

Análisis Estadístico de Datos

Guía de asignatura

Última actualización: agosto de 2024

1. Información general

Nombre de la asignatura	Análisis Estadístico de Datos
Código	
Tipo de asignatura	Obligatoria
Número de créditos	2
Tipo de crédito	A (teórico-práctico)
Horas de trabajo semanal con acompañamiento directo del profesor	Horas de trabajo con acompañamiento directo del profesor: 24
Horas semanales de trabajo independiente del estudiante:	Horas de trabajo independiente del estudiante: 48
Total, horas por periodo académico:	72
Prerrequisitos	Álgebra lineal y Estadística en una variable
Correquisitos	Ninguno
Horario	Sábado de 10:00am a 1:00pm
Salón	Sala Boole. Piso 4. Edificio El Tiempo

2. Información del profesor

Nombre profesor auxiliar o monitor	Andrés Nicolás López López
Perfil profesional	Profesional y maestro en Estadística de la Universidad Nacional de Colombia, científico de datos certificado con experiencia en investigación, consultoría y docencia

	universitaria a niveles de pregrado y posgrado. Estadístico integral con dominio en el análisis de información estructurada y no estructurada, junto a un gran interés en la aplicación interdisciplinar de la Estadística para la correcta toma de decisiones. Experiencia en la aplicación de metodologías estadísticas clásicas y modelamiento de vanguardia junto a visualizaciones efectivas de resultados tanto en roles de ejecución como de liderazgo. Motivado por el aprendizaje de lenguajes en programación, la enseñanza, el trabajo en equipo y la comunicación asertiva de resultados del análisis de información. Interesado en el análisis de información no estructurada, modelamiento bayesiano y funcional.
Correo electrónico institucional	andresn.lopez@urosario.edu.co
Lugar y horario de atención:	Atención virtual coordinada a través de eaulas.
Página web, Skype u otros medios (opcional)	https://anlopezl.github.io/files/AED/Contenido.html https://urosario.zoom.us/j/9945913391

3. Resumen y propósitos del curso

El estudio de fenómenos reales requiere considerar múltiples variables y sus posibles relaciones. Considerar múltiples variables o dimensiones suele ser una tarea difícil, así como extraer información útil de estos conjuntos de datos multidimensionales. Este curso presenta un conjunto de herramientas estadísticas para analizar datos multidimensionales y construir modelos con aplicación en un amplio espectro de áreas. El curso iniciará con un resumen de los fundamentos univariados del análisis de datos, junto a las bases del análisis multidimensional y la distribución normal multivariada. Luego se concentrará en la inferencia del vector de medias y el modelo de regresión múltiple. La reducción de dimensionalidad es el siguiente tema, seguido del análisis de correlaciones y métodos de clasificación y agrupamiento. El curso termina con el análisis de varianza, de correspondencias y tablas de contingencia.

4. Conceptos fundamentales

- Fundamentos univariados del análisis de datos,
- Introducción al análisis de datos multidimensionales.
- Reducción de dimensionalidad.
- Clasificación y agrupamiento.
- Regresión lineal en una y varias variables.
- Análisis de varianza y de correspondencias

5. Resultados de aprendizaje esperados (RAE)

De acuerdo con los propósitos de formación establecidos en el programa, se espera que el estudiante pueda alcanzar los siguientes resultados de aprendizaje (RAEs) con los respectivos niveles de profundidad:

