

SOFTWARE ESTADÍSTICO PARA ANÁLISIS DE DATOS EN PROYECTOS DE INVESTIGACIÓN

AUTOR: GUILLERMO RAMÓN FRANCO DEL RÍO

RESUMEN

En la era del big data y la ciencia abierta, el análisis estadístico es una competencia esencial para estudiantes, docentes e investigadores. Este artículo ofrece una revisión de los principales programas estadísticos utilizados en investigación científica, así como una guía práctica para elegir el software más adecuado según el tipo de proyecto, nivel de experiencia y recursos disponibles.

INTRODUCCIÓN

En todo proyecto de investigación, el análisis de datos es una etapa crítica que permite transformar observaciones en conocimiento. El software estadístico no solo facilita el cálculo de pruebas complejas, sino que también promueve la reproducibilidad, la calidad metodológica y la eficiencia en el manejo de grandes volúmenes de datos (1). Elegir el software adecuado puede marcar la diferencia entre un análisis sólido y uno lleno de errores o limitaciones técnicas.

Ejemplos de uso:

Análisis de ensayos clínicos:

SPSS es utilizado para analizar datos de ensayos clínicos, ya que permite realizar una amplia gama de pruebas estadísticas como t de Student, Chi Cuadrada, ANOVA, regresión logística y análisis de supervivencia entre muchos otros (2). Por ejemplo, en un ensayo clínico para evaluar la eficacia de un nuevo

medicamento contra la diabetes, SPSS puede ayudar a determinar si hay diferencias significativas en los niveles de glucosa en sangre entre los grupos de tratamiento y el grupo de control.

Estudios de cohortes:

SAS, este software es ideal para manejar grandes conjuntos de datos y realizar análisis complejos(3). En estudios de cohortes, como aquellos que investigan la relación entre la exposición a un factor de riesgo y la incidencia de enfermedades cardiovasculares, SAS puede manejar y analizar datos de miles de participantes a lo largo de varios años.

Investigación genómica:

Con su amplia gama de paquetes y funciones, R es muy utilizado en bioinformática y análisis genómico. Por ejemplo, análisis multivariado con modelación matemática y en estudios que buscan identificar variantes genéticas asociadas con el cáncer (4), R puede realizar análisis de asociación del genoma completo (GWAS) y visualizar los resultados en gráficos de Manhattan.

Epidemiología:

Stata es ampliamente utilizado en epidemiología para realizar análisis de datos de estudios observacionales (5). Por ejemplo, en un estudio que investiga la prevalencia de una enfermedad infecciosa en diferentes

regiones, Stata puede realizar análisis de regresión y ajustar por variables de confusión.

Estos ejemplos ilustran cómo el software estadístico puede ser una herramienta poderosa en la investigación médica, facilitando el análisis de datos complejos y mejorando la calidad y reproducibilidad de los resultados.

¿CÓMO ELEGIR EL SOFTWARE MÁS ADECUADO?

Seleccionar el software correcto depende de varios factores. Aquí se presentan algunos criterios clave:

- 1. Nivel de experiencia del usuario
 - Principiante: Excel, Jamovi, JASP, SPSS (con GUI)
 - Intermedio: SPSS (con sintaxis), Stata, SAS
 - Avanzado: R, Python, Google Colab
- 2. Tipo de análisis requerido
 - Descriptivo y básico: Excel, SPSS, Jamovi
 - Regresión, ANOVA, pruebas no paramétricas: SPSS, R, Stata, Jamovi
 - Modelos avanzados: R, Python, SAS, JASP
- 3. Acceso institucional o recursos económicos
 - Presupuesto limitado: R, Jamovi, JASP, Python, Google Colab
 - Licencias académicas: SPSS, Stata, SAS
- 4. Compatibilidad con estándares de publicación

Las revistas científicas favorecen entornos

OPCIONES DISPONIBLES

Software	Licencia	Dificultad	Fortalezas principales
R©	Libre	Alta	Flexible, potente, reproducible, ideal para investigación avanzada.
SPSS©	Comercial	Baja-media	Interfaz amigable, ampliamente usado en psicología, medicina y ciencias
Stata©	Comercial	Media	Muy utilizado en economía y salud pública; buen equilibrio entre código y GUI.
SAS©	Comercial	Alta	Robusto en análisis predictivo y grandes bases de datos.
Jamovi (6)	Libre	Baja	Intuitivo, interfaz amigable, basado en R.
JASP (7)	Libre	Baja	Ideal para estadística bayesiana, amigable para principiantes.
Python© (8)	Libre	Alta	Excelente para análisis integrados con machine learning. Basado en R.
Google Colab© (9)	Libre	Alta	Ejecución en la nube, ideal para colaboración y enseñanza.
Excel© (10)	Comercial	Muy baja	Accesible, útil para análisis descriptivos simples.

reproducibles como R, Python o Colab, que permiten compartir scripts completos.

