

El uso de lenguajes de programación R y Python en los grados de Sociología en España: un tren que se escapa

The Use of R and Python Programming Languages in Sociology Degrees in Spain: A Runaway Train

Javier Cantón-Correa¹

Resumen

La digitalización ha transformado múltiples disciplinas, incluida la sociología. Este estudio explora la integración de los lenguajes de programación R y Python en los programas de grado en sociología en universidades españolas, crucial para formar sociólogos y sociólogas competentes en análisis de datos. Se evalúa la inclusión de estos lenguajes a través de un análisis de planes de estudio de 17 universidades españolas, examinando 70 asignaturas técnicas. Los resultados indican que la enseñanza de R y Python es limitada y que predominan herramientas tradicionales como SPSS. Se analizan las barreras institucionales, económicas y culturales que dificultan su adopción y se proponen estrategias para su integración progresiva. Es esencial actualizar los currículos de sociología para incorporar estas herramientas desde los primeros cursos, diversificando las opciones de software y asegurando una capacitación docente adecuada.

Palabras clave

Sociología, lenguajes de programación, análisis de datos, R, Python, currículo académico.

Abstract

Digitalization has transformed multiple disciplines, including sociology. This study explores the integration of R and Python programming languages in sociology degree programs in Spanish universities, crucial to train sociologists competent in data analysis. The inclusion of these languages is evaluated through an analysis of curricula from 17 Spanish universities, examining 70 technical subjects. The results indicate that the teaching of R and Python is limited and that traditional tools such as SPSS predominate. Institutional, economic and cultural barriers to their adoption are analyzed and strategies for their progressive integration are proposed. It is essential to update sociology curricula to incorporate these tools from the first courses, diversifying software options and ensuring adequate teacher training.

Keywords

Sociology, programming languages, data analysis, R, Python, academic curriculum..

Cómo citar/Citation

Cantón-Correa, Javier (2025). El uso de lenguajes de programación R y Python en los grados de sociología en España: un tren que se escapa. *Revista de Sociología de la Educación - RASE*, 18(1), 250-263. <http://dx.doi.org/10.7203/RASE.18.1.29538>.

Recibido: 29-09-2024

Aceptado: 08-01-2025

¹ Javier Cantón-Correa, Universidad Internacional de La Rioja.

1. Introducción

En la era digital actual, el análisis de datos a gran escala se ha convertido en una herramienta indispensable para comprender los complejos fenómenos sociales (Gualda, 2021, 2022). Sin embargo, las enseñanzas universitarias en sociología han tardado en comprender la importancia del uso de herramientas computacionales en sociología, particularmente aquellas relacionadas con el análisis de datos y la modelización de fenómenos sociales mediante programación. Este trabajo investiga el papel que ocupan aquellos lenguajes de programación, específicamente R y Python, con capacidades no sólo para su aplicación metodológica para el análisis sino también para la modelización algorítmica mediante Inteligencia Artificial (Taboada, 2024), en los grados de Sociología en España.

R es un lenguaje de programación y un entorno de *software* especializado en estadísticas, análisis de datos y gráficos. Es ampliamente utilizado en el ámbito académico y de investigación por su gran capacidad para manejar, analizar y visualizar datos de manera eficiente. En las ciencias sociales, R se utiliza para realizar análisis estadístico, modelado de datos, investigación cuantitativa, y la creación de gráficas y visualizaciones complejas que ayudan a interpretar los fenómenos sociales a través de datos numéricos. Por otro lado, Python es un lenguaje de programación de alto nivel, interpretado y de propósito general, conocido por su legibilidad y simplicidad. Aunque es aplicable a múltiples dominios, en las ciencias sociales, Python se ha convertido en una herramienta esencial para el análisis de datos, la minería de textos (*text mining*), el procesamiento de lenguaje natural (NLP por sus siglas en inglés), y la visualización de datos.

Ambos lenguajes de programación son ampliamente utilizados en investigaciones sociales, de manera creciente, en los últimos años. Este cambio en la popularidad de los lenguajes de programación y paquetes estadísticos en las últimas dos décadas refleja una transformación significativa en las ciencias sociales. Herramientas como SPSS, que fueron ampliamente utilizadas en el pasado, han experimentado una notable disminución en favor de lenguajes de programación más versátiles y adaptables, como R y Python (Muenchen, 2012). Este cambio se debe, en parte, a las capacidades avanzadas de estos lenguajes para manejar grandes volúmenes de datos y su flexibilidad para personalizar y adaptar análisis complejos. Para los futuros sociólogos, el dominio de estas herramientas no solo representa una ventaja técnica, sino una competencia fundamental que permite acceder a metodologías más potentes y estar al nivel de los avances en investigación social.

R y Python se han convertido en estándares del análisis de datos y la programación estadística debido a su accesibilidad y la robustez de sus bibliotecas, lo que ha incrementado su uso en la investigación académica. Las cifras actuales de citas académicas muestran un ascenso continuo de estos lenguajes, mientras que el uso de programas tradicionales como SPSS sigue en declive. Esto evidencia que, para mantenerse al día en las prácticas de investigación, es fundamental que los programas de sociología incluyan la enseñanza de estos lenguajes. Incorporar R y Python en los currículos no solo prepara a los estudiantes para el mercado laboral, sino que les proporciona habilidades para realizar análisis complejos y adaptar su trabajo a las demandas de la era digital. Así, los futuros sociólogos podrán adoptar un enfoque empírico y basado en evidencia, mejorando su capacidad para abordar problemas complejos en la sociedad contemporánea.

Los objetivos de esta investigación son identificar la presencia y relevancia de estos lenguajes en los planes docentes de los Grados de Sociología en las Universidades españolas y examinar cómo se integran en el plan de estudios.