Propósito de formación general	Propósitos de formación específicos	RAES	nivel de profundidad I: Introdutorio M: medio A: avanzado
Formar profesionales que, con un alto sentido de responsabilidad profesional y ética, sean capaces de generar soluciones y conocimiento a problemas de la Ingeniería mediante el análisis de datos.	Formar profesionales capaces de entender el rango de aplicación de la estadística en la toma de decisiones.	1. Identificar problemas en los que las herramientas de análisis estadístico univariado y multivariado encuentran aplicación	M
	Formar profesionales capaces de comprender las bases teóricas del modelamiento estadístico, con el propósito de diferenciar sus ventajas y desventajas en la aplicación práctica de los mismos.	2. Comprender los fundamentos matemáticos que soportan los métodos de análisis estadístico de datos	I
		3. Construir y evaluar modelos multidimensionales y estocásticos a partir de datos	I
		4. Identificar las diferencias y similitudes entre los métodos de análisis estadístico.	I
	Formar profesionales capaces de interpretar los resultados obtenidos del análisis estadístico de datos.	Extraer conclusiones como resultado de emplear métodos de análisis estadístico sobre conjuntos de datos.	M
	Formar profesionales capaces de programar los conceptos aprendidos en clase para su ejecución de manera reproducible mediante un lenguaje propio en estadística.	6. Emplear el lenguaje R para aplicar los conceptos vistos en el curso.	I

6. Modalidad del curso

Presencial: La modalidad de la asignatura para el semestre 2024-2 para todos sus estudiantes será presencial. Hyflex para estudiantes de la maestría en ciudades inteligentes.

7. Estrategias de aprendizaje

Se realizarán cuadernos de clase programáticos como fundamento teórico-practico de clase. Esto permite desarrollar las habilidades técnicas y prácticas del curso. A su vez se desarrollan talleres grupales con el objetivo de afianzar el conocimiento aprendido y consolidar en un único

resultado la programación y análisis de datos. Finalmente, un trabajo final que permite apropiación de un tema avanzado relacionado con los vistos en clase, este da la oportunidad de afianzar lo aprendido tanto estadística como programáticamente: una exposición grupal que a su vez sirve como repositorio para futura referencia de los participantes del curso. Se desarrollan clases magistrales con un alto componente práctico de programación.

8. Actividades de evaluación

Tema	Actividad de evaluación	Porcentaje	Fechas
1 y 2	Taller 1	25	Sesión 3
3,4 y 5	Taller 2	25	Sesión 5
6 y 7	Taller 3	25	Sesión 7
Todos	Proyecto aplicado	25	Sesión 8

Programación de actividades

Fecha	Tema	Trabajo independiente del estudiante	Recursos que apoyan la actividad
Sesión 1	Herramientas estadísticas básicas. Sesión 1	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias
Sesión 2	Herramientas estadísticas básicas. Sesión 2	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias
Sesión 3	Introducción Normal Multivariada y Visualización	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias
Sesión 4	Introducción Normal Multivariada y Visualización	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias
Sesión 5	Análisis de componentes principales	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias
Sesión 6	Agrupamiento o Clustering	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias
Sesión 7	Regresión lineal en múltiples variables	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias
Sesión 8	Regresión lineal en múltiples variables	Leer los recursos y hacer todos los ejercicios del notebook	Notebook de clase y referencias

9. Factores de éxito para este curso

A continuación, se sugieren una serie de acciones que pueden contribuir, de manera significativa, con el logro de metas y consecuentemente propiciar una experiencia exitosa en este curso:

1. Planificar y organizar el tiempo de trabajo individual que le dedicará al curso.
2. Organizar el sitio y los materiales de estudios.
3. Tener un grupo de estudio, procurar el apoyo de compañeros.
4. Cultivar la disciplina y la constancia, trabajar semanalmente, no permitir que se acumulen temas ni trabajos.
5. Realizar constantemente una autoevaluación, determinar si las acciones realizadas son productivas o si por el contrario se debe cambiar de estrategias.

6. Utilizar la plataforma eaulas y sus foros para las preguntas.
7. Propiciar espacios para el descanso y la higiene mental, procurar tener buenos hábitos de sueño
8. Tener presente en todo momento valores como la honestidad y la sinceridad, al final no se trata solo de aprobar un examen, se trata de aprender y adquirir conocimientos. El fraude es un autoengaño.
9. Presentación del proyecto final de la asignatura, haciendo uso de las diferentes herramientas vistas en la asignatura.
10. Entender y extrapolar el alcance de la tecnología, su uso y aplicaciones en ambientes industriales.

