la imputación en un paso automático sin justificación ni sensibilidad.
Veinte escenarios docentes	Incluye bases sintéticas para cohorte, ensayo, casos y controles, encuestas, diagnóstico, longitudinal, mapas, predicción, causalidad y otros recorridos.	Permite aprender sin usar información real ni depender de descargas externas.

3. Rigor estadístico y epidemiológico

Beneficio	Qué hace SAmuel	Valor para el usuario
Cobertura descriptiva organizada	Agrupa frecuencias, tendencia central, dispersión, posición, forma, distribución, precisión, Kaplan-Meier, Tabla 1 y resumen completo.	Presenta cada medida dentro de su propósito y admite resumir una o varias variables.
Elección razonada de pruebas	Combina naturaleza, grupos, diseño, frecuencias esperadas, distribución y varianzas para orientar entre pruebas compatibles.	La recomendación explica su fundamento y no oculta las decisiones del analista.
Medidas epidemiológicas correctas	Integra OR, RR, RP, HR, tasas, diferencias, intervalos, tablas 1, supervivencia y ajuste multivariable con evento y referencia explícitos.	Facilita reportar una medida coherente con el diseño y evita llamar riesgo a una odds.
Explicación separada de predicción	La finalidad explicativa prioriza fuerza de asociación, IC, valor p y ajuste; la predictiva muestra coeficientes, discriminación, calibración, error y utilidad.	Evita mezclar asociación etiológica con rendimiento predictivo.
Protección de indicadores	Los indicadores derivados conservan faltantes; las categorías binarias numéricamente equivalentes se normalizan y las medidas binarias exigen exactamente dos categorías observadas.	Impide recodificaciones silenciosas o estimaciones con eventos inexistentes.
Tablas y efectos robustos	Protege una categoría real llamada Total, calcula Hedges g con corrección para muestras pequeñas y adapta la evaluación de normalidad al tamaño muestral.	Reduce colisiones de etiquetas y sesgos conocidos en efectos estandarizados.
Modelos con control de validez	Bloquea la publicación de logística no convergente, revisa diagnósticos y diferencia GEE, mixtos y modelos no gaussianos según la unidad de análisis.	Evita presentar como estable una estimación que el algoritmo no logró ajustar correctamente.
Predicción sin fuga evitable	Imputa dentro de cada pliegue, calcula importancia por permutación fuera del entrenamiento, admite particiones agrupadas por sujeto y usa AUC por rangos eficiente.	Produce una evaluación interna más honesta y escalable.

4. Métodos avanzados integrados

Área	Capacidades incorporadas	Propósito
Inferencia bayesiana	Proporciones, comparación de riesgos, medias, pruebas t, correlación, ANOVA, regresión lineal y logística; priors, posteriores, intervalos creíbles, sensibilidad y BF cuando corresponde.	Permite enseñar y aplicar Bayes sin ocultar el prior ni presentar aproximaciones como resultados exactos.
Causalidad y diseño	DAG, plan causal, IPTW, AIPW, emparejamiento, balance, soporte, sensibilidad y advertencias sobre temporalidad y sesgos.	Separa la pregunta causal del simple ajuste automático de covariables.
Dependencia y poblaciones	GEE, modelos mixtos gaussianos y no gaussianos, encuestas complejas, dominios y métodos de réplicas.	Respeta agrupación, repetición, pesos, estratos y UPM cuando el diseño lo exige.
Mediación y estructura	Mediación, moderación, CFA guiada y SEM de rutas observadas con estimaciones, intervalos y advertencias de identificación.	Explora mecanismos, modificación de efecto y estructura sin confundir mediación estadística con causalidad demostrada.
Metaanálisis	Síntesis básica y avanzada, heterogeneidad, forest/funnel, influencia, subgrupos, metarregresión y acumulación temporal.	Integra la síntesis principal con análisis de estabilidad y moderadores.
Predicción clínica	Elastic Net, random forest, gradient boosting y SVM, validación cruzada, importancia, calibración disponible y curvas de decisión.	Relaciona el rendimiento estadístico con explicabilidad y beneficio neto por umbral.
Tiempo, instrumentos y simulación	Series interrumpidas, ARIMA/SARIMA, alertas, alfa/omega, EFA, ICC, potencia por simulación y error Monte Carlo.	Extiende SAmuel a vigilancia, psicometría y diseños que requieren escenarios explícitos.
Calidad y experimentación	Cartas I-MR, Cp/Cpk/Pp/Ppk, Pareto y diseños factoriales con interacciones.	Combina control estadístico de procesos con diseño experimental dentro del mismo entorno.

5. Interpretación, comunicación y reporte

Beneficio	Qué hace SAmuel	Valor para el usuario
Lectura estructurada	Procura describir qué se realizó, qué variables y niveles se usaron, cuáles fueron los estadísticos principales y cuál es su interpretación breve.	Convierte una tabla técnica en una lectura comprensible sin emitir automáticamente una conclusión clínica.
Lenguaje epidemiológico	Interpreta referencias, eventos, razones y medidas ajustadas, nombrando las covariables del modelo cuando corresponde.	Reduce frases ambiguas y diferencia resultados crudos de los ajustados.
Supuestos y límites visibles	Presenta diagnósticos, exclusiones, precisión, validación aparente, confusión y advertencias contra inferencias causales injustificadas.	Ayuda a comunicar tanto el hallazgo como sus condiciones de validez.
Historial integral editable	Acumula interpretaciones, tablas, scripts y figuras; permite editar texto, guardar HTML/TXT y exportar selecciones.	Conserva la secuencia completa del análisis y facilita preparar informes.
Cita, licencia y terceros	Incluye referencia y BibTeX, licencia académica y aviso de componentes externos en ventanas consistentes.	Hace visibles autoría, condiciones de uso y atribuciones necesarias.

6. Gráficos, mapas y presentación científica

Beneficio	Qué hace SAmuel	Valor para el usuario
Editor de estilo científico	Permite modificar textos, series, ejes, elementos, fondo, tamaño, tipografía, leyenda, cuadrícula, líneas, marcadores y resolución.	Produce figuras sobrias y configurables sin depender de otro programa.
Formato ampliable y consistente	El panel puede maximizarse y utiliza las mismas opciones en Gráficos y en el Historial integral.	Evita trabajar en un área reducida y mantiene un único lenguaje visual.
Referencias horizontal y vertical	Añade líneas de referencia configurables en ambos ejes.	Facilita mostrar cortes, umbrales, metas o valores clínicamente importantes.
Gráficos estáticos e interactivos	Incluye visualizaciones 2D/3D, Plotly, supervivencia, ROC, forest, series y paneles; exporta SVG, PDF, EPS, PNG y TIFF.	Admite exploración, presentación y publicación con salidas vectoriales o raster.
Georreferenciación local	Integra Leaflet, marcadores, etiquetas, ventanas, jerarquías y recursos locales; cierra correctamente polígonos y reorganiza etiquetas estrechas.	Permite construir mapas epidemiológicos sin separar el análisis de la comunicación espacial.

7. Reproducibilidad, continuidad y programación

Beneficio	Qué hace SAmuel	Valor para el usuario
Proyecto integral `.samuel`	Guarda datos, diccionario, interfaz, scripts, imputación, resultados y recetas en un contenedor verificable.	Permite retomar un estudio con su contexto y no solo con una tabla aislada.
Guardado seguro	Usa reemplazo atómico, guardado de recuperación manual y, si se activa expresamente, periódico cada 60 segundos.	Reduce pérdida de trabajo y evita sustituir el proyecto anterior antes de comprobar la copia nueva.
Reejecución	Conserva recetas avanzadas y puede repetirlas en orden sobre los datos del proyecto.	Facilita actualizar resultados y comprobar decisiones analíticas.
Entornos R y Python	Incluye editores tipo IDE con consola, entorno, archivos, paquetes, ayuda, tablas, gráficos, autocompletado y menús compactos.	El usuario puede ampliar o estudiar el análisis sin abandonar la aplicación.
Código paralelo	Registra scripts R y Python con variables, momento y contexto del procedimiento guiado.	Conecta la interfaz con un registro auditable y reutilizable.
Expediente reproducible	Documenta estudio, responsable, protocolo, objetivo, semilla, versiones, huellas, decisiones y pendientes.	Apoya revisión, docencia, publicación y continuidad del equipo.

8. Privacidad, seguridad e integridad

Beneficio	Qué hace SAmuel	Valor para el usuario
Procesamiento local	Los análisis principales funcionan sin enviar la base a servidores externos y sin conexión, salvo descargas voluntarias o fondos de mapas en línea.	Reduce exposición de información y dependencia de servicios remotos.
Intercambio R protegido	Escribe temporales en una carpeta privada del perfil, restringe permisos y solo purga carpetas antiguas que llevan una marca propia.	Evita mezclar datos con temporales ajenos o eliminar archivos que no pertenecen a SAmuel.
HTML con minimización de datos	Leaflet y Plotly incorporan únicamente campos utilizados o autorizados e informan cuáles quedarán dentro del archivo.	Ayuda a evitar compartir columnas innecesarias en una visualización interactiva.
Proyectos defensivos	Antes de abrir un `.samuel` valida CRC, rutas, duplicados, cifrado, enlaces, esquema, tamaños, expansión y manifiesto.	Rechaza contenedores dañados o construidos para escribir fuera de su destino.
Excel protegido	Conserva como texto las cadenas que una hoja de cálculo podría ejecutar como fórmulas, sin transformar valores numéricos legítimos.	Reduce el riesgo de fórmulas involuntarias al abrir exportaciones.
Extensiones controladas	La arquitectura valida manifiesto, permisos, SHA-256 y firma Ed25519; no ejecuta código durante la inspección y bloquea paquetes no confiables.	Permite ampliar el catálogo sin convertir la extensión en acceso automático al núcleo.

9. Acceso, continuidad y mantenimiento

Beneficio	Qué hace SAmuel	Valor para el usuario
Cuenta autorizada	Usuario, contraseña y autenticador; gestión administrativa del acceso.	Control de acceso sin compartir claves entre personas.
Continuidad del proyecto	Guarda datos, resultados y recetas y permite retomar la configuración.	Facilita investigar en varias sesiones y revisar decisiones.
Recuperación cifrada	Complementa el proyecto con una copia por cuenta y confirmación antes de restaurar.	Reduce el riesgo de perder un avance reciente; no reemplaza su respaldo.
Respaldo verificable	Herramientas de copia cifrada e integridad, con ensayo de restauración separado.	Permite al operador comprobar la recuperación antes de intervenir el servicio.
Mantenimiento controlado	Entornos y bibliotecas versionados; comprobaciones previas de configuración.	Permite revisar cambios y planificar el servicio sin alterar datos de usuarios.

10. Rendimiento, verificación y confianza proporcionada

Evidencia	Resultado de la revisión	Qué demuestra y qué no
Cobertura funcional	Matriz de 40 módulos, cada uno vinculado con prueba, comprobación numérica o escenario de piloto.	Reduce omisiones de validación; no garantiza que todo conjunto de datos posible carezca de errores.
Pruebas automatizadas	Comprobaciones de rutas, motores, proyectos, recuperación y control de acceso con datos sintéticos.	Detectan regresiones en los casos cubiertos; no certifican todos los supuestos ni todas las variantes.
Capacidad de trabajo	Tablas paginadas, límites explícitos y configuración de recursos del servicio.	La carga y los tiempos deben comprobarse en el servidor y con el tamaño previsto; no hay un tiempo universal.
Adaptación visual	Interfaz comprobada en 1366 x 768 con escalas de 100 %, 125 %, 150 % y 200 %.	Reduce el riesgo de controles inaccesibles; todavía se recomiendan pruebas con tecnologías de asistencia.
Integridad y continuidad	Pruebas de rutas, integridad de proyectos, recuperación y configuración del servicio.	Comprueba la entrega exacta y su mantenimiento básico.
Revisión de interacción	Pruebas de DOM sintético para navegación, selección, ingreso, recuperación y controles.	No sustituyen la aceptación visual renderizada ni una evaluación de usabilidad con usuarios.

Propuesta integral. El beneficio distintivo de SAmuel no es una prueba aislada: es la continuidad entre comprender la base, escoger un método, calcularlo, revisar supuestos, interpretarlo, representarlo, conservarlo y reproducirlo dentro de un mismo entorno local. Esta integración está orientada especialmente a formación, investigación y práctica epidemiológica.

Un puente entre el menú y el código

El usuario elige variables	SAmuel revisa estructura y supuestos	Recomienda y ejecuta	Interpreta el resultado	Genera R y Python

Ejemplo: al seleccionar un desenlace cualitativo y una exposición cualitativa independiente, SAmuel construye la tabla de contingencia, revisa frecuencias esperadas, orienta entre Chi-cuadrado y Fisher, interpreta el resultado y añade el código equivalente a los editores de R y Python. La persona puede repetir el mismo procedimiento desde el código y comparar la salida.

Ventaja con responsabilidad. Las recomendaciones y explicaciones disminuyen barreras de uso, pero no convierten una selección automática en una decisión metodológica infalible. El usuario debe confirmar el diseño, la codificación, los supuestos, la pertinencia clínica y las limitaciones del estudio.

Nota sobre “gratis”: sin costo de uso no significa necesariamente software de código abierto ni que las herramientas externas instaladas tengan idénticas licencias. Revise las condiciones aplicables a R, Python y cada librería adicional.

SAmuel y las plataformas estadísticas de tercera generación

SAmuel pertenece a una generación de aplicaciones que busca llevar análisis estadísticos rigurosos a estudiantes, docentes y equipos de investigación mediante una interfaz gráfica comprensible, resultados disponibles inmediatamente después de cada decisión y una ruta gradual hacia la reproducibilidad.

Definición utilizada en este manual. “Plataforma estadística de tercera generación” no es una clasificación cronológica universal ni una certificación industrial. Aquí se usa como una **categoría funcional**: aplicación con interfaz gráfica o acceso web con acceso gratuito o académicamente asequible, interfaz gráfica que permite trabajar sin programar, retroalimentación rápida, resultados listos para comunicar y algún mecanismo de reproducibilidad o extensión. Bajo esta definición, SAmuel compite directamente con jamovi, JASP y BlueSky Statistics.

Alcance de la comparación. Revisada el 2 de octubre de 2026 con documentación oficial y literatura disponible. Es una descripción funcional, no una prueba independiente ni una afirmación de superioridad universal. Las funciones y licencias cambian: confirme versión, procedimiento y validación antes de elegir.

Por qué el grupo de comparación directa es pequeño

No debe afirmarse que existan menos de cinco programas estadísticos gratuitos en sentido absoluto. Una revisión académica de software gratuito comparó siete opciones, entre ellas hojas de cálculo, R, jamovi, JASP, PSPP y SOFA. El número se reduce cuando se exige simultáneamente: (1) interfaz moderna orientada a usuarios no programadores; (2) resultados inmediatos o claramente integrados; (3) cobertura general de análisis; (4) viabilidad académica gratuita; (5) desarrollo o documentación vigente; y (6) un flujo reproducible que no obligue a combinar varias aplicaciones. Con este filtro, los tres referentes externos más próximos son jamovi, JASP y BlueSky; SAmuel constituye una cuarta propuesta, especializada en medicina y epidemiología.

Esta selección no desconoce otros proyectos valiosos. GNU PSPP es una alternativa libre semejante a SPSS y ofrece GUI, sintaxis y procedimientos clásicos, aunque su propio manual reconoce que la cobertura de procedimientos continúa incompleta. SOFA Statistics prioriza facilidad y salidas visuales, pero su alcance analítico es menor. R Commander aporta una interfaz básica y extensible para R, aunque requiere instalar y administrar el ecosistema R. Orange es sobresaliente para programación visual, aprendizaje automático y visualización interactiva, pero no se diseñó como una estación epidemiológica integral. Son alternativas relacionadas, no equivalentes exactos.

Qué diferencia el recorrido de SAmuel

Revisión del 26 de octubre de 2026. El valor propuesto es conectar protocolo epidemiológico, aprobación de la relación entre observaciones, procedimientos guiados, lectura de resultados, figuras y territorio, proyectos retomables y código R/Python en un mismo recorrido. Esta combinación se presenta como propuesta de trabajo, no como superioridad estadística o clínica demostrada.

Beneficio verificable en SAmuel	Uso concreto	Comparación responsable
Diseño y unidad de análisis antes del cálculo	Documentar pregunta, población, exposición y desenlace; confirmar observaciones independientes o relacionadas.	El diseño no puede demostrarse con nombres de columnas ni normalidad; exige conocimiento del estudio.
Análisis guiado y programación en el contexto del proyecto	Alternar formularios y R/Python; conservar configuraciones y resultados para revisar el trabajo.	jamovi también ofrece análisis sin código, R y proyectos reproducibles. Esas funciones aisladas no son exclusivas.
Lectura epidemiológica, gráficos y territorio conectados	Revisar magnitud, precisión y supuestos junto a tablas, figuras y mapas.	JASP dispone de análisis clásicos y bayesianos; Epi Info ha integrado estadística, mapas y gráficos. Tener esas funciones no demuestra mayor precisión.

Fuentes oficiales: jamovi: funciones, JASP: funciones y CDC: Epi Info. No se ha realizado un ensayo comparativo de usabilidad, productividad o exactitud frente a todas estas plataformas.

Comparación de características, facilitadores y módulos

Plataforma	Acceso y motor	Experiencia y resultados	Áreas y módulos destacados	Facilitadores y especialidad
SAmuel V1	Tres planes: Académico gratuito, Individual COP $10.000/mes y Comercial COP $15.000.000/mes con contrato anual. Autorización administrativa; consulte la sección de activación para distinguir interfaz gráfica, piloto web y servicio central pendiente. Integra motores y editores R y Python.	Flujo guiado por naturaleza y nivel de medición; ejecuta, organiza tablas y gráficos e interpreta el resultado en la misma sesión. Proyecto `.samuel`, recuperación temporal e historial reproducibles.	40 módulos: datos, diccionario, imputación, análisis uni/bi/multivariado, Bayes, modelos estructurales, medidas epidemiológicas, supervivencia, GEE/mixtos, encuestas complejas, causalidad y DAG, predicción clínica, series temporales, psicometría, metaanálisis avanzado, simulación, calidad y diseño experimental, gráficos, mapas y gobernanza.	Asistente de prueba, separación entre intención explicativa y predictiva, interpretación epidemiológica, proyectos reanudables, priors y posteriores explícitos, bases docentes, gráficos editables, georreferenciación, código paralelo R/Python y funcionamiento local.
jamovi	Gratuito y de código abierto; interfaz gráfica y nube; motor R.	Hoja de datos familiar; los resultados se recalculan cuando cambian datos u opciones. Guarda datos, análisis y resultados en un archivo `.omv`.	Núcleo de descriptivos, pruebas t, ANOVA, regresión, frecuencias y factores; biblioteca de más de 70 módulos comunitarios para modelos mixtos, supervivencia, metaanálisis, SEM, tablas y otros métodos.	Muy accesible para docencia; gráficos publicables, edición/anotación de resultados, modo de sintaxis y editor Rj. Su mayor fortaleza diferencial es la extensibilidad comunitaria.
JASP	Gratuito, multiplataforma y de código abierto AGPL; análisis implementados principalmente sobre R.	Interfaz de hoja de cálculo, arrastrar y soltar, actualización dinámica, salida minimalista y divulgación progresiva de opciones.	Descriptivos, pruebas t, ANOVA, regresión, frecuencias, factores, SEM, modelos mixtos, supervivencia, metaanálisis, redes, potencia, predicción, calidad y diseño de experimentos.	Referencia para presentar en paralelo inferencia frecuentista y bayesiana; salida APA, anotaciones e integración con OSF. Su ventaja distintiva es la inferencia bayesiana accesible.
BlueSky Statistics	Edición Base “Free Forever” y edición Pro comercial; interfaz para R, con editor de código. La versión 7.x permanece publicada como código abierto.	Cuadrícula editable, diálogos, múltiples conjuntos de datos, ejecución de R y administración rica de resultados; exporta a HTML, PDF, R Markdown y LaTeX.	Estadística descriptiva e inferencial, modelado, supervivencia, longitudinal, potencia, aprendizaje automático, minería de datos, calidad, Six Sigma, SPC y diseño de experimentos.	Combina amplitud estadística y analítica industrial; facilita inspeccionar el código R y reorganizar o reutilizar resultados. Algunas funciones avanzadas y de productividad pertenecen a Pro.

Matriz funcional resumida

Dimensión	SAmuel	jamovi	JASP	BlueSky
Uso sin programar	Menús por nivel y asistentes.	Hoja estadística y módulos.	Arrastrar/soltar y divulgación progresiva.	Cuadrícula, cinta y diálogos.
Respuesta inmediata	Resultado, interpretación, tablas y figura tras ejecutar.	Recálculo dinámico al cambiar opciones.	Actualización dinámica del resultado.	Salida integrada después de ejecutar el diálogo.
Estadística clásica	Amplia, orientada a salud.	Amplia en núcleo y módulos.	Amplia en módulos oficiales.	Amplia, incluida ingeniería de calidad.
Bayesiano	Módulo guiado para proporción, comparación epidemiológica de dos riesgos, media y regresión lineal: posterior, ICr, probabilidades, sensibilidad de prior y BF exacto en binomial.	Disponible mediante módulos comunitarios.	Fortaleza central y cobertura más extensa de diseños bayesianos.	Depende de procedimientos y paquetes R disponibles.
Epidemiología aplicada	OR, RR, RP, HR, IC, diseños, ajuste, tablas 1, diagnóstico, tasas y supervivencia integrados.	Cobertura variable entre núcleo y módulos.	Regresión, supervivencia, metaanálisis y pruebas; no se presenta como flujo epidemiológico único.	Regresión, supervivencia, longitudinal y potencia; depende del procedimiento seleccionado.
Asistente e interpretación automática	Selección orientada por variables/supuestos y lectura estadística/epidemiológica.	Ayuda y anotación; la elección e interpretación recaen principalmente en el usuario.	Salida anotable y progresiva; la interpretación sustantiva recae en el usuario.	Diálogos y salida administrable; la interpretación recae en el usuario.
Predicción y aprendizaje automático	Predicción separada de explicación, validación, ROC y métricas.	Módulos comunitarios.	Módulo de analítica predictiva.	Más de 50 algoritmos y herramientas de ajuste/puntuación según su sitio oficial.
Causalidad, DAG y gobernanza	Integrados como módulos y advertencias metodológicas.	No figuran como un flujo central único en la descripción oficial; puede requerir módulos o R.	No figuran como un flujo central único en la descripción oficial.	No figuran como un flujo epidemiológico central en la descripción oficial.
Gráficos y mapas	Editor científico integral, Plotly, Leaflet, jerarquías, mapas locales y exportación vectorial/raster.	Gráficos publicables basados en ggplot2 y módulos.	Gráficos asociados a cada análisis; su guía docente describe un constructor más limitado.	Visualizaciones, gráficos y exportación; maximización desde la salida.
Puente hacia código	Editores completos R y Python y código equivalente acumulado.	Sintaxis R visible y editor Rj.	Cálculos reproducibles sustentados en R; experiencia principalmente guiada.	Editor y ejecución de R; extensibilidad mediante paquetes.
Privacidad y uso local	Procesamiento principal local; no necesita nube.	Escritorio local; nube opcional.	Escritorio local e integración opcional con OSF.	Aplicación de interfaz gráfica; servicios/licencias según edición.

Lectura experta de la comparación

jamovi establece una referencia muy alta en simplicidad, actualización instantánea y ecosistema de módulos. **JASP** es especialmente fuerte cuando el objetivo docente o científico exige análisis frecuentistas y bayesianos en paralelo. **BlueSky** destaca por amplitud, gestión de varios conjuntos de datos, editor R y herramientas de calidad, ingeniería y aprendizaje automático. Estas fortalezas se usan como referentes de mejora, no como permiso para declarar una superioridad universal sin una evaluación independiente.

La propuesta diferencial de **SAmuel** no consiste en afirmar que cada algoritmo sea exclusivo, sino en reunir en una sola experiencia local componentes que normalmente están dispersos: reconocimiento de variables y diseño, elección orientada de procedimientos, interpretación epidemiológica breve, separación explícita de asociación y predicción, medidas OR/RR/RP/HR, inferencia bayesiana con priors visibles, análisis de dependencia y encuestas, causalidad y DAG, validación, gráficos científicos profundamente editables, georreferenciación jerárquica, historial integral y transición simultánea hacia R y Python. Esta integración permite considerarlo competitivo dentro de la tercera

generación y particularmente amplio para enseñanza y práctica epidemiológica.

Mejoras incorporadas después de la comparación. SAmuel amplió Bayes sin programación a proporciones, comparación de riesgos, medias, pruebas t de una muestra, pareadas e independientes, correlación, ANOVA y regresiones lineal y logística. Además incorpora mediación, moderación, estructura factorial/SEM observado, metaanálisis avanzado, predicción clínica explicable con curvas de decisión, series temporales epidemiológicas, psicometría, potencia por simulación y control de calidad/diseño experimental. Presenta priors, posteriores, intervalos creíbles, probabilidad de dirección, semillas y análisis de sensibilidad; calcula BF10 exacto solo cuando existe una solución marginal implementada y rotula las aproximaciones. Esta amplitud no autoriza a afirmar superioridad universal: CFA/SEM se limita en V1 a estructura factorial exploratoria guiada y rutas observadas, y todavía no sustituye plataformas especializadas para redes bayesianas, metaanálisis bayesiano o modelos latentes complejos.

1. Propósito, alcance y filosofía de uso

Este manual enseña a transformar una pregunta clínica o epidemiológica en un análisis reproducible dentro de SAmuel. Combina explicación conceptual, instrucciones operativas, ejemplos ejecutables e interpretación.

Su estructura pedagógica toma como referencia la organización modular y autodidacta de la ayuda de Epidat: cada procedimiento se presenta con fundamento, datos requeridos, ruta en el programa, lectura de las salidas, advertencias y bibliografía. El contenido es original y está ajustado exclusivamente a la funcionalidad disponible en SAmuel.

Qué hace SAmuel. Importa y documenta datos; orienta la elección de procedimientos; calcula estadística descriptiva e inferencial; ajusta modelos; produce gráficos científicos; interpreta resultados; conserva un historial editable; y genera código equivalente en R y Python.

Qué no hace. No reemplaza el protocolo, el cálculo de tamaño muestral, el conocimiento clínico, la definición causal, la revisión de sesgos ni la asesoría bioestadística. Una asociación estadística no demuestra causalidad, efectividad ni utilidad clínica.

EPINex y cómo citar SAmuel

EPINex es la marca bajo la cual se presenta y comercializa SAmuel. Pulse **Citar SAmuel** para copiar la referencia editorial, su entrada BibTeX o guardarlas como texto. Esta referencia no sustituye los antecedentes legales de autoría.

Referencia sugerida: EPINex. (2026). *SAmuel: Sistema de análisis médico, epidemiológico y estadístico* (Versión 1) [Software].

```
@software{epinex_samuel_2026,
  author  = {{EPINex}},
  title   = {SAmuel: Sistema de análisis médico, epidemiológico y estadístico},
  year    = {2026},
  version = {1},
  note    = {Edición distribuida bajo la marca EPINex; referencia editorial}
}
```

La carpeta distribuida incluye además `CITATION.cff` y `CITAR_SAMUEL_V1.txt`. Si en el futuro existe DOI, URL oficial o nueva versión, use los metadatos de esa publicación en lugar de inventarlos.

Convenciones del manual

Elemento	Significado
Ruta	Secuencia exacta de pestañas, campos y botones.
Interpretación	Redacción epidemiológica apropiada para la salida del ejemplo.
Advertencia amarilla	Supuesto o decisión que debe comprobarse.
Advertencia roja	Uso potencialmente inválido o conclusión que debe evitarse.

Ilustraciones de esta edición: los esquemas EPINex sustituyen las capturas históricas que identificaban personalmente al autor. Son recorridos orientativos, no capturas exactas ni resultados calculados. Las instrucciones completas de cada módulo se mantienen en el texto.

Transición comercial a EPINex

La razón social, el NIT, los contactos empresariales y la formalización de la cesión patrimonial están pendientes. EPINex identifica la marca comercial; esta edición no acredita una sociedad constituida ni una cesión ya ejecutada. Los pagos continúan temporalmente a la cuenta personal identificada en la aplicación; no se presenta como cuenta empresarial. Las condiciones de acceso y las licencias firmadas existentes se conservan. No hay pautas de terceros ni temporizador publicitario; el patrocinio de una licencia no concede acceso a los datos del beneficiario.

1. Acceso, licencias y activación

Solicite autorización por el canal indicado en la portada. Revise el mensaje antes de enviarlo y no incluya contraseñas, semillas o datos clínicos. El administrador revisa declaración, plan y, cuando corresponde, comprobante de pago.

