

DATOS DE IDENTIFICACIÓN			
Nombre de la asignatura		Estadística para Ciencia de Datos	
Unidad Regional		Centro	
División		Ciencias Exactas y Naturales	
Departamento		Matemáticas	
Programa		Maestría en Ciencia de Datos	
Carácter		Obligatorio (X) Optativo ()	
Horas teoría	3	Horas práctica	4
Valor en créditos		10	
OBJETIVO GENERAL			
Adquirir los conocimientos estadísticos mínimos requeridos para la tarea de seleccionar, adaptar e implementar métodos estadísticos para el análisis de datos en relación a un problema en particular.			
OBJETIVOS ESPECÍFICOS			
<ol style="list-style-type: none"> 1. Conocer y aplicar diferentes técnicas estadísticas para el resumen, descripción y visualización de datos. 2. Construir y validar modelos estadísticos. 3. Desarrollar la habilidad de modelar procesos de adquisición de datos de manera probabilística, desde un enfoque tanto frecuentista como Bayesiano. 4. Analizar situaciones reales que permitan adquirir experiencia en la generación, prueba e interpretación de hipótesis estadísticas. 			
CONTENIDO SINTÉTICO			
Orden	Tema		
1	Introducción a la estadística para ciencia de datos.		
2	Variables aleatorias multivariadas, mezcla de distribuciones, muestreo de rechazo y covarianza.		
3	Procesos aleatorios: media, autocovarianza, sucesiones iid, ley de grandes números, procesos de Markov, etcétera.		
4	Análisis exploratorio de datos, componentes principales.		
5	Estimación: enfoque frecuentista: muestreo iid, error cuadrático medio, intervalos de confianza.		
6	Estimación: enfoque bayesiano: distribución a priori, función de verosimilitud, distribución a posteriori, intervalos de credibilidad.		
7	Pruebas de hipótesis paramétricas y no-paramétricas, con una y varias muestras.		
8	Aspectos éticos de la aplicación de la estadística en la ciencia de datos.		
MODALIDADES O FORMAS DE CONDUCCIÓN DE LOS PROCESOS DE ENSEÑANZA Y APRENDIZAJE			
<ol style="list-style-type: none"> 1. Lectura y discusión del rol de la estadística en ciencia de datos. 2. Análisis de aplicaciones de procesos aleatorios en ciencia de datos. 3. Presentación del uso de análisis multivariado en ciencia de datos. 4. Análisis de aplicaciones del enfoque frecuentista y Bayesiano en ciencia de datos. 5. Análisis de casos de estudio. 6. Presentación de proyectos individuales que involucren la aplicación de las técnicas vistas en clase. 			
MODALIDADES DE EVALUACIÓN Y ACREDITACIÓN			
Aspecto		Ponderación	
Ejercicios y tareas		30%	

Análisis de casos de estudio	20%
Desarrollo y reporte del proyecto individual	30%
Presentación del proyecto individual	20%

BIBLIOGRAFÍA, DOCUMENTACIÓN Y MATERIALES DE APOYO

Autor	Título	Editorial	Año
DasGupta, Anirban	Probability for Statistics and Machine Learning	Springer	2011
D'Agostini, Giulio	Bayesian Reasoning in Data Analysis: A Critical Introduction	World Scientific Pub	2003
Bruce, Peter	Practical Statistics for Data Scientists: 50 Essential Concepts	Wiley	2017

PERFIL ACADÉMICO DESEABLE DEL RESPONSABLE DE LA ASIGNATURA

Profesional con maestría o doctorado en estadística, matemáticas, ciencias de la computación o áreas afines. Se utilizará como criterio de selección la experiencia en la aplicación de la estadística en el análisis de datos y otras aplicaciones a la ciencia de datos.

NOMBRE Y FIRMA DE QUIEN DISEÑÓ LA CARTA DESCRIPTIVA

Olivia Carolina Gutú Ocampo

Gudelia Figueroa Preciado