



Universidad
de Alcalá

GUÍA DOCENTE

**Ciencia de Datos práctica:
resolución de problemas
ambientales mediante proyectos**

Asignatura Transversal

Universidad de Alcalá

Curso Académico 2021/22

GUÍA DOCENTE

Nombre de la asignatura: Ciencia de Datos práctica: resolución de problemas ambientales mediante proyectos

Código: 100246

Titulación en la que se imparte: Grados en Ciencias Biológicas, Ambientales, Grado en Criminalística: Ciencias y Tecnologías Forenses, Grado en Química, Grado en Biología Sanitaria

Departamento y Área de Conocimiento: Ciencias de la Vida (U.D. Ecología)

Carácter: Transversal

Créditos ECTS: 6

Curso: Segundo

Periodo: 2º Cuatrimestre

Profesorado: Ignacio Morales Castilla (Coordinador)
Paloma Ruiz Benito

Horario de Tutoría: Las tutorías se solicitarán mediante cita previa por e-mail o teléfono con el profesor

Idioma en el que se imparte: Español

1. PRESENTACIÓN

El aumento sin precedentes de la disponibilidad de datos derivado de la digitalización, hacen de la ciencia de datos y el “Big Data” sean herramientas imprescindibles para anticipar y enfrentar los desafíos del siglo XXI. Estas herramientas, pueden ser particularmente útiles en el ámbito de las ciencias naturales, ya que permiten solucionar problemas que eran inabordables hasta hace poco. Debido a la creciente demanda en la capacitación en estas disciplinas por parte de empresas e instituciones, el empleo de estas técnicas en el uso de datos es de especial interés las titulaciones de ciencias no técnicas.

Para proporcionar estas competencias altamente demandadas en carreras de ciencias se propone esta asignatura transversal de 6 ECTS que incluyen clases eminentemente basadas en proyectos, desarrollando las habilidades y competencias transversales de los estudiantes. Cabe **destacar el carácter práctico y aplicado de la asignatura**, orientada a que el alumno adquiera los conocimientos, destrezas y habilidades necesarias para **obtener, manejar, transformar, y analizar grandes volúmenes de datos, así como para visualizar y difundir la síntesis de resultados generados**. La adquisición de dichas aptitudes se vehicula mediante la resolución de problemas del ámbito académico y empresarial, transversales a carreras de la rama de Ciencias. Para ello se aprenderán nociones básicas de programación en R para el manejo, tratamiento y análisis de bases de datos, así como para la elaboración de figuras y mapas. En definitiva, esta asignatura fomentará que el alumno se familiarice y esté cómodo utilizando entornos de programación que son cada vez menos del dominio exclusivo de titulaciones técnicas.

Pre-requisitos y recomendaciones:

- Es necesario tener conocimientos básicos de informática y altamente recomendable tener nociones del lenguaje de programación *R*, que será utilizado en las prácticas y seminarios de la asignatura.
- Es altamente recomendable haber cursado y aprobado la asignatura *Estadística*, ya que una parte considerable de la asignatura se fundamenta en la aplicación de análisis estadísticos avanzados a conjuntos de datos.
- Se recomienda a los alumnos que cursen la asignatura por primera vez que lo hagan de forma presencial, ya que el carácter práctico de la asignatura hace que sea mucho más difícil adquirir las capacidades y competencias sin asistir a clase.

2. COMPETENCIAS

Competencias genéricas:

1. Desarrollo de habilidades para la **obtención de datos masivos**.

2. Adquisición de técnicas de **manejo, tratamiento y transformación de bases de datos masivos** mediante técnicas básicas de programación.
3. Capacidad de **análisis y síntesis de la información** contenida en las bases de datos utilizadas de forma crítica.
4. Aprendizaje de técnicas de visualización y presentación de la síntesis de información.
5. Desarrollar la capacidad de **trabajar en equipos multidisciplinares orientados a la resolución de problemas comunes** y también como generadores de sinergias que permitan diseñar soluciones originales y creativas.

Competencias específicas:

1. Adquisición de datos ambientales mediante búsqueda de información en los principales buscadores y repositorios de datos relacionados con las Ciencias Naturales, así como nociones de herramientas básicas de minado de texto y uso de APIs.
2. Analizar e integrar con precisión y rigor información sobre los problemas ambientales aplicando herramientas estadísticas, matemáticas y de análisis territorial.
3. Aprendizaje de rutinas de automatización de procesos analíticos mediante la aplicación de conceptos centrales de programación: diseño y explotación de bucles, uso de puertas lógicas, etc.
4. Conocer las tendencias actuales en ciencia de datos aplicado a las Ciencias Naturales.