5. Recomendaciones según el perfil del usuario

- Estudiante de licenciatura: Empieza con Jamovi o JASP. Usa Excel para análisis básicos. Luego avanza a R o Python.
- Tesista o investigador: SPSS y Stata si tu universidad tiene licencias. R si deseas habilidades reproducibles.
- Docente universitario: Jamovi o JASP para clases introductorias. R o Python para cursos avanzados.

CONCLUSIÓN

Elegir el software estadístico correcto puede influir en la calidad de los análisis y la aceptabilidad del trabajo académico. Hoy existen opciones potentes, gratuitas y accesibles para desarrollar competencias estadísticas sólidas. Lo importante es dar el primer paso.

RECURSOS Y ENLACES ÚTILES

R: <https://www.r-project.org>

RStudio: <https://posit.co/download/rstudio-desktop/>

Jamovi: <https://www.jamovi.org>

JASP: <https://jasp-stats.org>

SPSS: <https://www.ibm.com/products/spss-statistics/pricing>

Stata:

<https://www.stata.com/order/new/edu/grad/plans/>

SAS OnDemand:

https://www.sas.com/en_us/software/on-demand-for-academics.html

Python: <https://www.python.org>

Google Colab:

<https://colab.research.google.com>

Excel / Office para estudiantes:

<https://www.microsoft.com/education>

CRÉDITOS

R® es una marca registrada de la R Foundation for Statistical Computing; 2016 R Foundation (CC-BY-SA 4.0).

SPSS® es una marca registrada de IBM Corporation, registrada en múltiples jurisdicciones.

Stata® es una marca registrada de StataCorp LLC, College Station, TX, EE. UU.

SAS® y todos los nombres de productos o servicios de SAS Institute Inc. son marcas registradas de SAS Institute Inc. en EE. UU. y otros países.

Jamovi es software libre desarrollado por *The jamovi project* (AGPL-3.0). No existen indicios de que Jamovi sea una marca registrada.

JASP es software libre (AGPL-3.0) desarrollado por *The JASP Team*. Actualmente no existe registro de marca para JASP relacionado con software estadístico.

Python® and the two-snake logo are registered trademarks of the Python Software Foundation, used by permission.

Google Colab® es un servicio de Google LLC. Google y el logotipo de Google son marcas registradas de Google LLC.

Microsoft Excel® es una marca registrada de Microsoft Corporation en EE. UU. y otros países.

REFERENCIAS

1. Batmanabane G, Maiti R, Krishna MBN. Research Methodology and Biostatistics in the Postgraduate Curriculum: The AIIMS-Bhubaneswar Experience. J Assoc Physicians India. 2024;72(11):98-100.
- 2.2. Corp IBM. IBM SPSS Statistics for Windows, Version 29.0. Armonk, NY: IBM Corp.; 2024.
- 3.3. Inc SASI. SAS/STAT 15.3 User's Guide. Cary, NC: SAS Institute Inc.; 2023.
- 4.4. Team RC. R: A language and environment for statistical computing. Version 4.4.0. Vienna: R Foundation for Statistical Computing; 2024.
- 5.5. StataCorp. Stata Statistical Software: Release 18. College Station, TX: StataCorp LLC; 2023.
- 6.6. The jamovi p. jamovi (Version 2.4) [computer software]. Sydney: The jamovi project; 2025.
- 7.7. Team J. JASP (Version 0.18.3) [computer software]. Amsterdam: University of Amsterdam; 2025.
- 8.8. Python Software F. Python 3.12 documentation. Wilmington, DE: Python Software Foundation; 2024.
- 9.9. Bisong E. Google Colaboratory. Building machine learning and deep learning models on Google Cloud Platform. Berkeley, CA: Apress; 2019. p. 59-64.
- 10.10. Microsoft C. Microsoft Excel for Windows (Version 2403) [computer software]. Redmond, WA: Microsoft Corporation; 2024.