

CURSO: CIENCIA DE DATOS APLICADA AL TRANSPORTE



PROPUESTA
Versión 04



1 Objetivos

Adquirir nociones básicas de la ciencia de datos y el aprendizaje estadístico, métodos de visualización de datos, análisis de regresión y clasificación, para entender y formular modelos de movilidad y de transporte, y abordar el desarrollo de sistemas inteligentes de transporte.

Complementar el perfil del especialista en transporte con el manejo de bases de datos espaciales y los Sistemas de Información Geográfica, desarrollando métodos y herramientas para estimar y predecir fenómenos estadístico-espaciales de movilidad y demanda de transporte.

También:

- Entender modelos y asociación entre diferentes tipos de variables.
- Adquirir una mayor capacidad de análisis de datos a partir de diferentes herramientas estadísticas.
- Conocer y afianzar los elementos teóricos centrales de los Sistemas de Información Geográfica y su relación con entornos de programación y estadística.
- Crear, editar, interpretar y visualizar información georreferenciada asociada a la temática del transporte multimodal.
- Interpretar y generar indicadores para bases de datos espaciales estáticas y dinámicas en tiempo real.
- Generar la capacidad de trabajar con distintos formatos de Sistemas de Información Geográfica conociendo las ventajas y desventajas de cada uno de ellos.

2 Características del curso

2.1 Estructura (módulos, unidades, carga horaria por módulo o unidad)

Módulo	Materia (unidad)	Carga horaria
Transporte	Introducción al transporte	3 horas
Ciencia de datos	Introducción a la Ciencia de datos	24 horas
SIG + R	Introducción. Usos de archivos de SIG. Edición avanzada de SIG. Representación visual. Estadística espacial.	15 horas
Desarrollo de Casos	Caso 1. Ciencia de datos Caso 2. SIG	6 horas
Talleres	Consultas s/herramientas y desarrollo de casos	16 horas

CURSO: CIENCIA DE DATOS APLICADA AL TRANSPORTE



PROPUESTA
Versión 04



2.2 Contenidos de cada unidad o módulo

Semana	Contenidos
1	Conceptos básicos. Movilidad y demanda de transporte. Sistemas de transporte: componentes, características técnico económicas de los distintos modos, capacidad, velocidad, nivel de servicio. Análisis operacional de los sistemas de transporte. Sistemas programados, sistemas de organización espontánea. modelos de transporte.
2	Introducción a la ciencia de datos. ¿Qué campos maneja la ciencia de datos? ¿A qué nos referimos cuando hablamos de Machine learning, big data y data mining? Clasificación de variables. Visualización de datos a partir de los diferentes tipos de variables. Introducción al uso de R.
3	Métodos de visualización univariada: Histogramas, boxplot, y multivariada: mosaicplot, scatterplot. Medidas representativas para los diferentes tipos de variables, medidas de tendencia central y de dispersión. Repaso de conceptos estadísticos básicos: intervalos de confianza y test de hipótesis.
4	Aprendizaje supervisado. Análisis de regresión. Modelo lineal. Planteo del problema, interpretación. Cálculo de las estimaciones en función de la muestra utilizando R.
5	Correlación entre variables. Diferentes tipos de covariables y cómo interpretar su aporte en el modelo. Predicción a partir de un modelo de regresión múltiple.
6	Planteo de Casos de estudio
7	Clasificación. Métodos de Knn, regresión logística. Su uso para predicción de variables categóricas y para modelizar.
8	Continuación de clasificación. Linear discriminant análisis (LDA). Usos, interpretación de los resultados. Selección de modelos.

CURSO: CIENCIA DE DATOS APLICADA AL TRANSPORTE



PROPUESTA
Versión 04



Semana	Contenidos
9	Árboles de decisión. Árboles de regresión y clasificación, visualización de los resultados, interpretación y usos.
10	Métodos no supervisados. Introducción a Clusters. Clusters jerárquicos y particionantes. Visualización utilizando diferentes técnicas.
11	Introducción a los Sistemas de Información Geográfica Definiciones y conceptos básicos. Repaso histórico de los Sistemas de Información Geográfica. Proyecciones, escalas y su influencia en el manejo de datos georreferenciados. Bases de datos georreferenciadas. SIG y Big Data. Búsqueda de información, el rol de las API. Software SIG y enlace con lenguajes de programación libres.
12	Primeros usos de archivos de SIG Tipos de archivos con información georreferenciada. Importación de archivos shapefile, geoJSON, JSON, CSV, KML en R. Manejo básico de objetos SP y SF en R. Incorporación de datos espaciales a bases de datos y viceversa. Operaciones básicas de creación de mapas. Introducción a Leaflet.
13	Edición avanzada de SIG Tópicos avanzados de los paquetes sf, raster y rgdal. Reproyección de mapas. Uso de herramientas de corte, extensión, unión, recorte, intersección, disolución, adición y buffer. Cálculo de distancias y rutas entre puntos. Exportación de mapas y creaciones georreferenciadas. Concepto y creación de isócronas.
14	Representación visual. Operaciones avanzadas de Leaflet. Teoría de diseño visual de mapas y visualización de datos. Desarrollo del paquete ggplot como procesador de mapas y datos SIG. Concepto y usos de mapas de calor, densidad, intensidad y caracterización de variables según color, tamaño, opacidad o forma. Uso del hexágono como mejor unidad gráfica de representación de datos puntuales. Centroides. Introducción a la representación 3D.

CURSO: CIENCIA DE DATOS APLICADA AL TRANSPORTE



PROPUESTA
Versión 04



Semana	Contenidos
15	Estadística espacial Introducción a la estadística de datos georreferenciados. Construcción de indicadores estadísticos básicos. Análisis de datos de variación continua. Análisis de datos de variación discreta. Procesos estocásticos espaciales. Introducción a la predicción espacial gaussiana.
16	Resolución de casos – Evaluación final

2.3 Carga horaria total

Sesenta y cuatro horas (64) horas (48 horas de clases teórico-prácticas + 16 horas taller)

2.4 Requisitos de ingreso

Orientado a graduados universitarios de carreras afines al transporte y la movilidad, otros a consideración del director del curso, alumnos avanzados de la carrera de ingeniería civil y/o carreras afines.

Conocimientos previos: En este curso se asume que los participantes cuentan con algún conocimiento previo estadístico: variables aleatorias, conceptos y definiciones de probabilidad, distintos tipos de distribuciones, intervalos de confianza y test de hipótesis.

2.5 Modalidad de cursado

Virtual.

Dieciséis (16) encuentros de tres (3) horas (un encuentro semanal), para el desarrollo de clases teórico-prácticas, en las que se irá trabajando con datos reales y con momentos de trabajo individual para incorporar los conceptos a medida que se van desarrollando a lo largo del curso.

Dieciséis (16) encuentros de una (1) hora (un encuentro semanal) para el desarrollo de clases prácticas, de consulta y seguimiento de resolución de casos.

Herramientas y metodología de las clases: A lo largo de todos los encuentros se utiliza R para acompañar el aprendizaje de todos los temas, para una mejor comprensión de los mismos.

Es importante que los participantes cuenten con una computadora y el software instalado.

2.6 Nómina de colaboradores/docentes


Docentes:

- PAPAIZAN, Arturo
- GARCÍA, Jemina María
- FULPONI, Juan Ignacio

Colaboradores:

- PIANCIOLA, Gustavo
- ALLMI, Natalia

CURSO: CIENCIA DE DATOS APLICADA AL TRANSPORTE

 <p>.UBAfiuba FACULTAD DE INGENIERÍA CENTRO DE ESTUDIOS DE TRANSPORTE</p>	<p>PROPUESTA Versión 04</p>	 <p>depto. Transporte</p>
---	--	--

2.7 Modalidad de evaluación

La evaluación se realizará mediante la resolución de dos casos de estudio, y su exposición en la clase final. Se complementará un cuestionario sincrónico, con preguntas de respuesta cerrada, del tipo elección múltiple, verdadero-falso, apareamiento y/o mensaje para completar.

2.8 Requisitos de aprobación

Resolución de casos, y aprobación del cuestionario, dentro del plazo establecido.

2.9 Horario

Miércoles de 18 a 22 hs.

2.10 Fecha de inicio / finalización

13 de abril de 2022 / 3 de agosto de 2022.