2. Marco teórico

La integración de lenguajes de programación en la enseñanza de la sociología no es un fenómeno aislado, sino parte de una transformación más amplia impulsada por el auge del *big data* y la necesidad de adaptar la investigación social a nuevas fuentes de información (González, 2019; Gualda, 2021, 2022). Como sugieren Fernández Lizana (2020) y Taboada (2024), R y Python ofrecen herramientas versátiles para el análisis de datos cuantitativos y cualitativos, lo que las hace ideales para la modelización sociológica avanzada. Sin embargo, la incorporación de estos lenguajes en la enseñanza universitaria sigue enfrentando barreras que van más allá de las metodologías tradicionales, incluyendo factores institucionales y curriculares (Elosua, 2009). En este sentido, es clave considerar cómo otras disciplinas han integrado estas herramientas y qué estrategias pueden aplicarse en la sociología.

Dicha integración, especialmente en disciplinas tradicionalmente cualitativas como la sociología, ha sido objeto de un creciente interés académico. Elosua (2009) ya sugería hace 15 años que R es una alternativa robusta al uso predominante de SPSS, ofreciendo no solo capacidades avanzadas para el análisis de datos, sino también accesibilidad y una amplia gama de paquetes estadísticos que potencian tanto la docencia como la investigación. Fernández Lizana (2020) también apoya esta perspectiva, destacando que R facilita el análisis y la visualización de datos de manera flexible y personalizable, lo cual es crucial para la investigación social que implica el manejo de grandes volúmenes de datos.

Por otro lado, González (2019) subraya que el avance del *big data* y los algoritmos está redefiniendo las ciencias sociales, al permitir el análisis masivo de datos procedentes de redes digitales. Este nuevo paradigma no solo amplía las capacidades analíticas de los investigadores, sino que también plantea desafíos metodológicos y éticos importantes. En esta línea, Gualda (2021, 2022) argumenta que el desarrollo de la sociología computacional y el social *big data* ofrece nuevas oportunidades para el análisis sociológico, al combinar métodos tradicionales con técnicas avanzadas de minería de datos y modelado algorítmico. La integración de estas metodologías permite explorar fenómenos sociales complejos de una manera más sistemática y precisa.

El uso de plataformas como X (antes llamada Twitter) para el estudio de fenómenos migratorios es un ejemplo de cómo el *big data* puede aplicarse a la investigación social. Gualda y Díaz (2020) exploran esta área mediante el uso de técnicas de análisis de datos y *software* especializado para captar patrones de comportamiento y discurso en redes sociales. Estos estudios no solo amplían la comprensión de los procesos migratorios, sino que también demuestran la capacidad del *big data* para abordar problemas sociales contemporáneos.

Sin embargo, Maestro Cano (2019) advierte sobre las limitaciones del *big data* en la investigación social, tales como la calidad y representatividad de los datos, y los dilemas éticos asociados con el uso de información personal sin el debido consentimiento. Este autor aboga por una aproximación crítica y responsable al uso de estas tecnologías, resaltando la necesidad de políticas claras que regulen la obtención y uso de datos a gran escala.

Taboada (2024) aborda la automatización en el análisis de datos textuales a través del procesamiento de lenguaje natural y el aprendizaje automático. Estas tecnologías permiten a los sociólogos manejar y analizar grandes volúmenes de texto de forma eficiente, superando las limitaciones de los métodos cualitativos tradicionales y abriendo nuevas posibilidades para la investigación en ciencias sociales. La capaci-

dad de automatizar el análisis de datos textuales no solo mejora la eficiencia, sino que también permite a los investigadores descubrir patrones complejos en grandes corpus de datos.

Por tanto, el uso de lenguajes de programación y técnicas avanzadas de análisis de datos está transformando la práctica sociológica. La inclusión de estas herramientas en la formación académica de los sociólogos es esencial para enfrentar los desafíos de un mundo cada vez más digitalizado. A medida que las metodologías computacionales y las técnicas de *big data* e Inteligencia Artificial se integran más en la investigación social, es fundamental que los programas académicos se actualicen para incluir estas competencias, asegurando así que los futuros sociólogos estén preparados para aplicar enfoques analíticos avanzados en su práctica profesional.

La integración de lenguajes de programación en la enseñanza de la sociología no es un fenómeno exclusivo de España. Estudios realizados en el ámbito anglosajón han evidenciado la creciente importancia de estas herramientas en programas de formación en ciencias sociales (Boccardo & Ruiz, 2019; Salganik, 2019). En países como Estados Unidos o Reino Unido, la enseñanza de R y Python ha sido promovida mediante iniciativas institucionales que han incentivado la formación docente y la actualización de planes de estudio.

Por ejemplo, la Universidad de California, Berkeley, ha incorporado R y Python en sus programas de métodos cuantitativos en sociología y ciencias políticas, ofreciendo cursos específicos sobre programación para análisis de datos sociales (Healy, 2018). De manera similar, la London School of Economics (LSE) ha desarrollado módulos de formación en *data science* aplicados a la sociología, en los cuales se prioriza el uso de Python para la recopilación y análisis de datos provenientes de redes sociales y encuestas en línea (Salganik, 2019). La experiencia de estos países puede servir como referencia para la adaptación de estrategias similares en el contexto español, asegurando que la enseñanza sociológica se alinee con los estándares internacionales.