10. Bibliografía y recursos

- [1] Mendenhall, Beaver, Beaver (2010). Introducción a la probabilidad y estadística.
[2] Applied Multivariate Statistical Analysis, Johnson & Wichern, 6th edition, Pearson.

11. Bibliografía complementaria

- [3] Applied Multivariate Statistical Analysis, Härdle & Simar, 3rd edition, Springer.
[4] Bertsekas, Tsitsiklis. Introduction to Probability. Segunda Edición, 2008.
[5] Wackerly, Mendenhall, Scheaffer. Mathematical Statistics with Applications, 7th Ed. Versión en español: Estadística matemática con aplicaciones.

12. Acuerdos para el desarrollo del curso

Reglamento estudiantil de posgrado, que aplica en la Universidad del Rosario. En ese sentido, se contempla acuerdos como:

- La puntualidad a la hora de ingresar (presencial)
- El uso y consulta de la página del curso.
- Uso responsable de la plataforma Zoom (cuando sea necesario)
- La participación y proactiva sobre el contenido del curso en cada clase
- Cumplimiento de las actividades / laboratorios

13. Respeto y no discriminación

Si tiene alguna discapacidad, sea este visible o no, y requiere algún tipo de apoyo para estar en igualdad de condiciones con los(as) demás estudiantes, por favor informar a su profesor(a) para que puedan realizarse ajustes razonables al curso a la mayor brevedad posible. De igual forma, si no cuenta con los recursos tecnológicos requeridos para el desarrollo del curso, por favor informe de manera oportuna a la Secretaría Académica de su programa o a la Dirección de Estudiantes, de manera que se pueda atender a tiempo su requerimiento.

Recuerde que es deber de todas las personas respetar los derechos de quienes hacen parte de la comunidad Rosarista. Cualquier situación de acoso, acoso sexual, discriminación o matoneo, sea presencial o virtual, es inaceptable. Quien se sienta en alguna de estas situaciones puede denunciar su ocurrencia contactando al equipo de la Coordinación de Psicología y Calidad de Vida de la Decanatura del Medio Universitario (Teléfono o WhatsApp 322 2485756).

14. Rubrica de evaluación

RÚBRICA DE EVALUACIÓN PROYECTO FINAL: ANÁLISIS ESTADÍSTICO DE DATOS

Estudiante 1: _____

Estudiante 2: _____

Nota: _____

Explicación del método (30%)

Dimensión	Criterio	5	3	1
Comprende y explica los conceptos básicos del método seleccionado	1. Introducción ()	<input type="checkbox"/> El estudiante introduce claramente el método estadístico seleccionado.	<input type="checkbox"/> El estudiante introduce parcialmente el método estadístico seleccionado.	<input type="checkbox"/> El estudiante no introduce el método estadístico seleccionado.
	2. Relación con tema macro de clase	<input type="checkbox"/> El estudiante relaciona claramente el método estadístico seleccionado con el tema macro de clase mediante similitudes y diferencias.	<input type="checkbox"/> El estudiante relaciona parcialmente el método estadístico seleccionado con el tema macro de clase mediante similitudes y diferencias.	<input type="checkbox"/> El estudiante no relaciona el método estadístico seleccionado con el tema macro de clase mediante similitudes y diferencias.
	3. Fundamentos matemáticos básicos	<input type="checkbox"/> El estudiante explica claramente el fundamento matemático básico del método estadístico seleccionado.	<input type="checkbox"/> El estudiante explica parcialmente el fundamento matemático básico del método estadístico seleccionado.	<input type="checkbox"/> El estudiante no explica el fundamento matemático básico del método estadístico seleccionado.
	4. Visualización pedagógica del método.	<input type="checkbox"/> El estudiante visualiza pedagógicamente el método estadístico seleccionado.		<input type="checkbox"/> El estudiante no visualiza pedagógicamente el método estadístico seleccionado.
	5. Posibles fallos y supuestos	<input type="checkbox"/> El estudiante enumera los posibles fallos y supuestos del método seleccionado.	<input type="checkbox"/> El estudiante enumera parcialmente los posibles fallos y supuestos del método seleccionado.	<input type="checkbox"/> El estudiante no enumera los posibles fallos y supuestos del método seleccionado.

