

PROGRAMA DE CURSO

CIENCIA DE DATOS GEOGRÁFICOS

A. Antecedentes generales del curso:

Departamento	Ciencias de la Computación					
Nombre del curso	Ciencia de datos geográficos	Código	CC5216	Créditos	6	
Nombre del curso en inglés	Geographic Data Science					
Horas semanales	Docencia	3,0	Auxiliares	0	Trabajo personal	7,0
Carácter del curso	Obligatorio			Electivo	X	
Requisitos	CC4102: Diseño y Análisis de Algoritmos					

B. Propósito del curso:

La Ciencia de datos geográficos aplica técnicas de ciencias de datos (aprendizaje automático, minería de datos, visualización) en problemas que tienen una fuerte componente espacial, particularmente en contextos urbanos y geográficos. Estos problemas suelen provenir de necesidades en ciudades para todo tipo de plazo y aplicación, vale decir, corto, mediano o largo plazo, y en aplicaciones de gestión o planificación urbana. Como tal, son problemas que tienen directa incidencia en la calidad de vida de las personas, y cuyas partes interesadas provienen de otras disciplinas, como transporte, arquitectura, derecho, administración pública, logística, sustentabilidad, entre otras. Al ser problemas geográficos, es necesario considerar métodos y algoritmos que estén al tanto de la componente espacial (en inglés, que sean “spatially aware”), dado que se rompen los supuestos de los métodos tradicionales (donde las observaciones son independientes entre sí). Además, por tratarse de problemas de aplicabilidad urbana, es necesario comunicar, gestionar y transferir la solución o herramienta propuesta basada en datos a las partes interesadas.

Este curso estudia los conceptos, las herramientas y el proceso de un proyecto de ciencia de datos, desde su concepción hasta la transferencia y adopción de la solución propuesta. En él, los y las estudiantes aprenderán a:

- desarrollar habilidades para identificar problemas y situaciones urbanas de relevancia social que tienen una dimensión espacial.
- explicar los desafíos científicos que se enfrentan cuando se analizan datos espaciales y geográficos.

- aplicar un abanico de técnicas y enfoques analíticos adecuados para el análisis de datos espaciales y geográficos.
- desarrollar habilidades prácticas en Python y Javascript para implementar un amplio rango de herramientas analíticas espaciales y geográficas para problemas aplicados y de investigación.
- entender cómo la Ciencia de Datos Geográficos puede apoyar la toma de decisiones y diseño de políticas públicas.
- implementar proyectos de Ciencia de Datos Geográficos en problemas reales actuales utilizando fuentes de datos provenientes de sistemas digitales (sensores, registros de telefonía móvil, GPS, etc.).

El curso tributa a las siguientes competencias específicas (CE) y genéricas (CG):

CE1: Analizar problemas computacionales, construir modelos, expresándolos en representaciones y lenguajes formales adecuados.

CE2: Analizar, diseñar y/o adoptar, algoritmos y estructuras de datos que cumplan con las garantías requeridas de correctitud y eficiencia.

CE4: Extraer información relevante, utilizando el proceso de descubrimiento de conocimiento de datos.

CE9: Desarrollar soluciones computacionales de manera interdisciplinaria y colaborativa.

CG1: Comunicación académica y profesional

Comunicar en español de forma estratégica, clara y eficaz, tanto en modalidad oral como escrita, puntos de vista, propuestas de proyectos y resultados de investigación fundamentados, en situaciones de comunicación compleja, en ambientes sociales, académicos y profesionales.

CG5: Innovación

Concebir ideas viables y novedosas que generen valor para resolver necesidades latentes, materializadas en productos, servicios o en mejoras a procesos dentro de un sistema u organización, considerando el contexto sociocultural y económico y los beneficios para el usuario.