3. Metodología

El análisis de datos cuantitativos y cualitativos es esencial para la investigación sociológica y politológica (González, 2019; Gualda & Díaz, 2020). Tradicionalmente, SPSS ha sido el *software* más utilizado desde los primeros años del siglo XXI en los Grados de Sociología, pero, con el avance informático, lenguajes de programación como R o Python ofrecen potentes herramientas para el manejo, análisis y visualización de datos (Elosua, 2009; Fernández Lizana, 2020). Este estudio se basa en la revisión de planes de estudios de Grados de Sociología de universidades españolas, incluyendo un análisis de contenido de los planes de estudios universitarios y la comparación con las demandas actuales del mercado laboral en el campo de la sociología. Se presta especial atención a asignaturas metodológicas y técnicas que implican análisis de datos, como «Análisis y visualización de datos» (ya sean cuantitativos o cualitativos) o «Análisis de redes y *big data*», así como cursos complementarios y asignaturas optativas que puedan abordar metodologías de investigación aplicadas a entornos digitales y técnicas de minería de datos.

La metodología utilizada en este estudio se basó en el análisis de los planes de estudio de asignaturas dentro de los grados y dobles grados en Sociología. Para una primera extracción de datos, se utilizó la página web del Ministerio de Ciencia, Innovación y Universidades denominada QEDU². A partir de la

² <https://www.universidades.gob.es/donde-estudiar/>.

base de datos obtenida, se detectaron un total de 17 universidades españolas, públicas y privadas, que ofertaban 17 grados de Sociología y 10 dobles grados de una combinación de sociología y otras disciplinas. De ellos, a partir de la información disponible en sus respectivas páginas web, se revisaron los planes de estudio de un total de 70 asignaturas de carácter técnico, susceptibles de aplicar lenguajes de programación y *software* estadístico, para determinar el grado de utilización de dichas herramientas en la enseñanza de sociología.

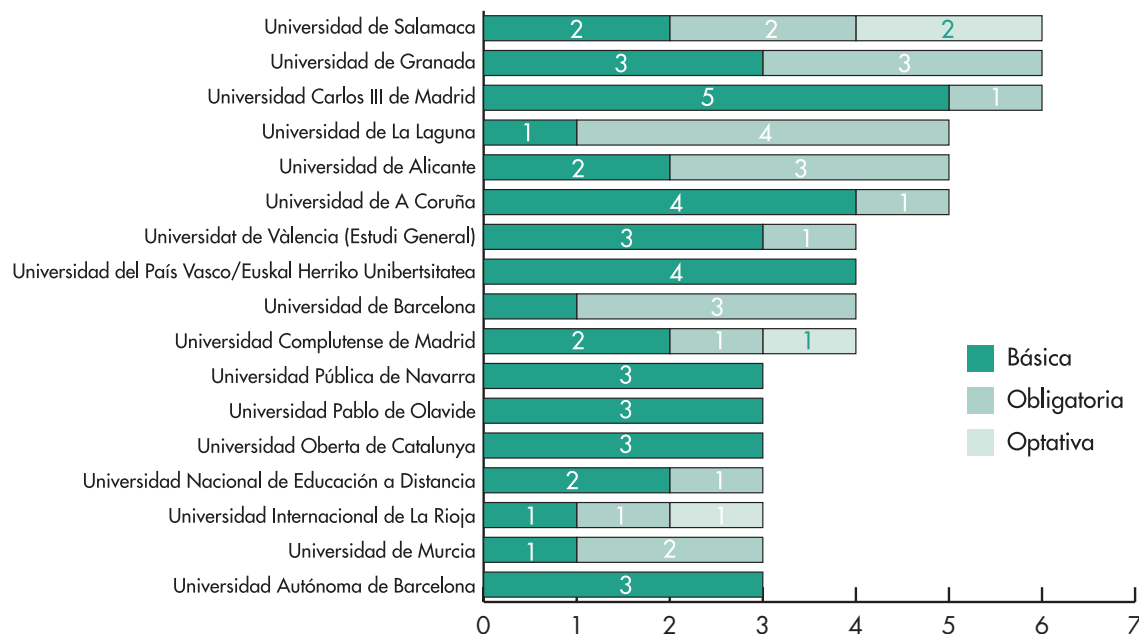
Identificadas las universidades y sus respectivos planes de estudio, se procedió a un análisis detallado de las asignaturas metodológicas y técnicas para evaluar la presencia de herramientas computacionales en la enseñanza. Para la recopilación de datos, se analizaron las guías docentes de cada asignatura, observando la presencia o ausencia de lenguajes de programación específicos, tales como R y Python, y otros *softwares* específicos de análisis como SPSS, Stata, PSPP o Gephi. El objetivo fue identificar no solo la frecuencia de uso de estos programas, sino también los contextos específicos en los que se emplean. También se extrajo información referente a las asignaturas, como el tipo de asignatura (básica, obligatoria u optativa), así como la URL donde está publicado el Plan de Estudios de dicha asignatura, el número de créditos y el curso y semestre en que se imparte esta asignatura.

Esta metodología permitió obtener una visión detallada de la situación actual de la enseñanza de herramientas computacionales en la sociología en España y de las oportunidades de mejora para incluir lenguajes de programación en asignaturas clave desde los primeros años de formación.

4. Resultados

En general, se observa una mayor presencia de asignaturas técnicas obligatorias, seguidas por las optativas y, en menor medida, por las básicas. El gráfico 1 muestra el número de asignaturas técnicas en los grados de Sociología de diversas universidades españolas, clasificadas según su tipo: básica, obligatoria y optativa.

Gráfico 1. Asignaturas técnicas en Grados de Sociología españoles según tipo de asignatura



Fuente: elaboración propia.