Aplicación del método (30%)

Dimensión	Criterio	5	3	1
Presenta el método seleccionado mediante el uso del software R	6. Claridad del código	<input type="checkbox"/> El código presentado de método estadístico seleccionado es legible.	<input type="checkbox"/> El código presentado de método estadístico seleccionado es medianamente legible.	<input type="checkbox"/> El código presentado de método estadístico seleccionado no es legible.
	7. Aplicación del método	<input type="checkbox"/> El código presentado aplica el método estadístico seleccionado de manera correcta.	<input type="checkbox"/> El código presentado aplica el método estadístico seleccionado de manera equivocada.	<input type="checkbox"/> El código presentado no aplica el método estadístico seleccionado.
	8. Visualización	<input type="checkbox"/> El código presentado visualiza el/los resultados del método estadístico seleccionado de manera correcta.	<input type="checkbox"/> El código presentado visualiza el/los resultados del método estadístico seleccionado con al menos un problema importantes.	<input type="checkbox"/> El código presentado no visualiza el/los resultados del método estadístico seleccionado.

Calidad de la presentación (20%)

Dimensión	Criterio	5	3	1
Explica de manera técnica el método seleccionado.	9. Lenguaje	<input type="checkbox"/> Los dos estudiantes utilizan un lenguaje técnico al momento de explicar el método seleccionado.	<input type="checkbox"/> Uno de los estudiantes utiliza un lenguaje técnico al momento de explicar el método seleccionado.	<input type="checkbox"/> Ninguno de los estudiantes utiliza un lenguaje técnico al momento de explicar el método seleccionado.
	10. Participación	<input type="checkbox"/> Los dos estudiantes tienen buen dominio durante la presentación.	<input type="checkbox"/> Uno de los estudiantes no tiene buen dominio durante la presentación.	<input type="checkbox"/> Ninguno de los estudiantes tiene dominio durante la presentación.
	11. Presentación	<input type="checkbox"/> Los dos estudiantes usan las diapositivas como soporte principalmente de visualización en su presentación.	<input type="checkbox"/> Un estudiante usa las diapositivas como soporte principalmente de visualización en su presentación, el otro como soporte de lectura.	<input type="checkbox"/> Los dos estudiantes usan las diapositivas como soporte principalmente de lectura en su presentación.
	12. Diapositivas	<input type="checkbox"/> Las diapositivas son claras y bien organizadas.		<input type="checkbox"/> Las diapositivas no son claras ni bien organizadas.

Video, preguntas y respuestas (20%)

Dimensión	Criterio	5	3	1
Contenido del video	13. Preguntas	<input type="checkbox"/> Una pregunta realizada a otro grupo, seleccionado de manera aleatoria, es precisa.	<input type="checkbox"/> Una pregunta realizada a otro grupo, seleccionado de manera aleatoria, es imprecisa.	<input type="checkbox"/> No hace una pregunta al grupo seleccionado de manera aleatoria.
	14. Respuestas	<input type="checkbox"/> La respuesta a una pregunta seleccionada de manera aleatoria es correcta.	<input type="checkbox"/> La respuesta a una pregunta seleccionada de manera aleatoria es regular.	<input type="checkbox"/> La respuesta a una pregunta seleccionada de manera aleatoria es incorrecta.
	15. Tiempo	<input type="checkbox"/> El video tiene una duración de máximo 15 minutos	<input type="checkbox"/> El video tiene una duración entre 15 y 16 minutos	<input type="checkbox"/> El video tiene una duración de más de 16 minutos