Plan	Condiciones
Académico	Gratuito e indefinido para estudiantes con declaración veraz y aprobación administrativa.
Individual	COP $10.000 mensuales desde la activación, sin período gratuito; requiere declaración de estudiante.
Comercial	Contrato anual con doce mensualidades de COP $15.000.000. Requiere declarar que no se es estudiante y seleccionar una categoría institucional válida.
Cuenta y sesión	Acceso personal con contraseña y autenticador. El servicio central permite una sesión simultánea por cuenta, sin cupo de computadores.
Patrocinio	Un patrocinador puede pagar por un beneficiario identificado sin adquirir acceso a sus datos o resultados.

Ingrese con su usuario y contraseña y confirme el código del autenticador. No debe cargar un archivo de activación en el formulario. Si pierde la sesión o su autenticador, solicite recuperación por el canal autorizado. El administrador debe verificar identidad y registrar los cambios; un cierre puede hacer perder el trabajo no guardado.

La activación y los pagos requieren revisión administrativa. El enlace Wompi de COP $10.000 corresponde a Individual; Comercial no tiene un enlace confirmado. Los datos empresariales pendientes se incorporarán cuando sean confirmados. Las condiciones completas están en Licencia; este resumen no las sustituye.

2. Inicio rápido

1. Cargar datos	2. Definir variables	3. Confirmar diseño	4. Analizar y diagnosticar	5. Interpretar y exportar

1. Abra SAmuel e ingrese. Utilice su acceso autorizado, escriba usuario y contraseña y confirme el código vigente de su autenticador.

2. Cargue una base. En Datos elija **Cargar base…**, **Usar datos sintéticos** o **Bases de ejemplo…**.

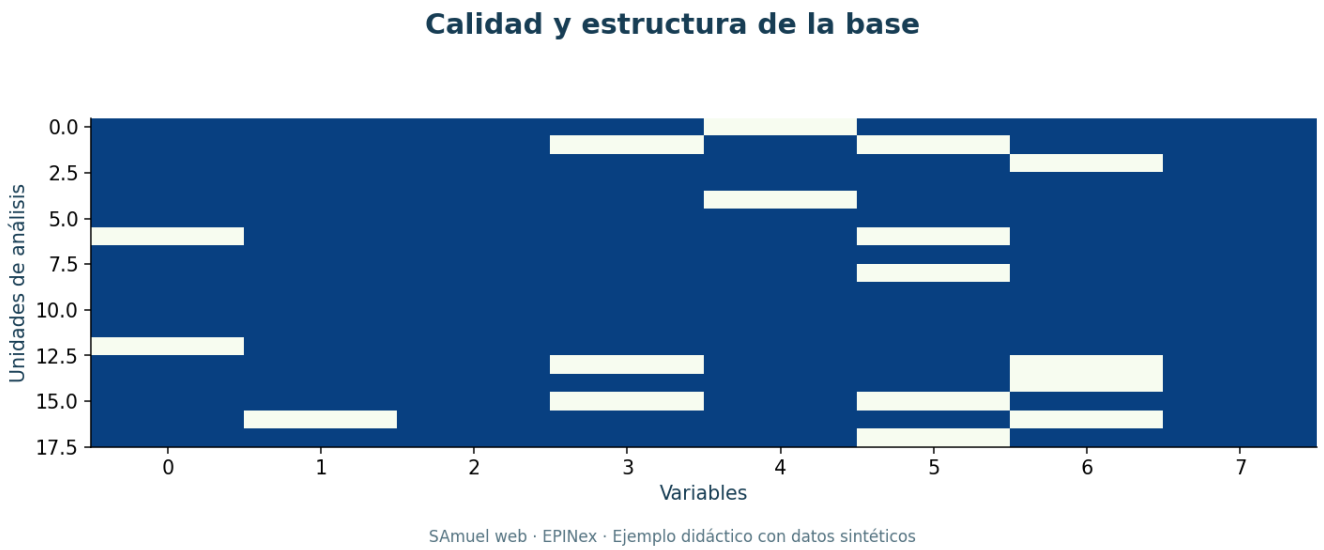
3. Revise el resumen superior. Confirme archivo, filas, variables, faltantes, tipos y estructura de muestras.

4. Corrija el diccionario. En Variables confirme naturaleza y nivel de medición.

5. Defina la pregunta. Abra Módulos y seleccione Básico, Intermedio o Avanzado; escriba fragmentos de nombres para encontrar variables.

6. Ejecute. Lea primero Interpretación, luego Tablas y Supuestos. Compruebe gráficos diagnósticos.

7. Conserve el trabajo. Guarde el proyecto y exporte el historial integral o una selección.



Ejemplo sintético de valores disponibles y faltantes; revise los patrones antes de transformar.

Guía de navegación del menú principal

La navegación se organiza por tarea y nivel de análisis. Las fichas siguientes describen familias metodológicas; una familia puede reunir varios procedimientos. Su número no equivale al número de botones de la portada ni al de pruebas disponibles.

Cómo moverse sin perder el contexto

1. **Datos y variables.** Cargue una base o un escenario sintético. Despliegue preparación, vista previa y diccionario cuando los necesite; revise naturaleza, nivel de medición y diseño.
2. **Confirme la relación entre observaciones.** La detección automática es orientativa. Identifique sujetos, grupos y momentos antes de aprobar el diseño.
3. **Seleccione el nivel.** Análisis básico incluye univariado cualitativo y cuantitativo. Análisis intermedio y Análisis avanzado agrupan métodos aplicados y especializados. Elija grupo y procedimiento, revise la orientación y seleccione las variables.
4. **Ejecute y revise.** Compruebe casos utilizados, supuestos, advertencias, estimación, intervalo e interpretación. Resultados y proyecto conserva el historial y permite ampliar tablas y figuras.
5. **Guarde y retome.** Guarde el proyecto .samuel y los archivos que necesite conservar. Consulte Recuperación temporal al final del menú; no sustituye un respaldo externo.

Acceso	Propósito
Datos y variables	Importación, preparación, vista previa, diccionario y revisión del diseño.
Análisis básico	Univariado, comparación, asociación y modelos básicos según el catálogo.
Análisis intermedio / Análisis avanzado	Seleccionar la familia metodológica y el procedimiento disponible.
Georreferenciación	Mapas por coordenadas, capas y herramientas espaciales. Los fondos en línea requieren internet y las condiciones de su proveedor.
Gráficos	Visualizaciones organizadas por propósito; formato y exportación.
Resultados y proyecto / Exportar informes	Historial, interpretación, continuidad del trabajo y entregables.
R / Python	Editor, consola, datos y resultados. Las plantillas se revisan antes de ejecutar; instalar paquetes depende de compatibilidad y permisos.
Guía de uso / Recuperación temporal	Consultar instrucciones y recuperar avances disponibles.

Lectura de las fichas. Busque el nombre del procedimiento dentro del nivel indicado. Los formularios pueden cambiar según método y datos; un nombre de familia no implica que todas sus variantes tengan los mismos requisitos. Las rutas antiguas como PROGRAMADOR se corresponden ahora con R / Python.

Mismo contenido en HTML y PDF. Ambas ediciones comparten los capítulos metodológicos. El PDF añade paginación e índice; la ayuda integrada aporta los controles de la versión en uso. Consulte Novedades para los ajustes posteriores.

BÁSICO · ¿pa' que más?

Reúne el recorrido esencial: cargar, documentar y depurar la base; describir variables; comparar grupos; y construir modelos ajustados. Es el punto de entrada recomendado para cualquier análisis.

Datos

Ruta. Datos y variables; seleccione el apartado correspondiente.

Para qué usarlo. Cargar archivos, inspeccionar filas y columnas, buscar registros, elegir columnas activas, definir el diseño de muestras y ajustar decimales por variable.

Por qué usarlo. Una estructura incorrecta, una unidad de análisis mal definida o un identificador usado como predictor puede invalidar todo lo que sigue.

Entregables. Vista paginada de la base, resumen de filas, variables, faltantes y tipos, selección de columnas activas y diseño de muestras confirmado.

Ejemplo concreto. Abra Cohorte cardiovascular, confirme 5.000 filas, oculte id_participante del análisis y declare muestras independientes antes de comparar tratamientos.

Variables

Ruta. Datos y variables; seleccione el apartado correspondiente.

Para qué usarlo. Construir el diccionario analítico: etiqueta, naturaleza, nivel de medición, categoría de referencia y habilitación de cada variable.

Por qué usarlo. La clasificación controla qué pruebas, resúmenes, modelos y formatos son válidos; el significado clínico no puede inferirse solo desde el tipo de archivo.

Entregables. Diccionario actualizado, variables activas, referencias explícitas, guía visual de tipo y nivel, y formato monetario cuando corresponde.

Ejemplo concreto. Clasifique ingreso_mensual como cuantitativa con nivel Dinero para verlo como $2.742.564,50 sin convertirlo en texto ni alterar los cálculos.

Imputación

Ruta. Análisis básico → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Explorar faltantes y aplicar, cuando sea justificable, imputación simple por media, mediana, moda o valor constante.

Por qué usarlo. Eliminar todos los registros incompletos reduce información; imputar sin revisar MCAR, MAR o MNAR también puede introducir sesgo y falsa precisión.

Entregables. Tabla de faltantes, método sugerido con justificación completa, copia de datos imputada y registro reproducible de la decisión.

Ejemplo concreto. En Cohorte cardiovascular seleccione IMC; si la distribución es asimétrica, revise la recomendación de mediana, despliegue la justificación y aplíquela solo tras aceptar sus límites.

Univariado

Ruta. Análisis básico → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Describir una variable mediante frecuencias, medidas de centro y dispersión, intervalos, normalidad y gráficos; también crear barras con error.

Por qué usarlo. Antes de comparar o modelar se debe conocer distribución, categorías, atípicos, faltantes y escala de cada variable.

Entregables. Tabla descriptiva, interpretación, prueba orientativa de normalidad, supuestos y figura formateable.

Ejemplo concreto. Seleccione edad en la cohorte: obtenga n, media, DE, mediana, RIC e histograma; luego verifique si los extremos justifican reportar mediana además de media.

Bivariado

Ruta. Análisis básico → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Comparar dos variables con pruebas para grupos independientes o pareados, tablas de contingencia, correlación y análisis de supervivencia Log-rank.

Por qué usarlo. La prueba correcta depende de naturaleza, número de grupos, dependencia, distribución y frecuencias esperadas; no basta con escoger por costumbre.

Entregables. Estimación, IC, valor p, tamaño del efecto, supuestos, tabla con n y porcentaje, medida cruda seleccionada y gráficos como barra con error.

Ejemplo concreto. Compare no_respuesta_12m entre marca_imatinib Innovador y Genérico; coloque primero la categoría de referencia y solicite RR con IC 95% y porcentaje por fila.

Multivariado

Ruta. Análisis básico → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Ajustar modelos lineales, logísticos, de conteo o Cox con varios predictores y categorías de referencia explícitas.

Por qué usarlo. Permite estimar asociaciones ajustadas y separar el efecto de covariables medidas, siempre que el modelo y el conjunto de ajuste estén justificados.

Entregables. Coeficientes o medidas exponenciadas, IC 95%, valores p, diagnósticos, interpretación ajustada, código R y Python, y figura cuando aplica.

Ejemplo concreto. Modele no_respuesta_12m con marca_imatinib, edad, sexo y multimorbilidad; interprete el OR ajustado y revise colinealidad y calibración antes de concluir.

INTERMEDIO · El atrevido

Agrupa trece herramientas aplicadas en Preparación y diseño, Comparación y modelado, Evaluación y población y Reporte científico. Es el nivel adecuado para métodos especializados de uso frecuente que requieren decisiones explícitas, pero no un marco causal o computacional avanzado.

Preparación de datos

Ruta. Datos y variables; seleccione el apartado correspondiente.

Para qué usarlo. Calcular variables derivadas, recodificar categorías o filtrar filas sobre una copia con vista previa.

Por qué usarlo. Las transformaciones deben ser transparentes, reversibles y reproducibles; editar manualmente una hoja rompe la trazabilidad.

Entregables. Base propuesta, confirmación antes de aplicar, posibilidad de deshacer y código reproducible de la transformación.

Ejemplo concreto. Cree IMC con peso_kg/(talla_m*talla_m), revise los primeros resultados y aplique únicamente si las unidades y valores son coherentes.

Tamaño de muestra

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Calcular participantes para medias, proporciones, comparaciones, cohortes, casos y controles, ANOVA, correlación, diagnóstico o supervivencia; opcionalmente seleccionar una muestra.

Por qué usarlo. Hace explícitos potencia, precisión, pérdidas, asignación y efecto de diseño antes de recolectar o seleccionar datos.

Entregables. Supuestos documentados, tamaño requerido, escenarios y archivo de muestra seleccionada cuando se solicita.

Ejemplo concreto. Planifique dos proporciones de 20% y 30%, alfa 0,05, potencia 80% y 10% de pérdidas; compare el n final con la capacidad real de reclutamiento.

Modelos longitudinales

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Analizar observaciones repetidas o agrupadas mediante GEE o modelos mixtos con familias gaussianas, binomiales, Poisson, Gamma o binomial negativa.

Por qué usarlo. Las mediciones de una misma persona u hospital están correlacionadas; tratarlas como independientes suele subestimar la incertidumbre.

Entregables. Efectos poblacionales o condicionales, componentes de agrupación, diagnósticos de convergencia, interpretación y código.

Ejemplo concreto. En Seguimiento crónico modele HbA1c por visita y tratamiento, use id_paciente como agrupador y compare una interacción tratamiento por visita preespecificada.

Imputación múltiple

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Crear varias realizaciones plausibles mediante PMM, MICE mixta o multinivel y combinar las estimaciones.

Por qué usarlo. Representa la incertidumbre de los valores faltantes mejor que rellenar una sola cifra y evita aparentar una precisión inexistente.

Entregables. Bases imputadas exportables, modelos combinados, trazas, estabilidad y diagnóstico de la imputación.

Ejemplo concreto. Impute colesterol e IMC con edad, sexo y desenlace como predictores, genere 20 realizaciones y revise las trazas antes de usar el modelo combinado.

Validación predictiva

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Evaluar un modelo lineal o logístico fuera de los datos con que fue entrenado mediante partición o validación cruzada.

Por qué usarlo. El rendimiento aparente suele ser optimista; separar correctamente entrenamiento y validación reduce fuga de información.

Entregables. Predicciones fuera de muestra, métricas de error o discriminación, particiones y resumen reproducible.

Ejemplo concreto. Prediga no_respuesta_12m con edad, adherencia e IMC usando validación cruzada de 10 pliegues; si hay visitas repetidas, mantenga cada paciente en un solo pliegue.

Modelos flexibles y diagnósticos

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Incorporar no linealidad, interacciones y revisar residuos, influencia, heterocedasticidad y forma funcional.

Por qué usarlo. Una relación curvada o modificada por otra variable puede quedar mal representada por una línea y un único efecto promedio.

Entregables. Coeficientes, términos flexibles, errores robustos, VIF, residuos e influencia con figuras diagnósticas.

Ejemplo concreto. Modele presión sistólica con un spline de edad y una interacción sexo por tratamiento definida antes del análisis; compruebe si la complejidad mejora el ajuste sin sobreinterpretar.

Comparaciones múltiples

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Comparar varios grupos con Welch-Holm, Tukey, Dunnett o Dunn con ajustes Holm o Benjamini-Hochberg.

Por qué usarlo. Realizar muchas pruebas sin corrección aumenta la probabilidad de falsos positivos dentro de la familia de contrastes.

Entregables. Contrastes por pares o contra control, diferencias o rangos, IC y valores p ajustados.

Ejemplo concreto. Compare HbA1c entre cuatro tratamientos; use Dunnett si el objetivo es contrastar cada uno con el control y declare el control antes de ejecutar.

Tabla 1 y balance

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Describir características basales por grupo y evaluar balance con diferencias medias estandarizadas, con peso opcional.

Por qué usarlo. La DME muestra magnitud del desequilibrio sin depender del tamaño muestral como lo hace el valor p.

Entregables. Tabla 1, DME por covariable y gráfico de balance antes o después de ponderar.

Ejemplo concreto. Compare Innovador y Genérico en edad, sexo, IMC y multimorbilidad; revise qué DME superan 0,10 antes de interpretar el desenlace.

Encuestas complejas

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Estimar medias, proporciones o regresiones respetando pesos, estratos, unidades primarias y dominios.

Por qué usarlo. Ignorar el diseño probabilístico altera estimaciones, varianzas, intervalos y grados de libertad.

Entregables. Estimaciones poblacionales ponderadas, IC con diseño, efectos de diseño y código reproducible.

Ejemplo concreto. En Encuesta poblacional estime prevalencia de hipertensión usando peso_muestral, estrato y UPM; defina el dominio urbano sin filtrar previamente la base.

Pruebas diagnósticas

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Evaluar una prueba frente a referencia mediante sensibilidad, especificidad, valores predictivos, razones de verosimilitud, ROC, AUC y comparación DeLong o bootstrap.

Por qué usarlo. La exactitud depende del evento, dirección, corte, prevalencia y diseño; una AUC aislada no describe todas las consecuencias clínicas.

Entregables. Matriz diagnóstica, métricas con IC, curva ROC, corte y comparación entre pruebas cuando corresponde.

Ejemplo concreto. En Exactitud diagnóstica use enfermedad_referencia como verdad y biomarcador_A como prueba; fije dirección y corte, luego compare su AUC con biomarcador_B.

Epidemiología poblacional

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Estimar tasas, estandarizar poblaciones y cuantificar desigualdades territoriales o sociales con incertidumbre.

Por qué usarlo. Las tasas crudas pueden ser engañosas cuando las poblaciones difieren en edad u otros estratos; además, todo conteo necesita un denominador y periodo explícitos.

Entregables. Tasas con IC de Poisson, tasa estandarizada, SMR, diferencia o razón entre grupos y gráfico de intervalos.

Ejemplo concreto. Compare mortalidad entre dos territorios por 100.000 habitantes, estandarice por edad con una población común y reporte tanto la diferencia absoluta como la razón de tasas.

Psicometría e instrumentos

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Examinar fiabilidad, dimensionalidad y acuerdo de escalas o evaluadores.

Por qué usarlo. Una medida poco reproducible o mal estructurada compromete cualquier asociación posterior.

Entregables. Alfa, omega, ítem-total, cargas, comunalidades, scree plot, ICC y comparaciones entre evaluadores.

Ejemplo concreto. Revise alfa, omega e ítem-total de una escala antes de usar su puntaje como desenlace clínico.

Asistente de publicación

Ruta. Análisis intermedio → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Organizar el reporte según STROBE, CONSORT, TRIPOD, RECORD, STARD o PRISMA.

Por qué usarlo. Una lista contextual ayuda a evitar omisiones que impiden evaluar validez, reproducibilidad y aplicabilidad.

Entregables. Checklist contextualizado, responsables, evidencia, pendientes y flujo inicial de reporte.

Ejemplo concreto. Para una cohorte observacional seleccione STROBE, registre título, objetivo y protocolo, asigne quién documentará pérdidas y adjunte la evidencia de cada ítem.

AVANZADO · Un semiDIOS

Organiza diecisiete herramientas de alta complejidad en Causalidad y robustez, Predicción y tiempo, Síntesis de evidencia, Bayes y modelos estructurales, Reproducibilidad y gobernanza y Simulación y experimentación. Exige mayor claridad sobre protocolo, identificación, validación y supuestos.

DAG y plan causal

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Representar exposición, desenlace, temporalidad y roles causales para definir un conjunto de ajuste antes de modelar.

Por qué usarlo. Evita seleccionar confusores solo por valor p y reduce el riesgo de ajustar mediadores o colisionadores.

Entregables. DAG editable, nodos y flechas formateables, roles, advertencias y plan de ajuste auditable.

Ejemplo concreto. Para estimar el efecto de tratamiento sobre no respuesta, declare edad y gravedad basal como confusores preexposición, adherencia posterior como posible mediador y documente cada flecha.

Expediente reproducible

Ruta. Resultados y proyecto.

Para qué usarlo. Congelar la identidad, versiones, decisiones y productos de una sesión analítica.

Por qué usarlo. Permite que otra persona sepa qué archivo, diccionario, semilla, análisis y scripts originaron un informe meses después.

Entregables. Huella SHA-256, inventario, metadatos del estudio, historial, versiones y lista de pendientes; no copia valores individuales.

Ejemplo concreto. Antes de enviar un manuscrito, registre protocolo CARDIO-01, responsable, objetivo y semilla 202609; genere el expediente y adjúntelo con el historial exportado.

Validación predictiva avanzada

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Evaluar discriminación, calibración, error, incertidumbre y utilidad clínica mediante remuestreo y curvas de decisión.

Por qué usarlo. Un modelo puede discriminar bien y producir probabilidades mal calibradas o decisiones clínicamente inútiles.

Entregables. IC de rendimiento, calibración, ROC, diagnósticos, curva de decisión y evaluación por umbrales.

Ejemplo concreto. Valide un riesgo de no respuesta con 1.000 bootstrap; revise AUC corregida por optimismo, pendiente de calibración y beneficio neto entre umbrales de 10% y 30%.

Supervivencia avanzada

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Analizar RMST, riesgos competitivos, incidencia acumulada y diagnóstico de riesgos proporcionales.

Por qué usarlo. El HR puede ser difícil de interpretar o inadecuado cuando los riesgos no son proporcionales o existe un evento competidor.

Entregables. RMST con IC, curvas de incidencia acumulada, pruebas y diagnósticos PH, y residuos de Schoenfeld.

Ejemplo concreto. Compare tratamientos por tiempo libre de evento a 24 meses mediante diferencia de RMST cuando las curvas se cruzan; trate muerte como competidor si impide observar el evento clínico.

Auditoría de calidad de datos

Ruta. Datos y variables; seleccione el apartado correspondiente.

Para qué usarlo. Detectar faltantes, duplicados, rangos imposibles, fechas incoherentes y señales atípicas sin modificar la fuente.

Por qué usarlo. Un error de captura o una duplicación puede dominar una asociación y no se corrige con un modelo sofisticado.

Entregables. Inventario de incidencias localizables, perfil visual y reglas documentadas para resolverlas en la fuente.

Ejemplo concreto. Defina id_paciente como único, edad entre 18 y 100 y fecha_alta posterior a fecha_ingreso; revise cada incumplimiento antes de imputar o excluir.

Inferencia causal aplicada

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Estimar ATE o ATT mediante IPTW, AIPW o emparejamiento usando confusores definidos por el DAG.

Por qué usarlo. Separa el diseño de la comparación del modelo final y obliga a revisar positividad, solapamiento y balance.

Entregables. Efecto causal estimado, propensity score, pesos o pares, balance antes y después, e IC.

Ejemplo concreto. Estime ATT de Innovador frente a Genérico sobre no respuesta, ajuste por edad, sexo y gravedad basal, recorte pesos extremos y compruebe DME posteriores.

Metaanálisis

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Combinar efectos de estudios comparables con modelos fijos o aleatorios y cuantificar heterogeneidad.

Por qué usarlo. La síntesis aumenta precisión, pero debe hacer visible cuánto difieren los estudios y si la combinación es defendible.

Entregables. Efecto combinado, Q, I^2, tau^2, intervalos, forest plot y funnel plot.

Ejemplo concreto. Cargue una fila por estudio con log(OR) y error estándar, use efectos aleatorios si fue preespecificado y explique un I^2 alto antes de citar el promedio.

Análisis de sensibilidad

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Examinar cómo cambia la conclusión ante supuestos alternativos, faltantes MNAR, puntos de corte o confusión no medida.

Por qué usarlo. Una conclusión robusta debe sobrevivir escenarios plausibles, no solo el conjunto de decisiones más favorable.

Entregables. Tabla completa de escenarios, punto de quiebre, E-value o trayectoria gráfica cuando aplica.

Ejemplo concreto. Repita la asociación principal suponiendo que el riesgo entre faltantes es 1,0; 1,5 y 2,0 veces el observado e informe desde qué escenario cambia la conclusión.

Gobernanza de datos

Ruta. Datos y variables; seleccione el apartado correspondiente.

Para qué usarlo. Planear minimización, acceso, custodia, retención y riesgo de reidentificación de datos sensibles.

Por qué usarlo. Los datos de salud requieren controles organizacionales y técnicos adicionales al análisis estadístico.

Entregables. Inventario de identificadores y cuasi-identificadores, perfil k, responsables, retención y plan de controles.

Ejemplo concreto. Antes de compartir una base, marque fecha de nacimiento y municipio como cuasi-identificadores, revise clases con k menor de 5 y defina supresión o generalización institucional.

Inferencia bayesiana

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Estimar proporciones, riesgos, medias, pruebas t, correlación, ANOVA y regresiones lineal/logística mediante distribuciones posteriores y priors explícitos.

Por qué usarlo. Permite responder directamente qué valores son plausibles y con qué probabilidad posterior, usando ROPE y factores de Bayes exactos solo donde existe cálculo marginal implementado.

Entregables. Posteriores, intervalos creíbles, probabilidades de dirección, ROPE, BF10 exacto o aproximado correctamente rotulado, medidas epidemiológicas, betas, OR y gráficos.

Ejemplo concreto. Compare el riesgo de evento entre Expuesto y Referencia con prior objetiva, ICr 95% y semilla declarada; repita con prior débil y escéptica para documentar sensibilidad antes de concluir.

Mediación, moderación, CFA y SEM

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Estimar efectos indirectos, interacciones, cargas factoriales y rutas observadas en un flujo guiado.

Por qué usarlo. Conecta hipótesis sobre mecanismos y medición con estimaciones, bootstrap, ajuste y advertencias causales explícitas.

Entregables. Efectos a, b, directo e indirecto, interacción, cargas, unicidad, índices de ajuste aproximados, rutas y figuras.

Ejemplo concreto. Estime si adherencia media parcialmente el efecto del tratamiento sobre el desenlace, ajustando por gravedad basal y conservando la temporalidad definida en el DAG.

Metaanálisis avanzado

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Evaluar influencia, subgrupos, metarregresión y evolución acumulada de una síntesis.

Por qué usarlo. El efecto combinado puede ocultar heterogeneidad sistemática, dependencia de un estudio o cambios de evidencia en el tiempo.

Entregables. Análisis dejar-uno-fuera, efectos por subgrupo, coeficientes de metarregresión, síntesis acumulada e intervalos.

Ejemplo concreto. Después del modelo principal, excluya sucesivamente cada estudio y compruebe si la conclusión depende de uno solo.

Predicción clínica y explicabilidad

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Validar Elastic Net, random forest, gradient boosting o SVM y cuantificar contribución predictiva y utilidad clínica.

Por qué usarlo. La exactitud aparente y la importancia interna pueden ser optimistas si no se calculan fuera de muestra.

Entregables. AUC/Brier o RMSE/MAE/R^2 validados, importancia por permutación, predicciones y curva de decisión.

Ejemplo concreto. Prediga evento con cinco particiones, compare logística Elastic Net y random forest y revise beneficio neto antes de escoger modelo.

Series temporales epidemiológicas

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Analizar intervenciones, tendencia, estacionalidad, pronósticos y señales anómalas en vigilancia.

Por qué usarlo. El orden y la autocorrelación temporal invalidan comparaciones transversales simples.

Entregables. Regresión segmentada HAC, pronóstico ARIMA/SARIMA, diagnósticos, línea base y alertas verificables.

Ejemplo concreto. Evalúe el cambio inmediato y de tendencia mensual después de una política, controlando estacionalidad anual.

Arquitectura de extensiones

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Permitir crecimiento modular mediante formularios declarativos y motores firmados sin modificar el núcleo.

Por qué usarlo. Un contrato con permisos, huellas, pruebas y firma protege estabilidad, datos y autoría.

Entregables. Contrato API 1, inventario obligatorio y ciclo de descubrimiento, validación, firma, prueba y habilitación.

Ejemplo concreto. Mantenga bloqueada una extensión nueva hasta comprobar su SHA-256, permisos, pruebas y firma Ed25519 del titular.

Potencia por simulación

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Estimar potencia Monte Carlo para medias, proporciones y ensayos por conglomerados.

Por qué usarlo. Diseños complejos y supuestos inciertos no siempre tienen una fórmula cerrada adecuada.

Entregables. Potencia estimada, intervalo de error Monte Carlo, distribución de estimaciones, semilla y escenario completo.

Ejemplo concreto. Simule un ensayo con 20 conglomerados, ICC 0,05 y varios efectos para elegir un tamaño que no dependa de un único supuesto.

Control de calidad y diseño experimental

Ruta. Análisis avanzado → seleccione el grupo y busque el procedimiento por su nombre.

Para qué usarlo. Evaluar estabilidad, capacidad, prioridades y efectos principales/interacciones de diseños factoriales de dos niveles.

Por qué usarlo. La variación común, las causas especiales y los límites de especificación responden preguntas distintas.

Entregables. Cartas I-MR, Cp/Cpk/Pp/Ppk, Pareto y modelo factorial con interacciones.

Ejemplo concreto. Verifique estabilidad con I-MR antes de calcular capacidad respecto de los límites de especificación.