La Universidad Carlos III de Madrid presenta la mayor cantidad de asignaturas técnicas, con un fuerte enfoque en las obligatorias, lo que indica una orientación significativa hacia la formación técnica en su currículo de sociología. La Universidad del País Vasco y la Universidad de A Coruña también destacan con un número notable de asignaturas obligatorias, lo que sugiere un compromiso similar con la inclusión de competencias técnicas en sus programas. Por otro lado, universidades como la Universidad de La Laguna y la Universidad de Barcelona ofrecen un número limitado de asignaturas técnicas, concentradas principalmente en las obligatorias, lo que refleja una menor integración de estas competencias en su currículo.

Estos datos resaltan la disparidad en la oferta de asignaturas técnicas entre universidades y evidencian el predominio de asignaturas obligatorias frente a las básicas y optativas. Esto sugiere que, aunque hay un esfuerzo por incluir habilidades técnicas en la formación sociológica, la implementación varía considerablemente, lo que podría generar diferencias en las competencias técnicas adquiridas por los estudiantes según la universidad en la que estudien.

La mayoría de estas asignaturas se concentran en el segundo y tercer año del programa. El 41,03 % de las asignaturas técnicas se imparten en tercer curso, seguido de un 35,9 % en segundo curso, lo que indica que la formación técnica intensiva se prioriza en la parte intermedia del grado. En el primer curso, solo un 10,26 % de las asignaturas son técnicas, lo que sugiere que los estudiantes reciben una base más general o introductoria al inicio de la carrera, y la formación técnica se introduce progresivamente. Finalmente, en el cuarto curso, el porcentaje de asignaturas técnicas disminuye al 12,82 %, lo que podría estar relacionado con una orientación hacia asignaturas más especializadas o aplicadas, o la realización de prácticas y proyectos de fin de grado en lugar de cursos formativos estrictamente técnicos.

Esta distribución refleja una estructura curricular que busca consolidar las competencias técnicas en los años intermedios, preparándolos adecuadamente antes de enfrentar asignaturas avanzadas o aplicaciones prácticas en el último año del programa.

A partir de la información recabada sobre las 70 asignaturas, se procedió a su clasificación mediante el análisis de las materias seleccionadas, agrupándolas en cinco bloques temáticos que reflejan los diferentes enfoques y competencias técnicas necesarias para la formación en investigación sociológica. Esta organización permite entender mejor el tipo de formación que reciben los estudiantes en relación con el análisis de datos y las técnicas cuantitativas, fundamentales en la sociología contemporánea:

1. **Estadística y análisis de datos:** Este bloque incluye asignaturas como «Socioestadística I», «Socioestadística II», «Análisis de Datos I», «Análisis de Datos II», «Análisis Multivariado de Datos», «Estadística Aplicada a las Ciencias Sociales I» y II, entre otras. Estas materias están diseñadas para desarrollar habilidades estadísticas y de análisis de datos, esenciales para interpretar grandes volúmenes de información y realizar investigaciones sociológicas robustas.
2. **Métodos cuantitativos y técnicas de investigación:** En este grupo se encuentran asignaturas como «Métodos Cuantitativos de Investigación Social», «Métodos y Técnicas de Investigación Social Cuantitativos», y «Técnicas de Investigación Social Cuantitativa I» y II. Estas materias proporcionan una base sólida en los métodos y técnicas de investigación cuantitativa, capacitándolos para llevar a cabo estudios empíricos y análisis cuantitativo en el ámbito de la sociología.

3. **Software y herramientas de análisis:** Este bloque incluye asignaturas enfocadas en el manejo de herramientas informáticas y software específico para el análisis de datos, tales como «Hojas de Cálculo. Nivel Intermedio», «Ciencia y Minería de Datos», «Análisis de Datos Asistido por Ordenador en Sociología», «*Data Science*» y «Análisis y Visualización de Datos Cuantitativos». Estas asignaturas dotan a los estudiantes de competencias prácticas en el uso de software estadístico, minería de datos y herramientas de visualización.
4. **Análisis multivariado y avanzado:** Este conjunto abarca materias como «Análisis Multivariado de Datos», «Introducción al Análisis Multivariado» y «Análisis Multivariado Aplicado a la Sociología». Estas asignaturas están orientadas hacia el dominio de técnicas avanzadas de análisis multivariado, esenciales para examinar relaciones complejas entre múltiples variables y realizar análisis profundos en estudios sociológicos.
5. **Estadística aplicada y métodos descriptivos:** Este bloque contiene asignaturas como «Estadística Descriptiva Aplicada a las Ciencias Sociales», «Estadística Inferencial Aplicada a las Ciencias Sociales» y «Estadística Aplicada a la Investigación Mediante Encuestas». Estas materias se enfocan en la aplicación de métodos estadísticos descriptivos e inferenciales en contextos sociales, así como en el diseño y análisis de encuestas, una herramienta crucial para la recolección de datos en sociología.
6. Además de estos bloques temáticos, se identifican algunos ejemplos específicos de asignaturas que reflejan **enfoques y aplicaciones especializadas**, como «Análisis de Redes y *Big Data*», que muestra la creciente importancia del análisis de redes sociales y el *big data* en la investigación actual, o «Metodología de la Investigación por Encuestas», que se centra en el diseño, implementación y análisis de encuestas, fundamentales en la sociología empírica.