Resultados del aprendizaje:

1. Capacidad de identificación de datos necesarios para la resolución de problemas aplicados comunes en distintas disciplinas de las Ciencias Naturales.
2. Capacidad autónoma para la búsqueda de información, selección y adquisición de datos de diversa naturaleza (p.ej. datos de medio ambiente, biomedicina, ecología, etc.).
3. Capacidad autónoma de analizar y sintetizar datos para obtener solución a los problemas planteados.
4. Capacidad para resolver problemas aplicados en el ámbito de las Ciencias Naturales usando nociones de Ciencia de Datos comunes a los problemas ambientales.
5. Aplicar técnicas programáticas al análisis de datos permitiendo automatizar y acelerar procesos.
6. Desarrollo de habilidades para la visualización, presentación y difusión de resultados en medios y repositorios digitales.

3. CONTENIDOS

Bloque	Contenido	Total de horas
Bloque I. Introducción a la programación para la solución de problemas ambientales.	Entornos de programación. Introducción a R y R studio. Programación aplicada a las Ciencias Naturales	2
Bloque 2. Programación básica.	Tipos de variables y datos. Operaciones aritméticas y lógicas. Creación de vectores, matrices, listas y tablas de datos. Selección y gestión de datos. Recomendaciones de programación. Directorios de trabajo. Librerías. Estructura. Funciones. Estructuras de programación.	6
Bloque 3. Trabajo reproducible: desde programación colaborativa a control de versiones.	Trabajo reproducible. Git y Github. Rmarkdown	2
Bloque 5. Obtención y tratamiento aplicado a la Ciencia de Datos	Análisis exploratorio de datos y generación de gráficos enfocados a la resolución de problemas ambientales.	5
Bloque 6. Visualización de datos	Análisis exploratorio de datos y generación de gráficos para la resolución de problemas ambientales. R como gis. Mapas estáticos y dinámicos.	5
Bloque 7. Difusión de resultados	Publicación y difusión de información	2
VII. Seminarios	Cada bloque se acompañará con dos horas de seminarios donde se expandirán y practicarán los contenidos mediante resolución de problemas, ejemplos clave y ejemplos prácticos.	7 h presenciales
VIII. Prácticas	Resolución de problemas reales desde la generación de un equipo de trabajo colaborativo y reproducible, hasta la exploración y análisis de datos con diferentes técnicas para la solución del problema propuesto.	17 h presenciales

4. METODOLOGÍAS DE ENSEÑANZA-APRENDIZAJE- ACTIVIDADES FORMATIVAS

4.1. Distribución de créditos (especificar en horas)

Número de horas presenciales: 46	Clases teóricas basadas en resolución de casos prácticos (22 h). Seminarios con resolución de problemas ambientales (7 h) y clases prácticas (17 h).
Número de horas del trabajo propio del estudiante: 100	Preparación individual o en grupo de temas de la asignatura: 40 horas Lectura comprensiva de trabajos científicos y de guías metodológicas: 20 horas Elaboración de trabajos de forma escrita y oral: 40 horas
Tutorías en grupos reducidos: 4h	4 h
Total horas: 150	

4.2. Estrategias metodológicas, materiales y recursos didácticos

Estrategias metodológicas	<p>Exposición de contenidos y resolución de problemas: Presentación de contenidos de la asignatura, ayudando con material gráfico y/o escrito, realizando la resolución de problemas propuestos por el profesorado y discusión de casos y ejemplos reales.</p> <p>Basado en proyectos: El trabajo final y uno de los ejercicios prácticos se basa en el desarrollo de un supuesto práctico.</p>
Material	<p>Referencias bibliográficas: se facilitarán referencias que pueden consultarse en internet o la biblioteca.</p> <p>Software y datos necesarios para el desarrollo de la asignatura, generalmente abiertos y de libre acceso (R, Rstudio, Github y Git, datos espaciales abiertos).</p> <p>Aula virtual.</p>

5. EVALUACIÓN: Procedimientos, criterios de evaluación y de calificación¹

Esta sección muestra la relación entre criterios, instrumentos y calificación tanto en convocatoria ordinaria como en convocatoria extraordinaria.

Convocatoria ordinaria

(a) Modalidad evaluación continua

Todos los alumnos serán evaluados por un sistema de evaluación continua a excepción de aquéllos que, previa solicitud por escrito al coordinador de la asignatura durante las dos primeras semanas de impartición de la asignatura, hayan sido autorizados para acogerse a la evaluación final. La evaluación continua implica la obligatoriedad de asistir a un mínimo del 75% de las clases presenciales para poder aprobar la asignatura.

Se considerará que los alumnos han agotado la convocatoria correspondiente cuando concurren en una o más de las siguientes situaciones:

- * No asistan al mínimo de clases establecido.
- * No participen en las actividades programadas sin causa justificada.
- * No entreguen los trabajos en los plazos establecidos sin causa justificada.