PROGRAMADOR · The Goat

Permite pasar del recorrido guiado al código sin abandonar la aplicación. Ambos entornos muestran ayudas, consola, tablas y gráficos, y pueden reutilizar el historial generado por los demás módulos.

Modo R

Ruta. R / Python; seleccione el lenguaje.

Para qué usarlo. Revisar, editar y ejecutar scripts R en una interfaz de cuatro paneles inspirada en RStudio: Source; Console/Terminal/Jobs; Environment/History/Connections; y Files/Plots/Packages/Help/Viewer. Una sola barra agrupa Nuevo, Archivo, Código, Datos y paquetes y Vista para ampliar verticalmente los paneles.

Por qué usarlo. Hace visible la traducción de cada decisión del menú y permite ampliar el análisis con un lenguaje estadístico reproducible sin abandonar el proyecto.

Entregables. R Script, R Notebook, R Markdown, Shiny, Plumber, texto, Python, SQL, Stan, D3, Sweave, R HTML y documentación; además de consola, historial, tablas, gráficos, archivos, paquetes, Addins y ayuda. La ruta técnica solo aparece al abrir Estado de R….

Ejemplo concreto. Abra Nuevo → R Markdown, ejecute con Ctrl+Enter una regresión, revise Console y Jobs, recupere el comando desde History y use Vista para mostrar u ocultar los paneles derechos.

Modo Python

Ruta. R / Python; seleccione el lenguaje.

Para qué usarlo. Revisar, editar y ejecutar código Python en una mesa científica inspirada en Spyder: editor; IPython Console/Terminal/Jobs; Variable Explorer/DataFrame/History; y Files/Plots/Help. Una sola barra compacta agrupa Nuevo, Archivo, Código, Datos y entorno y Vista.

Por qué usarlo. Facilita automatización, aprendizaje y extensiones analíticas, y permite inspeccionar variables y resultados sin perder el contexto de la base activa ni espacio vertical.

Entregables. Python Script, cuaderno por celdas, módulo, texto o SQL; consola, historial, explorador de variables, tablas, figuras, archivos, trabajos, autocompletado y ayuda. El diagnóstico técnico solo aparece al abrir Estado de Python….

Ejemplo concreto. Abra Nuevo → Jupyter Notebook, ejecute una celda con Ctrl+Enter, compruebe nombres y tipos en Variable Explorer, vea la figura en Plots y consulte Estado de Python… solo cuando necesite diagnosticar el motor.

ACCESOS INDEPENDIENTES · siempre visibles

Gráficos y Resultados permanecen fuera de los cuatro desplegables para recuperar productos de cualquier análisis sin cambiar el modo seleccionado.

Gráficos

Ruta. Gráficos; seleccione la familia de visualización.

Para qué usarlo. Crear y administrar figuras base, barras con error, georreferenciación y visualizaciones Plotly; recuperar las producidas por otros módulos.

Por qué usarlo. Una figura científica debe responder a la pregunta, mostrar incertidumbre y conservar formato, escala, etiquetas y procedencia ajustables.

Entregables. Figura editable, panel de formato minimizado, exportación y, en georreferenciación y Plotly, interacción, etiquetas y popup configurables.

Ejemplo concreto. En Vigilancia geográfica cree una georreferenciación con latitud y longitud, radio proporcional a tasa_100000, color por evento_salud y popup con ciudad, casos y tasa.

Resultados

Ruta. Resultados y proyecto.

Para qué usarlo. Centralizar interpretaciones, tablas, supuestos, figuras, código y el historial producido por todos los módulos.

Por qué usarlo. Revisar estimación, incertidumbre, supuestos y trazabilidad en un mismo lugar reduce conclusiones basadas solo en el valor p.

Entregables. Historial integral, selección exportable, tablas con decimales por columna, figuras recuperables y código R y Python asociado.

Ejemplo concreto. Después de una logística y una validación predictiva, seleccione ambas entradas, muestre OR con dos decimales y valores p con cinco, y exporte el informe HTML o PDF.

3. Bases demostrativas incluidas

SAmuel incluye una cohorte breve de 240 registros para seguir los ejemplos de este manual y un catálogo de veinte bases sintéticas con 82.820 registros en total. Ningún archivo representa pacientes reales. El acceso único al catálogo está en **Datos → Bases de ejemplo…**. Busque por texto o filtre por nivel, tema y módulo; cada ficha incluye pregunta de práctica, estructura, variables clave y funciones sugeridas. La ventana se puede maximizar; la vista previa permite elegir 25, 50, 100 o 250 filas, muestra todas las columnas y dispone de barras vertical y horizontal. La base activa no cambia hasta que pulse Cargar o haga doble clic.

Escenario	Estructura	Formato	Prácticas destacadas
Cohorte cardiovascular	5.000 participantes	CSV	Tabla 1, supervivencia, regresión, predicción y ROC.
Ensayo de hipertensión	1.800 sujetos, pre/post ancho	XLSX	t/Wilcoxon pareado, McNemar, cambios y modelos con una fila por sujeto.
Seguimiento crónico	1.000 sujetos × 6 visitas, largo	TSV	GEE, mixtos e imputación multinivel.
Casos y controles	4.000 unidades	Stata DTA	OR crudo/ajustado, logística e interacciones.
Encuesta poblacional	6.000 unidades	SPSS SAV	Pesos, estratos, conglomerados y dominios.
Exactitud diagnóstica	3.500 evaluaciones	JSON	Sensibilidad, especificidad, ROC/AUC y DeLong.
Hospitales	5.200 pacientes en 80 hospitales	SAS XPT	Multinivel, GLMM y desenlaces no gaussianos.
Vigilancia geográfica	8.000 puntos	XLSX	Leaflet, tasas, círculos, rutas y polígonos.
Biomarcadores 3D	6.000 participantes	XLSX	3D, burbujas, contorno y coordenadas paralelas.
Monitoreo hospitalario	10.000 horas	XLSX	Plotly, series, áreas, calor y panel interactivo.

3.1 Veinte recorridos de práctica

1. **Cohorte cardiovascular:** cargue CSV, revise Tabla 1, ajuste regresión y Cox, construya forest plot y valide ROC.
2. **Ensayo de hipertensión:** confirme muestras dependientes, compare pre/post con t o Wilcoxon pareado y dos desenlaces binarios con McNemar.
3. **Seguimiento crónico:** identifique sujeto y visita, seleccione varios predictores y ajuste GEE o mixto; practique imputación multinivel.
4. **Casos y controles:** estime OR crudo y ajustado, revise confusión e interacción.
5. **Encuesta poblacional:** declare peso, estrato y UPM; compare Taylor y métodos de réplicas sin filtrar dominios previamente.
6. **Exactitud diagnóstica:** defina referencia, dirección y corte; compare dos AUC pareadas mediante DeLong o bootstrap.
7. **Hospitales:** declare hospital como agrupador, ajuste GLMM binomial/Poisson/binomial negativo/Gamma y revise convergencia.
8. **Vigilancia geográfica:** use latitud, longitud, ciudad, casos y tasa para practicar todos los mapas Leaflet.
9. **Biomarcadores 3D:** combine glucosa, insulina y HbA1c en dispersión 3D, burbujas, contorno y coordenadas paralelas.
10. **Monitoreo hospitalario:** explore series horarias, ocupación, espera y demanda con líneas, áreas y paneles Plotly.

11 **Plan causal y DAG:** declare exposición, desenlace, confusores, mediadores y temporalidad antes de modelar.

12 **Validación predictiva:** compare desarrollo, validación temporal y externa mediante discriminación, calibración y utilidad.

13 **Auditoría de calidad:** detecte faltantes, duplicados, rangos imposibles y fechas incoherentes sin corregir automáticamente.

14 **Inferencia causal:** estime ATE o ATT con IPTW, AIPW o emparejamiento y revise balance y soporte.

15 **Supervivencia avanzada:** practique RMST, riesgos competitivos y proporcionalidad.

16 **Metaanálisis:** combine una fila por estudio y examine heterogeneidad, forest y funnel plots.

17 **Sensibilidad a faltantes:** contraste conclusiones bajo mecanismos MCAR, MAR y MNAR simulados.

18 **Gobernanza:** evalúe riesgo de reidentificación, consentimiento y restricciones de uso compartido.

19 **Expediente reproducible:** documente lotes, versiones, responsables, semilla y huellas de procedencia.

20 **Muestreo y potencia:** calcule tamaño y seleccione unidades por método aleatorio, sistemático, estratificado o por conglomerados.

La cohorte breve `cohorte_sintetica.csv` continúa siendo la referencia de los ejemplos numéricos del manual:

Variable	Naturaleza sugerida	Rol en los ejemplos
`id_sintetico`	Identificador nominal	Unidad de análisis; no predictor clínico.
`marca_imatinib`	Cualitativa dicotómica	Exposición: Innovador/Genérico.
`edad`	Cuantitativa continua	Descripción y covariable.
`sexo`, `adherencia`	Cualitativas	Frecuencias y estratificación.
`multimorbilidad`	Dicotómica 0/1	Covariable.
`no_respuesta_12m`	Dicotómica 0/1	Desenlace; 1 = no respuesta.
`tiempo_meses`	Tiempo cuantitativo ≥0	Seguimiento.
`evento_no_respuesta`	Evento 0/1	1 = evento; 0 = censura.
`calidad_vida_0m`, `6m`, `12m`	Cuantitativas	Mediciones repetidas.

No use estos resultados como evidencia clínica. Los valores fueron simulados para docencia y validación de la interfaz.

4. Flujo analítico seguro y reproducible

4.1 Antes de abrir el programa

- Formule una pregunta y especifique población, exposición/intervención, comparador, desenlace y tiempo.
- Identifique la unidad de análisis: persona, consulta, hospital, medición o episodio.
- Defina variables, codificación, categorías de referencia y ventana temporal antes de mirar asociaciones.
- Preespecifique análisis primario, secundarios y de sensibilidad.

4.2 Dentro de SAmuel

1. Conserve intacto el archivo fuente.
2. Revise tipos, etiquetas, rangos imposibles, duplicados y faltantes.
3. Confirme si las observaciones son independientes o pareadas.
4. Describa toda variable antes de contrastarla.
5. Revise estimación, IC 95%, tamaño del efecto, supuestos y gráficos; no solo el valor p.
6. Documente exclusiones y decisiones en el Historial integral.

4.3 Después del análisis

Reproduzca el resultado con el código generado cuando el análisis sea crítico, sométalo a revisión y reporte siguiendo la guía pertinente (por ejemplo STROBE para estudios observacionales o TRIPOD para modelos predictivos).

5. Carga, exploración y diseño de muestras

5.1 Formatos y estructura

SAmuel admite CSV, TSV, TXT delimitado, JSON, JSON Lines, Excel (`.xlsx`, `.xlsm`, `.xls`), SPSS (`.sav`, `.zsav`), Stata (`.dta`) y SAS (`.sas7bdat`, `.xpt`). En Excel se carga la primera hoja utilizable; se retiran filas/columnas completamente vacías y los nombres duplicados reciben un sufijo.

Regla	Recomendación
Una fila	Una unidad de análisis o una medición claramente identificada.
Una columna	Una variable; no mezcle número y unidad en una misma celda.
Nombres	Únicos, breves y estables; evite nombres vacíos.
Faltantes	Celdas vacías o una codificación única documentada; nunca 999 sin recodificar.
Fechas	Formato coherente y origen temporal explícito.

5.2 Paso a paso

1. Pulse **Cargar base…** y seleccione el archivo, o abra **Bases de ejemplo…** para practicar.

2. Compruebe filas, variables y porcentaje de faltantes en el encabezado.

3. Use **Buscar** para filtrar cualquier fragmento visible.

4. Pulse un encabezado para ordenar. Ajuste las filas por página y use **Decimales…** si necesita una precisión diferente en cada variable.

5. En **Columnas…** o en **Variables → Habilitada para análisis** desactive identificadores o campos administrativos. Desaparecerán de pruebas, modelos y gráficos, pero seguirán en la base y podrán reactivarse.

6. Use **Exportar base a Excel…** para obtener un libro con columnas listas, encabezados, filtro y fila superior fija, sin importar texto delimitado.

7. Abra **Diseño de muestras…** y confirme independencia o dependencia.

> **Independientes**: cada fila aporta una unidad distinta. **Pareadas en formato ancho**: una fila contiene al mismo sujeto y las mediciones relacionadas están en columnas, por ejemplo `presion_pre`/`presion_post`. **Longitudinales en formato largo**: el identificador se repite en varias filas y una variable indica tiempo/visita. En este último caso, GEE o modelos mixtos deben conservar la correlación intrapersonal. Tratar filas repetidas como sujetos independientes suele subestimar la incertidumbre.

Ejemplo. Cargue la cohorte sintética, busque “Innovador” y compruebe que la tabla se filtra sin modificar la base. Limpie la búsqueda para recuperar las 240 filas.

5.3 Decimales independientes por columna

Las tablas de **Datos, Variables, Imputación, Resultados, Modo R y Modo Python** incluyen el botón **Decimales…**. En el diálogo, asigne **Auto** o de 0 a 10 cifras a cada encabezado de manera independiente y pulse Aplicar. El número entre paréntesis del botón indica cuántas columnas tienen formato personalizado. Por ejemplo, puede mostrar `Valor p` con 5 cifras y una estimación clínica con 2. La configuración modifica únicamente la presentación: el valor almacenado, los análisis, el orden y los filtros permanecen intactos.

6. Diccionario, naturaleza y nivel de medición

La clasificación determina qué procedimientos se recomiendan y cómo se interpretan. La inferencia automática es orientativa: el significado clínico prevalece sobre el tipo de almacenamiento.

Naturaleza y nivel de medición

Variable	Nivel	Qué revisar
Identificador	Nominal	Distingue unidades
Etapa de enfermedad	Ordinal	Tiene orden
Temperatura °C	Intervalo	Diferencias interpretables
Edad en años	Razón	Cero interpretable
Fecha de seguimiento	Temporal	Estructura en el tiempo

SAmuel web · EPINex · Ejemplo didáctico con datos sintéticos

Ejemplos sintéticos de variables y su nivel de medición; el investigador confirma su significado.

6.1 Habilitar o inhabilitar variables para análisis

1. Seleccione la variable en el campo **Variable**.

2. Desmarque **Habilitada para análisis** si corresponde a un identificador, folio, consecutivo, código administrativo o texto libre no analítico.

3. Pulse **Aplicar metadatos**. La fila permanece en el diccionario con **uso_en_análisis = No**, pero desaparece de las listas de pruebas, modelos, imputación y gráficos.

4. Para recuperarla, selecciónela de nuevo, marque la casilla y aplique los metadatos. La base original no pierde la columna.

Regla inicial de SAmuel: nombres explícitos como `id`, `identificación`, `id_paciente`, `sujeto` o `participante` se conservan, pero comienzan inhabilitados. Un identificador sirve para reconocer unidades o vincular registros; normalmente no representa una característica que deba contrastarse. La exclusión automática se basa en el nombre y debe confirmarse en el protocolo.

Naturaleza	Ejemplo	Resumen apropiado	Error frecuente
Cuantitativa continua	Edad, creatinina	Media/DE o mediana/RIC	Crear categorías sin justificación.
Cuantitativa discreta	Número de hospitalizaciones	Mediana/RIC o modelo de conteo	Usar regresión lineal con conteos muy asimétricos.
Cualitativa dicotómica	Evento 0/1	n y %	No declarar cuál valor es el evento.
Cualitativa politómica nominal	Tratamiento A/B/C	n y % por categoría	Asignar orden inexistente.
Ordinal	Estadio I-IV	Mediana/rangos o frecuencias	Suponer distancias iguales.
Temporal	Tiempo hasta evento	Kaplan-Meier/Cox	Ignorar censura u orígenes distintos.

6.2 Guía visual de tipo y nivel

La primera columna del diccionario combina un pictograma con texto para reconocer de inmediato la naturaleza y el nivel. La leyenda superior identifica continua, discreta, dicotómica, politómica y fecha, junto con nominal, ordinal, intervalo, razón, dinero y temporal. El texto permanece visible aunque el sistema no dibuje un pictograma; la guía no cambia la clasificación ni sustituye su revisión.

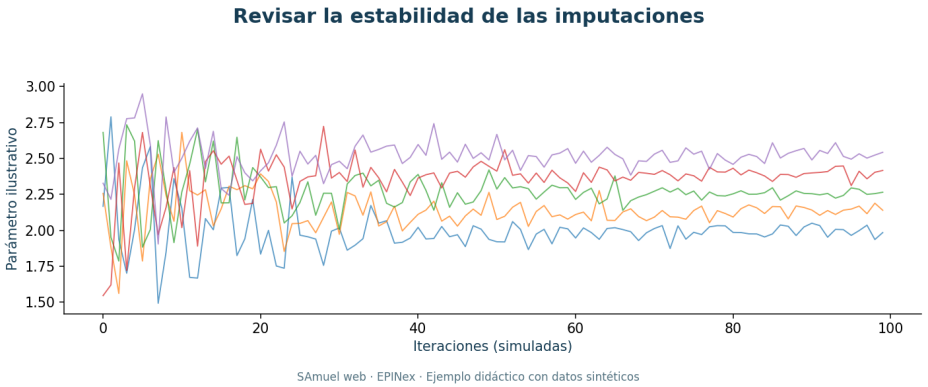
Nivel Dinero. Solo está disponible para variables cuantitativas. Al guardar, SAmuel muestra el signo **$**, punto como separador de miles y coma para decimales: `$2.742.564` o `$2.742.564,50`. Es un formato visual aplicado en Datos, el ejemplo del diccionario y las tablas compatibles; internamente el valor continúa siendo numérico, de modo que los cálculos no cambian.

6.3 Ejercicio guiado: edad

Variables → escriba “edad” → Naturaleza: Cuantitativa continua → Nivel: Razón/intervalo → Guardar.

Si la edad está en años cumplidos, una diferencia de un año es interpretable; documente la unidad. No la cambie a cualitativa solo porque tenga valores enteros.

7. Datos faltantes e imputación



Trazas simuladas con propósito didáctico; no certifican convergencia ni el mecanismo de ausencia.

7.1 Conceptos

Mecanismo	Idea	Implicación
MCAR	La ausencia no depende de datos observados ni no observados.	El análisis de casos completos puede ser insesgado, pero pierde precisión.
MAR	La ausencia puede explicarse por variables observadas.	La imputación múltiple requiere incluir predictores adecuados de ausencia y resultado.
MNAR	La ausencia depende del valor no observado incluso tras condicionar.	Exige supuestos explícitos y análisis de sensibilidad.

7.2 Uso

1. Revise número y porcentaje de celdas faltantes.

2. Seleccione una fila. La sección **Justificación completa de la variable seleccionada** inicia minimizada para dar más espacio a la tabla; pulse **Ver justificación** para consultar el motivo íntegro y vuelva a contraerla cuando termine.

3. Elija método; si usa valor constante, escríbalo y justifíquelo.

4. Pulse **Aplicar a esta variable** o revise individualmente antes de aplicar recomendaciones.

5. Use **Deshacer última imputación** si la decisión fue incorrecta.

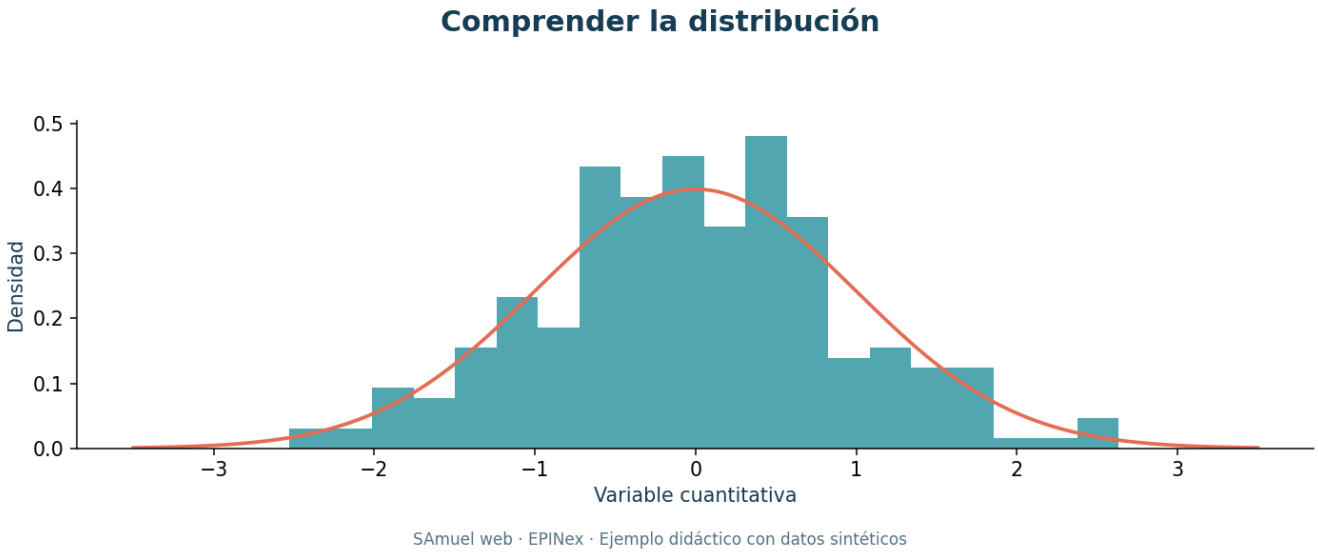
> **Imputación simple.** Media, mediana, moda, LOCF o NOCB no incorporan adecuadamente la incertidumbre y pueden reducir artificialmente la variabilidad. Son útiles para exploración o contextos muy específicos, no una solución universal para inferencia confirmatoria.

Por qué se sugiere cada método: el diccionario prevalece: una cualitativa codificada 0/1 recibe moda, no media. En cuantitativas, la distribución orienta media o mediana. SAmuel no propone interpolación por el orden de las filas: exige tiempo real, trayectoria plausible y separación por sujeto. La justificación íntegra permanece en la celda y se lee en el panel inferior.

> **Ejemplo.** `calidad_vida_12m` tiene 4 faltantes (1,7%). Antes de imputar, compare participantes con y sin medición, revise el patrón temporal y haga un análisis de sensibilidad de casos completos. Si usa mediana para una exploración, documente que fue imputación simple.

8. Análisis univariado

Describe una o varias variables sin relacionarlas con otras. Es el paso obligatorio para detectar errores, comprender la distribución y seleccionar resúmenes compatibles.



Distribución cuantitativa sintética, referencia normal y resumen gráfico.

8.1 Medidas disponibles

Familia	Opciones de SAmuel	Uso
Tendencia central	Media, mediana, moda, geométrica, armónica, recortada 10%	Centro de la distribución según escala y forma.
Dispersión	DE, varianza, rango, mínimo, máximo, RIC, MAD, CV, error estándar	Variabilidad de observaciones o precisión del estimador.
Posición	Percentiles, deciles, cuartiles, mediana	Ubicación relativa.
Forma	Asimetría, curtosis, Shapiro-Wilk/D'Agostino, resumen completo	Diagnóstico de distribución.
Precisión	EE, IC 95% de media, IC 95% bootstrap de mediana	Incertidumbre del estimador.
Cualitativas	Distribuciones separadas de una o varias variables	Conteos y frecuencias relativas de cada variable, sin cruzarlas.
Tabla 1 científica	Todas las variables o subconjunto; total y estratificación por grupos	Características generales, basales o demográficas en formato de artículo.
Tiempo	Kaplan-Meier	Función de supervivencia con censura.

8.2 Ejemplo: mediana de edad

Univariado → Medida de tendencia central → Mediana → añada `edad` → Ejecutar análisis.

En la cohorte sintética la mediana es 55 años (Q1=42; Q3=66). El 50% de las observaciones tiene 55 años o menos y el otro 50% 55 años o más. El 50% central se encuentra aproximadamente entre 42 y 66 años. Esto no implica que la mayoría tenga exactamente 55 años.

Redacción sugerida: "La mediana de edad fue 55 años (RIC: 42-66; n=240)". Incluya unidades, n y faltantes.

8.3 Media, DE, mediana y RIC

La edad tuvo media 54,25 años (DE 16,21), mediana 55 y Shapiro-Wilk p=0,094. La calidad de vida a 12 meses tuvo media 74,83 (DE 15,13), mediana 74,75 y p=0,0019. Un valor p pequeño orienta a desviación de normalidad, pero el diagnóstico debe integrar histograma, Q-Q, asimetría, atípicos y tamaño muestral.

DE ≠ EE. La desviación estándar describe variabilidad entre observaciones; el error estándar describe precisión de una estimación y disminuye con n. No reporte media ± EE como si describiera la dispersión de pacientes.

8.4 Frecuencias cualitativas múltiples

Univariado → Frecuencias para variables cualitativas → Distribuciones de frecuencias separadas → añada `sexo` y `adherencia` → Ejecutar.

SAmuel entrega ambas distribuciones en el mismo resultado, pero cada variable se analiza por separado. No genera cruces ni asociaciones en Univariado. Para relacionar categorías pulse **Crear tabla de contingencia en Bivariado** →. Seleccione acumulados solo si el orden de las categorías tiene significado.

8.5 Tabla 1 científica: toda la base en una ejecución

Univariado → Tabla 1 científica → Características basales → añada únicamente las variables que desea presentar al panel “Variables elegidas” → seleccione el diseño y la variable de grupos → Ejecutar análisis. Para incluir variables en grupo, use “Añadir por tipo” y elija “Cualitativas”, “Cuantitativas” o “Todas”. Sin variables elegidas no se ejecuta la tabla.

SAmuel presenta una sola tabla de salida. Su orden es: variable, categoría, columnas de los grupos, total, valor p, prueba estadística y presentación. Las cuantitativas compatibles con normalidad se muestran como media (DE); las no compatibles, como mediana (Q1-Q3). Cada categoría cualitativa se informa como n (%). El valor p y la prueba se repiten en todas las categorías de su variable para evitar celdas ambiguas. Los faltantes aparecen explícitamente y los identificadores o textos libres con al menos 90% de valores únicos se excluyen con una advertencia.

Situación	Prueba automática orientativa	Presentación
Cuantitativa normal, 2 grupos	t de Student o Welch según Levene	Media (DE)
Cuantitativa no normal, 2 grupos	Mann-Whitney	Mediana (RIC)
Cuantitativa, 3 o más grupos	ANOVA/Alexander-Govern o Kruskal-Wallis	Según distribución
Cualitativa	Chi-cuadrado; Fisher si es 2×2 dispersa	n (%) por categoría

Cómo leerla: un valor p menor de 0,05 indica evidencia de diferencias globales entre grupos para esa característica bajo los supuestos de la prueba. No demuestra que la selección haya sido adecuada o inadecuada, no mide importancia clínica y no confirma equivalencia cuando es mayor de 0,05. Revise magnitud, distribución, faltantes y diseño.

Ensayos aleatorizados. CONSORT desaconseja las pruebas de significación de diferencias basales, porque las diferencias observadas tras una aleatorización válida son producto del azar. SAmuel desactiva inicialmente el valor p para este diseño; puede activarlo de forma explícita, pero debe justificarlo. En estudios observacionales, los valores p siguen siendo auxiliares y no sustituyen el análisis de confusión o sesgo de selección.

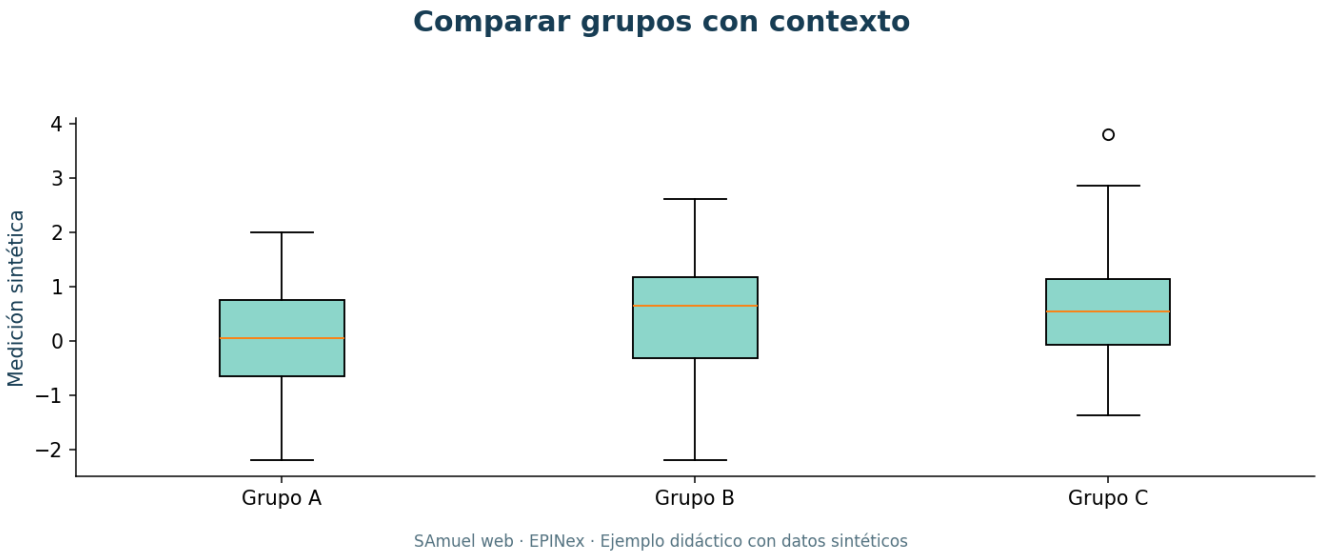
Tabla 1 y dependencia: las comparaciones automáticas entre columnas de grupo presuponen sujetos independientes entre grupos. En una base pareada ancha, una fila por sujeto permite describir variables y comparar brazos independientes, pero no usar las columnas pre/post como si fueran grupos distintos. En datos longitudinales largos con el identificador repetido, el valor p se desactiva; use un modelo longitudinal o prepare una tabla basal con una fila por sujeto.