A partir de estos bloques, se ha procedido a categorizar las asignaturas en tres tipos principales según su enfoque y nivel de profundidad:

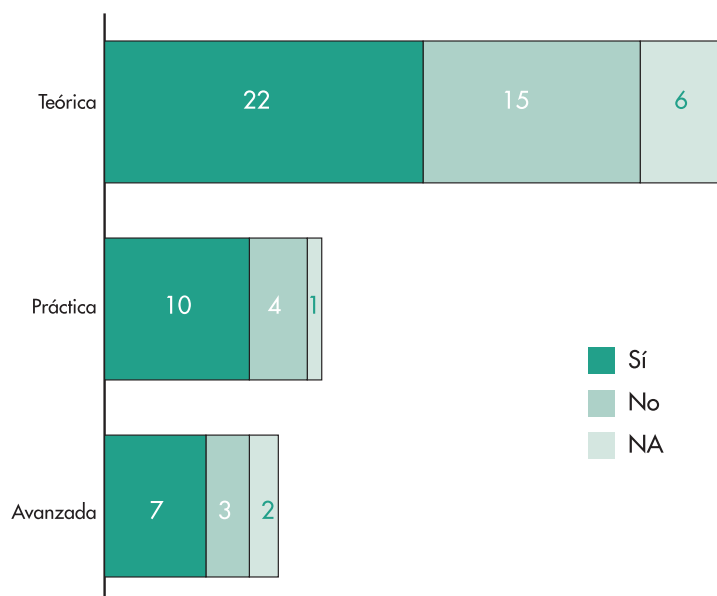
- **Teóricas:** Estas asignaturas cubren los principios y teorías subyacentes en estadística, métodos cuantitativos y técnicas de análisis, ofreciendo una base conceptual fundamental para los estudiantes.
- **Prácticas:** En esta categoría se agrupan las asignaturas que incluyen el uso de software y herramientas específicas para el análisis de datos, proporcionando habilidades prácticas y aplicadas.
- **Avanzadas:** Estas asignaturas abordan técnicas avanzadas, como el análisis multivariado, la minería de datos y el análisis de redes, que son necesarias para investigaciones avanzadas y permiten un análisis profundo de fenómenos complejos en la sociología.

La clasificación de las asignaturas en teóricas, prácticas y avanzadas se realizó considerando su enfoque metodológico y los contenidos específicos que imparten. Se definieron como teóricas aquellas asignaturas que presentan conceptos estadísticos o metodológicos sin un componente aplicado explícito. Las asignaturas prácticas incluyen el uso de *software* y herramientas computacionales específicas para el análisis de datos, proporcionando habilidades aplicadas. Finalmente, las asignaturas avanzadas se caracterizan por el empleo de técnicas de análisis multivariado, minería de datos o *big data*, esenciales para investigaciones más complejas. Esta distinción permite evaluar en qué medida los programas de sociología preparan a

los estudiantes en competencias técnicas aplicadas. Para definir esta clasificación, se revisaron las guías docentes de cada asignatura y se categorizó cada una en función de la proporción de contenidos teóricos y aplicados. Se consideraron ‘teóricas’ aquellas asignaturas donde más del 70 % del contenido se enfocaba en la enseñanza conceptual o metodológica sin aplicación directa en *software*. Las ‘prácticas’ incluían al menos un 50 % de sesiones aplicadas con *software* o herramientas de análisis. Finalmente, las ‘avanzadas’ se identificaron como aquellas que, además de incluir herramientas computacionales, introducían técnicas de modelado, análisis multivariado o enfoques complejos de minería de datos.

A partir de la categorización en teóricas, prácticas y avanzadas, esta gráfica muestra la distribución de asignaturas técnicas en los grados de Sociología según la inclusión de herramientas de análisis cuantitativo, clasificada en «Sí», «No» y «NA» (no aplicable). De las 70 asignaturas analizadas, se observa una preponderancia de asignaturas teóricas, representando un total de 43 asignaturas, que equivalen al 61,4 % del total. Dentro de este grupo, 22 asignaturas (31,4 % del total) incluyen herramientas de análisis cuantitativo, mientras que 15 (21,4 %) no lo hacen y 6 (8,6 %) están marcadas como «NA».

Gráfico 2. Asignaturas técnicas en Grados de Sociología españoles según categoría de asignatura



Fuente: elaboración propia.

En el caso de las asignaturas prácticas, se cuentan 15 asignaturas en total, lo que representa el 21,4 % del total de asignaturas analizadas. De estas, 10 asignaturas (14,3 % del total) incluyen herramientas de análisis cuantitativo, mientras que 4 (5,7 %) no las incluyen y 1 (1,4 %) está clasificada como «NA».

Por último, en las asignaturas avanzadas, que suman un total de 12 (17,1 % del total), se observa que 7 de ellas (10 % del total) integran herramientas de análisis cuantitativo, 3 (4,3 %) no lo hacen y 2 (2,9 %) están marcadas como «NA».

Este análisis sugiere que, aunque existe un esfuerzo en incluir herramientas cuantitativas en las asignaturas técnicas de sociología, especialmente en las prácticas y avanzadas, todavía hay una proporción significativa de asignaturas teóricas y prácticas que no utilizan estos recursos, lo cual podría limitar el desarrollo de habilidades cuantitativas en los estudiantes.

El análisis de las frecuencias de palabras en los títulos de las asignaturas técnicas en los grados de Sociología muestra una clara orientación hacia el análisis de datos y el uso de técnicas estadísticas. Los términos más comunes son «análisis» (28), «datos» (19), «investigación» (17) y «estadística» (15), lo cual indica que estas áreas representan el núcleo de la formación técnica en sociología. Esta alta frecuencia sugiere que gran parte de las asignaturas están diseñadas para proporcionar a los estudiantes habilidades en la recopilación, procesamiento y análisis de datos, elementos clave para la investigación en ciencias sociales.