C. Resultados de aprendizaje:

Competencias específicas	Resultados de aprendizaje
CE1	RA1: Identifica y analiza problemas y situaciones urbanas de relevancia social y que tienen una dimensión espacial, considerando el uso de conceptos y técnicas de la Ciencia de Datos Geográficos.
CE2	RA2: Utiliza conceptos de las Ciencias de la Computación y Ciencia de Datos para explicar un problema urbano en el que la componente espacial es crítica para llegar a una solución que tenga potencial de aplicación en contextos reales.
CE4	RA3: Concibe una solución a un problema geográfico acotado, a través de la integración de múltiples herramientas para ejecutar análisis de datos especiales que den respuesta a la necesidad planteada.
CE9	RA4: Resuelve un problema geográfico complejo de extensión urbana, utilizando un amplio rango de herramientas (específicamente implementadas con Python y Javascript) para proponer una solución concreta y funcional. RA5: implementa un proyecto de Ciencia de Datos Geográficos mediante avances incrementales, utilizando técnicas de análisis y visualización geográfica para dar respuesta a las necesidades y objetivos del proyecto.
Competencias genéricas	Resultados de aprendizaje
CG1	RA6: Explica, de manera clara y coherente, en modalidad oral los desafíos científicos que se enfrentan cuando se analizan datos espaciales y geográficos respecto de un proyecto y sus resultados, considerando las soluciones propuestas, limitaciones y líneas de investigación a seguir.
CG5	RA7: Utiliza aplicaciones novedosas como fuentes de datos provenientes de sistemas digitales (sensores, registros de telefonía móvil, GPS, etc.) y técnicas de análisis espacio-temporal y clasificación espacial para recoger información con la cual resolver una necesidad, considerando las ventajas competitivas de las soluciones a proponer.

D. Unidades temáticas:

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
1	RA1, RA2, RA7	Fundamentos de Ciencia de Datos Geográficos	3 semanas
Contenidos		Indicador de logro	
1.1. Fundamentos de Ciencia de Datos. 1.2. Introducción a la Geografía. 1.3. Aplicaciones innovadoras de Ciencia de Datos Geográficos en problemas urbanos. 1.4. Fuentes de datos y trazas digitales. 1.5. La componente espacial. 1.6. Ecosistema de trabajo en Python. 1.7. Ecosistema de trabajo con tecnologías Web. 1.8. Estrategias de definición de contexto y problemática a resolver en un proyecto.		El/la estudiante: - Identifica las etapas generales a ejecutar en el desarrollo de sus proyectos de Ciencia de Datos Geográficos, desde el planteamiento del problema hasta la evaluación de la solución, con énfasis en procesos innovadores. - Elabora la versión inicial de un proyecto de ciencia de datos en el contexto de problemáticas urbanas o geográficas, considerando el tipo de problema a analizar. - Examina múltiples fuentes de información geográfica para determinar necesidades latentes, considerando beneficios para la sociedad y creación de espacios de ejecución. - Establece relaciones relevantes entre lo leído respecto del uso de datos para problemáticas urbanas (refs. [1,5]) sobre las cuales investigar. - Reconoce los aspectos específicos de la problemática abordada y los desafíos presentes desde el punto de vista de análisis, organización y procesamiento de datos con componente espacial. - Reconoce, como parte de su reflexión, los logros esperados y las dificultades que deberá abordar a lo largo del desarrollo del curso. - Intercambia con sus pares, profesores, otros profesionales y actores relevantes conocimientos en el ámbito de la Ciencia de Datos Geográficos en la FCFM, ideas sobre diferentes desafíos profesionales en torno a los proyectos que trabajará.	
Bibliografía de la unidad		[1], [5]	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
2	RA3, RA4, RA5, RA6	Modelamiento Espacial y Aprendizaje de Máquinas	5 semanas
Contenidos		Indicador de logro	
2.1. Pesos espaciales. 2.2. Análisis exploratorio de datos geográficos. 2.3. Autocorrelación espacial. 2.4. Econometría y regresión espacial. 2.5. Clustering espacial. 2.6. Análisis espacio-temporal. 2.7. Clasificación espacial. 2.8. Visualización espacial en Python. 2.9. Visualización espacial en Javascript.		El/la estudiante: 1. Implementa algoritmos y técnicas de análisis exploratorio de datos geográficos para diagnosticar el problema de carácter urbano o geográfico. 2. Utiliza herramientas recientes y/o fuentes de datos no tradicionales para proponer alternativas de solución al problema detectado desde el paradigma del modelamiento espacial y aprendizaje de máquinas. 3. Implementa técnicas de econometría y regresión espacial aplicados a distintas fuentes de datos (tanto tradicionales como trazas digitales provenientes de sensores u otros sistemas digitales). 4. Aplica técnicas de análisis espacio-temporal y clasificación espacial a distintas fuentes de datos (tanto tradicionales como trazas digitales). 5. Evalúa las diferencias entre resultados de técnicas tradicionales de aprendizaje de máquinas (que no consideran la componente espacial) y las técnicas que sí realizan análisis espacial. 6. Comunica de manera efectiva y eficiente los resultados de la exploración del espacio de soluciones al problema a resolver, considerando herramientas y técnicas de visualización de información.	
Bibliografía de la unidad		[2], [4], [6]	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
3	RA3, RA4, RA5, RA6	Redes y flujos urbanos	4 semanas
Contenidos		Indicador de logro	
3.1. Redes urbanas. 3.2. Flujos Origen-Destino. 3.3. Análisis de accesibilidad. 3.4. Visualización de flujos y redes en Python. 3.5. Visualización de flujos y redes en Javascript.		El/la estudiante: 1. Implementa algoritmos y técnicas de construcción y análisis de redes urbanas, utilizando grafos dirigidos y no dirigidos. 2. Implementa técnicas de análisis de flujos origen-destino en contextos urbanos (por ej., movilidad, logística). 3. Implementa técnicas de análisis de datos para problemas de accesibilidad y equidad urbana en ciudades. 4. Analiza las posibles soluciones que se han propuesto desde el paradigma de las redes y flujos urbanos para el problema a resolver. 5. Comunica de manera efectiva y eficiente los resultados de la exploración del espacio de soluciones desde el paradigma de las redes para el problema a resolver, considerando herramientas y técnicas de visualización de redes y flujos urbanos.	
Bibliografía de la unidad		[3], [4], [6]	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
4	RA6, RA7	Comunicación y Transferencia de Resultados de GDS	3 semanas
Contenidos		Indicador de logro	
4.1. Estrategias de colaboración con partes interesadas: contrapartes, data-holders, know-how I+D. 4.2. Etapas en la ejecución y adopción de un proyecto de Ciencia de Datos Geográficos.		El/la estudiante: 1. Determina la solución adecuada para el problema a partir de las soluciones previas desde los paradigmas de aprendizaje de máquinas/modelamiento espacial o redes y flujos urbanos. 2. Evalúa la solución elegida, explicando los criterios de aplicación tecnológica para Ciencia de Datos Geográficos usados y criterios multivariantes (correctitud teórica pero también factibilidad de implementación). 3. Desarrolla estrategias de comunicación de datos y resultados adecuadas para la diversidad de partes interesadas en el proceso de transferencia tecnológica de un proyecto de Ciencia de Datos Geográficos (por ej., planificadores/as, arquitectos/as, ciudadanos/as, etc.). 4. Explica el proceso completo dentro del cual se desarrolló el proyecto, considerando las decisiones tomadas, utilizando un registro de habla formal, según las diferentes audiencias a las que se dirige. 5. Reconoce que un proyecto de Ciencia de Datos Geográficos no termina con el análisis y visualización de los datos, sino que requiere fomentar su transferencia y adopción tecnológica a través de un proceso de innovación que considera el contexto socio-cultural y económico de la contraparte beneficiada por la solución. 6. Intercambia con sus pares, profesores, conocimientos en el ámbito de la Ciencia de Datos en la FCFM, e ideas sobre diferentes desafíos profesionales en torno a lo desarrollado como proyecto. 7. Cumple obligaciones y acuerdos, respetando los compromisos adquiridos en sus actividades académicas.	