Ejemplo de cohorte: elija `marca_imatinib` como grupo. La fila de edad mostrará el resumen por cada cohorte, el total, el valor p, la prueba y la presentación. Las filas de sexo mostrarán Mujer y Hombre con n (%), repitiendo la prueba global y su valor p en ambas categorías. La información aparece en una única tabla; los diagnósticos se incorporan a la interpretación automática.

8.6 Contraste frente a un valor

La t de una muestra contrasta una media frente a un valor de referencia bajo independencia y distribución razonable de la variable. Wilcoxon de una muestra contrasta localización/rangos y requiere interpretar con cuidado el supuesto de simetría.

9. Asociación, comparación y correlación



Grupos sintéticos: compare distribución y magnitud, además del contraste.

9.1 Dos variables cualitativas

Bivariado → Resultado: `no_respuesta_12m` → Independiente: `marca_imatinib` → Chi-cuadrado/Fisher → configure referencias, porcentaje y medida cruda → Ejecutar.

En **Tabla de contingencia y medida cruda** elija la categoría de referencia del desenlace y de la exposición. La referencia aparece primero y la categoría de comparación después. Cada celda muestra simultáneamente **n (porcentaje)**; el denominador se puede cambiar entre fila, columna y total. Para tablas 2×2 puede añadir OR, RR o RP, con IC 95% y valor p. HR también está disponible cuando se indica el tiempo hasta el evento, la categoría natural que significa evento y la referencia del grupo.

La tabla observada del ejemplo contiene: Genérico 89 sin evento y 37 con evento; Innovador 88 sin evento y 26 con evento. Las frecuencias esperadas mínimas son 29,93, por lo que Pearson es razonable.

Resultado: $\chi^2(1)=1,33$; $p=0,249$; V de Cramér=0,074. No se obtuvo evidencia estadística suficiente de asociación entre marca y no respuesta en esta muestra. El tamaño de asociación observado fue pequeño. Esto no demuestra equivalencia, ausencia de efecto ni causalidad.

9.2 Cuantitativa entre dos grupos

Procedimiento	Supuestos/objetivo	Lectura
t de Student	Independencia, medias, residuos razonables, varianzas comparables.	Diferencia de medias e IC 95%.
t de Welch	Como t, sin exigir igualdad de varianzas.	Preferible ante heterocedasticidad.
Mann-Whitney	Independencia; comparación de distribuciones/rangos.	No es automáticamente una prueba de medianas.

9.3 Tres o más grupos

ANOVA de una vía compara medias; Welch admite varianzas distintas; Kruskal-Wallis compara rangos/distribuciones. Una prueba global significativa no identifica qué pares difieren: revise los contrastes post hoc, multiplicidad e IC.

9.4 Correlación

Método	Cuándo	No significa
Pearson	Relación aproximadamente lineal entre cuantitativas.	Causalidad ni acuerdo.
Spearman	Relación monótona, ordinales o sensibilidad a atípicos/no normalidad.	Linealidad.
Kendall	Orden, muestras pequeñas o muchos empates.	Que el efecto sea clínicamente grande.

9.5 Muestras relacionadas

Use t pareada sobre diferencias aproximadamente normales; Wilcoxon para rangos con signo; McNemar para desenlaces dicotómicos pareados; Friedman para tres o más mediciones relacionadas. La unidad de análisis sigue siendo el sujeto/par, no cada medición aislada.

Protección frente a combinaciones incompatibles: si las dos variables pareadas no corresponden a dos cuantitativas ni a dos dicotómicas, SAmuel muestra “Revisión metodológica requerida” en lugar de forzar una prueba inválida. Para respuestas politómicas pareadas, conteos repetidos o más de dos momentos pueden requerirse métodos marginales, mixtos o pruebas específicas no implementadas en este panel.

9.6 Cómo se construye la sugerencia

Estructura	Datos	Orientación automática	Comprobación del usuario
Independiente	Dos cualitativas	Chi-cuadrado; Fisher solo en 2×2 con esperados bajos.	Categorías excluyentes, independencia y frecuencias esperadas.
Independiente	Cuantitativa y 2 grupos	Student si normalidad por grupo y varianzas comparables; Welch si normal con varianzas distintas; Mann-Whitney si no normal.	Gráficos, atípicos, forma, tamaño muestral y pregunta sobre medias o distribuciones.
Independiente	Cuantitativa y 3+ grupos	ANOVA, ANOVA de Welch o Kruskal-Wallis con el mismo criterio.	Contraste global, post hoc, multiplicidad y tamaños de efecto.
Independiente	Dos cuantitativas	Pearson si ambas son compatibles con normalidad; Spearman en caso contrario.	La decisión final exige dispersión, linealidad/monotonicidad, atípicos y empates.
Pareada	Dos cuantitativas	t pareada si las diferencias son compatibles con normalidad; Wilcoxon en caso contrario/no evaluable.	Q-Q y atípicos de las diferencias; simetría, ceros y empates para Wilcoxon.
Pareada	Dos dicotómicas	McNemar.	Pares correctamente vinculados y discordancias suficientes.

Transparencia: debajo de la prueba sugerida aparece el motivo con naturaleza, diseño, número de grupos y diagnósticos calculados. La normalidad nunca debe evaluarse solo con un valor p: complemente con histograma, gráfico Q-Q, forma y valores atípicos.

10. RR, OR y razón de prevalencias en tablas 2×2

Medida	Fórmula conceptual	Interpretación
Riesgo relativo (RR)	Riesgo en expuestos / riesgo en no expuestos	Razón de incidencias acumuladas; requiere temporalidad y seguimiento apropiado.
Odds ratio (OR)	Odds en expuestos / odds en no expuestos	Razón de odds; útil en casos y controles y logística. No equivale a RR si el evento es frecuente.
Razón de prevalencias (RP)	Prevalencia en expuestos / prevalencia en no expuestos	Comparación transversal de prevalencias.

Ejemplo: RR crudo

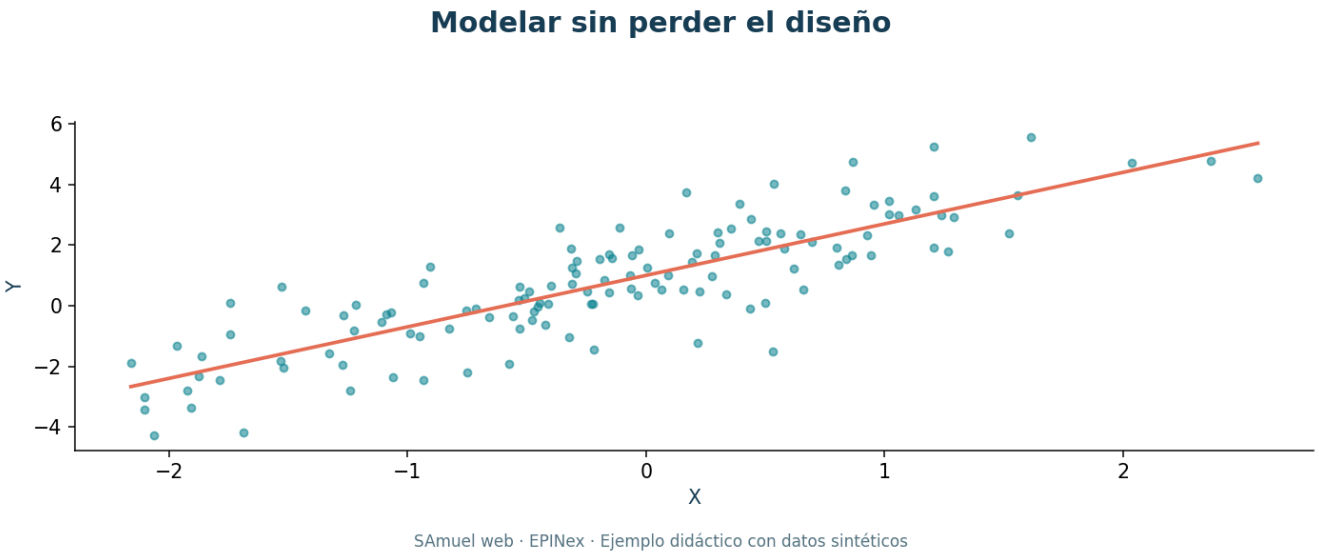
Bivariado → Resultado `no_respuesta_12m` → Independiente `marca_imatinib` → Chi-cuadrado/Fisher → referencia del desenlace 0 → referencia de exposición Innovador → medida cruda RR.

Riesgo en Genérico: 37/126=29,4%. Riesgo en Innovador: 26/114=22,8%. RR=1,29 (IC 95% 0,84-1,99).

Interpretación cruda: la incidencia acumulada observada de no respuesta fue 1,29 veces en Genérico respecto de Innovador. El IC incluye 1 y es compatible tanto con reducción moderada como con aumento; la estimación es imprecisa. No está ajustada por edad, multimorbilidad u otros factores.

No escriba “Genérico aumenta 29% el riesgo” como conclusión causal sin diseño, temporalidad, control de confusión y evaluación de sesgos.

11. Modelos multivariados: explicación y ajuste



Ejemplo sintético de relación y ajuste lineal; no implica causalidad.

11.1 Modelos disponibles

Desenlace	Modelo	Salida principal
Continuo	Lineal	β: cambio medio ajustado por unidad/categoría.
Continuo positivo asimétrico	Gamma con enlace log	Razón de medias.
Conteo	Poisson o binomial negativa	IRR.
Dicotómico	Logística	OR ajustado.
Dicotómico	Poisson robusto	RR o RP ajustada, según diseño/interpretación.
Politómico nominal	Logística multinomial	OR por categoría frente a referencia.
Ordinal	Logística ordinal	OR proporcional acumulado.
Tiempo-evento	Cox	HR ajustado.

El panel explica por qué propone la familia a partir de la naturaleza y nivel del desenlace, y enumera el supuesto crítico que debe comprobarse. La sugerencia no selecciona covariables por valor p ni decide causalidad; el usuario puede cambiarla antes de ajustar.

11.2 Modelos cuando las muestras son pareadas o dependientes

El botón **Ajustar modelo** permanece activo, con texto blanco sobre fondo verde, y la validación se realiza al ejecutarlo. En formato ancho, con una fila por sujeto, puede modelar un único desenlace -por ejemplo abandono- usando predictores basales; las comparaciones pre/post continúan requiriendo pruebas pareadas. En formato largo, donde el identificador se repite por visita, los modelos clásicos no deben tratar cada fila como independiente: abra la pestaña integrada **3. Modelos longitudinales** y seleccione GEE o un modelo mixto con identificador de sujeto/conglomerado.

> **No confunda estructuras.** Tener columnas pre/post no convierte en dependientes a sujetos distintos entre filas. Tener varias filas del mismo sujeto sí exige modelar la correlación. SAmuel distingue ambas estructuras y conserva el botón accesible para explicar o ejecutar el procedimiento compatible. Si el botón aparece gris, cierre esa copia y abra el ejecutable contenido en la carpeta completa de la distribución vigente.

11.3 Paso a paso: RR ajustado

Multivariado → Resultado `no_respuesta_12m` → Explicativa/asociación → RR · Poisson robusto log → Evento 1 → covariables `marca_imatinib`, `edad`, `multimorbilidad` → Ajustar modelo.

Para Innovador frente a Genérico, RR ajustado=0,76 (IC 95% 0,49-1,17; p=0,213). Edad: RR por año=1,00 (IC 95% 0,98-1,01). Multimorbilidad: RR=1,32 (IC 95% 0,84-2,09).

Redacción ajustada: “Tras ajustar simultáneamente por edad y multimorbilidad, el riesgo estimado de no respuesta en el grupo Innovador fue 0,76 veces el del grupo Genérico (IC 95%: 0,49-1,17). El intervalo incluye 1 y no permite descartar ausencia de asociación; tampoco prueba equivalencia”.

11.4 OR ajustado

En la logística con las mismas variables, Innovador tuvo OR ajustado=0,69 (IC 95% 0,38-1,24; p=0,213). Diga siempre qué variables permanecen constantes. El OR compara odds y no riesgo; con eventos frecuentes, puede alejarse del RR.

11.5 Confusión, mediación y colisión

Las covariables no deben seleccionarse únicamente por valor p. Use conocimiento temporal y un diagrama causal: ajuste confusores preexposición; evite ajustar mediadores si busca efecto total y colisionadores que pueden inducir sesgo. Reporte la estrategia preespecificada.

11.6 Diagnósticos mínimos

- Codificación y categoría de referencia.
- Casos completos y pérdidas por faltantes.
- Número de eventos frente a parámetros.
- Forma funcional de predictores continuos.
- Colinealidad, categorías escasas, separación e influencia.
- Convergencia, residuos y ajuste del modelo.
- Coherencia temporal y plausibilidad causal.

12. Kaplan-Meier, Log-rank y Cox

El análisis incorpora seguimiento incompleto mediante censura. Cada sujeto necesita un tiempo no negativo, una variable de estado y una variable de grupo. El estado no tiene que estar codificado 0/1: en Bivariado seleccione la categoría observada que significa evento, por ejemplo `Fallo terapéutico`, `Falleció` o `Sí`; las demás categorías válidas se tratan como censura.

12.1 Kaplan-Meier

La función estimada es el producto de probabilidades condicionales de permanecer sin evento. La mediana es el tiempo cuando la supervivencia cruza 0,5; puede no ser estimable.

12.2 Log-rank

Compara globalmente curvas bajo la hipótesis nula de igualdad. En Bivariado, use como resultado la variable de estado, como exposición la variable de grupo y, en **Configuración de supervivencia**, seleccione el tiempo, la categoría que significa evento y la referencia del grupo. Es más sensible a diferencias persistentes y no expresa por sí mismo una magnitud.

12.3 Cox

Modela el hazard instantáneo relativo. HR>1 indica mayor tasa instantánea del evento, condicionada a covariables y supuestos; no es una razón de riesgos acumulados.

Ejemplo. Con `tiempo_meses`, `evento_no_respuesta` y grupo `marca_imatinib`, las medianas estimadas fueron 17,0 meses (Innovador) y 14,5 (Genérico); Log-rank p=0,341. El Cox ajustado por edad y multimorbilidad estimó HR=0,85 (IC 95% 0,62-1,18) para Innovador.

Interpretación: no hubo evidencia estadística suficiente de diferencia entre las curvas. El HR ajustado fue compatible con una tasa instantánea menor o mayor. La prueba de Schoenfeld para edad tuvo p=0,0056 en el ejemplo, señal de posible incumplimiento de proporcionalidad que impide aceptar el resumen Cox sin investigación adicional.

Revise origen temporal común, censura no informativa, riesgos proporcionales, forma funcional y eventos competidores. Si las curvas se cruzan, el log-rank y un HR único pueden resumir mal el patrón.

13. Desarrollo y evaluación de modelos predictivos

Seleccione **Predictiva** solo cuando el objetivo sea estimar resultados futuros en individuos. Una asociación ajustada no constituye automáticamente un modelo predictivo.

Estadístico	Qué resume	Interpretación prudente
AUC	Discriminación entre casos y no casos.	0,5≈azar; no informa calibración ni utilidad clínica.
Brier	Error cuadrático de probabilidades.	Menor suele ser mejor, pero depende de prevalencia y comparación.
Sensibilidad/especificidad	Rendimiento en un umbral.	Dependen del umbral y consecuencias de errores.
Exactitud	Proporción correctamente clasificada.	Puede engañar con clases desbalanceadas.
Concordancia (C-index)	Orden predictivo en supervivencia.	Desempeño aparente si se calcula en desarrollo.

Ruta y reporte

Multivariado → seleccione desenlace → Intencionalidad Predictiva → medida/modelo → predictores → Ajustar modelo.

- Justifique tamaño muestral y eventos por parámetro.
- Evite selección automática no validada y dicotomización de continuas.
- Modele no linealidad e interacciones plausibles.
- Use validación interna (bootstrap o validación cruzada) y, antes de aplicación, validación externa.
- Reporte calibración, discriminación e incertidumbre.
- Defina umbrales por consecuencias clínicas y análisis de utilidad, no por exactitud aparente.

El desempeño mostrado por SAmuel es aparente en la muestra analizada. No demuestra transportabilidad ni beneficio clínico. Para publicación, siga TRIPOD y valore herramientas de riesgo de sesgo como PROBAST.

14. Cómo elegir la prueba o el modelo

Pregunta	Datos	Opción orientativa	Comprobación clave
Describir centro	Una cuantitativa	Media/DE o mediana/RIC	Distribución, atípicos, escala.
Comparar dos grupos	Cuantitativa + grupo binario	Student/Welch/Mann-Whitney	Independencia, varianzas, forma.
Comparar ≥3 grupos	Cuantitativa + grupo	ANOVA/Welch/Kruskal-Wallis	Global vs post hoc, multiplicidad.
Asociar cualitativas	Tabla r×c	Chi-cuadrado/Fisher	Esperadas, independencia.
Dos mediciones	Mismos sujetos	t pareada/Wilcoxon/McNemar	Emparejamiento real.
Relacionar continuas	Dos cuantitativas	Pearson/Spearman/Kendall	Dispersión, linealidad/monotonía.
Evento binario ajustado	Desenlace 0/1	Logística o Poisson robusto	Interpretar OR vs RR/RP.
Conteo ajustado	Enteros ≥0	Poisson/binomial negativa	Dispersión y exposición-tiempo.
Tiempo-evento	Tiempo + evento	KM/Log-rank/Cox	Censura y proporcionalidad.

Regla de oro: la pregunta, el diseño y la unidad de análisis eligen el método; la normalidad por sí sola no lo hace.

15. Figuras científicas y editor de formato

SAmuel inicia con fondo y rellenos blancos, contornos negros y sin color, grises, degradados, tramas, cuadrícula ni título interno. El usuario puede activar independientemente color, degradado, tramas y cuadrícula, además de editar texto, relleno, líneas, ejes, leyenda, dimensiones y resolución. **Generar gráfico** y **Exportar gráfico…** se ofrecen una sola vez en la barra inferior fija: siguen visibles aunque el constructor o el panel de formato requieran desplazamiento.

15.1 Catálogo

Familia	Opciones
Univariado	Histograma+density, Q-Q, ECDF, puntos, boxplot, violin, barras, barras con error, Pareto y nube de palabras.
Bivariado	Dispersión+tendencia, hexagonal, box/violin por grupo, barras apiladas/agrupadas, barras con error por grupo, calor de contingencia, trayectorias pareadas, Bland-Altman, Kaplan-Meier y burbujas.
Multivariado	Matriz de correlación, PCA, varianza explicada, coordenadas paralelas, forest plot y burbujas.
Georreferenciación	Mapas interactivos Leaflet con marcadores, círculos proporcionales, rutas y polígonos; 10 fondos detallados en línea, 12 capas locales, medición, etiquetas y fichas configurables por registro.
Especiales Plotly	Dispersión, líneas/áreas, barras, histograma, caja, violín, burbujas, calor, contorno, 3D y coordenadas paralelas interactivas; etiquetas, lectura al pasar y ficha al hacer clic configurables.

15.1.1 Barras de error

En **Univariado → Barras con error** seleccione de una a cuatro variables cuantitativas. En **Bivariado → Barras con error por grupo** seleccione grupo y resultado cuantitativo. El control **Barras de error** permite IC 95%, error estándar o desviación estándar. La presentación inicial muestra únicamente el punto de la media y su barra de error con topes, sin columnas rellenas; informe en el pie qué representa el intervalo y el tamaño de cada grupo.

15.1.2 Georreferenciación (motor Leaflet)

1. Seleccione latitud y longitud; SAmuel valida los rangos geográficos. Añada una etiqueta y, opcionalmente, una variable cuantitativa no negativa para el tamaño.
2. Las opciones de formato se abren minimizadas. Pulse **Mostrar opciones de formato** para elegir grupos, campos del popup, forma e imagen del marcador. El gráfico ocupa la ventana; ancho y alto en cero significan tamaño adaptable.
3. El fondo inicial es **OpenStreetMap · en línea**. El selector situado arriba a la derecha ofrece diez fondos públicos: OpenStreetMap, Humanitario, OpenTopoMap y mapas Esri de imágenes, calles, topografía, terreno, océanos, National Geographic y lienzo gris. Requieren Internet y conservan la atribución y disponibilidad del proveedor.
4. SAmuel abre el mapa desde una dirección privada de este equipo (`127.0.0.1`), no directamente como `file://`. Esto evita el bloqueo que algunos servidores de mapas aplican a archivos locales. No publica los datos ni los envía a un servidor de SAmuel; el navegador solicita al proveedor únicamente las teselas visibles del fondo elegido.
5. Sin Internet, seleccione **Natural Earth · local**. Hay doce capas mundiales incorporadas: países, superficie terrestre, océanos, costas, lagos, ríos, ciudades, límites regionales, Ecuador y trópicos, glaciares, regiones geográficas y regiones marinas. Están generalizadas a escala 1:110 millones y no incluyen calles detalladas ni fotografía satelital.
6. En **Elegir grupos y campos del popup**, añada variables en el orden de la jerarquía y pulse **Actualizar niveles seleccionados** para asignar, si lo desea, una imagen distinta a cada nivel. Cada grupo conserva un color determinista y los puntos con la misma combinación de características comparten color. Al hacer clic, el nivel principal permanece visible, las unidades hijas se despliegan alrededor unidas por líneas jerárquicas y el mapa amplía progresivamente la selección. **Subir un nivel** permite regresar. Los conglomerados acercan primero la vista y, en el máximo zoom, separan los puntos superpuestos.

7 En **Formato del gráfico**, cambie título, letra, forma, relleno, borde, grosor, línea, opacidad y dimensiones. Puede utilizar una imagen PNG, JPEG, WebP o GIF, incorporada como PNG en el HTML.

8 Con una variable de tamaño, el área de los símbolos crece proporcionalmente al valor. El cero tiene radio cero; si falta el valor, se conserva el tamaño base. En modo **Radio geográfico**, el radio en metros equivale al valor multiplicado por los metros por unidad elegidos.

9 En **Etiquetas y popup por paciente**, seleccione la variable principal y escriba plantillas con `{nombre de columna}`. Como en Plotly, puede añadir variables una a una, incorporar todas las columnas, restaurar la selección inicial, quitar campos, ordenarlos, cambiar sus nombres visibles, ocultar datos vacíos y comprobar la vista previa. Cada punto muestra exclusivamente los valores de su propia fila. Las etiquetas ajustan ancho y saltos de línea para evitar texto vertical o ilegible. Puede usar etiquetas permanentes o desactivar la ficha.

10 **Dibujar y medir** calcula distancia, rumbo del último tramo, perímetro, área y radio en sistema métrico o imperial. Permite quitar el último vértice, terminar con doble clic, copiar la medición, borrar dibujos y exportarlos como GeoJSON; los círculos conservan **radius_m**.

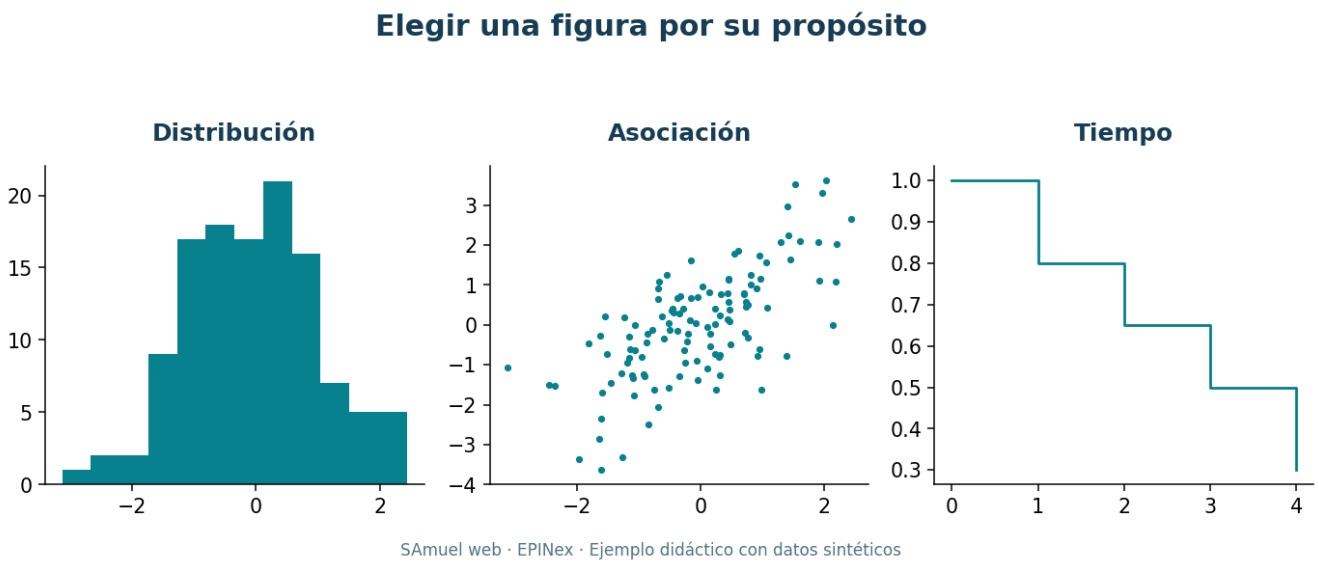
11 En **Vista, coordenadas y fondos** puede añadir un servicio XYZ HTTPS propio con su atribución. No escriba claves privadas si compartirá el HTML. También puede mostrar la cuadrícula, imprimir o guardar el mapa formateado.

12 Si un proveedor no responde, aparece un aviso dentro del mapa. Cambie a otro fondo o a Natural Earth local; los datos y controles no se pierden.

Los gráficos especiales Plotly también abren ampliados y con **Formato y tamaño** plegado. El panel lateral aparece debajo de la barra nativa de Plotly para dejar libres sus controles de cámara, zoom, selección y descarga. Permite editar título, ejes X/Y/Z, ángulos, tipografía y dimensiones; el valor cero adapta cada dimensión a la ventana. En mapas, los ejes corresponden a latitud y longitud: puede mostrar su cuadrícula desde **Vista, coordenadas y fondos**.

15.1.3 Especiales Plotly

Esta familia genera HTML interactivo local con zoom, selección, desplazamiento, lectura al pasar el cursor, edición de títulos y descarga desde la barra del gráfico. Al cambiar de tipo, solo quedan visibles los roles que necesita: X, Y, color/etiqueta y tamaño/Z. Antes de generar, **Elegir información de etiquetas y popup…** define la selección inicial. Dentro del gráfico, **Formato y tamaño** y **Etiquetas y popup** comparten un único panel lateral y funcionan como acordeón: al abrir uno se cierra el otro, por lo que nunca se superponen. El segundo permite añadir una variable, incorporar todas las columnas, restaurar la selección, escribir plantillas, renombrar y ordenar campos, ocultar vacíos, mostrar texto junto a los puntos y activar una ficha amplia y redimensionable por registro al hacer clic. Tema, paleta, leyenda, barra de herramientas, zoom y edición permanecen en el formato guiado; **Mostrar formato** controla tipografía, colores, opacidad, grosor, ángulos, ejes, cuadrícula, fondo y tamaño. El HTML incluye Plotly.js y conserva la interacción sin Internet ni instalación de R, Python o bibliotecas.



Familias de gráficos ilustradas con datos sintéticos.