Además, palabras como «sociales» (13), «técnicas» (12), «social» (12) y «aplicada» (12) refuerzan el enfoque práctico y aplicado de estas materias en el contexto de las ciencias sociales. Otros términos relevantes como «ciencias» (11), «cuantitativas» (6) y «sociología» (5) sugieren una conexión explícita entre estas competencias técnicas y su aplicación en el análisis sociológico. Asimismo, términos como «multivariante» (4), «encuestas» (5) y «métodos» (4) subrayan el interés en técnicas avanzadas y metodologías específicas, como el análisis multivariante y el diseño y aplicación de encuestas, que son fundamentales en la investigación empírica en sociología. Este análisis de frecuencias revela, en conjunto, una estructura curricular enfocada en dotar a los estudiantes de habilidades cuantitativas y de análisis de datos aplicadas al estudio de los fenómenos sociales.

Gráfico 3. Términos usados en el título de las asignaturas técnicas en Grados de Sociología españoles, en formato de nube de palabras

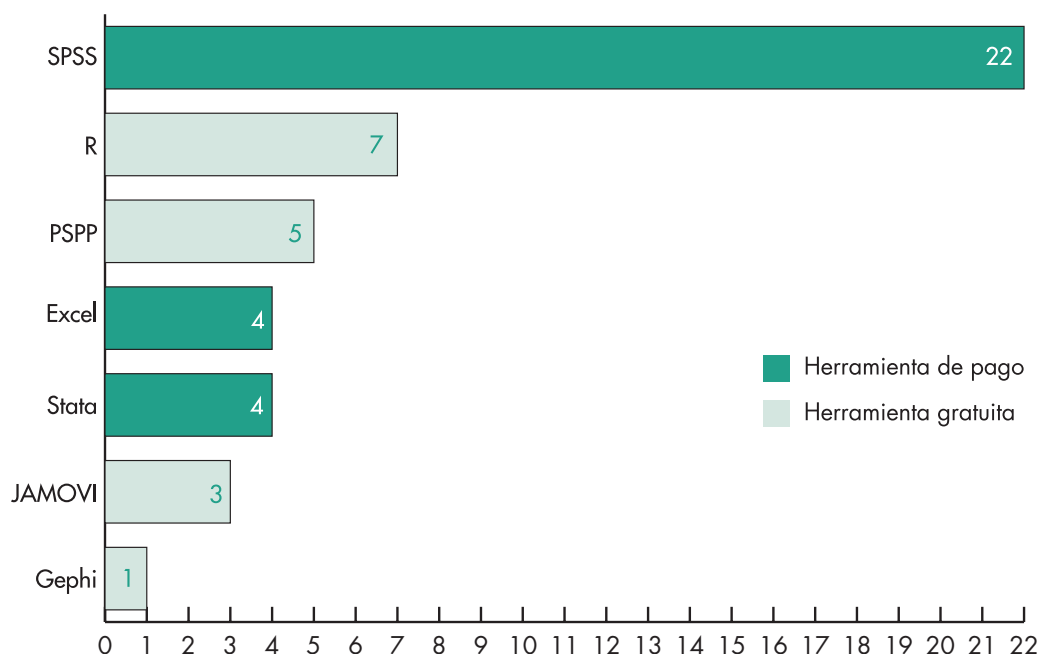


Fuente: elaboración propia.

4.1. Lenguajes de programación y software utilizados

Centrándonos en las herramientas usadas dentro de estas asignaturas técnicas, de las 70 asignaturas detectadas, sólo en 46 de ellas, el 65,7% del total, se detectó el uso de algún tipo de herramienta, ya sea *software* específico o lenguaje de programación. Como puede verse en el gráfico 4, la herramienta más utilizada es SPSS (22 asignaturas), *software* considerado estándar para el análisis estadístico y de gran implantación en los planes de estudios de las universidades españolas, probablemente por su accesibilidad y familiaridad en el entorno académico.

Gráfico 4. Herramientas usadas en asignaturas técnicas de los Grados de Sociología en España



Fuente: elaboración propia.

En comparación, otras herramientas presentan frecuencias de uso mucho menores. El lenguaje de programación R se usa en 7 asignaturas, seguido por PSPP, una alternativa gratuita a SPSS, en 5. Excel y Stata, ambas de pago, se utilizan en 4 asignaturas respectivamente, mientras que JAMOV, otro *software* gratuito enfocado en la accesibilidad, se emplea en 3 asignaturas. Finalmente, Gephi, una herramienta específica para el análisis de redes, aparece solo en 1 asignatura.

El análisis de los datos obtenidos arroja varias conclusiones: la primera es que solo el 10% de las asignaturas de los grados de Sociología en las universidades españolas incluyen algún lenguaje de programación, siendo R el único mencionado y utilizado principalmente en asignaturas avanzadas. En contraposición, SPSS mantiene una posición predominante, ya que es usado en casi el 50% de las asignaturas técnicas, lo cual refuerza su rol tradicional en la formación sociológica. Otros programas, como Stata, PSPP, JAMOV y Gephi, también están presentes en los planes de estudio, aunque con menor frecuencia, lo que indica una limitada diversidad en el uso de herramientas de análisis y programación dentro de los grados de Sociología. Cabe destacar además la inexistencia de Python en el currículo, lo que priva a los estudiantes de una de las herramientas de programación más versátiles y ampliamente utilizadas en la ciencia de datos actual, con amplias potencialidades de uso en campos emergentes como la Inteligencia Artificial. A diferencia de R, que se ha consolidado como un estándar en la estadística aplicada, Python ofrece un ecosistema más amplio que abarca desde el análisis de datos hasta el aprendizaje automático, *web scraping* y procesamiento de lenguaje natural. Esta versatilidad hace que su exclusión de los programas de sociología limite las oportunidades de los estudiantes para abordar problemáticas sociales desde enfoques computacionales más avanzados.