E. Estrategias de enseñanza-aprendizaje:

La metodología de enseñanza y aprendizaje fomenta la participación de los y las estudiantes en el aula. Las clases son principalmente de dos tipos:

- **Clase expositiva**, en donde los y las estudiantes identifican los problemas fundamentales en PLN así como modelos y técnicas para abordarlos, en base a casos de estudio.
- **Resolución de problemas**: En cada unidad los y las estudiantes son expuestos a herramientas computacionales que permiten implementar los conceptos de las unidades correspondientes relacionados con Ciencia de Datos Geográficos.

La resolución de problemas se da en el contexto de un proyecto aplicado y colaborativo que involucra el uso intensivo de herramientas para el procesamiento, análisis y visualización de datos geográficos provenientes de fuentes tradicionales (como el censo) y fuentes no tradicionales, principalmente trazas digitales provenientes de sensores u otros sistemas tecnológicos (por ej., registros de telefonía móvil). Cada equipo de trabajo de 3 o 4 personas resolverá un problema urbano propuesto por el equipo docente. Cada problemática a resolver se presenta a los (as) estudiantes al inicio del curso, definiendo claramente el problema, los objetivos mínimos del proyecto, los datos pertinentes para la resolución del mismo, y su relación con las unidades del curso. Cada equipo de trabajo contará con el apoyo de un integrante del equipo docente a lo largo del proyecto con el fin de guiar el trabajo a realizar, apoyando en el uso de herramientas y la literatura.

