



Guía docente

Guía docente					
Datos Identificativos				2022/23	
Asignatura (*)	Big Data para las Ciencias Sociales			Código	615545008
Titulación	Máster Universitario en Métodos Avanzados de Investigación e Innovación na Análise Social				
Descriptores					
Ciclo	Periodo	Curso	Tipo		Créditos
Máster Oficial	2º cuatrimestre	Primero	Obligatoria		6
Idioma	Castellano				
Modalidad docente	Presencial				
Prerrequisitos					
Departamento	Socioloxía e Ciencias da Comunicación				
Coordinador/a	León Medina, Francisco J.	Correo electrónico	f.leon@udc.es		
Profesorado	León Medina, Francisco J.	Correo electrónico	f.leon@udc.es		
Web					
Descripción general	Las tecnologías digitales han revolucionado el modo en que vivimos y el modo en que obtenemos y analizamos datos sobre el comportamiento social. Por un lado, la creciente capacidad computacional ha hecho posible la elaboración y testado de modelos de procesos sociales que hasta el momento eran matemáticamente imposibles. Por otro, el acceso a los registros de la huella digital que deja nuestro comportamiento online ha supuesto la posibilidad de trabajar con bases de datos masivos, comportamentales, no-reactivos, relacionales y dinámicos. Ambas posibilidades han confluído en la llamada ciencia social computacional, que a su vez entronca con una larga tradición empírica, analítica, computacional y generativista que ha sido troncal en aquella ciencia social más orientada a la explicación científica y rigurosa de los fenómenos sociales. En esta asignatura, se introducirá al alumnado a los principales debates epistemológicos, teóricos y éticos de la ciencia social computacional y, especialmente, aquellos que se producen alrededor del uso del Big Data para la investigación social. Además, ofrecerá las herramientas y conocimientos básicos para diseñar y ejecutar una investigación basada en datos masivos: se introducirá al alumno a las técnicas y procedimientos orientados a la obtención o producción, gestión y análisis de bases de datos masivos, así como a las herramientas más novedosas para la presentación y visualización de resultados.				

Competencias del título

Código	Competencias del título
A1	C1. Conocer al detalle los principales avances teóricos y prácticos en materia y técnicas de investigación cualitativa y cuantitativa y su aplicación - crítica, reflexiva y contextual - para comprender, definir y explicar fenómenos de alta complejidad e incertidumbre asociados a los retos sociales, económicos y políticos actuales.
A6	C6. Manejar los fundamentos teóricos y científicos necesarios para llevar a cabo una investigación social de tipo cuantitativo y cualitativo empleando herramientas tecnológicas avanzadas (R).
A7	C7. Reconocer las distintas opciones de presentación y visualización de resultados de investigación y su fundamentación metodológica para transmitir de forma clara los resultados de la investigación a los distintos públicos.
A8	C8. Conocer los modos y potencialidades de uso de datos masivos y de minería de datos en el contexto de las ciencias sociales.
A11	H1. Evaluar y seleccionar el paradigma metodológico adecuado (cuantitativa, cualitativa, mixta, prospectiva, participativa) para formular juicios a partir de información incompleta o limitada sobre fenómenos asociados a los retos sociales, económicos y políticos.
A14	H4 Transmitir los resultados de una investigación básica y aplicada, así como los fundamentos más relevantes sobre los que se sustentan, de modo claro y sin ambigüedades atendiendo a la diversidad de contextos, públicos y necesidades, usando distintas opciones de presentación y visualización.
A15	H5 Llevar a cabo una investigación de tipo cualitativo y cuantitativo en R.
A16	H6 Desarrollar capacidad para planificar la investigación, eligiendo correctamente las fases y los procedimientos que garantizan la máxima rigurosidad y que sean adecuados a distintos contextos dentro del ámbito público o privado.
A17	H7 Adquirir la capacidad de continuar profundizando en los métodos de investigación y análisis de un modo que habrá de ser en gran medida auto-dirigido y autónomo.
A19	H9 Desarrollar una capacidad avanzada para evaluar la calidad y los resultados de los procesos de investigación.



A20	CP1 Contrastar la importancia de la diversidad metodológica existente para el análisis de la realidad social, política y económica, abordando distintos problemas complejos de las sociedades actuales a través de soluciones creativas y novedosas
A22	CP3. Elaborar e implementar proyectos de investigación en entornos nuevos o poco conocidos.
A24	CP5. Ser capaz de trabajar con fuentes de datos, metodologías y técnicas de investigación científica y herramientas informáticas avanzadas propias de las ciencias sociales.
A27	CP8. Diseñar y ejecutar un proyecto de investigación basado en el uso de Big Data con propósitos científicos, realizando aportaciones originales al campo de estudio.
A28	CP9. Obtener, gestionar y analizar bases de datos masivos empleando las herramientas informáticas y lenguajes de programación necesarios para ello, especialmente el lenguaje avanzado R.