5. Discusión

Los resultados muestran una notable dependencia de herramientas tradicionales como SPSS en la formación de los grados de Sociología en España, con un 50% de las asignaturas técnicas utilizando este *software*, mientras que lenguajes de programación más flexibles y versátiles, como R y Python, tienen una presencia marginal o inexistente. Esto evidencia una brecha significativa entre las competencias técnicas ofrecidas en los planes de estudio y las habilidades demandadas en la investigación actual y el mercado laboral. La presencia limitada de R en solo el 10% de las asignaturas y la ausencia total de Python sugieren que los programas de sociología no están integrando herramientas que son cada vez más relevantes para el análisis de datos avanzado y el *big data*, áreas que se están volviendo fundamentales en la sociología computacional y en investigaciones basadas en datos masivos.

A pesar de la flexibilidad y gratuidad de R y Python, el predominio de SPSS en la enseñanza de la sociología en España podría responder a varios factores. En primer lugar, la familiaridad y consolidación de SPSS en los planes de estudio a lo largo de décadas han generado una inercia institucional difícil de modificar. En segundo lugar, los perfiles docentes juegan un papel clave, ya que muchos profesores se han formado en metodologías tradicionales y no han recibido capacitación suficiente en lenguajes de programación. Además, SPSS es una herramienta de uso algo más intuitivo, lo que reduce la curva de aprendizaje en comparación con R o Python, que requieren una alfabetización en programación. Por otro lado, las barreras económicas también influyen, ya que la enseñanza de R y Python implica una mayor inversión en formación docente y en el desarrollo de materiales didácticos adaptados. Estas razones explican la lenta adopción de estas herramientas en comparación con otras disciplinas más orientadas a la computación y el análisis de datos.

La rigidez de los planes de estudio y la burocracia académica podrían dificultar también la introducción de nuevas asignaturas o la actualización de programas existentes. La creación de cursos específicos en R y Python requiere la aprobación institucional y, en algunos casos, el desarrollo de nuevas infraestructuras o materiales, lo que implica un proceso largo y sujeto a restricciones presupuestarias. El uso extendido de SPSS en la enseñanza universitaria responde, en gran parte, a su consolidación en los planes de estudio y a la disponibilidad de licencias institucionales. En este contexto, la incorporación de R y Python requiere una mayor inversión en formación docente y adaptación curricular.

Asimismo, se observa una distribución desigual de las asignaturas técnicas a lo largo del plan de estudios. La concentración de estas asignaturas en los años intermedios indica que las habilidades técnicas son introducidas después de los cursos iniciales de teoría sociológica, lo cual puede tener la ventaja de aportar una base conceptual sólida antes de introducir el análisis de datos. Sin embargo, la escasez de formación técnica en los primeros años limita el desarrollo temprano de competencias prácticas, y la reducción de estas asignaturas en el último curso puede dificultar la consolidación de habilidades avanzadas antes de que los estudiantes concluyan sus estudios. Además, la predominancia de asignaturas teóricas sobre prácticas y la baja frecuencia de asignaturas avanzadas, que abordan análisis multivariado o técnicas de minería de datos, apunta a una formación centrada en conocimientos abstractos más que en habilidades aplicadas, lo cual podría ser insuficiente en una disciplina que está incorporando crecientemente métodos cuantitativos y computacionales.

En conjunto, estos hallazgos ponen de relieve un enfoque formativo que, aunque incluye técnicas estadísticas básicas y *software* de análisis, no responde completamente a las exigencias de una sociología

moderna orientada al análisis de grandes volúmenes de datos y al uso de herramientas de programación flexibles y escalables. Una de las principales limitaciones de este estudio es la dependencia de información obtenida de fuentes públicas en línea, lo que puede generar sesgos en la recopilación de datos sobre los planes de estudio. Asimismo, no se han realizado entrevistas a docentes o estudiantes para comprender su percepción sobre la enseñanza de R y Python en sociología, lo que sería un enfoque complementario útil para estudios futuros. Además, la ausencia de datos longitudinales impide evaluar si la adopción de estos lenguajes ha cambiado en los últimos años, lo que representa una oportunidad para investigaciones posteriores. Otro aspecto a considerar es la falta de información directa sobre la percepción de docentes y estudiantes respecto a la integración de R y Python en sus programas académicos. La inclusión de entrevistas cualitativas o encuestas en futuras investigaciones podría proporcionar un panorama más detallado sobre las barreras percibidas y las estrategias más adecuadas para su implementación en la enseñanza de la sociología.

6. Conclusiones

Los resultados de este estudio confirman que la formación en herramientas de programación y análisis de datos en los grados de Sociología en España es limitada y está desactualizada en comparación con las demandas actuales del mercado laboral y de la investigación social basada en datos. El predominio de SPSS, frente a la escasa presencia de lenguajes de programación más versátiles como R y la total ausencia de Python, revela una brecha entre los contenidos que se enseñan y las competencias que los futuros sociólogos necesitarán para afrontar los desafíos de una sociedad cada vez más orientada al *big data* y al análisis cuantitativo. Como se señaló al principio, la integración de lenguajes de programación en la enseñanza sociológica es fundamental para dotar al estudiantado de habilidades técnicas avanzadas que faciliten su adaptación a un entorno profesional en constante cambio.

En relación con los objetivos planteados, este estudio ha puesto de manifiesto la necesidad de revisar y actualizar los planes de estudio de Sociología en las universidades españolas, de forma que se incorpore la enseñanza de lenguajes de programación desde las etapas iniciales de formación, especialmente en las asignaturas básicas de estadística. Sin embargo, se observa que estas asignaturas rara vez incluyen programación, lo que representa una oportunidad perdida para familiarizar a los estudiantes con herramientas de análisis de datos desde el inicio de su formación. Integrar la programación en estas asignaturas permitiría un aprendizaje progresivo y fortalecería las competencias cuantitativas necesarias para enfrentar asignaturas avanzadas en etapas posteriores.

