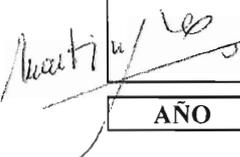


UNIVERSIDAD NACIONAL DEL SUR						1	4
BAHIA BLANCA				ARGENTINA			
DEPARTAMENTO DE MATEMÁTICA							
PROGRAMA DE: ANALISIS EXPLORATORIO DE DATOS MULTIVARIADOS						CÓDIGO: 8128	
						ÁREA N°: IV	
HORAS DE CLASES: 96 horas				PROFESOR RESPONSABLE			
TEÓRICAS		PRÁCTICAS		Dra. María Cristina MARTÍN			
Por semana	Por cuatrim.	Por semana	Por cuatrim.				
1	16	1	16				
ASIGNATURAS CORRELATIVAS PRECEDENTES							
CARRERA		APROBADA			CURSADA		
Licenciatura en Matemática		Probabilidad y Estadística			Probabilidad y Estadística		
DESCRIPCIÓN							
<p>El curso se centrará en la presentación y desarrollo teórico de las principales técnicas propias del Análisis de Datos Multivariantes. Para cada técnica multivariada introducida se proporcionarán una serie de ejercicios y experiencias con datos reales a efectos de que el estudiante pueda introducirse a problemas multivariantes concretos. Las clases teóricas tendrán cuatro (4) horas de duración semanal y las prácticas, incluyendo fundamentalmente una guía para manejo de rutinas de R, tendrán dos (2) de duración semanal, de manera que, posteriormente el estudiante cuente con los recursos teóricos de cada técnica y esté capacitado para ingresar y analizar las salidas ofrecidas por el paquete computacional.</p>							
OBJETIVOS							
<ol style="list-style-type: none"> 1. Presentar y discutir las técnicas exploratorias propias del Análisis de Datos Multivariado que conducen a la reducción e interpretación del número de casos y del número de variables. 2. Comprender conceptos claves de las técnicas estadísticas del Análisis Multivariado de Datos. 3. Ser capaz de distinguir qué técnica multivariante es la más adecuada, en función tanto del tipo de datos como del uso que se quiera hacer de ellos. 4. Discutir los Usos Adecuados e Inadecuados de los distintos Procedimientos. 5. Adquirir los conocimientos necesarios para la aplicación mediante el manejo de distintos Paquetes Estadísticos, en particular R. 							
PROGRAMA SINTÉTICO SEGÚN PLAN DE ESTUDIOS							
<ol style="list-style-type: none"> 1. Introducción a las Técnicas de Análisis de Datos Multivariados 2. Medidas Descriptivas y Técnicas Gráficas de Datos Multivariados 3. Análisis de Agrupamiento 4. Análisis de Componentes Principales 5. Análisis Discriminante y Clasificación 6. Análisis Factorial 7. Escalonamiento Multivariado 8. Análisis de Correspondencias 							
							