Al finalizar el semestre, cada equipo de trabajo debe realizar una exposición oral final presentando el proceso a lo largo del semestre, las soluciones propuestas, la justificación de por qué son soluciones innovadoras en su conceptualización y en su materialización usando fuentes de datos no tradicionales, cuáles son sus limitaciones y las líneas de investigación a seguir, resultados obtenidos en el proyecto, y una estrategia de adopción de la solución propuesta por parte de una contraparte. Finalmente, incluyendo los comentarios y correcciones resultantes de la presentación oral, el equipo de trabajo debe presentar el informe final escrito que resumen el trabajo realizado y resultados del proyecto.

F. Estrategias de evaluación:

El curso se evalúa a partir de hitos de avance del proyecto que miden tanto el cumplimiento de los objetivos propuestos al comienzo del semestre para los grupos de trabajo, como la aplicación de los contenidos de cada unidad. Cada hito de avance puede corresponder a una presentación oral o a un informe escrito. Al finalizar el semestre, cada equipo deberá entregar un informe escrito con los resultados obtenidos y realizar una presentación oral de sus resultados. La evaluación del informe considerará la estrategia de materialización y adopción de la solución propuesta, y la evaluación de la presentación considerará cómo se justifica el aspecto innovador de cada una. En particular, la innovación se evaluará en la mayoría de los proyectos a través del uso de fuentes no tradicionales de datos provenientes de trazas digitales para resolver un problema urbano o de contexto geográfico, lo que requiere desarrollar una comprensión del problema y probar distintos enfoques para llegar a una solución. En otros casos, la innovación se evaluará a través del uso de métodos previamente no explorados para la resolución del problema, dados los avances recientes en Ciencia de Datos Geográficos.

El cálculo de esas notas se efectúa de la siguiente forma:

NH = Promedio de los hitos de avance.

NP = Nota de presentación oral final.

NI = Nota de informe escrito final.

$$\text{Nota Final} = 0,5 * \text{NH} + 0,25 * \text{NP} + 0,25 * \text{NI}$$

La condición para aprobar el curso es: $\text{NH} \geq 4.0$, $\text{NP} \geq 4.0$, $\text{NI} \geq 4.0$

Tipo de evaluación	RA asociado a la evaluación
Hito 1: Presentación del proyecto a realizar:	Evalúa RA1, RA2
Hito 2: Análisis espacial y exploratorio de los datos del proyecto	Evalúa RA2, RA3, RA4, RA5
Hito 3: Análisis y visualización de flujos u otros aspectos urbanos del proyecto	Evalúa RA2, RA3, RA4, RA5
Presentación Final	Evalúa RA2, RA6, RA7
Informe Escrito	Evalúa RA2, RA6, RA7

G. Recursos bibliográficos:

Bibliografía obligatoria:

- [1] Shi, Goodchild, M. F., Batty, M., Kwan, M.-P., & Zhang, A. (2021). Urban Informatics. Springer Singapore Pte. Limited. [\[biblioteca\]](#)
- [2] Anselin, & Rey, S. J. (2012). Spatial econometrics in an age of CyberGIScience. International Journal of Geographical Information Science: IJGIS, 26(12), 2211–2226. <https://doi.org/10.1080/13658816.2012.664276>. [\[biblioteca\]](#)
- [3] Barthelemy. (2022). Spatial networks : a complete introduction : from graph theory and statistical physics to real-world applications. Springer International Publishing AG. [\[biblioteca\]](#)
- [4] Lima. (2011). Visual complexity : mapping patterns of information. Princeton Architectural Press. [\[biblioteca\]](#)
- [5] Cattuto, & Lapucci, M. (2021). Data science for social good : philanthropy and social impact in a complex world (Cattuto & M. Lapucci, Eds.). Springer. [\[biblioteca\]](#)
- [6] Graells-Garrido, Serra-Burriel, F., Rowe, F., Cucchiatti, F. M., & Reyes, P. (2021). A city of cities: Measuring how 15-minutes urban accessibility shapes human mobility in Barcelona. PloS One, 16(5), e0250080–e0250080. <https://doi.org/10.1371/journal.pone.0250080> [\[biblioteca\]](#)
- [7] Graells-Garrido, Peña-Araya, V., & Bravo, L. (2020). Adoption-Driven Data Science for Transportation Planning: Methodology, Case Study, and Lessons Learned. Sustainability (Basel, Switzerland), 12(15), 6001–. <https://doi.org/10.3390/su12156001> [\[biblioteca\]](#)

H. Datos generales sobre elaboración y vigencia del programa de curso:

Vigencia desde:	Otoño, 2023
Elaborado por:	Eduardo Graells-Garrido
Validado por:	Validación académico par: Iván Sipirán, Benjamín Bustos CTD de Ciencias de la Computación
Revisado por:	Área de Gestión Curricular