Estos hallazgos corroboran la visión expresada en estudios previos sobre la importancia de que las asignaturas de estadística y métodos cuantitativos no se limiten al uso de herramientas tradicionales, sino que incluyan opciones *open source* y lenguajes de programación que permitan a los estudiantes desarrollar una mayor flexibilidad y adaptabilidad en el análisis de datos.

Además, la distribución de las asignaturas técnicas a lo largo del grado muestra que el enfoque práctico se concentra en los cursos intermedios, dejando los primeros años sin una formación técnica sólida y limitando la consolidación de habilidades cuantitativas en el último año. Como sugieren diversos autores, una formación progresiva que combine teoría y práctica es clave para desarrollar competencias técnicas robustas. La falta de asignaturas avanzadas y prácticas, así como la predominancia de asignaturas teóricas, limita la capacidad de los estudiantes para aplicar sus conocimientos en entornos reales y complejos.

Para facilitar la adopción de R y Python en la enseñanza de la sociología, proponemos una estrategia progresiva en tres etapas, de cara a minimizar resistencias y buscar una transición efectiva. Primero, en las asignaturas introductorias de estadística y metodología, donde se puede incorporar R y Python mediante ejercicios guiados que complementen el uso de *software* tradicional. Segundo, en asignaturas intermedias de análisis de datos, estos lenguajes pueden utilizarse en conjunto con SPSS, mostrando la resolución de ejercicios con distintas herramientas, mostrando así las diferencias entre el *software* tradicional y los lenguajes de programación y permitiendo una transición gradual. Finalmente, en asignaturas avanzadas y en los trabajos de fin de grado, R y Python pueden emplearse de manera autónoma, promoviendo proyectos aplicados que preparen a los estudiantes para el mercado laboral. Este enfoque escalonado facilitaría la adaptación tanto de estudiantes como de docentes, minimizando la resistencia al cambio y permitiendo una transición gradual sin necesidad de reemplazar por completo las herramientas actualmente en uso.

Por tanto, este estudio subraya la urgencia de una transformación curricular que responda a los retos de una sociología moderna, basada en datos y técnicas computacionales.

La actualización de los planes de estudio, incorporando lenguajes de programación como R y Python junto con una mayor diversidad de herramientas de análisis, garantizará que los futuros sociólogos estén mejor preparados para enfrentar los desafíos del análisis de datos en la era digital y puedan contribuir con metodologías más avanzadas en la investigación social.

Referencias bibliográficas

- Boccardo, Giorgio y Ruiz, Felipe (2019). *Desafíos de la enseñanza de R en contextos hispanoparlantes*. LatinR 2019 Conference. DOI: <https://doi.org/10.13140/RG.2.2.10887.96160>.
- Elosua, Paula (2009): “¿Existe vida más allá del SPSS? Descubre R”. *Psicothema*, 21 (4), 652-655.
- Fernández Lizana, Miguel Ignacio (2020): “Ventajas de R como herramienta para el análisis y visualización de datos en ciencias sociales”. *Revista Científica de la UCSA*, 7 (2), 97-111.
- González, Felipe (2019): “Big data, algoritmos y política: Las ciencias sociales en la era de las redes digitales”. *Cinta de moebio*, 65, 267-280. DOI: <http://dx.doi.org/10.4067/s0717-554x2019000200267>.
- Gualda, Estrella (4 de junio de 2021). *Sociología Computacional y Social Big Data: Desafíos teóricos y metodológicos*. Presentación y materiales en el Máster de Sociología Aplicada, Universidad de Málaga. <http://rabida.uhu.es/dspace/handle/10272/19925>.
- Gualda, Estrella (2022): “Social Big Data, sociología y ciencias sociales computacionales”. *Empiria: Revista de Metodología de Ciencias Sociales*, 53, 147-177. DOI: <https://doi.org/10.5944/empiria.53.2022.32631>.
- Gualda, Estrella y Rebollo Díaz, Carolina (2020): “Big data y Twitter para el estudio de procesos migratorios: métodos, técnicas de investigación y software”. *Empiria: Revista de Metodología de Ciencias Sociales*, 46, 147-177.
- Healy, Kieran (2018). *Data Visualization: A Practical Introduction*. Nueva Jersey: Princeton University Press.
- Maestro Cano, Ignacio C. (2019): “Sobre las limitaciones de big data en ciencias sociales”. *Sociología y Tecnología*, 9 (2), 77-98.

- Muenchen, Robert A. (2012): “The Popularity of Data Science Software”. *r4stats.com*. <https://r4stats.com/articles/popularity/>.
- Salganik, Matthew J. (2019). *Bit By Bit: Social Research in the Digital Age*. Nueva Jersey: Princeton University Press.
- Taboada, Alba (2024): “*Big data* en ciencias sociales. Una introducción a la automatización de análisis de datos de texto mediante procesamiento de lenguaje natural y aprendizaje automático”. *Revista CEN-TRA de Ciencias Sociales*, 3 (1). DOI: <https://doi.org/10.54790/rccs.51>.

Nota biográfica

Javier Cantón-Correa es Docente e investigador en UNIR, UGR, UOC y UI1. Es Doctor en Ciencias Sociales, Máster en Inteligencia Artificial y en Problemas Sociales y Licenciado en Sociología y Comunicación Audiovisual. Experto en redes sociales, análisis y visualización de datos, ha sido periodista en RTVE.es. Actualmente trabaja sobre desinformación e IA e imparte clases en diversos Másters y Grados. ORCID: <https://orcid.org/0000-0002-8466-1679>.