AÑO	2018						

UNIVERSIDAD NACIONAL DEL SUR							2	4
BAHIA BLANCA				ARGENTINA				
DEPARTAMENTO DE MATEMÁTICA								
PROGRAMA DE: ANALISIS EXPLORATORIO DE DATOS MULTIVARIADOS						CÓDIGO: 8128		
						ÁREA N°: IV		
PROGRAMA ANALÍTICO Y METODOLOGÍA DE ENSEÑANZA								
CAPÍTULO	CONTENIDO TEMÁTICO						METODOLOGÍA	
1-	Introducción a las Técnicas de Análisis de Datos Multivariados: Un poco de historia. Descripción de Datos Multivariados. Tipos de Variables. Matriz de Datos. El vector de medias. Matriz de Varianzas y Covarianzas. Medidas globales de variabilidad. Variabilidad y distancias. Presentación y Ejemplos de las Técnicas más usuales.						- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Se presentan típicos problemas reales. - Manejo de R (carga de datos)	
2-	Medidas Descriptivas y Técnicas Gráficas de Datos Multivariados: Análisis previo de datos. Análisis exploratorio y gráfico múltiple: histograma de frecuencias, gráfico de tallo y hoja, gráfico de caja y bigotes, gráfico de simetría, gráfico de dispersión, gráfico de estrellas (Star Symbol Plot), gráfico de las Caras de Chernoff. Análisis de Datos Ausentes. Análisis y detección de valores atípicos. Comprobación de los supuestos del análisis multivariado: normalidad, heterocedasticidad, linealidad, multicolinealidad, autocorrelación, análisis de los residuos. Ejemplos y Ejercicios. Uso de Paquetes Estadísticos.						- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Ejercicios y Aplicaciones con datos reales - Uso de R. -Práctico N°1	
3-	Análisis de Agrupamiento: Principios y esquema general del Análisis de Conglomerados o Clusters. Medidas de Similitud y Distancia. Métodos de Agrupamientos Jerárquicos (Single Linkage, Complete Linkage, Average Linkage, Método del Centroide, Método de Ward, entre otros), Métodos de Agrupamientos No Jerárquicos (Método de las k-medias). El análisis de Conglomerados en dos fases. Ejemplos y Ejercicios. Uso de Paquetes Estadísticos.						- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Ejercicios y Aplicaciones con datos reales - Uso de R -Práctico N°2.	
4-	Análisis de Componentes Principales: Objetivo. Obtención de las Componentes Principales. Varianza de las componentes. Estructura factorial de las componentes principales. Puntuaciones o medición de las componentes. Contrastes sobre el número de componentes principales a retener (Criterio de la Media Aritmética, Contraste sobre las raíces características no retenidas, Prueba de Anderson, Prueba del bastón roto de Frontier, entre otras). Interpretación Geométrica. Propiedades muestrales. Ejemplos y Ejercicios. Uso de Paquetes Estadísticos.						- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Ejercicios y Aplicaciones con datos reales - Uso de R -Práctico N°3.	
5-	Análisis Discriminante y Clasificación: Introducción. Separación y Clasificación para dos poblaciones. Clasificación con dos poblaciones normales multivariadas. Evaluación de las funciones de clasificación. Función Discriminante de Fisher -Separación de poblaciones. Clasificación con varias poblaciones. Método de						- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Ejercicios y	
								
AÑO	2018							

UNIVERSIDAD NACIONAL DEL SUR		3	4
BAHIA BLANCA	ARGENTINA		

DEPARTAMENTO DE MATEMÁTICA	
PROGRAMA DE:	ANÁLISIS EXPLORATORIO DE DATOS MULTIVARIADOS
	CÓDIGO: 8128
	ÁREA N°: IV

	Fisher para discriminar entre varias poblaciones. Breve descripción de otros Métodos de Clasificación: Diagrama de Árboles, Regresión Logística y Redes Neuronales. Ejemplos y Ejercicios. Uso de Paquetes Estadísticos.	Aplicaciones con datos reales - Uso de R -Práctico N°4.
6-	Análisis Factorial: Objetivos, Conceptos Básicos y Terminología del Análisis Factorial. Técnicas del Análisis Factorial: Método de Turstone, Método del Factor Principal, Método Alpha, Método del centroide, Método de las componentes principales, Métodos Minres, ULS y GLS, entre otros. Contrastes en el Modelo Factorial. Interpretación Geométrica del Análisis Factorial. Rotación de los Factores, Rotaciones ortogonales y oblicuas. Puntuaciones o Medición de los Factores. Análisis Factorial exploratorio y confirmatorio. Ejemplos y Ejercicios. Uso de Paquetes Estadísticos.	- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Ejercicios y Aplicaciones con datos reales - Uso de R -Práctico N°5.
7-	Escalonamiento Multivariado: Concepto de Escalonamiento Multidimensional: Mapas perceptuales; solución, ajuste y preferencias en el escalonamiento multivariado; Interpretación y validación de los resultados. Modelos de escalonamiento: escalonamiento métrico, escalonamiento no-métrico, escalonamiento de diferencias individuales, escalonamiento desdoblado, escalonamiento vertical. Interpretación de los resultados obtenidos (dimensional, por agrupamientos, por regiones). Aplicaciones y su relación con otras técnicas de análisis de datos. Uso de Paquetes Estadísticos.	- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Ejercicios y Aplicaciones con datos reales - Uso de R -Práctico N°6.
8-	Análisis de Correspondencias: Análisis Factorial de Correspondencias. Análisis de Correspondencias Simples: formación de las nubes y definición de distancias, Ejes Factoriales: Análisis en R^p y en R^n ; Relaciones entre los Análisis en R^p y en R^n ; Reconstrucción de la tabla de frecuencias. Análisis de Correspondencias Múltiple: Obtención de los factores, tabla de Burt. Ejemplos y Ejercicios. Uso de Paquetes Estadísticos	- Exposición de los contenidos teóricos, con ayuda de cañón multimedia. - Ejercicios y Aplicaciones con datos reales - Uso de R -Práctico N°7.