En caso de no superar la evaluación continua o figurar como “no presentado”, los alumnos tienen derecho a optar a la convocatoria extraordinaria.

¹ *Es importante señalar los procedimientos de evaluación: por ejemplo evaluación continua, final, autoevaluación, co-evaluación. Instrumentos y evidencias: trabajos, actividades. Criterios o indicadores que se van a valorar en relación a las competencias: dominio de conocimientos conceptuales, aplicación, transferencia conocimientos. Para el sistema de calificación hay que recordar la **Normativa del Consejo de Gobierno del 16 de Julio de 2009**: la calificación de la evaluación continua representará, **al menos, el 60%**. Se puede elevar este % en la guía.*

Procedimiento de evaluación

Procedimientos de evaluación continua	Criterios calificación (peso)	Criterios de evaluación (grado de consecución de las competencias)
Memorias de prácticas y seminarios	60 %	<ul style="list-style-type: none">• Se conocen los principales conceptos y estructuras de programación, aplicando técnicas para el trabajo reproducible y colaborativo.• Se documentan adecuadamente los procedimientos de adquisición, manejo, análisis y publicación de información mediante lenguajes de programación.• Se aplican los conceptos y estructuras de programación en el ámbito de ejercicios para la resolución de problemas ambientales.
Trabajo final	40 %	<ul style="list-style-type: none">• Se aplican conceptos y estructuras de programación para la resolución de problemas ambientales.• Se integran los conocimientos adquiridos para formular soluciones y comunicar las conclusiones de trabajo.• Se demuestran habilidades de aprendizaje para estudiar autodirigida y autónomamente en la resolución del problema planteado.

En caso de no superar la evaluación continua, los alumnos tienen derecho a optar a un examen final en la convocatoria extraordinaria.

Los estudiantes no deben realizar prácticas de copia o plagio, ya sea en los trabajos encomendados por el profesorado o en las pruebas finales. En el caso de realizarlas serán suspendidos pudiéndose estudiar posibles acciones disciplinarias.

(b) Modalidad evaluación final

En la modalidad de evaluación final se realizará un examen práctico que constará de información desarrollada por el estudiante similar al trabajo final con contenidos teóricos, otra sobre los contenidos prácticos y seminarios. La parte del ejercicio final

será de un 40% y los ejercicios de evaluación del material de las prácticas y seminarios del 60%, ponderándose para la nota final. Para la ponderación en cada una de las partes debe obtenerse un mínimo de 4.

Convocatoria extraordinaria

(a) Modalidad evaluación continua

Se realizará un ejercicio escrito consistente en supuestos prácticos relacionados con los contenidos desarrollados en la asignatura y la entrega de las asignaciones requeridas a los estudiantes de evaluación continua durante el desarrollo del curso. La parte del ejercicio final será de un 40% y los ejercicios de evaluación del material de las prácticas y seminarios del 60%, ponderándose para la nota final. Para la ponderación en cada una de las partes debe obtenerse un mínimo de 4.

(b) Modalidad evaluación final

Se mantendrá el mismo tipo de evaluación que en la convocatoria ordinaria de examen final.

6. BIBLIOGRAFÍA

Bernabé-Poveda MA, López-Vázquez CM (2012) *Fundamentos de las Infraestructuras de datos espaciales (IDE)*. Universidad Politécnica de Madrid.

Bivand RS, Pebesma E, Gómez-Rubio V (2013) *Applied Spatial Data Analysis with R*. Springer.

Brudson C, Comber L (2015) *An Introduction to R for spatial analysis and mapping*. SAGE publications.

Crawley MJ (2007) *The R Book*. Chichester, UK, John Wiley & Sons Ltd.

Iniesto M, Núñez A (2014) *Introducción a las Infraestructuras de Datos Espaciales*. Centro Nacional de Información Geográfica (CNIG) Dirección General del Instituto Geográfico Nacional (IGN).

Lacovella S (2014) *GeoServer CookBook*. Packt publishing.

Matloff N (2011) *The art of R programming: a tour of statistical software design*. No starch press.

Wickham, H., & Grolemund, G. (2016). *R for data science: import, tidy, transform, visualize, and model data*. " O'Reilly Media, Inc."

La Universidad de Alcalá garantiza a sus estudiantes que, si por exigencias sanitarias las autoridades competentes impidieran la presencialidad total o

parcial de la actividad docente, los planes docentes alcanzarían sus objetivos a través de una metodología de enseñanza-aprendizaje y evaluación en formato online, que retornaría a la modalidad presencial en cuanto cesaran dichos impedimentos.