15.2 Paso a paso: forest plot

1. Familia **Multivariado** y tipo **Forest plot**.

2. Modelo **Regresión logística binaria · OR**.

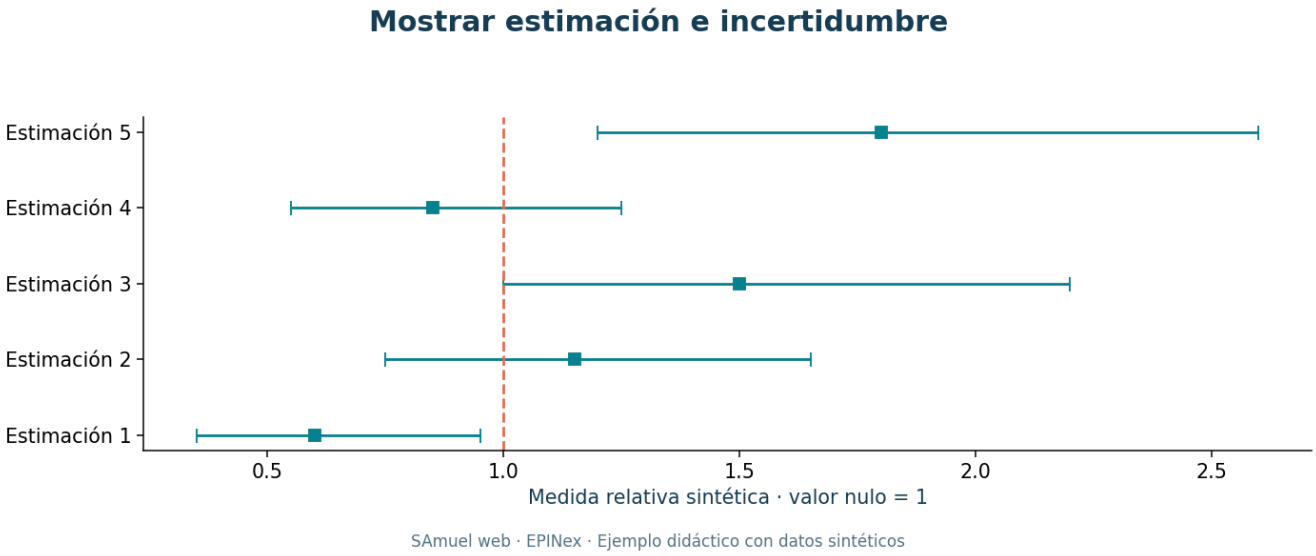
3. Desenlace `no_respuesta_12m`, evento 1.

4. Añada `marca_imatinib`, `edad` y `multimorbilidad`.

5. Pulse Generar. Cada fila conserva a la izquierda el nombre de la variable/categoría. Junto al intervalo aparecen solo estimación, IC, valor p y asterisco si aplica; no se repite el nombre.

6. Para reescribir los nombres, abra Formato → Texto → “Editar etiquetas del forest plot…”. Edite cada fila y pulse “Aplicar etiquetas”. Cambia solo la presentación del gráfico y sus nuevos guiones R/Python, no los datos ni las estimaciones. Al reconstruir el modelo se generan nuevamente las etiquetas del modelo.

7. El gráfico inicia blanco, con contornos negros y puntos independientes, sin una línea que conecte las variables. Para conectarlos voluntariamente, abra Formato → Series y active “Conectar estimaciones / marcadores”. La referencia nula discontinua puede mostrarse u ocultarse en Formato → Ejes.



Ejemplo sintético de estimaciones con intervalos; títulos, ejes y referencias deben ser claros.

15.3 Menú completo de formato

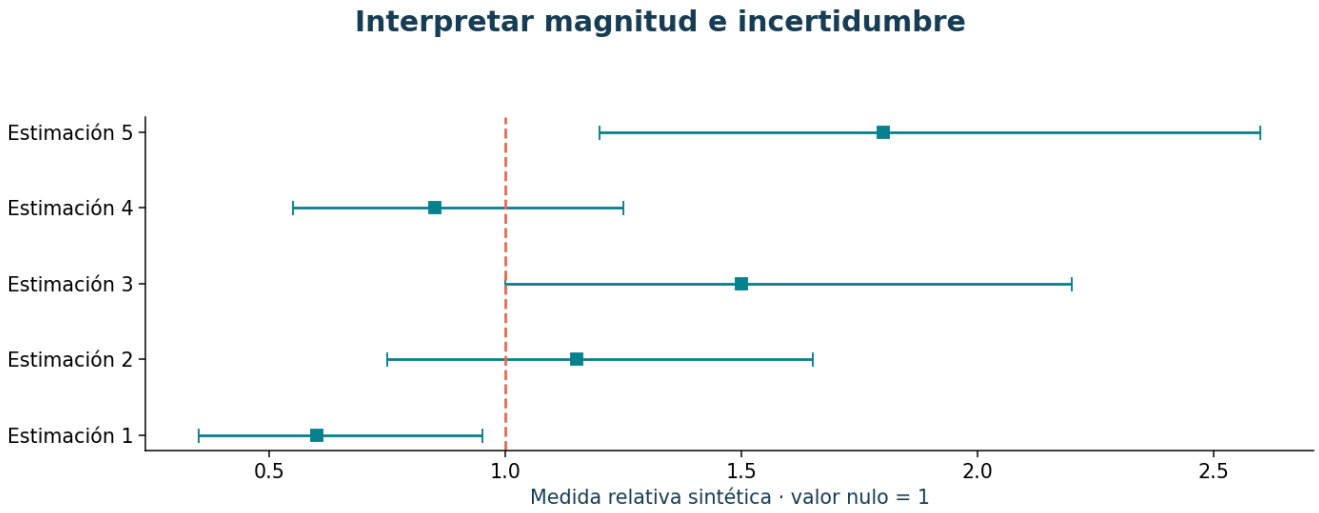
Pestaña	Controles	Uso
Texto	Título, ejes, pie; familia y tamaño; negrita, cursiva; ángulo X/Y y alineación.	Legibilidad y adaptación editorial; las etiquetas del forest plot pueden reescribirse sin alterar estimaciones.
Series	Color, degradado, marcador, relleno/hueco, tamaño, tipo/grosor/remate de línea, borde, transparencia, tramas y conexión.	Diferenciar series sin exagerar diferencias. El forest plot no conecta puntos por defecto.
Ejes	Mostrar/ocultar ejes, títulos, etiquetas y marcas; dirección, decimales, cuadrículas, líneas de referencia vertical X y horizontal Y con valores independientes, escala lineal/log y límites.	Definir exactamente la escala, las graduaciones y las referencias clínicas; verifique que los límites no recorten IC o datos.
Elementos	Título, leyenda, etiquetas de hasta 40 datos y tendencia lineal opcional.	Añada solo elementos que respondan la pregunta; la tendencia no prueba causalidad ni linealidad.
Fondo	Fondos de figura/trazado/leyenda, texto y bordes; tamaños tipográficos.	Contraste accesible y consistencia visual.
Tamaño	Una o dos columnas, ancho/alto personalizados y dpi.	Ajustar a revista o presentación antes de exportar.

15.4 Calidad de publicación

Elemento	Recomendación
Formato	SVG/PDF/EPS para líneas y barras; PNG/TIFF para mapa de bits. Evite JPEG.
Resolución	300 dpi para imágenes; 600-1200 dpi para línea pura. SAmuel inicia en 600 dpi.
Ancho	≈85 mm una columna; ≈180 mm doble columna.
Texto	Arial/Helvetica legible, 6-8 pt al tamaño final.
Ejes	Nombre y unidad; Y desde cero en barras salvo justificación explícita.
Relleno	Inicialmente blanco. Active color, degradado o tramas solo si aportan información y compruebe su legibilidad.
Pie	Fuera del gráfico; explique n, abreviaturas, barras de error y análisis.

Aplicar ajustes reconstruye desde una copia limpia del gráfico para evitar superposiciones. El forest plot mantiene su geometría y no conecta estimaciones salvo decisión del usuario. La línea nula es un control independiente; si cambia escalas, verifique intervalos y valor de referencia.

16. Interpretación, tablas e historial integral



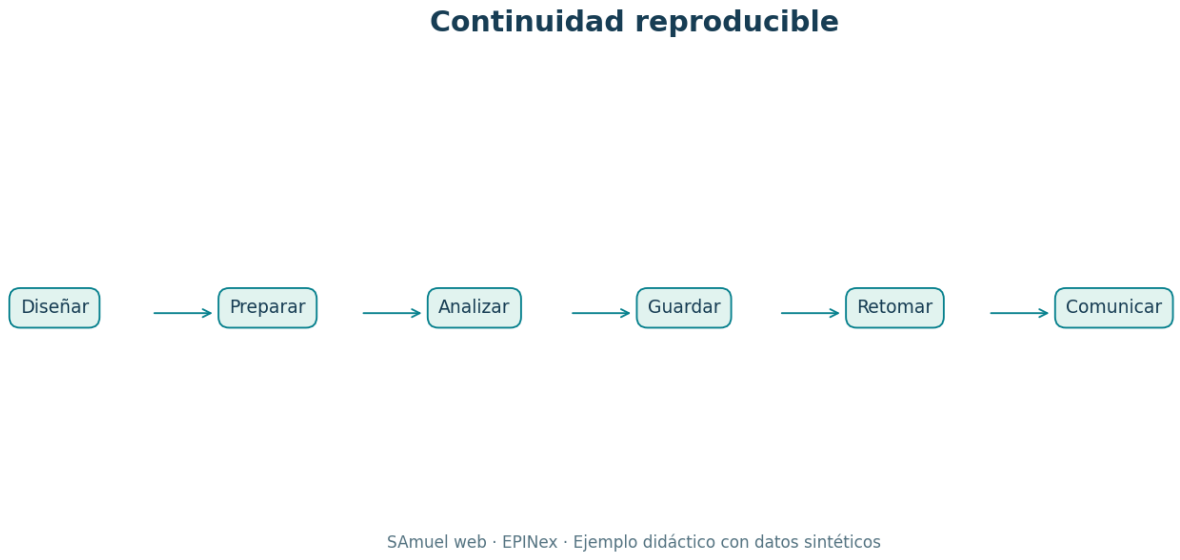
Forest plot sintético: el intervalo y el valor nulo forman parte de la lectura.

16.1 Tres pestañas

1 **Interpretación:** lectura epidemiológica y bioestadística del resultado activo.

2 **Tablas:** estimaciones, diagnósticos y resultados auxiliares con búsqueda, columnas visibles y decimales configurables por encabezado.

3 **Historial integral:** registro acumulado y editable de análisis, tablas, supuestos, código y figuras. Al pulsar **Mostrar figuras**, el panel derecho incorpora las mismas seis pestañas del módulo Gráficos: Texto, Series, Ejes, Elementos, Fondo y Tamaño. Use **Maximizar** para abrir un espacio ampliado con vista previa, barra de navegación y el editor completo; **Aplicar formato** actualiza una copia de la figura sin alterar las estimaciones.



Esquema conceptual de guardar, retomar, revisar y comunicar una investigación.

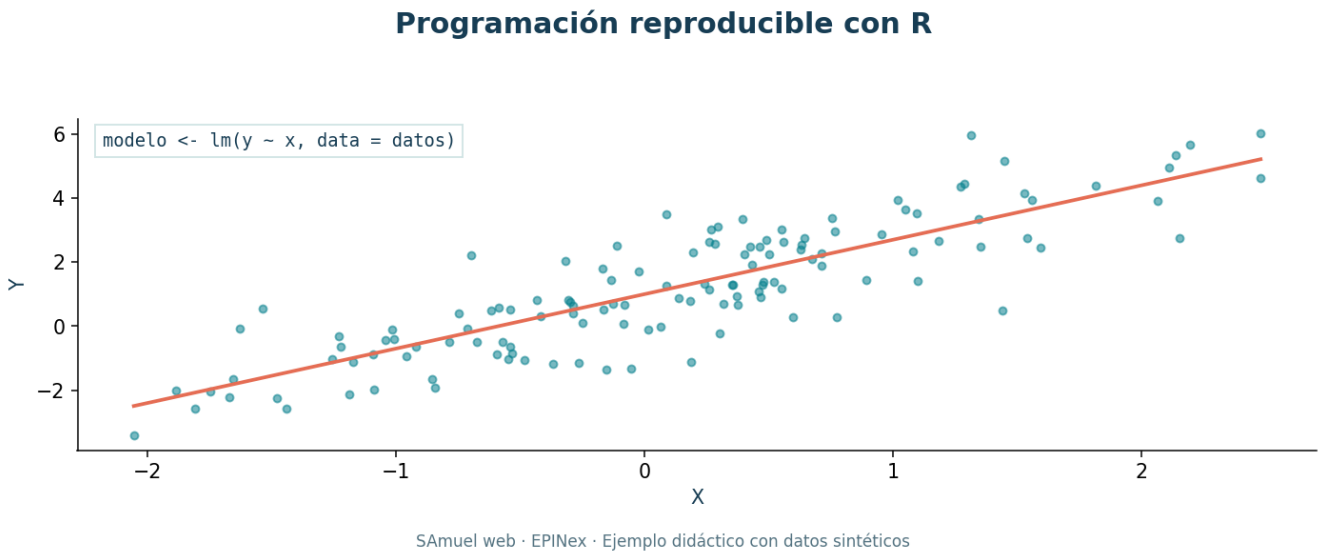
16.2 Redactar un resultado

Incluya método, población/n, estimación, IC 95%, valor p si procede, unidades, referencias, ajuste, faltantes, supuestos y una conclusión limitada al diseño.

Plantilla: “En [n] participantes, [exposición] se asoció con [desenlace] con [medida]=[estimación] (IC 95%: [LI-LS]; p=[p]), ajustada por [covariables]. [Diagnóstico principal]. Debido a [diseño/sesgo], el hallazgo no establece causalidad”.

Edite el historial con títulos, viñetas, negrita, cursiva, subrayado, resaltado, color y tamaño. Use la galería minimizable para conservar figuras sin ocupar la vista principal. Al pulsar **Mostrar figuras**, la galería usa todo el espacio del historial; **Ampliar** abre una ventana maximizada con navegación, zoom, desplazamiento y guardado.

17. Reproducibilidad y extensión con R



Ejemplo didáctico de fórmula y resultado; no es una captura de la consola.

- La base activa se llama `datos`.
- **Ctrl+Enter** ejecuta selección o línea actual.
- **Nuevo, Abrir y Guardar** permiten trabajar con scripts; **Plantillas** inserta ejemplos de descripción, diagnóstico de modelos y gráficos con IC.
- El autocompletado muestra funciones y el popup explica argumentos.
- Los análisis guiados añaden bloques completos al historial de código.
- Las tablas y figuras creadas por R aparecen en sus paneles; sus decimales pueden configurarse por columna.

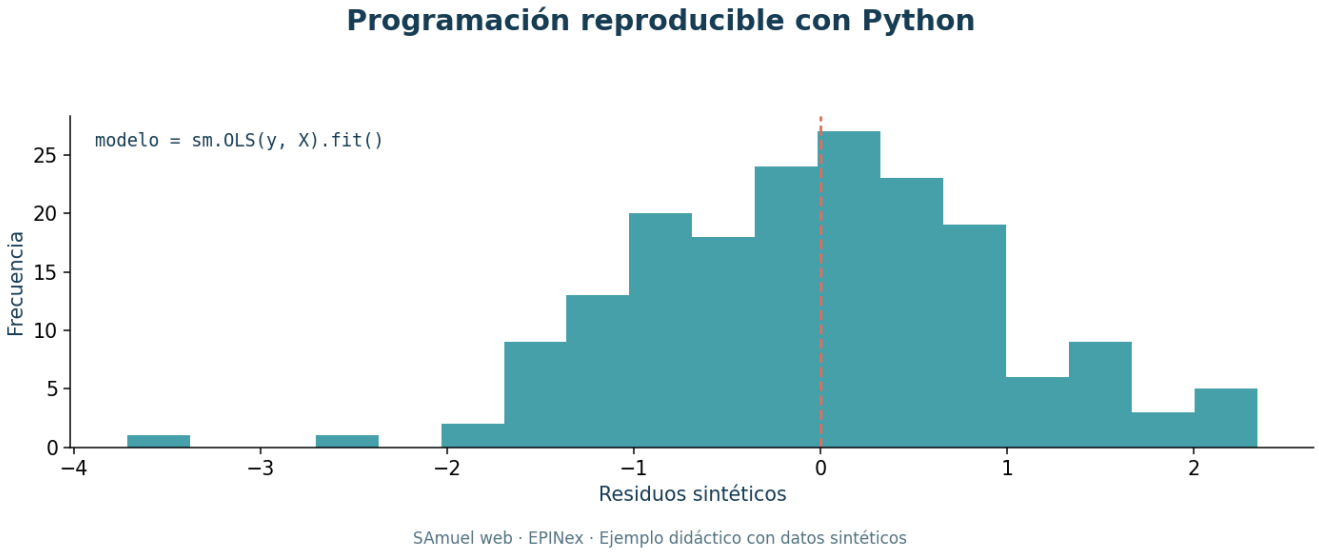
```
# Resumen reproducible
summary(datos$edad)
median(datos$edad, na.rm = TRUE)
tab <- table(datos$marca_imatinib, datos$no_respuesta_12m)
addmargins(tab)
prop.table(tab, margin = 1) * 100
```

Instalar paquetes

Abra **Librerías y dependencias**, escriba los nombres y pulse **Preparar instalación en el editor**. Revise el script y sus permisos. La instalación requiere ejecutar y autorizar el código; abrir el panel no inicia descargas.

> Instalar R o paquetes requiere internet y puede solicitar permisos del sistema. Revise procedencia, versión y licencia. SAmuel no garantiza compatibilidad de cualquier paquete externo.

18. Reproducibilidad y extensión con Python



Ejemplo didáctico de ajuste y revisión de residuos; no es una captura de la consola.

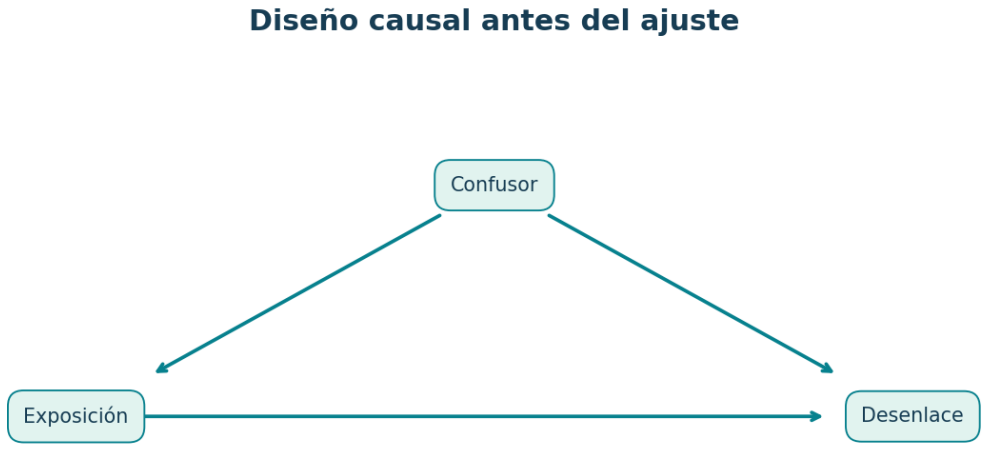
```
# La base activa es un DataFrame llamado datos
print(datos['edad'].describe())
tabla = pd.crosstab(datos['marca_imatinib'],
                    datos['no_respuesta_12m'], margins=True)
print(tabla)
datos['edad'].plot.hist(bins='auto')
plt.show()
```

Nuevo, Abrir, Guardar y Plantillas ofrecen el mismo recorrido de trabajo por proyectos que el editor R; las plantillas incorporan descripción, diagnóstico y figuras con incertidumbre. Las librerías integradas principales incluyen pandas, NumPy, SciPy, statsmodels y matplotlib. Las tablas devueltas permiten elegir decimales por columna sin modificar el DataFrame. Para extensiones, pulse **Instalar librerías…**; el buscador autocompleta nombres de PyPI, la barra muestra progreso por paquete y la aplicación informa la finalización.

> Código escrito por el usuario puede leer, modificar o eliminar archivos a los que tenga acceso. Revise cada línea y conserve copias de los datos. La ejecución de código amplía las posibilidades, pero también la responsabilidad.

19. Módulos estadísticos especializados

SAmuel reúne sus **Catálogo unificado por nivel y tema**, precedido por **Selecciona tu modo:**. Por defecto solo se muestran **BÁSICO · Datos y análisis**, **INTERMEDIO · Diseño aplicado**, **AVANZADO · Modelos expertos**, **R/Python · programación autorizada**, **Gráficos** y **Resultados**. Básico contiene Datos, Variables, Imputación, Univariado, Bivariado y Multivariado; Intermedio reúne trece módulos aplicados en cuatro áreas; Avanzado contiene diecisiete módulos de alta complejidad en seis áreas; Programador contiene Modo R y Modo Python. INTERMEDIO y AVANZADO usan submenús temáticos para evitar listas largas. **Síntesis de evidencia** → **Metaanálisis** añade un tercer nivel con **Básico** y **Avanzado**. Los nombres visibles no llevan numeración. Al elegir una opción se cierra el desplegable y queda visible un solo módulo; el menú no cambia de lugar. El indicador muestra primero el modo seleccionado y después el módulo utilizado. Gráficos y Resultados permanecen como accesos independientes. A la derecha, el menú **Proyecto** es el único punto para abrir, guardar, guardar como, reejecutar recetas y consultar la recuperación temporal. En cada módulo puede desplegar la orientación para consultar **para qué**, **por qué**, **cuándo usar**, **cuándo no usar** y **qué entrega**. Cada campo incorpora un botón **?** con ayuda contextual.



DAG conceptual: un confusor puede influir en exposición y desenlace. Las flechas representan supuestos.

Opción	Configuración esencial	Resultado y comprobación
Preparación de datos	Elija calcular, recodificar o filtrar; complete solo los campos visibles; previsualice.	Crea una copia, solicita confirmación y permite deshacer. Nunca sobrescribe el archivo fuente.
Tamaño de muestra	Objetivos: dos medias, proporción/media con precisión, medias o proporciones pareadas, dos proporciones, cohorte/transversal, casos-controles, ANOVA, correlación, exactitud diagnóstica y log-rank. Ajuste pérdidas, población finita o efecto de diseño cuando corresponda.	Entrega n y puede seleccionar unidades por muestreo aleatorio simple, sistemático, estratificado proporcional/equilibrado o conglomerados, con semilla reproducible.
Modelos longitudinales	Seleccione desenlace, varios predictores e ID repetido. Elija GEE gaussiana/binomial/Poisson o mixto lineal/binomial/Poisson/binomial negativo/Gamma; pendiente, segundo agrupador, exposición y cuadratura son opcionales según familia.	Informe poblacional (GEE) o condicional (mixto/GLMM). Revise convergencia, agrupación, sobredispersión y enlace.
Imputación múltiple	PMM numérica, MICE mixta o multinivel con jomo; defina desenlace, predictores, m, iteraciones/semilla e ID/nivel 2 si aplica.	Realizaciones exportables y combinación lineal/logística. Revise MAR, trazas, estabilidad y constancia de variables de nivel 2.
Validación predictiva	Modelo lineal/logístico; predictores; validación cruzada simple o anidada; ID para mantener visitas juntas o archivo externo.	Métricas fuera del entrenamiento. Evita fuga al aprender el preprocesamiento solo en cada conjunto de entrenamiento.
Modelos flexibles	Lineal/logístico con varios predictores; interacción y spline cúbico natural opcionales.	Coeficientes y diagnósticos; en lineal incluye HC3, residuos, VIF e influencia. No interprete bases del spline aisladamente.

Opción	Configuración esencial	Resultado y comprobación
Comparaciones múltiples	Variable cuantitativa, grupo y Welch-Holm, Tukey, Dunnett o Dunn con Holm/BH.	Comparaciones por pares con ajuste. Declare control en Dunnett y distinga error familiar de FDR.
Tabla 1 y balance	Pase solo las características deseadas, seleccione dos grupos y peso opcional.	Tabla 1, diferencias medias estandarizadas y gráfico de balance; el peso afecta DME ponderada, no los resúmenes no ponderados.
Encuestas complejas	Resumen o regresión; peso, estrato, UPM, dominio/FPC opcionales; Taylor, Jackknife, BRR, Fay o Bootstrap.	Estimaciones ponderadas e IC con diseño. BRR/Fay exige dos UPM por estrato; no filtre dominios antes de estimar.
Pruebas diagnósticas	Referencia, puntuación, evento, dirección y corte; segunda prueba opcional; DeLong o Bootstrap.	Sensibilidad, especificidad, valores predictivos, razones de verosimilitud, ROC/AUC e IC; DeLong compara AUC pareadas con IC y p.
DAG y plan causal	Exposición, desenlace, estimando, roles causales e instrumentos propuestos. Puede añadir nodos/conexiones y editar forma, colores, tamaño, texto y flechas.	DAG básico por defecto, conjunto de ajuste, rutas de sesgo, evaluación preliminar de instrumentos, sintaxis compatible con DAGitty y plan auditable.
Expediente reproducible	Estudio, responsable, protocolo, objetivo y semilla.	Huella SHA-256, diccionario, historial, versiones y lista de pendientes; no copia valores individuales.
Validación predictiva avanzada	Modelo, predictores, validación, remuestreos, sujeto, archivo externo y umbrales clínicos.	IC de rendimiento, calibración, ROC/diagnósticos y curva de decisión.
Supervivencia avanzada	Tiempo, evento, grupo y RMST, riesgos competitivos o diagnóstico PH.	RMST e IC, incidencia acumulada Aalen-Johansen o residuos de Schoenfeld.
Auditoría de calidad	Variables, ID, obligatoriedad, fechas y rangos del protocolo.	Perfil no destructivo de faltantes, duplicados, incoherencias y señales atípicas.
Inferencia causal aplicada	ATE/ATT, IPTW/AIPW/emparejamiento, exposición, desenlace y confusores definidos en el DAG.	Efecto, puntajes/pesos y balance antes/después.
Metaanálisis	Una fila por estudio, efecto, error/IC o correlación con tamaño muestral; modelo fijo o aleatorio.	Efecto combinado, transformación z de Fisher para correlaciones, Q, I^2, τ^2, forest y funnel plots.
Asistente de publicación	STROBE, CONSORT, TRIPOD, RECORD, STARD, PRISMA o preparación local de protocolo PROSPERO/revisión viva.	Checklist contextualizado, pregunta, elegibilidad, búsqueda, síntesis, cronograma, enmiendas, responsables y pendientes. No realiza el registro externo.
Análisis de sensibilidad	E-value, escenarios MNAR o cortes alternativos.	Robustez, cuadrícula/punto de quiebre y trayectoria completa de escenarios.
Gobernanza de datos	Identificadores, cuasi-identificadores, datos sensibles, k, custodio y retención.	Inventario, riesgo básico de reidentificación y plan de controles; no modifica la base.
Epidemiología poblacional	Eventos, población o tiempo-persona, estratos, grupo, estándar, multiplicador e IC.	Tasa cruda, estandarización directa, SMR o desigualdad absoluta/relativa con IC de Poisson y gráfico.
Inferencia bayesiana	Procedimiento, desenlace, pares/grupos, evento, predictores, valor nulo, prior, ICr, muestras y semilla.	Proporciones y riesgos; pruebas t; correlación; ANOVA; regresiones lineal/logística; posterior, ICr, ROPE y BF correctamente rotulado.
Mediación, moderación, CFA y SEM	Procedimiento, desenlace, exposición, mediador/moderador, indicadores o rutas observadas y covariables.	Efectos directo/indirecto con bootstrap, interacción, cargas/estructura factorial o coeficientes de rutas observadas.
Metaanálisis avanzado	Efecto y error estándar por estudio; subgrupo, moderador u orden opcional.	Dejar-uno-fuera, síntesis por subgrupos, metarregresión y metaanálisis acumulado.
Predicción clínica y explicabilidad	Desenlace, predictores, algoritmo, particiones, semilla y umbrales clínicos.	Métricas fuera de muestra, importancia por permutación, predicciones y curva de decisión.
Series temporales epidemiológicas	Fecha/orden, desenlace, intervención, estacionalidad y horizonte.	Serie interrumpida HAC, ARIMA/SARIMA, pronóstico, diagnóstico y detección de anomalías.
Psicometría e instrumentos	Ítems o evaluadores, procedimiento y número de factores cuando corresponde.	Alfa, omega, ítem-total, EFA, comunalidades, scree plot e ICC/acuerdo.

Opción	Configuración esencial	Resultado y comprobación
Arquitectura de extensiones	Carpeta de extensión, manifiesto, permisos, huella y firma.	Catálogo y validación segura; una extensión sin firma o no confiable permanece bloqueada.
Potencia por simulación	Diseño, efecto, tamaño, alfa, simulaciones, semilla e ICC si hay conglomerados.	Potencia Monte Carlo, incertidumbre de simulación y distribución de estimaciones.
Control de calidad y diseño experimental	Variable, método, límites de especificación, frecuencia o factores/respuesta.	Cartas I-MR, capacidad, Pareto o modelo factorial de dos niveles con interacciones.

Figuras de los módulos especializados. Cuando el resultado incluye una figura, pulse **Formatear última figura**. SAmuel abre el editor científico con tipografía, títulos, colores, líneas, marcadores, ejes, tamaños y exportación. El DAG añade controles semánticos propios y admite nodos y conexiones que no se declararon en la clasificación inicial.