SISTEMA DE EVALUACIÓN

Se resolverán problemas y ejercicios durante el desarrollo de la cursada.
 Periódicamente, se requerirá la entrega de soluciones a problemas teórico prácticos y con datos reales previamente asignados y que formarán parte de la evaluación final.
 Se tomarán por lo menos dos (2) exámenes parciales.
 La nota final del curso se calculará como el promedio simple de las notas obtenidas por cada instancia de evaluación.

lectiva

AÑO	2018								
-----	------	--	--	--	--	--	--	--	--

UNIVERSIDAD NACIONAL DEL SUR		4	4
BAHIA BLANCA		ARGENTINA	
DEPARTAMENTO DE MATEMÁTICA			
PROGRAMA DE:	ANÁLISIS EXPLORATORIO DE DATOS MULTIVARIADOS		CÓDIGO: 8128
			ÁREA N°: IV

BIBLIOGRAFÍA

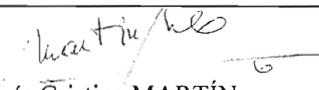
Bibliografía Básica

- [1] CUADRAS, C. (2014): "Nuevos Métodos de Análisis Multivariante". Barcelona, ES: CMC Editions. 304 p.
- [2] CUADRAS C.M. y FORTIANA J., OLIVA F. (1996): "Análisis de Datos Multivariados". Departamento de Estadística de la Universidad de Barcelona, 99 páginas.-
- [3] de OLIVERA BUSSAB W., MIAZAKI E.S. e de ANDRADE D. (1990): "Introducao a Análise de Agrupamentos". Asociación Brasileira de Estadística, ABE, 9° Simposio Nacional de Probabilidad y Estadística, San Pablo, 105 páginas.
- [4] JOHNSON R and WICHERN D. (1992): "Applied Multivariate Statistical Analysis", Editorial Prentice Hall, 3^{ra} Edición, 642 páginas.
- [5] HAIR, J.F. Jr., ANDERSON, R.E., TATHAM, R.L. y BLACK, W.C. (1999): "Análisis Multivariante". Prentice Hall, 5ta. Edición. (traducido de "Multivariate Data Analysis", Fifth edition. Prentice Hall Internationa, Inc.). 799 páginas.
- [6] PENHA, D. (2002): "Análisis de Datos Multivariantes". Mc Graw Hill/Interamericana de España, S.A.U., 539 paginas.
- [7] PÉREZ, C. (2004): "Técnicas de Análisis Multivariante de Datos: Aplicaciones con SPSS". Madrid, ES: Pearson Educación. 672 p.
- [8] R CORE TEAM. (2016): "R: A language and environment for statistical Computing". R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

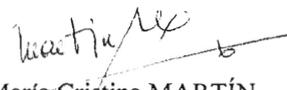
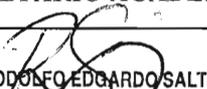
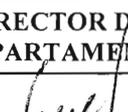
Bibliografía Complementaria

- [9] KAUFMAN L. and ROUSSEUW, P.J. (1990): "Finding Groups in Data: An Introduction to Cluster Analysis". Wiley, New York.
- [10] MANLY, B.F.J. (1997): "Multivariate Statistical Methods. A primer". Editorial Chapman & Hall, 2^{da} Edición, 214 páginas.
- [11] SEBER, G.A.F. (1984): "Multivariate Observation". Wiley & Sons, 686 páginas.
- [12] SHARMA, S. (1996): "Applied Multivariate Techniques". John Wiley & Sons, Inc., New York. 493 páginas.

VIGENCIA DE ESTE PROGRAMA

AÑO	PROFESOR RESPONSABLE (Firma aclarada)	AÑO	PROFESOR RESPONSABLE (Firma aclarada)
2018	 Dra. María Cristina MARTÍN		

VISADO

COORDINADOR ÁREA	SECRETARIO ACADÉMICO	DIRECTOR DEL DEPARTAMENTO
 Dra. María Cristina MARTÍN	 Lic. RODOLFO EDUARDO SALTHÚ SECRETARIO ACADÉMICO Departamento de Matemática	 Lic. ENELDY JAVIER OMBROSI DIRECTOR DECAÑO Departamento de Matemática
FECHA:	FECHA:	FECHA: 28 FEB 2018
AÑO	2018	