Resultados de aprendizaje			
Resultados de aprendizaje		Competencias del título	
El alumnado será capaz de diseñar y ejecutar un proyecto de investigación basado en el uso de Big Data con propósitos científicos, realizando aportaciones originales y siendo capaz de resolver problemas teóricos y prácticos empleando estrategias innovadoras.		AM1	
		AM6	
		AM7	
		AM8	
		AM11	
		AM15	
		AM16	
		AM17	
		AM20	
		AM22	
		AM24	
		AM27	
		AM28	
El alumnado será capaz de obtener, gestionar y analizar bases de datos masivos empleando las herramientas informáticas y lenguajes de programación necesarios para ello, especialmente el lenguaje avanzado R, así como las técnicas estadísticas apropiadas para su análisis.		AM8	
		AM24	
		AM28	
El alumnado será capaz de evaluar de forma crítica la calidad del conocimiento científico producido a través del análisis de datos masivos y de reflexionar sobre su importancia e impacto socioeconómico.		AM19	
		AM20	
El alumnado será capaz de emplear las herramientas más avanzadas para la visualización de datos y para la presentación de los resultados de investigación.		AM7	
		AM14	
		AM24	

Contenidos	
Tema	Subtema
TEMA 1 : Fundamentos de la ciencia social computacional	Discusiones teóricas, metodológicas y éticas alrededor del Big Data
TEMA 2: Obtención, gestión y almacenamiento de datos masivos.	Tipos y arquitecturas de datos masivos. Obtención de datos masivos: web scrapping, APIs, etc.
TEMA 3: Procesamiento y análisis de datos masivos	Análisis de sentimiento, análisis de los social media, social network analysis, machine learning, etc.
TEMA 4: Herramientas avanzadas de visualización de datos masivos	R (ggplot2), Gephi, etc.

Planificación				
Metodologías / pruebas	Competencias	Horas presenciales	Horas no presenciales / trabajo autónomo	Horas totales



Portafolio del alumno	A6 A7 A11 A14 A15 A16 A17 A19 A20 A22 A24 A27 A28	0	53	53
Taller	A1 A6 A7 A8 A11 A14 A15 A16 A17 A19 A20 A22 A24 A27 A28	21	21	42
Presentación oral	A7 A14	8	24	32
Sesión magistral	A1 A6 A8 A11 A16 A17 A19 A20	21	0	21
Atención personalizada		2	0	2
(*)Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos				

Metodologías	
Metodologías	Descripción
Portafolio del alumno	Los distintos ejercicios que se irán elaborando en las sesiones de "Taller", así como los ejercicios propuestos para realizarse de manera autónoma, se irán incorporando a una "carpeta del estudiante" que se entregará a final de curso y que será objeto de evaluación.
Taller	A lo largo del curso se realizarán distintas sesiones prácticas en las que se realizarán ejercicios donde el alumnado aprenderá los procedimientos básicos para la obtención, estructuración, almacenamiento, análisis y visualización de datos masivos.
Presentación oral	Al final del curso se dedicará alguna sesión a la presentación de algunos de los ejercicios que forman parte del portafolio del estudiante, especialmente los orientados a la visualización y presentación de resultados.
Sesión magistral	Se realizarán sesiones de clase magistral en las que el peso de la actividad recae fundamentalmente sobre el profesor. Este método pedagógico se complementa con la introducción de otras actividades y métodos que permiten un rol más activo de los estudiantes. Dado el contenido eminentemente orientado a la práctica investigadora de esta materia, las sesiones magistrales frecuentemente tendrán un contenido técnico, orientado al aprendizaje de los procedimientos de investigación basada en Big Data más habituales. Se complementará esta orientación técnica con una reflexión sobre los fundamentos epistemológicos, éticos y teóricos de este tipo de investigación.

Atención personalizada	
Metodologías	Descripción
Sesión magistral Portafolio del alumno Taller Presentación oral	La atención personalizada de esta asignatura se desarrollará de forma presencial y no presencial. Para la comprensión de las lecturas y la correcta elaboración de los ejercicios prácticos, los alumnos contarán con la presencia del profesor en el aula durante las sesiones de clase y, además, podrán programar las tutorías presenciales que consideren necesarias en el horario que se comunicará al inicio de curso. El profesor también estará disponible de forma no presencial, preferiblemente a través del correo electrónico (y también a través del Campus Virtual).

Evaluación			
Metodologías	Competencias	Descripción	Calificación
Portafolio del alumno	A6 A7 A11 A14 A15 A16 A17 A19 A20 A22 A24 A27 A28	A lo largo del curso se realizarán distintas sesiones prácticas en las que se realizarán ejercicios donde el alumnado aprenderá los procedimientos básicos para la obtención, estructuración, almacenamiento, análisis y visualización de datos masivos. Los distintos ejercicios que se irán elaborando en estas sesiones, así como los ejercicios propuestos para realizarse de manera autónoma, se irán incorporando a una "carpeta del estudiante" que se entregará a final de curso y que será objeto de evaluación.	90



Presentación oral	A7 A14	Al final del curso se dedicará alguna sesión a la presentación de algunos de los ejercicios que forman parte del portafolio del estudiante, especialmente lo orientados a la visualización y presentación de resultados.	10
-------------------	--------	--	----

Observaciones evaluación

Fuentes de información

Básica	- Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. Science, 346(6213), 1063-1064. - Lazer D, et al. (2009). Computational social science. Science, 323(5915):721?723.. - Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt - Marr, B. (2016). Big data in practice: how 45 successful companies used big data analytics to deliver extraordinary results. John Wiley & Sons - Salganik, M. J. (2019). Bit by bit: Social research in the digital age. Princeton University Press - Spaiser, V. (2021). Digital data and methods. In Research Handbook on Analytical Sociology (pp. 352-363).. Edward Elgar Publishing - Flache, A., Mäs, M., & Keijzer, M. A. (2022). Computational approaches in rigorous sociology: agent-based computational modeling and computational social science. In Handbook of Sociological Science (pp. 57-72).. Edward Elgar Publishing
Complementaria	

Recomendaciones

Asignaturas que se recomienda haber cursado previamente

Análisis Estadístico Descriptivo e Inferencial/615545001

Fundamentos del Análisis Social en R/615545003

Análisis Estadístico Avanzado/615545007

Asignaturas que se recomienda cursar simultáneamente

Asignaturas que continúan el temario

Otros comentarios

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías