

Curso

Introducción al pensamiento y metodología de la ciencia de datos con R

Nombre del Profesor:	Mtro. Miguel David Alvarez Hernández
Breve semblanza del Profesor:	Investigador y científico de datos con más de 7 años de experiencia laboral en el sector público y en la academia, dedicado a resolver problemas relacionados con el análisis estadístico, así como la modelización de fenómenos sociales, y en el diseño y gestión de proyectos de investigación que implican trabajo teórico y de campo. Su formación académica ha transitado entre el campo de las ciencias naturales (Licenciatura en Física en la UAM) y el campo de la matemática aplicada (Complex Systems Modelling en el King's College London y Ciencia de Datos en INFOTEC). Asimismo, ha publicado diversos artículos de investigación en revistas prestigiadas e indexadas, y ha colaborado en diversos proyectos de investigación que cubren diversas áreas del conocimiento (modelado de tráfico vehicular, modelos de crecimiento económico, complejidad electoral).
Fecha:	Del 3 al 12 de octubre de 2022
Horario:	20 horas divididos en 5 sesiones. 3, 5 y 7 de octubre, 16 a 20 hrs 10 y 12 de octubre, 10 a 14 hrs
Modalidad:	En línea
Breve resumen del curso:	Este curso proporciona una introducción, a través del uso de R, a las metodologías fundamentales de la ciencia de datos. En particular se hará un repaso de los objetivos de la ciencia de datos, de los fundamentos del lenguaje R, de los métodos de análisis exploratorio y visualización de datos, y de los métodos para facilitar la reproducibilidad y publicación de los resultados de investigación.
Objetivos de aprendizaje:	El objetivo es lograr un entendimiento general del funcionamiento del lenguaje R, y de algunos procedimientos estadísticos ampliamente utilizados para el análisis de datos.

Temario:	<ol style="list-style-type: none"> 1. Sesión del 3 de octubre Introducción a la ciencia de datos y al lenguaje de <i>R</i> <ul style="list-style-type: none"> - Introducción y motivación - Instalación de <i>R</i> e IDE's (RStudio, RCloud, Jupyter Lab/Notebook) - Elementos básicos de <i>R</i> (Sintaxis, <i>shortcuts</i>, comandos básicos) 2. Sesión del 5 de octubre Análisis exploratorio: procesamiento y visualización de datos <ul style="list-style-type: none"> - Tipos de datos y formulación de preguntas - Principios y técnicas de visualización de datos - <i>Tidyverse</i> 3. Sesión del 7 de octubre Estadística y ciencia de datos <ul style="list-style-type: none"> - Pensamiento estadístico para la ciencia de datos - Tipos de estudios, sus datos, y sus alcances - Elementos básicos de estadística descriptiva e inferencial 4. Sesión del 10 de octubre Análisis reproducible con <i>R</i> <ul style="list-style-type: none"> - <i>RMarkdown</i> y <i>Shiny</i> 5. Sesión del 12 de octubre Casos de estudio
Dirigido a:	Científicos de datos, estudiantes o profesionistas en otras áreas que deseen entender el funcionamiento básico de <i>R</i> y los procedimientos estadísticos que sustentan los análisis de datos, también está dirigido a cualquiera que quiera refrescar sus conocimientos de <i>R</i> .
Requisitos:	Contar con equipo de cómputo y conexión a internet.