19.1 Secuencia común

1. Seleccione el módulo y lea la ayuda visual.

2. Añada al panel derecho únicamente las variables necesarias; los campos cambian con el método.

3. Pulse **Ejecutar** o **Previsualizar**, según el módulo.

4. Revise interpretación, tablas, advertencias, figura y código reproducible antes de concluir.

19.2 Mapa de cobertura funcional

Área	Capítulos	Incluye
Entrada y preparación	3, 5-7, 19	Veinte bases, formatos, variables activas, diseño, faltantes, transformación e imputación simple/múltiple/multinivel.
Análisis	8-14, 19	Descriptivo, pruebas independientes/pareadas, 2×2, regresión, supervivencia, predicción, longitudinal, encuestas, diagnóstico, tasas, estandarización y supuestos.
Comunicación	15-16, 20-23	Gráficos base, georreferenciación y Plotly; formato guiado; historial exacto; exportación HTML/PDF/tablas; interpretación y control de calidad.
Reproducibilidad	17-20	Motores portátiles, código R/Python, proyectos, recursos web incorporados y uso sin conexión salvo fondos cartográficos en línea o instalaciones autorizadas.

19.3 Guía técnica completa de los treinta módulos especializados

Estas fichas forman parte del único manual de SAmuel. Cada una explica propósito, método, recorrido, interpretación y límites.

19.4 Preparación de datos

Para qué sirve. Calcular variables derivadas, recodificar categorías o filtrar filas sobre una copia, con previsualización, confirmación y deshacer en la sesión. Importa CSV/Excel y archivos SPSS, Stata y SAS mediante pyreadstat.

Cómo funciona. El formulario muestra únicamente los campos de la operación elegida. Las fórmulas usan col("variable"), aritmética, comparaciones y funciones permitidas, no código arbitrario. Las recodificaciones se escriben como original = nuevo. Se rechaza sobrescribir una columna existente y se detiene el cálculo ante divisiones por cero, valores infinitos o dominios inválidos.

Recorrido recomendado

1 Seleccione Preparación de datos y la operación Calcular variable; observe que los campos no aplicables desaparecen.

2 Elija la variable principal. Puede usar los botones para insertar su fórmula o crear un indicador por punto de corte. Para IMC escriba imc y use el ejemplo incluido.

3 Para recodificar escriba, por ejemplo, 0 = No; 1 = Sí. Las demás categorías se conservan.

4 Pulse Ejecutar / previsualizar y compruebe filas, columnas y unidades.

5 Acepte Aplicar preparación solo si la propuesta es correcta; Deshacer preparación recupera el estado anterior durante esta sesión.

Interpretación. Ejemplo: 80 kg y 2 m producen IMC=20 kg/m². El filtro cambia la muestra analizada, no corrige sesgo de selección. Los archivos originales permanecen intactos.

Límites y precauciones. Los códigos importados deben revisarse: no suponga que todas las etiquetas o convenciones del programa de origen fueron trasladadas al diccionario. Para dominios de encuesta no filtre filas previamente: use el campo Dominio del módulo 9.

Referencia metodológica: https://ofajardo.github.io/pyreadstat_documentation/_build/html/index.html

19.5 Tamaño de muestra

Para qué sirve. Planificar estimación de medias/proporciones, comparaciones independientes o pareadas, cohortes, casos y controles, ANOVA, correlación, precisión diagnóstica o supervivencia. También permite seleccionar las unidades desde la base mediante muestreo probabilístico, sin modificarla.

Cómo funciona. Los campos se adaptan al objetivo. Incluye pruebas t, h de Cohen, McNemar, ANOVA balanceado, transformación de Fisher, precisión normal y Schoenfeld. DEFF=1+(m-1)×CCI ajusta aproximadamente por conglomerados. La selección ofrece aleatorio simple, sistemático, estratificado proporcional/equilibrado y conglomerados completos.

Recorrido recomendado

1 Abra Tamaño de muestra y elija el objetivo; solo verá los supuestos que intervienen.

2 Para el ejemplo elija Dos medias independientes, alfa=0.05, potencia=0.80, efecto=0.5 y asignación=1.

3 Deje pérdidas=0 y tamaño del conglomerado=1: se requieren 64 personas por grupo, 128 en total.

4 En Después de calcular elija Solo calcular o un método de selección. Para seleccionar necesita una base cargada y, según el método, ID, estrato o conglomerado.

5 Ejecute y revise la primera tabla de tamaño y, si seleccionó unidades, la segunda tabla con la muestra reproducible mediante semilla.

Interpretación. El efecto para medias independientes es diferencia relevante / DE común. Para medias pareadas usa la DE de las diferencias y el resultado cuenta pares, no mediciones. El redondeo es hacia arriba.

Límites y precauciones. No existe un único método válido para todo protocolo. Equivalencia, no inferioridad y diseños adaptativos requieren cálculos especializados. Potencia no interviene al calcular precisión. El muestreo sistemático depende del orden del marco; estratificación y conglomerados requieren variables correctas. DEFF no sustituye el diseño completo de un ensayo por conglomerados.

Referencia metodológica: https://www.statsmodels.org/stable/stats.html#power-and-sample-size-calculations

19.6 Modelos longitudinales y mixtos

Para qué sirve. Analizar observaciones repetidas o agrupadas. Incluye GEE gaussiana, binomial y Poisson; mixto lineal REML; y GLMM binomial, Poisson, binomial negativo y Gamma mediante lme4.

Cómo funciona. GEE estima efectos poblacionales con correlación intercambiable y covarianza robusta. GLMM estima efectos condicionales al grupo, con intercepto aleatorio, pendiente opcional y segundo agrupador. Cuadratura=1 usa Laplace; más puntos mejoran la aproximación para un único intercepto aleatorio.

Recorrido recomendado

1. En Modelos longitudinales elija la familia según el desenlace.
2. En Predictores seleccione una o varias variables del panel izquierdo y pulse Añadir; las incluidas quedan visibles en el panel derecho. El ID se elige aparte y nunca es predictor.
3. Seleccione el ID del sujeto y, si corresponde, la categoría exacta de evento.
4. Para un intercepto aleatorio puede usar cuadratura=9. Con pendiente o segundo agrupador use 1. En conteos con tiempo de exposición, selecciónelo como offset.
5. Ejecute y revise efectos, varianzas aleatorias, convergencia, singularidad y registros excluidos.

Interpretación. En el GLMM binomial se presenta OR condicional, no riesgo relativo. En conteos sin exposición, razón de conteos esperados; con exposición, razón de tasas. En Gamma con enlace log, razón de medias. Los IC son Wald 95% y el contraste de efectos fijos es Z bilateral.

Límites y precauciones. No se elige familia por normalidad del desenlace: considere su escala y mecanismo. Se rechazan ajustes no convergentes o no identificados. Singularidad requiere simplificar o justificar la estructura. Pocos grupos comprometen la inferencia asintótica. Los efectos GEE y GLMM no son intercambiables.

Referencia metodológica: https://lme4.github.io/lme4/reference/glmer.html

19.7 Imputación múltiple y multinivel

Para qué sirve. Completar valores ausentes en múltiples realizaciones y combinar estimaciones, preservando la base original. PMM numéricas conserva el procedimiento anterior; MICE mixtas admite categorías nominales; Multinivel (jomo) considera agrupación y variables de nivel 1 y 2.

Cómo funciona. MICE usa PMM para cantidades, logística para variables binarias y multinomial para nominales. jomo/mitml utiliza un modelo conjunto multinivel, con representación latente de categorías. Se combina regresión lineal/logística en un nivel o su modelo mixto de intercepto aleatorio en multinivel, mediante Rubin y grados de libertad Barnard-Rubin.

Recorrido recomendado

1. Revise en Variables cuáles son cantidades y cuáles categorías; no codifique nominales como cantidades para forzar un método.
2. En Imputación múltiple seleccione Multinivel (jomo), desenlace, predictores e ID de grupo completo.
3. Indique como nivel 2 solo variables incluidas en el análisis que realmente sean constantes dentro del grupo. El desenlace del modelo mixto permanece en nivel 1.
4. Comience con 10 imputaciones, 5000 iteraciones iniciales y 100 entre realizaciones; registre semilla. Estos valores no garantizan convergencia.
5. Revise estimaciones combinadas, trazas, realizaciones completas y coeficientes por imputación. Repita con más iteraciones y otra semilla para evaluar estabilidad.

Interpretación. La varianza total combina incertidumbre dentro y entre imputaciones: $T=U+(1+1/m)B$. En logística se combinan log-odds antes de exponenciar. En multinivel se usan grados de libertad completos conservadores: grupos menos parámetros fijos. No se promedian valores p.

Límites y precauciones. MAR debe defenderse con conocimiento del proceso de ausencia; el algoritmo no demuestra MAR ni resuelve MNAR. Examine trazas y sensibilidad. Las variables de nivel 2 no pueden variar dentro del grupo. El modelo de imputación debe ser compatible con el análisis: esta interfaz no configura estructuras de tres niveles ni todas las interacciones posibles.

Referencia metodológica: https://search.r-project.org/CRAN/refmans/mitml/html/jomoImpute.html

19.8 Validación predictiva

Para qué sirve. Evaluar modelos lineales y logísticos en observaciones no usadas para ajustarlos. Ofrece validación cruzada simple, anidada y con archivo externo, respetando grupos cuando se indica un ID.

Cómo funciona. Imputación, escalado y codificación se ajustan dentro de entrenamiento. En anidada se selecciona regularización Ridge o C logístico usando particiones internas; las externas estiman rendimiento. La rejilla es 0.01, 0.1, 1, 10 y 100. Sin selección, logística utiliza L2 con C=1.

Recorrido recomendado

1. Seleccione desenlace y predictores definidos antes de examinar resultados.
2. Elija Cruzada anidada si desea seleccionar regularización; indique por ejemplo 5 particiones externas y 4 internas.
3. Indique ID cuando varias filas pertenezcan al mismo sujeto: permanecerán juntas en ambos niveles de partición.
4. Para validación externa, elija un archivo con las mismas variables y sujetos distintos; se reserva para evaluación final.
5. Revise RMSE, MAE y R^2 en lineal, o AUC, Brier y calibración en logística; consulte predicciones y parámetros elegidos por partición.

Interpretación. AUC mide discriminación, no calibración. Brier evalúa el error cuadrático de probabilidades. El rendimiento fuera de entrenamiento es más informativo que el aparente, pero puede cambiar en otra población.

Límites y precauciones. No seleccione variables ni puntos de corte mirando el conjunto de evaluación. Las particiones deben contener suficiente información; se detiene una partición logística inválida. Validación interna, incluso anidada, no sustituye validación externa prospectiva. Modelos predictivos penalizados no equivalen a inferencia causal.

Referencia metodológica: https://scikit-learn.org/stable/auto_examples/model_selection/plot_nested_cross_validation_iris.html

19.9 Modelos flexibles y diagnósticos

Para qué sirve. Ajustar regresión lineal con covarianza HC3 o logística, añadiendo una interacción numérica o un spline cúbico natural centrado. Permite revisar residuos, colinealidad e influencia.

Cómo funciona. La interacción incluye ambos efectos principales. El spline representa una asociación no lineal mediante varias funciones base. En lineal se calculan VIF, residuos, apalancamiento, distancia de Cook y Breusch-Pagan.

Recorrido recomendado

1. Elija Lineal para un resultado cuantitativo o Logística para uno binario.
2. Seleccione desenlace y todos los predictores del modelo.
3. Para interacción seleccione dos predictores numéricos ya incluidos. Para spline elija un predictor numérico y entre 3 y 8 grados de libertad.
4. Ejecute y revise primero advertencias de rango, convergencia e información disponible.
5. Consulte diagnósticos antes de interpretar los coeficientes. Si identifica un registro influyente, revise su origen; no lo elimine automáticamente.

Interpretación. Una interacción modifica la asociación de una variable según otra. Los coeficientes individuales de una base spline no son efectos clínicos aislados. HC3 modifica la incertidumbre, pero no corrige una forma funcional equivocada.

Límites y precauciones. El programa no selecciona variables ni justifica causalidad. Preespecifique flexibilidad según conocimiento y tamaño de muestra. Las pruebas diagnósticas no sustituyen la inspección de datos y del diseño. Los módulos originales de regresión se conservan.

Referencia metodológica: https://www.statsmodels.org/stable/examples/notebooks/generated/linear_regression_diagnostics_plots.html

19.10 Comparaciones múltiples

Para qué sirve. Comparar grupos independientes sin presentar todos los contrastes como si fueran una única prueba. Incluye Welch-Holm, Tukey, Dunnett y Dunn con Holm o Benjamini-Hochberg.

Cómo funciona. Tukey compara todas las medias bajo varianza común; Dunnett compara con un control especificado. Welch-Holm no exige igualdad de varianzas. Dunn utiliza rangos y corrige empates. Holm controla el error familiar; Benjamini-Hochberg controla FDR bajo sus condiciones.

Recorrido recomendado

1. Seleccione una variable cuantitativa y la variable de grupos.
2. Confirme independencia solo si los grupos representan unidades independientes.
3. Elija el procedimiento según objetivo y supuestos, no según cuál produzca menor valor p.
4. Si usa Dunnett, indique la categoría exacta del control.
5. Ejecute y examine la familia completa de contrastes y los valores p ajustados.

Interpretación. El valor p ajustado corresponde a la familia de comparaciones mostrada. Una diferencia estadística no describe por sí sola su importancia práctica. Para rangos, interpretar como diferencia de medianas requiere condiciones adicionales sobre las distribuciones.

Límites y precauciones. No use estos procedimientos para medidas repetidas o pares. Holm y FDR responden a objetivos distintos. No quite contrastes después de observar resultados para reducir artificialmente el ajuste.

Referencia metodológica: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.dunnett.html

19.11 Tabla 1 y balance entre grupos

Para qué sirve. Añadir diferencias de medias estandarizadas (DME) y un gráfico de balance a una Tabla 1 con las variables seleccionadas. La Tabla 1 habitual del módulo univariado permanece intacta.

Cómo funciona. Compara dos grupos. Para cantidades usa diferencias estandarizadas; para cualitativas, indicadores por categoría. Puede comparar DME sin ponderar y ponderadas con un denominador común sin ponderar para evaluar cambios en balance.

Recorrido recomendado

1. Seleccione Tabla 1 y balance.
2. Elija únicamente las características que desea describir y una variable con dos grupos.
3. Revise tipos en el diccionario para que cantidades y categorías se presenten correctamente.
4. Si ya dispone de pesos justificados, seleccione la columna opcional; el módulo no crea pesos por sí mismo.
5. Revise la tabla única de resultados y el gráfico de balance; los datos y análisis originales no se modifican.

Interpretación. La DME cuantifica desequilibrio descriptivo en unidades estandarizadas. Una magnitud pequeña no demuestra ausencia de confusión ni equivalencia de poblaciones. Los valores p no son una medida suficiente de balance.

Límites y precauciones. Los pesos afectan la DME ponderada y su gráfico, no convierten automáticamente los resúmenes y valores p de esta Tabla 1 en estimadores de encuesta. Para inferencia ponderada con diseño use Encuestas complejas. No incluye emparejamiento o estimación automática de puntajes de propensión.

Referencia metodológica: https://www.stata.com/features/overview/treatment-effects-balance/

19.12 Encuestas complejas

Para qué sirve. Estimar medias, proporciones y regresiones lineales/logísticas considerando pesos, estratos y unidades primarias de muestreo (UPM). Incorpora Taylor, Jackknife, BRR, Fay y Bootstrap de diseño; FPC y dominios opcionales.

Cómo funciona. Taylor utiliza linealización por UPM y estrato. Las réplicas se construyen desde el diseño mediante survey, no remuestreando filas sin estructura. FPC requiere una columna con el número total de UPM poblacionales por estrato. Un dominio se selecciona dentro del diseño para preservar la información de varianza.

Recorrido recomendado

1. Seleccione procedimiento, peso, estrato y UPM, según el protocolo de muestreo.
2. Elija método de varianza: BRR/Fay exige exactamente dos UPM por estrato; con Fay indique su coeficiente.
3. Si corresponde FPC, seleccione el total entero de UPM, constante por estrato y no menor que el número observado.
4. Para una subpoblación, indique aquí variable y categoría de dominio; mantenga la base completa.
5. Seleccione variables para resumen o desenlace/predictores para regresión. Ejecute y revise IC, grados de libertad y datos de diseño.

Interpretación. Los IC usan distribución t según los grados de libertad del diseño o modelo. En regresión logística las estimaciones exponenciadas son OR. Las proporciones se presentan entre 0 y 1; sus IC Wald pueden exceder ese intervalo y no se recortan silenciosamente.

Límites y precauciones. Los pesos deben reflejar selección y ajustes de muestreo. No son automáticamente pesos de frecuencia o pesos de tratamiento. Se rechazan estratos solitarios. Esta interfaz especifica una etapa de UPM; no reconstruye por sí sola diseños multietápicos desconocidos.

Referencia metodológica: https://r-survey.r-forge.r-project.org/survey/html/as.svrepdesign.html

19.13 Pruebas diagnósticas y DeLong

Para qué sirve. Evaluar una prueba frente a referencia binaria: tabla 2×2, sensibilidad, especificidad, valores predictivos, razones de verosimilitud y ROC/AUC. Compara dos puntuaciones medidas en los mismos sujetos mediante DeLong o IC Bootstrap.

Cómo funciona. Las proporciones usan IC Wilson. Las razones de verosimilitud usan aproximación logarítmica cuando las celdas permiten estimarla. DeLong calcula la covarianza de AUC empíricas considerando empates y correlación entre pruebas; entrega diferencia, IC 95%, estadístico Z y valor p bilateral.

Recorrido recomendado

1. Seleccione referencia, categoría positiva y puntuación de la primera prueba.
2. Indique punto de corte y si valores mayores o menores representan positividad.
3. Seleccione una segunda puntuación solo si corresponde a los mismos sujetos; puede indicar su dirección por separado.

4 Elija DeLong para IC y contraste pareado. Bootstrap ofrece IC percentil, sin valor p en esta implementación.

5 Revise AUC de ambas pruebas, diferencia e IC, no solo el valor p; compruebe cuántos sujetos completos entraron al análisis.

Interpretación. Con 8 verdaderos positivos de 10 enfermos y 7 verdaderos negativos de 10 no enfermos: sensibilidad=0.80, especificidad=0.70 y VPP=8/11. AUC resume discriminación en todos los umbrales, no utilidad clínica. Un p no significativo no demuestra equivalencia.

Límites y precauciones. DeLong aquí compara AUC completas, no suavizadas ni parciales, en muestras pareadas. No use puntuaciones de entrenamiento para validar un modelo. Varianzas nulas y separación perfecta generan advertencias. Los valores predictivos de una muestra de casos y controles no se extrapolan automáticamente a la población.

Referencia metodológica: https://search.r-project.org/CRAN/refmans/pROC/html/roc.test.html

19.14 DAG y plan causal

Para qué sirve. Hacer explícitas las premisas causales del estudio antes del modelado: exposición, desenlace, confusores, mediadores, colisionadores y modificadores de efecto. Produce un diagrama, un conjunto de ajuste propuesto y una auditoría de las variables involucradas.

Cómo funciona. El usuario declara los roles y SAmuel construye una estructura causal canónica. Para efecto total propone ajustar confusores preexposición y evitar mediadores y colisionadores; para efecto directo controlado incorpora mediadores y advierte los supuestos adicionales. El gráfico no aprende flechas de los datos.

Recorrido recomendado

1 Seleccione exposición y desenlace diferentes.

2 Elija el estimando: efecto total, efecto directo controlado o asociación descriptiva.

3 Asigne cada covariable a un solo rol causal y documente el orden temporal.

4 Ejecute y revise una por una las flechas, el conjunto de ajuste y las variables que no deben condicionarse.

5 Congele este plan antes del modelo final y registre cualquier desviación del protocolo.

Interpretación. Un conjunto de ajuste es válido solo bajo el DAG y los supuestos declarados. Los modificadores de efecto requieren interacciones preespecificadas y presentación estratificada; no se eligen por el menor valor p.

Límites y precauciones. El módulo no descubre causalidad, no descarta confusión no medida y no prueba la temporalidad. Ajustar un colisionador puede introducir sesgo. Los efectos directos requieren supuestos adicionales sobre la relación exposición-mediador-desenlace.

Referencia metodológica: https://www.dagitty.net/learn/index.html

19.15 Expediente reproducible del análisis

Para qué sirve. Reunir identificación del estudio, objetivo, protocolo, huella de la base activa, diccionario auditado, historial de análisis, versiones del entorno y una lista de verificación reproducible.

Cómo funciona. La huella SHA-256 combina esquema, tipos, valores e índice de la base activa. El expediente resume filas, variables, faltantes, duplicados, variables habilitadas y los resultados registrados en la sesión, sin copiar valores individuales a sus tablas.

Recorrido recomendado

1 Escriba título, responsable, protocolo o versión y objetivo analítico.

2 Declare una semilla maestra para procedimientos estocásticos.

3 Ejecute el módulo y conserve la huella junto con la fuente protegida.

4 Revise pendientes en la lista de verificación y complete el análisis.

5 Desde Resultados guarde el historial integral para reunir expediente, tablas, figuras, decisiones y scripts.

Interpretación. Dos bases idénticas producen la misma huella; cualquier cambio en contenido o esquema genera otra. El historial documenta lo disponible en la sesión al momento de crear el expediente.

Límites y precauciones. Una huella no anonimiza ni protege el archivo original. La reproducibilidad computacional no elimina sesgos de diseño, medición, selección o análisis. Registre también versiones del protocolo y decisiones fuera del software.

Referencia metodológica: https://www.nature.com/articles/s41562-016-0021

19.16 Validación predictiva avanzada

Para qué sirve. Extender la validación básica con incertidumbre del rendimiento, calibración detallada, visualización diagnóstica y, en modelos logísticos, análisis de curva de decisión para un intervalo de umbrales clínicos.

Cómo funciona. Conserva imputación, codificación y escalado dentro de cada entrenamiento. Puede usar validación cruzada simple o anidada y archivo externo. Calcula IC bootstrap percentiles sobre predicciones fuera de entrenamiento. En logística estima intercepto y pendiente de calibración, ROC, Brier y beneficio neto; en lineal informa RMSE, MAE, R^2 y residuos.

Recorrido recomendado

1 Preespecifique desenlace, predictores y categoría de evento.

2 Use validación anidada si seleccionará regularización; mantenga juntas las filas del mismo sujeto mediante ID.

3 Indique al menos 50 remuestreos y un intervalo de umbrales clínicamente defendible.

4 Si dispone de una cohorte independiente, cárguela como archivo externo y no la use para ajustar decisiones.

5 Revise discriminación, calibración, incertidumbre, utilidad y subgrupos antes de proponer uso clínico.

Interpretación. AUC evalúa ordenamiento, Brier combina calibración y discriminación, y el beneficio neto pondera falsos positivos según el umbral. Intercepto 0 y pendiente 1 son referencias de calibración ideal. En lineal, patrones residuales pueden revelar errores que una métrica promedio oculta.

Límites y precauciones. Los IC sobre predicciones no corrigen falta de transportabilidad. Elegir variables, hiperparámetros o umbrales mirando la evaluación produce optimismo. La curva de decisión depende de consecuencias clínicas y no sustituye validación externa prospectiva ni evaluación de impacto.

Referencia metodológica: https://www.bmj.com/content/384/bmj-2023-074819

19.17 Supervivencia avanzada

Para qué sirve. Responder preguntas de tiempo hasta evento que una razón de hazards aislada no resuelve: tiempo medio restringido (RMST), incidencia acumulada con riesgos competitivos y diagnóstico de proporcionalidad.

Cómo funciona. RMST integra Kaplan-Meier hasta un horizonte tau común. La incidencia acumulada utiliza Aalen-Johansen y conserva los eventos competitivos. El diagnóstico explora residuos de Schoenfeld frente al tiempo.

Recorrido recomendado

1 Seleccione tiempo y estado/tipo de evento.

2 Elija RMST, riesgos competitivos o diagnóstico de proporcionalidad.

3 Para RMST seleccione grupo y justifique tau; para riesgos competitivos declare causa y competidor.

4 Ejecute y revise seguimiento, eventos, estimación, IC y figura.

5 Informe por qué eligió el horizonte y cómo trató cada evento.

Interpretación. RMST es el tiempo medio libre del evento hasta tau. La incidencia acumulada estima la probabilidad de la causa en presencia de competidores. Los residuos ayudan a detectar un efecto que cambia en el tiempo.

Límites y precauciones. No use 1 menos Kaplan-Meier para una causa específica con competidores. Un diagnóstico no significativo no demuestra proporcionalidad. Evite horizontes con seguimiento insuficiente.

Referencia metodológica: https://cran.r-project.org/web/packages/survival/vignettes/compete.pdf

19.18 Auditoría de calidad de datos

Para qué sirve. Localizar faltantes, duplicados, constantes, rangos imposibles, fechas incoherentes y señales de valores atípicos antes del análisis.

Cómo funciona. El perfil es no destructivo. Resume completitud y unicidad, usa IQR como señal exploratoria, aplica reglas explícitas variable=mínimo:máximo y verifica orden temporal entre dos fechas.

Recorrido recomendado

1 Seleccione las variables a auditar.

2 Declare ID único y variables obligatorias cuando correspondan.

3 Añada reglas del diccionario y, si aplica, fecha inicial/final.

4 Ejecute y revise primero el resumen y después cada incidencia localizable.

5 Corrija en la fuente o documente la decisión; repita la auditoría.

Interpretación. Una marca identifica algo que requiere revisión, no un registro que deba eliminarse. El gráfico prioriza variables por proporción faltante.

Límites y precauciones. IQR no define error clínico. Las reglas deben proceder del protocolo, unidades y sistema de captura. No modifique datos sin conservar fuente y trazabilidad.

Referencia metodológica: https://www.who.int/publications/i/item/9789241564991

19.19 Inferencia causal aplicada

Para qué sirve. Estimar ATE o ATT con ponderación IPTW, estimador doblemente robusto AIPW o emparejamiento por puntaje de propensión, mostrando el balance antes y después.

Cómo funciona. El puntaje se estima mediante regresión logística de la exposición sobre confusores definidos previamente en el DAG. Se recortan probabilidades extremas y se informa DME, pesos y efecto en escala de diferencia.

Recorrido recomendado

1 Complete primero el DAG y congele el estimando.

2 Seleccione exposición, categoría expuesta, desenlace y confusores preexposición.

3 Elija ATE/ATT y método; justifique el recorte.

4 Revise solapamiento, pesos y DME antes/después.

5 Interprete el efecto solo bajo consistencia, positividad e intercambio condicional.

Interpretación. ATE resume el efecto medio poblacional; ATT, entre expuestos. AIPW combina modelos de exposición y desenlace, pero no protege si ambos están mal especificados.

Límites y precauciones. No seleccione confusores por valor p. El método no resuelve confusión no medida, mala medición, interferencia ni falta grave de positividad.

Referencia metodológica: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

19.20 Metaanálisis

Para qué sirve. Combinar estimaciones comparables, cuantificar heterogeneidad y producir forest y funnel plots.

Cómo funciona. Acepta efecto con error estándar o IC 95%. Ofrece modelo fijo, DerSimonian-Laird y REML; informa Q, I^2 y tau^2.

Recorrido recomendado

1. Cargue una fila por estudio.
2. Seleccione identificador, efecto y error estándar o IC.
3. Declare medida y modelo antes del cálculo.
4. Revise pesos, forest, Q, I^2, tau^2 y funnel.
5. Explique comparabilidad y sensibilidad.

Interpretación. El combinado resume el modelo elegido. I^2 expresa inconsistencia relativa, no importancia clínica ni su causa.

Límites y precauciones. Un funnel asimétrico no demuestra sesgo de publicación. No combine estudios clínica o metodológicamente incompatibles.

Referencia metodológica: https://training.cochrane.org/handbook/current/chapter-10

19.21 Asistente de publicación

Para qué sirve. Organizar el reporte y sus evidencias según STROBE, CONSORT, TRIPOD, RECORD, STARD o PRISMA.

Cómo funciona. Genera una lista contextualizada con responsable, evidencia y pendientes. Resume un flujo inicial desde la base y reconoce evidencia parcial disponible en el historial, sin certificar cumplimiento.

Recorrido recomendado

1. Seleccione la guía compatible con el diseño.
2. Registre título, objetivo y protocolo/registro.
3. Genere la lista y asigne responsable a cada pendiente.
4. Vincule cada ítem con manuscrito, protocolo, tablas, figuras o anexos.
5. Contraste la salida con la declaración oficial vigente antes del envío.

Interpretación. La salida es un tablero de trabajo: Revisión humana requerida significa que la evidencia no puede inferirse de la base o del historial.

Límites y precauciones. Las listas resumidas no sustituyen el texto oficial, explicaciones narrativas ni requisitos específicos de la revista.

Referencia metodológica: https://www.equator-network.org/

19.22 Análisis de sensibilidad

Para qué sirve. Evaluar robustez ante confusión no medida, datos faltantes MNAR o definiciones alternativas del desenlace.

Cómo funciona. E-value cuantifica la asociación mínima de un confusor no medido en escala de razón. El punto de quiebre asigna probabilidades de evento a faltantes por exposición. Los cortes alternativos repiten RR crudos a través de un rango completo.

Recorrido recomendado

1. Elija la amenaza principal a la conclusión.
2. Defina valores o escenarios clínicamente plausibles antes de ejecutar.
3. Para MNAR/cortes seleccione variables y categorías exactas.
4. Revise la cuadrícula o trayectoria completa y localice el punto de quiebre.
5. Reporte todos los escenarios, incluso los que contradicen la conclusión principal.

Interpretación. Una conclusión estable en escenarios plausibles es más robusta; cruzar el nulo bajo cambios pequeños indica sensibilidad importante.

Límites y precauciones. E-value no corrige sesgo y aplica a medidas en escala de razón. La cuadrícula MNAR no es imputación múltiple. La exploración post hoc de cortes puede inflar falsos positivos.

Referencia metodológica: https://www.evalue-calculator.com/

19.23 Gobernanza de datos

Para qué sirve. Clasificar identificadores, cuasi-identificadores y datos sensibles; evaluar k-anonimato básico y preparar controles de minimización, acceso, retención y eliminación.

Cómo funciona. El perfil cuenta tamaños de clases definidos por los cuasi-identificadores y registros por debajo del k objetivo. No transforma la base y conserva visibles los pendientes de custodia y seguridad.

Recorrido recomendado

1. Identifique proyecto, custodio y plazo de retención.
2. Clasifique cada variable en un solo rol; incluya identificadores aunque estén inhabilitados para análisis.
3. Defina k objetivo y ejecute.
4. Revise clases pequeñas, unicidad y acciones sugeridas.
5. Aplique seudonimización, generalización, cifrado y controles conforme a la política institucional.

Interpretación. k pequeño señala combinaciones raras que facilitan vinculación; k alto no garantiza anonimización frente a conocimiento externo o homogeneidad de atributos.

Límites y precauciones. La herramienta no anonimiza ni cifra. Requiere base jurídica, evaluación institucional, control de acceso y gestión segura de claves/copias.

Referencia metodológica: https://www.iso.org/standard/63553.html

19.24 Epidemiología poblacional

Para qué sirve. Calcular tasas crudas, estandarización directa o indirecta y desigualdades absolutas y relativas entre poblaciones.

Cómo funciona. Usa eventos y población o tiempo-persona con intervalos de Poisson. La estandarización directa aplica una población estándar común; la indirecta compara observados con esperados mediante SMR.

Recorrido recomendado

1 Defina evento, denominador, periodo y multiplicador antes de analizar.

2 Para estandarizar, seleccione estrato y población o tasa estándar compatible.

3 Si compara grupos, confirme que las definiciones y periodos sean equivalentes.

4 Revise estimación, IC y denominadores; no interprete solo el orden de las tasas.

5 Conserve en el informe la fuente del estándar y la definición de cada evento.

Interpretación. La tasa expresa eventos por la unidad elegida. La diferencia cuantifica desigualdad absoluta; la razón y el SMR, desigualdad relativa frente al grupo o estándar de referencia.

Límites y precauciones. La estandarización controla únicamente la estructura de los estratos incluidos. No corrige confusión residual, cambios diagnósticos, migración, subregistro ni denominadores incompatibles.

Referencia metodológica: https://www.sergas.gal/Saude-publica/epidat?idioma=es

19.25 Inferencia bayesiana

Para qué sirve. Estima proporciones, comparaciones epidemiológicas, medias y regresiones lineales mediante distribuciones posteriores, con priors y semillas visibles.

Cómo funciona. Los procedimientos binomiales combinan una prior Beta con la verosimilitud y calculan factores de Bayes exactos mediante verosimilitudes marginales. La comparación de dos grupos simula riesgos posteriores y deriva diferencia de riesgos, RR/RP y OR. Media y regresión usan modelos conjugados Normal-Inversa-Gamma; la regresión informa betas estandarizadas y R^2 posterior.

Recorrido recomendado

1 Seleccione proporción, comparación de dos proporciones, media o regresión lineal.

2 Defina desenlace, evento, referencia, predictores y valor nulo únicamente cuando correspondan.

3 Elija y justifique el perfil de prior, el nivel creíble, las muestras posteriores y la semilla.

4 Repita con priors plausibles alternativos y compare posterior, intervalo creíble, probabilidades y BF10 exacto cuando esté disponible.

Interpretación. Un ICr 95% contiene el parámetro con probabilidad posterior 0,95 bajo el modelo, datos y prior especificados. P(parámetro > nulo | datos) expresa directamente dirección posterior. BF10 compara H1 con H0 en el modelo binomial; no es un valor p.

Límites y precauciones. Preespecifique el prior, documente sensibilidad y no seleccione el perfil que favorezca la conclusión. Bayes no corrige confusión, selección, medición deficiente ni dependencia ignorada. Las betas son estandarizadas y el R^2 es aparente, no validación externa. SAmuel no informa factores aproximados como si fueran exactos.

Referencia metodológica: https://doi.org/10.3758/s13423-020-01798-5

19.26 Mediación, moderación, CFA y SEM

Para qué sirve. Examina mecanismos, modificación de efecto, estructura de ítems y relaciones dirigidas entre variables observadas.

Cómo funciona. Mediación estima rutas a, b, efecto directo e indirecto con bootstrap; moderación añade el producto exposición×moderador; CFA ofrece una aproximación factorial guiada y SEM estima rutas observadas declaradas.

Recorrido recomendado

1 Defina la hipótesis y la temporalidad antes de seleccionar variables.

2 Elija el procedimiento y complete únicamente los roles requeridos.

3 Revise estimaciones, intervalos, figura y advertencias de identificación.

Interpretación. Un efecto indirecto compatible con valores distintos de cero describe mediación estadística bajo el modelo; una interacción indica que el efecto cambia con el moderador.

Límites y precauciones. La mediación no demuestra causalidad sin supuestos causales; CFA/SEM de V1 no reemplaza un motor de variables latentes complejas.

19.27 Metaanálisis avanzado

Para qué sirve. Investiga influencia, subgrupos, moderadores y acumulación temporal después de una síntesis principal.

Cómo funciona. Recalcula un modelo aleatorio DerSimonian-Laird al retirar cada estudio, estratifica por subgrupo, ajusta metarregresión ponderada o acumula estudios en el orden indicado.

Recorrido recomendado

1 Confirme comparabilidad y escala de los efectos.

2 Ejecute primero el modelo principal.

3 Use un análisis avanzado preespecificado y compare estabilidad.

Interpretación. Una conclusión que cambia al retirar un solo estudio es sensible; la metarregresión describe asociación entre moderador y efecto, no causalidad.

Límites y precauciones. Pocos estudios producen estimaciones inestables; los subgrupos post hoc elevan riesgo de hallazgos espurios.

19.28 Predicción clínica y explicabilidad

Para qué sirve. Compara modelos supervisados, estima desempeño fuera de muestra y relaciona predicción con decisiones clínicas.

Cómo funciona. Aplica validación cruzada estratificada o continua a Elastic Net, random forest, gradient boosting o SVM; calcula importancia por permutación y beneficio neto por umbral.

Recorrido recomendado

1 Declare desenlace, predictores y horizonte clínico.

2 Seleccione algoritmo, particiones y semilla.

3 Compare discriminación/error, calibración disponible, importancia y curva de decisión.

Interpretación. La importancia por permutación cuantifica pérdida predictiva al alterar una variable; no es un efecto causal. Beneficio neto positivo frente a tratar a todos o a nadie apoya utilidad en ese umbral.

Límites y precauciones. La validación interna no prueba transportabilidad; requiere muestra externa, revisión de sesgo y utilidad clínica.

19.29 Series temporales epidemiológicas

Para qué sirve. Estima cambios tras intervenciones, modela tendencia/estacionalidad y detecta señales inusuales.

Cómo funciona. Ofrece regresión segmentada con errores HAC, ARIMA/SARIMA y una línea base robusta para alertas de brote.

Recorrido recomendado

1. Ordene periodos y verifique intervalos regulares.
2. Declare intervención o estacionalidad según la pregunta.
3. Revise residuos, autocorrelación y plausibilidad de las alertas.

Interpretación. El cambio de nivel es inmediato y el cambio de pendiente modifica la tendencia posterior; una alerta es una señal para investigar, no un brote confirmado.

Límites y precauciones. Cambios concurrentes, estacionalidad omitida y modificaciones de vigilancia pueden sesgar una serie interrumpida.

19.30 Psicometría e instrumentos

Para qué sirve. Evalúa consistencia, estructura dimensional y acuerdo entre mediciones.

Cómo funciona. Calcula alfa, omega e ítem-total; EFA con cargas/comunalidades y scree; o ICC/acuerdo para evaluadores.

Recorrido recomendado

1. Seleccione ítems que deban medir el mismo constructo.
2. Revise codificación inversa y faltantes.
3. Interprete fiabilidad junto con dimensionalidad y contenido.

Interpretación. Alfa u omega altos indican consistencia bajo sus supuestos, no validez; una carga representa relación del ítem con el factor.

Límites y precauciones. No persiga un alfa alto eliminando ítems esenciales; la estructura debe confirmarse en otra muestra.

19.31 Arquitectura de extensiones

Para qué sirve. Permite ampliar SAmuel mediante paquetes declarativos controlados sin alterar el núcleo registrado.

Cómo funciona. Descubre manifiestos API 1, valida permisos, archivos, SHA-256 y firma Ed25519; V1 cataloga y bloquea todo paquete no firmado o no confiable y no ejecuta código durante la inspección.

Recorrido recomendado

1. Prepare manifiesto y archivos en una carpeta independiente.
2. Revise permisos y huellas.
3. Firme, pruebe y habilite solo mediante una distribución autorizada.

Interpretación. “Válida” significa que estructura, integridad y firma cumplen el contrato; no garantiza calidad científica del método.

Límites y precauciones. Nunca habilite código de origen desconocido; una firma identifica al firmante pero no reemplaza auditoría, pruebas ni revisión metodológica.

19.32 Potencia por simulación

Para qué sirve. Estima potencia para diseños donde conviene representar explícitamente variabilidad y estructura de conglomerados.

Cómo funciona. Simula dos medias, dos proporciones o ensayos por conglomerados con efecto, alfa, tamaño, ICC y semilla definidos; informa error Monte Carlo.

Recorrido recomendado

1 Justifique un rango de efectos plausibles.

2 Ejecute suficientes simulaciones con semilla fija.

3 Repita escenarios de tamaño, pérdidas e ICC.

Interpretación. La potencia es la proporción simulada de estudios que rechazan H0 bajo el escenario; no es una propiedad universal del estudio.

Límites y precauciones. Una simulación precisa de supuestos irreales sigue siendo engañosa; documente generador, escenarios y error Monte Carlo.

19.33 Control de calidad y diseño experimental

Para qué sirve. Distingue estabilidad del proceso, capacidad frente a especificaciones, prioridades de defectos y efectos factoriales.

Cómo funciona. Construye cartas I-MR, índices Cp/Cpk/Pp/Ppk, Pareto y regresión factorial de dos niveles con interacciones.

Recorrido recomendado

1 Verifique el orden temporal y la calidad de medición.

2 Evalúe estabilidad antes de capacidad.

3 En experimentos, conserve aleatorización, réplica y jerarquía de interacciones.

Interpretación. Un punto fuera de control sugiere causa especial; Cpk relaciona dispersión y centrado con especificaciones, no con límites de control.

Límites y precauciones. Límites de especificación deben provenir del proceso clínico/técnico; un factorial sin réplica limita la estimación independiente del error.

Alcance del 100%. Este manual documenta todas las rutas y controles disponibles en la versión 1. No sustituye un protocolo, la documentación original de cada método ni la revisión clínica/bioestadística.

20. Guardar proyectos, copiar y exportar

En SAmuel web, utilice **Resultados y proyecto** para conservar y retomar análisis. La descarga crea un archivo en el equipo; no modifica silenciosamente una descarga anterior. Compruebe el nombre, la ubicación y la fecha del archivo conservado.

Acción	Qué conserva	Uso recomendado
Guardar proyecto .samuel	Instantánea compatible de datos, metadatos, scripts, configuración, resultados y recetas.	Conservar una versión del estudio antes de salir o hacer cambios importantes.
Abrir proyecto web .samuel	Restaura los componentes compatibles después de validar el archivo; los scripts importados no se ejecutan.	Continuar un proyecto conservado y revisar sus resultados.
Retomar y ajustar configuración	Recupera los parámetros del resultado elegido; una nueva ejecución añade otro resultado.	Comparar revisiones sin sustituir el análisis previo.
Reejecutar resultado seleccionado	Reconstruye el resultado cuya receta esté disponible; las vistas interactivas pueden requerir esta acción.	Actualizar un cálculo después de revisar datos o configuración.
Recuperación temporal	Copia cifrada por cuenta. Guardado manual o periódico cada 60 segundos si se activa expresamente; restaurar exige confirmación.	Complementar la descarga del proyecto. No se guarda por defecto.
Sesión y repositorio de archivos	Guardados voluntarios del archivo adjunto o del archivo junto con el proyecto, dentro de los límites admitidos.	Reutilizar un archivo sin cargarlo ni analizarlo automáticamente.
Guardar resultado JSON	Representación estructurada del resultado seleccionado.	Intercambiar o auditar la salida; no sustituye al proyecto completo.
Guardar informe HTML	Contenido del resultado en un documento consultable; puede imprimirse o guardarse como PDF desde el navegador.	Preparar un entregable revisado.
Exportar informes	Acciones de exportación de tablas, figuras y documentos disponibles en ese apartado.	Elegir el formato admitido y comprobar el archivo descargado.
Biblioteca técnica	Manual, guía de uso, guía de estudio, referencias, licencia y ayuda.	Consultar la metodología sin abandonar la vista de trabajo.

Alcance del proyecto web. Un archivo aceptado no garantiza que toda receta importada pueda reejecutarse en cualquier entorno. SAmuel valida compatibilidad y límites antes de restaurar. Las vistas interactivas se descargan por separado; una imagen guardada no equivale a parámetros editables. Revise los mensajes de la importación y conserve el archivo de origen.

Control de versiones. Guarde copias identificables con fecha, decisiones y versión de SAmuel. Los límites del proyecto y del adjunto son diferentes: si la base supera el límite de instantánea, conserve el adjunto y documente los análisis. La recuperación temporal no reemplaza un respaldo externo comprobado.

Manual en PDF: abra Manual PDF desde la biblioteca. Para regresar use Volver a SAmuel dentro de la ventana de documentación; no es necesario iniciar otra sesión.

21. Guía de interpretación de estadísticos

Salida	Cómo leerla	Evite
Media	Promedio aritmético; sensible a extremos.	Interpretarla como valor típico con asimetría grave.
Mediana	50% de observaciones a cada lado.	Decir que 50% tiene exactamente ese valor.
DE	Dispersión alrededor de la media.	Confundir con precisión.
IC 95%	Valores compatibles con datos y modelo en repeticiones del procedimiento.	“95% de probabilidad de que el parámetro esté dentro” en enfoque frecuentista.
Valor p	Compatibilidad de datos con H0 bajo supuestos.	Probabilidad de que H0 sea verdadera o tamaño del efecto.
V de Cramér	Magnitud estandarizada de asociación en cualitativas.	Interpretarla causalmente.
β lineal	Cambio medio ajustado por unidad/categoría.	Extrapolar fuera del rango.
OR	Razón de odds; ajustada si el modelo tiene otras variables.	Llamarla riesgo relativo.
RR/RP	Razón de riesgos/prevalencias.	Usarlas sin definir diseño, evento y referencia.
IRR	Razón de tasas/conteos esperados.	Ignorar tiempo de exposición o sobredispersión.
HR	Razón de hazards instantáneos.	Llamarla razón de probabilidades acumuladas.
AUC	Capacidad de ordenar casos frente a no casos.	Concluir utilidad clínica sin calibración/validación.

Significativo no equivale a importante

Una estimación pequeña puede ser estadísticamente precisa pero clínicamente irrelevante; una estimación grande puede ser imprecisa. Presente magnitud, IC, umbral clínico y contexto.

No significativo no demuestra equivalencia

Para equivalencia o no inferioridad se requieren márgenes predefinidos y un diseño/análisis apropiado. Un p>0,05 puede reflejar baja potencia o gran incertidumbre.

22. Problemas frecuentes

Problema	Comprobación	Solución
No aparece una variable	Columnas activas en Datos.	Reactive la variable; escriba un fragmento en el selector.
Prueba bloqueada	Diseño confirmado y naturaleza.	Corrija el diccionario o confirme conscientemente el diseño.
Ajustar modelo parece inactivo	Diseño pareado ancho frente a longitudinal largo.	En la versión actual el botón permanece activo. Con una fila por sujeto puede ajustar un desenlace; con filas repetidas abra la pestaña integrada 3. Modelos longitudinales.
No se ven covariables/botones	Panel expandido supera la pantalla.	Use la barra vertical exterior; minimice paneles que no necesita.
Gráfico sin formato visible	Formato está minimizado.	Pulse Mostrar formato o Maximizar; desplácese al botón Aplicar.
R no ejecuta	Rscript configurado y paquetes.	Solicite al administrador la revisión del entorno y sus dependencias antes de volver a ejecutar.
Python no encuentra módulo	Nombre de paquete y entorno externo.	Instale desde el diálogo; confirme finalización y reinicie el código.
Barra parece quieta	Un paquete puede tardar sin informar porcentaje interno.	Lea “paquete actual” y contador completado; no cierre el diálogo.
Excel no abre CSV en columnas	Configuración regional.	Importe como UTF-8 delimitado por coma; el archivo sí contiene comas.
Resultado inesperado	Evento, referencia, faltantes, tipos.	Revise diccionario, tabla cruda, exclusiones y código generado.

Para reportar un problema, anote versión, módulo, pasos exactos, archivo mínimo sin datos sensibles, mensaje completo y resultado esperado.

23. Lista de verificación del análisis

Datos y diseño

- Pregunta y desenlace predefinidos.
- Unidad de análisis correcta.
- Independencia/pareamiento confirmado.
- Rangos, duplicados y codificación revisados.
- Faltantes descritos y manejados con justificación.

Método

- Prueba compatible con escala y diseño.
- Supuestos comprobados numérica y gráficamente.
- Referencias y evento declarados.
- Covariables elegidas por criterio sustantivo.
- Multiplicidad y sensibilidad consideradas.

Resultados

- n analizado y exclusiones informados.
- Estimación, IC y unidades reportados.
- Crudo y ajustado diferenciados.
- Diagnósticos y limitaciones presentes.
- No se confunde asociación con causalidad.

Reproducibilidad

- Proyecto, código e historial guardados.
- Figura exportada en calidad adecuada.
- Versiones de software documentadas.
- Datos personales protegidos.
- Revisión clínica/bioestadística realizada.

24. Glosario esencial

Término	Definición
Ajustado	Estimación condicionada a las demás variables incluidas en el modelo.
Censura	El evento no se observó durante el seguimiento disponible.
Confusor	Causa común de exposición y desenlace que puede distorsionar asociación.
Colinealidad	Alta dependencia entre predictores que inestabiliza estimaciones.
Calibración	Concordancia entre probabilidades predichas y observadas.
Discriminación	Capacidad de separar u ordenar personas con distinto desenlace.
Estimación cruda	Comparación sin control simultáneo de otras variables.
Evento	Resultado de interés codificado explícitamente.
Heterocedasticidad	Varianza no constante de errores/residuos.
Intervalo de confianza	Rango construido por un procedimiento con cobertura nominal.
Referencia	Categoría contra la cual se comparan las demás.
Residuo	Diferencia o contribución entre lo observado y lo ajustado.
Validación externa	Evaluación del modelo en datos independientes de otra población/entorno.

25. Fichas técnicas de los procedimientos disponibles

Estas fichas complementan el paso a paso. Resumen la pregunta que responde cada procedimiento, la hipótesis nula, los supuestos que deben comprobarse y la forma correcta de comunicar la salida.

25.1 Shapiro-Wilk y D'Agostino K²

Aspecto	Shapiro-Wilk	D'Agostino K²
Pregunta	¿La distribución observada se aparta de una normal?	¿Asimetría y curtosis se apartan conjuntamente de normalidad?
H0	Los datos proceden de una distribución normal.	Asimetría y curtosis son compatibles con normalidad.
Uso en SAmuel	n de 3 a 5000.	Alternativa operativa cuando n>5000.
Informe	W, p, n y diagnóstico gráfico.	K², p, n y diagnóstico gráfico.

Con n grande, desviaciones triviales producen p pequeño; con n pequeño, la prueba puede no detectar desviaciones importantes. Nunca elija un método únicamente por este valor p.

25.2 t y Wilcoxon de una muestra

t de una muestra: contrasta H0: μ=μ□. El estadístico estandariza la diferencia entre media observada y referencia por su error estándar. Requiere observaciones independientes y distribución de la media razonable; inspeccione extremos. Reporte media, diferencia frente a μ□, IC 95%, t, grados de libertad y p.

Wilcoxon de una muestra: contrasta localización mediante rangos con signo. Bajo simetría puede interpretarse respecto de una mediana/diferencia central; sin simetría, la interpretación es más general. Reporte valor de referencia, estadístico, p y un resumen robusto.

> **Ejemplo operativo:** Univariado → Contraste con valor de referencia → t de una muestra → `calidad_vida_12m` → referencia 70. La conclusión debe decir si la media observada es compatible con 70 y mostrar diferencia e IC; no basta "p<0,05".

25.3 Chi-cuadrado de Pearson y prueba exacta de Fisher

Ambas evalúan independencia entre variables cualitativas. Pearson compara frecuencias observadas y esperadas mediante Σ(O−E)²/E. Fisher calcula una probabilidad exacta condicionada, especialmente útil en tablas 2×2 pequeñas. Las observaciones deben ser independientes y cada sujeto aportar a una sola celda.

Salida	Función
Tabla observada y porcentajes	Muestra patrón y denominador clínicamente interpretable.
Frecuencia esperada mínima	Orienta si la aproximación Chi-cuadrado es adecuada.
p	Compatibilidad con independencia.
V de Cramér	Magnitud estandarizada de asociación; 0 indica ausencia en la tabla.

25.4 Student, Welch y Mann-Whitney

Student contrasta igualdad de medias con varianzas comparables. **Welch** contrasta medias sin imponer igualdad de varianzas y suele ser una opción robusta por defecto. **Mann-Whitney** compara rangos/distribuciones; solo bajo formas similares puede describirse como desplazamiento de localización. Las tres requieren grupos independientes.

Plantilla: “La diferencia media fue X unidades (IC 95% LI-LS; prueba de Welch, p=…). Las distribuciones y atípicos fueron…”. Para Mann-Whitney, informe medianas/RIC por grupo y evite llamarlo prueba de medianas sin justificar formas comparables.

25.5 ANOVA, ANOVA de Welch y Kruskal-Wallis

ANOVA evalúa H0 de igualdad de medias en todos los grupos y supone independencia, residuos aproximadamente normales y varianzas comparables. Welch relaja igualdad de varianzas. Kruskal-Wallis evalúa diferencias de rangos/distribuciones y exige independencia. Un resultado global no identifica pares: se necesitan comparaciones post hoc con control de multiplicidad.

F = variabilidad entre grupos / variabilidad dentro de grupos

25.6 Pearson, Spearman y Kendall

Los coeficientes varían entre −1 y 1. Pearson resume asociación lineal; Spearman, relación monótona por rangos; Kendall, concordancia ordinal de pares y suele ser útil con muchos empates o muestras pequeñas. Siempre inspeccione dispersión: una curva fuerte puede tener Pearson cercano a cero y un atípico puede dominarlo.

Correlación no es acuerdo. Dos métodos pueden correlacionarse casi perfectamente y diferir sistemáticamente. Para acuerdo de mediciones, use un análisis específico como Bland-Altman cuando corresponda.

25.7 t pareada, Wilcoxon pareado, McNemar y Friedman

Prueba	Datos	Unidad analizada	H0 aproximada
t pareada	Dos cuantitativas del mismo sujeto	Diferencias individuales	Media de diferencias=0.
Wilcoxon pareado	Dos mediciones ordinales/cuanti	Rangos con signo de diferencias	Distribución de diferencias centrada en 0.
McNemar	Dos dicotómicas pareadas	Pares discordantes	Probabilidades de cambio opuesto iguales.
Friedman	≥3 mediciones relacionadas	Rangos dentro de sujeto/bloque	Distribuciones/rangos iguales.

En todos los casos, elimine/gestione faltantes preservando pares o bloques. No mezcle sujetos distintos como si fueran mediciones repetidas.

25.8 Regresión lineal

Modela E(Y|X)=β□+β□X□+…+βpXp. Cada β es el cambio medio esperado en Y por unidad o categoría del predictor, manteniendo constantes los demás. Compruebe linealidad, independencia, homocedasticidad, residuos, influencia, especificación y colinealidad. La normalidad relevante es la de errores para inferencia, no necesariamente la de Y sin condicionar.

Ejemplo de redacción: “Cada año adicional de edad se asoció con un cambio medio ajustado de β unidades en calidad de vida (IC 95%…), manteniendo constantes tratamiento y multimorbilidad”.

25.9 Regresión logística binaria, multinomial y ordinal

La logística binaria modela el logit de P(Y=1). Al exponentiar β se obtiene un OR. La multinomial crea comparaciones para cada categoría no referente; la ordinal usa logits acumulativos y exige odds proporcionales. Compruebe codificación, independencia, forma funcional en continuas, separación, categorías escasas, colinealidad, convergencia y, para ordinal, proporcionalidad.

$$\text{logit}[P(Y=1)] = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p; \quad OR = \exp(\beta)$$

25.10 Poisson robusto, Poisson de conteos y binomial negativa

Para un desenlace binario, Poisson con enlace log y varianza robusta permite estimar RR/RP directamente; la corrección sándwich ajusta la incertidumbre porque la varianza Poisson no corresponde al binario. Para conteos, Poisson modela tasas o conteos esperados; la binomial negativa incorpora sobredispersión. Declare si existe tiempo-persona u offset cuando sea necesario.

La elección entre **RR** y **RP** es epidemiológica: cohorte/incidencia acumulada frente a estudio transversal/prevalencia. El mismo algoritmo no corrige un diseño mal interpretado.

25.11 Gamma con enlace log

Es apropiada para desenlaces continuos estrictamente positivos y asimétricos cuando la varianza crece con la media. Exp(β) se interpreta como razón multiplicativa de medias condicionadas. Verifique ausencia de ceros/negativos, especificación del enlace, residuos y observaciones influyentes.

25.12 Kaplan-Meier, Log-rank y Cox

Kaplan-Meier estima supervivencia por productos de probabilidades condicionales; Log-rank compara curvas; Cox modela $h(t|X)=h_0(t)\exp(\beta X)$. Exp(β) es HR. Evalúe censura no informativa, origen temporal, proporcionalidad con residuos de Schoenfeld, linealidad de continuas e influencia.

$$h(t|X) = h_0(t)\,\exp(\beta_1 X_1 + \ldots + \beta_p X_p)$$

25.13 PCA y matriz de correlación

El análisis de componentes principales resume variables cuantitativas correlacionadas en combinaciones lineales. Estandarice si las escalas son diferentes; interprete cargas, varianza explicada y propósito. PCA es exploratorio, sensible a atípicos y no identifica factores causales. La matriz de correlación debe acompañarse de tamaño muestral y consideración de multiplicidad.

25.14 Bland-Altman

Representa diferencia entre dos métodos/mediciones frente a su media y muestra sesgo promedio y límites de acuerdo. Evalúe si la variabilidad de diferencias cambia con la magnitud y si los límites son clínicamente aceptables. No sustituya estos límites por correlación.

25.15 Cómo usar una ficha en SAmuel

1. Localice la pregunta en los capítulos 8-14.
2. Ejecute el recorrido en la pestaña correspondiente.
3. Vuelva a esta ficha y marque cada supuesto con evidencia tabular o gráfica.
4. Compare la interpretación automática con la plantilla técnica.
5. Documente cualquier desviación o análisis de sensibilidad en el historial.

26. Referencias bibliográficas y guías

1 Dirección Xeral de Saúde Pública. Epidat 4: Ayuda general. Xunta de Galicia; 2016.

2 Dirección Xeral de Saúde Pública. Epidat 4: Ayuda de análisis descriptivo. 2014.

3 Dirección Xeral de Saúde Pública. Epidat 4: Ayuda de regresión logística. 2014.

4 von Elm E, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. *PLoS Med*. 2007;4:e296. doi:10.1371/journal.pmed.0040296.

5 Hopewell S, et al. CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials. *BMJ*. 2025;389:e081124. doi:10.1136/bmj-2024-081124.

6 Collins GS, et al. TRIPOD+AI statement: updated guidance for reporting prediction models. *BMJ*. 2024;385:e078378.

7 Wasserstein RL, Lazar NA. The ASA Statement on p-Values. *Am Stat*. 2016;70:129-133. doi:10.1080/00031305.2016.1154108.

8 Altman DG, Bland JM. Statistics notes: the normal distribution. *BMJ*. 1995;310:298.

9 McHugh ML. The chi-square test of independence. *Biochem Med*. 2013;23:143-149.

10 Zou G. A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159:702-706.

11 Barros AJD, Hirakata VN. Alternatives for logistic regression in cross-sectional studies. *BMC Med Res Methodol*. 2003;3:21.

12 Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. 3rd ed. Wiley; 2013.

13 Kleinbaum DG, Klein M. *Survival Analysis: A Self-Learning Text*. 3rd ed. Springer; 2012.

14 Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Wiley; 1987.

15 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med*. 2011;30:377-399.

16 NIST/SEMATECH. e-Handbook of Statistical Methods: Normal Probability Plot. Consulta 2026.

17 SciPy. Statistical functions: independent, paired, association and normality tests. Manual de referencia.

18 Riley RD, et al. Minimum sample size for developing a multivariable prediction model. *Stat Med*. 2019;38:1276-1296.

19 Moons KGM, et al. PROBAST: a tool to assess risk of bias and applicability of prediction models. *Ann Intern Med*. 2019;170:51-58.

20 jamovi project. jamovi: características, resultados dinámicos y reproducibilidad; y biblioteca oficial de módulos. Consulta: 2 de octubre de 2026.

21 JASP Team. JASP: lista oficial de módulos y características; descarga y licencia. Consulta: 2 de octubre de 2026.

22 Love J, Selker R, Marsman M, et al. JASP: Graphical Statistical Software for Common Statistical Designs. *J Stat Softw*. 2019;88(2):1-17.

23 van Doorn J, van den Bergh D, Böhm U, et al. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon Bull Rev*. 2021;28:813-826.

24 Kruschke JK. Bayesian Analysis Reporting Guidelines. *Nat Hum Behav*. 2021;5:1282-1291.

25 BlueSky Statistics. Descripción oficial del producto, planes y funciones Base/Pro y misión y versiones. Consulta: 2 de octubre de 2026.

26 Shepherd MA, Richardson EJ. Opting for open-source? A review of free statistical software programs. *Teaching Statistics*. 2024;46(1):53-63.

27 Muenchen RA. A Comparative Review of the BlueSky Statistics GUI for R. Actualizado el 23 de julio de 2026.

28 GNU Project. GNU PSPP y manual oficial. Consulta: 2 de octubre de 2026.

29 SOFA Statistics. Statistics Open For All. Consulta: 2 de octubre de 2026.

30 Fox J. Rcmdr: R Commander. CRAN. Consulta: 2 de octubre de 2026.

31 Orange Data Mining. Visual programming, machine learning and data visualization. Consulta: 2 de octubre de 2026.

Los enlaces sirven para profundizar. La presencia de una referencia no implica que SAmuel implemente todos los procedimientos descritos en ella.

27. Límites, privacidad y responsabilidad

- SAmuel procesa localmente la base dentro del equipo; el usuario controla ubicación, copias y acceso.
- No cargue información identificable en servicios externos ni la comparta sin autorización.
- Las recomendaciones automáticas son orientativas y dependen del diccionario y diseño confirmados.
- Revise resultados con una persona experta cuando influyan en atención, política, publicación o decisiones de alto impacto.
- Conserve trazabilidad de datos, transformaciones, exclusiones, paquetes y versiones.

SAmuel busca hacer el análisis médico y epidemiológico más accesible, transparente y reproducible. Su mejor resultado surge cuando la facilidad de uso se combina con pensamiento crítico, conocimiento clínico y rigor metodológico.

28. Investigación integral en SAmuel web

Plantillas temáticas y programación reproducible

El catálogo ampliado de la revisión del 27 de octubre de 2026 contiene 190 plantillas editables, derivadas de los módulos habilitados y sus variantes principales. Incluye las 24 medidas univariadas, los 27 tipos de gráficos estadísticos y las familias interactivas. Se carga cuando se abre R/Python para evitar solicitudes innecesarias al iniciar otros módulos.

Las plantillas Python reutilizan los motores de la web. Las ampliadas para R se identifican como puente R/Python y requieren `reticulate`; no constituyen implementaciones nativas independientes de cada fórmula en R. El usuario también puede usar plantillas R nativas y código disponible de resultados anteriores. Revise los marcadores de variables, las opciones y el diseño antes de ejecutar. Insertar una plantilla no ejecuta análisis ni instala paquetes.

La instalación se prepara para CRAN, Bioconductor o PyPI por nombre, sin limitarse a los paquetes utilizados por SAmuel. La compatibilidad depende de versiones, sistema operativo, compiladores, licencias y permisos. Se conservan tablas y objetos compatibles; los punteros internos del puente R/Python se retiran antes de guardar el entorno de la sesión.

Referencias técnicas: reticulate: llamadas de Python desde R y Bioconductor: instalación y compatibilidad. Disponer de una plantilla o paquete no demuestra que todas las combinaciones estadísticas hayan sido validadas.

Navegación y lectura del espacio de trabajo

El menú ofrece Análisis básico, Análisis intermedio y Análisis avanzado. El análisis univariado, cualitativo y cuantitativo, pertenece al nivel básico. Georreferenciación y Gráficos tienen accesos separados; la recuperación temporal se consulta al final del menú. Los apartados minimizados deben ocupar únicamente su cabecera.

Antes de analizar, revise el protocolo y confirme expresamente la relación entre observaciones. La detección es orientativa; aprobar el diseño requiere conocer cómo se obtuvieron los datos. Resultados y proyecto dispone de un visualizador independiente para ampliar tablas y figuras, incluido Forest plot.

La biblioteca técnica se abre dentro de la aplicación. Volver a SAmuel recupera la vista de trabajo. Acerca de SAmuel presenta el video y los beneficios del recorrido integrado. R y Python ofrecen Editor, Consola, Datos y Resultados; en pantalla completa se distribuyen en cuatro paneles equivalentes, inicialmente abiertos. La instalación de paquetes adicionales depende de compatibilidad, licencias y requisitos del sistema.

La acción visible Cerrar sesión retira la autorización y los recursos de la cuenta. No debe detener el servicio de otros usuarios en un servidor compartido. La X de una ventana anfitriona cierra el proceso de esa instancia cuando se utiliza el lanzador del equipo.

Cerrar la ventana y liberar recursos

Guarde su proyecto antes de cerrar. Al confirmar el cierre de la ventana propia de SAmuel en Windows, el lanzador apaga el servicio y las consolas. Un grupo privado de procesos permite a Windows terminar los hijos restantes de SAmuel, incluidos los motores Python/R y WebView, sin cerrar programas ajenos. Si Windows no puede preparar este control, el lanzador informa el problema y no abre una instancia sin protección. Al volver a abrir el acceso mientras SAmuel está activo, se vuelve a su ventana existente para evitar duplicados. Esta conducta corresponde al acceso de ventana local; no debe confundirse con apagar un servidor central compartido. Cerrar una pestaña de navegador independiente no ordena apagar el servidor.

Portada, movimiento y orientación por tema

Continuidad de proyectos. La importación permite proyectos nativos sencillos con todas sus columnas activas, sin diseño confirmado ni ajustes adicionales incompatibles. Recupera datos, resultados y scripts de R/Python como texto; no ejecuta esos scripts. Antes de reemplazar el trabajo actual valida la compatibilidad. Si encuentra formatos, registros o ajustes no migrados, mantiene intacto el trabajo abierto y rechaza el archivo. Conserve siempre su original. Esta adaptación no garantiza abrir todos los proyectos anteriores.

Audio de la presentación. Las calidades 720p y 1080p comparten una narración nivelada por tramos y música atenuada durante el habla. Se conservan las palabras y pausas naturales. El reproductor permite pausar, avanzar y ajustar el volumen; cambiar la calidad conserva la posición. Si tiene abierta una reproducción anterior, cierre y vuelva a abrir SAmuel para cargar la mezcla actualizada.

Las figuras estadísticas del fondo y del área libre de presentación tienen movimiento decorativo propio: las barras cambian de altura, las curvas varían su forma, los intervalos se abren y cierran y las cohortes se reorganizan. La animación se suspende cuando la página queda oculta para reducir trabajo innecesario. Use **Pausar animación** en el pie para detenerlo; **Activar animación** lo reanuda. Se conserva esta preferencia en el navegador. Al iniciar sin preferencia propia, se respeta el movimiento reducido del sistema. La decoración no representa resultados de su base.

Las ventanas informativas organizan el contenido con cabecera azul y tarjetas de color. Cada uno de los 24 grupos explica su propósito y permite desplegar **Datos e interpretación**: qué variables necesita y qué supuestos o límites revisar. Estas orientaciones acompañan el criterio metodológico; no sustituyen la comprobación del diseño ni validan por sí solas una conclusión.

SAmuel web reúne diseño, análisis médico y epidemiológico, interpretación, visualización y reporte reproducible. El catálogo se organiza en niveles y familias para que el investigador pueda elegir por la pregunta de investigación y no por un listado extenso de nombres. La aplicación utiliza motores estadísticos y bibliotecas versionadas; cada procedimiento mantiene sus requisitos y supuestos.

Lectura y presentación. La portada comparte paneles compactos de igual altura, con márgenes visibles. Ctrl − reduce la vista, Ctrl + o Ctrl = la amplía y Ctrl 0 restablece el tamaño original mediante el zoom nativo. Las ventanas de información conservan desplazamiento interno cuando su contenido lo requiere. En el panel de presentación, seis botones resumen beneficios y alcance: diseño, análisis, interpretación, gráficos y mapas, continuidad del proyecto y aprendizaje e informes. Cada uno abre una ventana informativa. El reproductor usa sus controles nativos de reproducción, pausa, volumen y pantalla completa; los destacados breves acompañan los capítulos.

Explorar el potencial. La portada presenta 24 categorías, con Gráficos, Mapas y Programación separados. Conserva los 55 procedimientos públicos sin duplicarlos entre grupos. Las tarjetas de beneficios usan el mismo formato del catálogo y resumen aportes, ventajas prácticas del flujo y alcance. La vista previa del video y las vistas interactivas ofrecen acceso explícito a pantalla completa.

Una ruta de investigación, no solo una calculadora

Empiece por la pregunta, población, exposición, desenlace, tiempo y unidad de análisis. Revise la propuesta de muestras relacionadas o independientes y apruébela según su protocolo. La detección automática no demuestra el diseño ni sustituye el conocimiento de cómo se obtuvieron los datos. En Datos y variables, prepare y previsualice, revise la base, declare naturaleza y nivel de medición y seleccione variables. Los apartados comienzan plegados para que abra únicamente lo necesario.

Univariado pertenece a Básico e incluye frecuencias cualitativas, medidas cuantitativas, distribución y normalidad. Intermedio y Avanzado reúnen comparación, regresión, supervivencia, epidemiología, diagnóstico, causalidad, Bayes, síntesis de evidencia, predicción, series, encuestas, psicometría, modelos estructurales, imputación y diseño. La normalidad aislada no decide todos los supuestos; revise también independencia, estructura del estudio, residuos, tamaño muestral y objetivo.

Valor de la integración. SAmuel conecta orientación epidemiológica, cálculo guiado, interpretación, mapas, programación y continuidad del proyecto. La ventaja práctica es mantener estas tareas dentro de un mismo trabajo. No son prestaciones necesariamente exclusivas: jamovi documenta análisis y proyectos reproducibles; JASP, métodos frecuentistas y bayesianos; y Epidat, análisis epidemiológico. La elección depende del alcance del método y su validación. Fuentes oficiales: jamovi: Features, JASP: Features y Epidat (consulta: 25 de octubre de 2026).

Validar con otra población

Los módulos de validación predictiva permiten cargar una base externa en el formulario. Compruebe nombres, tipos, categorías, unidades y desenlace antes de ejecutar. La validación externa evalúa el desempeño del modelo en datos distintos: no equivale a volver a entrenarlo sin declarar el cambio. Considere discriminación, calibración, incertidumbre y utilidad, además del ajuste. El archivo elegido se conserva en la receta del proyecto para reproducir y retomar el análisis; por ello el proyecto debe custodiarse como información confidencial.

Tablas, figuras y mapas

Las tablas inician con diez filas y permiten buscar, ordenar y exportar. En resultados grandes, el visor consulta páginas: su exportación corresponde a la página cargada; la exportación completa es una acción independiente. Mapas y gráficos organiza familias por intención de análisis. Forest plot puede usar columnas de estimación e intervalos o un modelo compatible del historial. Los cambios de formato se aplican a copias para conservar el resultado previo.

Leaflet y los gráficos interactivos disponen de capas, etiquetas y recursos propios del entorno. La cartografía local puede consultarse sin pedir teselas; elegir un fondo externo implica solicitudes al proveedor y sus condiciones. Una función que exige clave, servicio o contexto Shiny requiere esos recursos. Verifique que los campos visibles en una ficha o exportación sean los autorizados para compartir.

Guardar, continuar y revisar

Guarde el archivo **.samuel** para conservar datos, resultados compatibles, recetas, protocolo epidemiológico, naturaleza y niveles declarados y scripts. Retomar configuración permite modificarla y producir un resultado nuevo sin sustituir el anterior. Una imagen PNG importada es una representación del resultado: reejecute su receta para recuperar parámetros editables cuando estén disponibles. El proyecto nunca ejecuta scripts importados por sí solo.

La recuperación temporal mantiene una copia cifrada por cuenta. No se guarda por defecto: use Guardar recuperación ahora o active expresamente el guardado cada 60 segundos; en este último caso se actualiza si hay cambios y la página está visible. Una copia previa requiere su elección antes de sustituirse; restaurar requiere confirmar el reemplazo del trabajo en memoria. Esta protección complementa la descarga del proyecto, no es un historial ilimitado ni reemplaza copias verificadas. La recuperación depende de la custodia del servidor o perfil donde se creó. Conserve una copia propia y revise los respaldos antes de un cambio de entorno.

Programación autorizada y extensiones verificables

R y Python disponen de editor con resaltado, búsqueda, selección, temas, scripts, salida y objetos compatibles conservados durante la sesión. El código propio requiere autorización y dependencias. Una consola no es una sandbox por tener un editor: usa los permisos del proceso. En servicios multiusuario, la ejecución libre necesita workers aislados. El administrador debe configurar ese entorno antes de habilitarla; los análisis guiados no exigen programar.

El catálogo de extensiones explica el contrato de manifiesto, versión API, permisos, pruebas, huellas y firma. También permite cargar un paquete para inspección. Se verifican archivos y firma sin importar o ejecutar su código. Un paquete inspeccionado no queda instalado o habilitado automáticamente. Las restricciones preservan el control de integridad y los datos del investigador.

Cuenta, documentación y uso responsable

Solicite aprobación por el canal indicado y acceda mediante usuario, contraseña y autenticador. El servicio central aplica una sola sesión simultánea por cuenta y separa los espacios de trabajo. Si pierde una sesión, coordine su recuperación: el cierre administrativo puede hacer perder trabajo no guardado. Las condiciones completas de Académico, Individual y Comercial están en Licencia; un comprobante se revisa administrativamente y no activa una cuenta por sí mismo.

La guía metodológica conecta propósito, datos requeridos, supuestos, cálculo, ejemplos, interpretación y referencias. Lea magnitud, intervalo e incertidumbre junto con el diseño; no deduzca causalidad solo por una asociación o un valor p. Cite SAmuel y las fuentes metodológicas utilizadas. Novedades documenta cambios disponibles y cómo utilizarlos. La puesta en servicio central exige alojamiento, HTTPS, custodia, permisos, respaldos y evaluación de carga; no surge de abrir una ventana del navegador.

SAmuel · Ayuda y formato documental

Versión 1 · EPINex · Revisión de formato: 27/10/2026

Formato EPINEX-DOC-1

La biblioteca utiliza una identificación consistente de producto y versión, una fecha de revisión editorial diferenciada de las versiones de fuentes o acuerdos, un único título principal y navegación común. Las páginas se adaptan a escritorio y móvil, permiten saltar al contenido y muestran foco al usar el teclado.

Dónde consultar

Desde Ayuda en el pie de la portada se accede a las guías, al manual PDF, a Cómo citar, Licencia y Novedades. SAmuel conserva también las rutas de seguridad, componentes de terceros y guía de estudio. La cita de RADAR remite a las condiciones de la edición web que muestra la aplicación.

Conservación y alcance

Los acuerdos de licencia, la cita de SAmuel en BibTeX y las atribuciones se conservan con el mismo SHA256 que antes de esta revisión. Las tablas clínicas existentes no se sustituyen. La actualización del formato no constituye certificación jurídica o clínica ni amplía las autorizaciones de uso. Este anexo documenta el formato vigente; las páginas históricas mantienen su presentación original.

Verificación documental

Se comprobaron 20 rutas de documentos, estilos y PDF mediante los controladores reales en memoria: todas respondieron 200. Dieciséis documentos HTML cumplen la estructura y navegación comunes, sin identificadores duplicados ni anclas internas rotas. Se verificaron 32 diseños, a 390 y 1440 píxeles, y se corrigió el desbordamiento del encabezado en móvil.

Apertura y cierre

Tres ciclos nativos por producto registraron 17 procesos propios al estar ambas aplicaciones abiertas y cero de esos procesos después de cerrarlas. Repetir los iconos conservó los PID de las instancias existentes. Se comprobó que cerrar una aplicación conserva los PID y la respuesta HTTP de la otra. Se cerró mediante la interfaz, sin finalizar tareas a la fuerza.

Límite de la prueba

Los ciclos se realizaron desde la portada, sin cargar fuentes, iniciar análisis ni ejecutar consolas. No apareció un proceso R durante esos ciclos; esto no equivale a probar la salida de un cálculo R activo. Guarda tus proyectos y scripts antes de cerrar. En RADAR, el trabajo pendiente puede interrumpirse y deberá retomarse; lo confirmado se conserva.

SAmuel de EPINex - Versión 1

28/10/2026 - Bienvenida y despedida en el video

La misma identidad de la aplicación

La presentación abre con el logo que aparece por piezas, utilizando los recursos originales de la aplicación. Al finalizar, el logo aparece por franjas para acompañar la despedida. El símbolo mantiene también un movimiento suave durante el recorrido central.

Saludos completos con la voz de referencia

La bienvenida dice Bienvenido a SAmuel. El cierre dice Espero que disfrutes de SAmuel. Se conserva la voz de referencia Leslie en español, con velocidad 0,92. Cada saludo se generó como frase completa y se montó sin dividir palabras ni acelerar la grabación.

Recorrido aprobado conservado

Las escenas y la narración del recorrido central se mantienen. El audio original se conserva muestra por muestra en el máster continuo, y el video conserva sus gráficos, textos y tiempos; la codificación facilita la navegación. Las dos resoluciones comparten el mismo audio final. Los tiempos de navegación incorporan los 3,2 segundos de bienvenida.

Cómo verlo

Abre la presentación desde la portada. Puedes reproducir desde el comienzo, ajustar el volumen, ampliar a pantalla completa y elegir 720p o Full HD. Si la aplicación estaba abierta durante la actualización, guarda primero tu trabajo y vuelve a abrirla cuando esté libre.

Verificación y alcance

La edición dura 213.23 segundos. Ambas resoluciones se decodificaron completas, y el reproductor se comprobó con reproducción continua en navegador de prueba. Se revisaron las uniones, las frases completas, los cambios de calidad y el margen de audio sin saturación. La prueba técnica no sustituye la percepción individual del oyente. Esta actualización afecta a la presentación audiovisual y conserva los datos, análisis, licencias y controles de acceso.

SAmuel de EPINex - Versión 1

28/10/2026 - Transiciones tecnológicas del video

Un puente entre el logo y el recorrido

La bienvenida enlaza con la presentación mediante un portal radial de luz: circuitos, anillos holográficos y partículas en profundidad. El mismo lenguaje visual une la finalización del recorrido con la despedida. Una gráfica de datos distingue las uniones de SAmuel.

Voz y tiempos conservados

Cada transición ocupa 1,2 segundos dentro del tiempo existente. El audio AAC se copia íntegro de la edición aprobada, sin recortes, aceleración, nuevas pausas ni efectos sonoros sobre la narración. Se mantienen completas las frases de bienvenida y despedida y los tiempos de capítulos, temas y subtítulos.

Cómo ver esta edición

Abre el video de la portada. Puedes elegir 720p o Full HD, ajustar el volumen y ampliar a pantalla completa. Si la aplicación estaba abierta durante la actualización, guarda el proyecto o trabajo pendiente y vuelve a abrirla normalmente cuando esté libre.

Verificación y alcance

Se decodifican completos ambos archivos, se comprueba el recuento de fotogramas y la igualdad SHA256 del audio frente a la edición previa. La reproducción completa y los controles se prueban en un navegador aislado. Se revisan visualmente las cuatro transiciones de ambos productos. Esta prueba técnica no afirma una escucha humana completa ni una observación dentro de las ventanas nativas. El cambio afecta solo al material audiovisual; conserva datos, análisis, proyectos, acceso y licencias. Las versiones anteriores se conservan.

SAmuel de EPINex - Versión 1

29/10/2026 - Transiciones entre todos los temas del video

Continuidad visual durante el recorrido

La presentación incorpora 12 transiciones internas de 0,8 segundos. Un portal radial de circuitos, anillos holográficos y partículas enlaza cada tema con el siguiente. Cada unión incorpora un símbolo del tema que comienza: diseño, calidad de datos, análisis, incertidumbre, mapas, proyectos y bibliografía. Se conservan también los efectos de bienvenida y despedida.

La narración permanece completa

Se copia íntegra la pista AAC aprobada, sin volver a grabar, recortar o acelerar palabras. No se añaden pausas ni efectos sonoros sobre la voz. Las transiciones ocupan tiempo de imagen ya existente; los capítulos, temas sincronizados y subtítulos mantienen sus tiempos. Ambos archivos se generan a 30 fotogramas por segundo.

Cómo ver la actualización

Abre la presentación desde la portada. El reproductor permite 720p, Full HD, volumen y pantalla completa. Si la aplicación estaba abierta durante la actualización, guarda el proyecto o trabajo pendiente y vuelve a abrirla normalmente cuando esté libre.

Comprobaciones y límites

Las dos resoluciones se decodifican completas y se comprueba el recuento de fotogramas. Se compara por SHA256 que el audio siga idéntico, se revisa cada cambio de tema y se verifica que las imágenes fuera de las uniones conserven su contenido. La reproducción completa y los controles se prueban en navegador aislado. El control de Windows no expuso las ventanas nativas, por lo que no se afirma una observación dentro de ellas ni una escucha humana completa. Esta edición afecta solo al video; conserva datos, análisis, acceso y licencias. Las versiones anteriores se mantienen.

SAmuel de EPINex - Versión 1

30/10/2026 - Lectura ajustable y espacio de trabajo

Ajustar el tamaño

Ctrl + aumenta la lectura, Ctrl - la disminuye y Ctrl 0 restablece el tamaño original. También puedes mantener Ctrl y girar la rueda del ratón. El ajuste pertenece a la ventana: no abre un segundo navegador ni carga datos.

Activar esta actualización

Guarda tu trabajo y cierra normalmente la aplicación. Vuelve a abrirla desde su icono. Una ventana que ya estaba abierta conserva el código anterior hasta reabrirla. No cierres ni recargues durante una operación que quieras conservar en curso.

Controles y alcance

El zoom conserva la autenticación, los permisos y el cierre independiente de procesos. No se habilitan herramientas de desarrollo ni atajos generales del navegador. Los atajos de teclado se comprobaron en ventanas nativas sintéticas sin datos reales; Ctrl con rueda queda habilitado por el motor WebView2, sin prueba automatizada del gesto físico combinado.

SAmuel · Suplemento de uso y perfiles

Guía vigente · 2 de noviembre de 2026

2026-11-02 · Perfiles personalizados por módulos y recorrido completo de uso

Administración puede crear varios perfiles, seleccionar módulos y funciones y asignarlos a cuentas individuales. Se comprueban permisos en el servidor; ocultar un botón no es la única protección. El perfil funcional es independiente del alcance de datos, licencia y autenticador. Las cuentas existentes conservan su acceso; el administrador debe revisar y asignar los perfiles nuevos.

Administración de perfiles

Dentro de la aplicación, usa Mi perfil y permisos, junto a Recorrido de uso. Administración puede crear o editar un nombre de perfil, marcar funciones y procedimientos y asignarlo a cuentas ya registradas. Consultar es requisito básico; Preparar, Analizar, Exportar y Consola son permisos separados. Cada procedimiento también tiene su casilla. Administrador conserva todos sus permisos y debe quedar uno. Los cambios se comprueban en las siguientes solicitudes; recarga para actualizar listas visibles. El primer registro administrativo existente conserva administración inicial; el orden de inicio de sesión no concede ese privilegio.

Journey: de la entrada al cierre

1. Preparar y entrar. Define tu tarea e ingresa con una cuenta previamente registrada y licenciada, contraseña y autenticador. Los perfiles no crean licencias ni reemplazan MFA.

2. Definir la pregunta. Especifica población, variables, desenlace, comparación y diseño. Consulta la guía metodológica antes de elegir el cálculo.

3. Preparar la base. Carga un archivo compatible o ejemplo sintético; revisa tipos, unidades, faltantes y duplicados. Conserva el original y documenta transformaciones.

4. Elegir el procedimiento. Selecciona grupo y procedimiento habilitado, variables y parámetros; lee requisitos, ayuda y supuestos. Previsualiza la preparación antes de aplicar.

5. Analizar e interpretar. Ejecuta y revisa tablas, figuras, magnitud, incertidumbre y advertencias. Distingue asociación de causalidad y significación de relevancia.

6. Usar código cuando corresponda. R/Python requiere permiso Consola y consentimiento de ejecución. Revisa el código; seleccionar un procedimiento no concede automáticamente ejecución libre.

7. Guardar y comunicar. Guarda proyecto o repositorio disponible. Exporta resultados, tablas, figuras e informes únicamente con permiso y revisa su contenido antes de compartirlo.

8. Finalizar y retomar. Verifica el guardado y cierra sesión. Abre tu proyecto y revisa historial al volver. Las cuentas, licencias y autenticador conservan sus controles previos.

Ayuda durante el trabajo

Pulsa Recorrido de uso para consultar una ruta de ocho pasos con Qué hacer, Ayuda para decidir y Antes de continuar. Anterior y Siguiente permiten revisar la ruta sin cambiar datos; Volver a la aplicación cierra la ayuda en la misma ventana. Usa además la ayuda específica de cada módulo.

Límites de esta revisión

Edición web local. No se modificó Zoho, no se publicaron servicios y no se reinició RADAR en producción. Estos cambios no certifican validación clínica completa, rendimiento con cualquier volumen ni entrega real de mensajes; se mantienen las limitaciones documentadas de cada aplicación. Una función visible no otorga permisos de otras funciones. Los datos ya descargados no se revocan al cambiar un perfil.

SAmuel | Consolas de R y Python en web

Anexo de candidata · 2026-11-05 · No publicada

Alcance confirmado

Ambos lenguajes se incluyen en la primera versión web de SAmuel. El adaptador está implementado y sus rutas probadas con un ejecutor simulado. No hay un servicio de ejecución real aprobado todavía.

Cómo se usa

Abre la consola, escribe o importa tu script, revisa la advertencia y autoriza la ejecución. Se envía una copia CSV de tu base como datos (también df); revisa sus tipos de fechas, categorías e identificadores. Importar el script o abrir la pantalla no lo ejecuta.

Separación y límites

Cada trabajo utiliza un contenedor temporal sin red ni carpetas de las aplicaciones. Límites iniciales: 60 segundos, 512 MiB, una CPU y 64 procesos; dos trabajos por broker y uno por sesión. La copia admite 8 MiB de CSV, 100.000 filas y 200 columnas. El exceso se rechaza sin muestreo automático.

Resultados y cancelación

Se reciben texto, tablas y PNG. Asigna la tabla a resultado para devolverla; la vista previa admite 200 filas y 50 columnas. No se exporta una vista previa incompleta como si fuese toda la tabla. Cancelar o cerrar sesión afecta al trabajo propio. Una sesión nueva requiere nuevo consentimiento.

Qué no está habilitado

No hay red, instalación de paquetes bajo demanda, entorno persistente ni HTML interactivo de código propio. El operador debe incluir las librerías necesarias en una imagen revisada. No existe ejecución nativa de respaldo si el broker falla.

Validación pendiente

Las pruebas actuales no ejecutaron contenedores reales. Antes de activar: comprobar R/Python, gVisor y límites impuestos por Linux, aislamiento, resultados, cancelación, agotamiento de recursos y limpieza de trabajos tras una caída. El recolector externo está implementado con etiquetas propias, pero falta aceptarlo en el host.

Estado de la entrega

Candidata separada, no publicada ni aprobada para Internet. No modifica la consola local ni los proyectos del PC. Detalles del operador: CONSOLA_WEB_SAMUEL.md en la entrega web.