

PROGRAMA DE ESTUDIOS Análisis Avanzado de Datos II

I. Identificación.

Código:	SOC9035
Créditos:	5
Período Académico:	1º semestre 2024
Requisito:	Análisis Avanzado de Datos
Duración:	Semestral
Horario:	Lunes módulos 2 y 3 10.00 a 11.20 y 11.30 a 12.50
Sala:	Laboratorio de computación
Asistencia:	No es obligatoria
Docente:	Profesor: Gabriel Sotomayor gabriel.sotomayor@mail.udp.cl
Ayudante:	Patricio Alarcón patricio.alarcon@mail.udp.cl Fernanda Hurtado fernanda.hurtado@mail.udp.cl

II. Presentación.

El curso profundiza las principales técnicas multivariantes, que permiten analizar en forma conjunta la interacción de múltiples factores, lo cual da la posibilidad de estudiar de forma más compleja los problemas sociales.

Se enfatizan la comprensión de procedimientos e interpretación de resultados, así como el uso de herramientas computacionales, principalmente en R y RStudio. Se trabajará en la aplicación de los métodos sin necesidad de estudiar en profundidad los fundamentos matemáticos de todos ellos.

Al final del curso el alumno deberá ser capaz de realizar análisis descriptivo de los datos tanto primarios como secundarios y poder seleccionar la técnica estadística adecuada a utilizar cuando se tienen datos multivariantes.

III. Resultados de Aprendizaje

III.1. Resultado general:

Entregar herramientas al alumno para que pueda abordar el análisis de datos desde la estadística multivariante descriptiva e inferencial, elaborando y probando hipótesis y desarrollando modelos explicativos de determinados fenómenos de estudio.

III.2. Resultados específicos:

1. Manejar bases de datos complejas.
2. Plantear y estructurar adecuadamente un problema de análisis multivariante.
3. Identificar el tipo de técnica adecuada, o combinación de ellas, a aplicar en cada situación concreta.
4. Comprender las herramientas utilizada en cada una de las técnicas estudiadas.
5. Aprender a utilizar los programas computacionales necesarios para aplicar los distintos modelos estudiados, particularmente R.
6. Interpretar los resultados del análisis.
7. Redactar correctamente un informe con los resultados del análisis.

IV. Contenidos.

1. Gestión de datos con R

- Introducción a R y Rstudio
- Gestión de datos con tidyverse.
- Cálculo de estadísticos descriptivos.
- Visualización de datos con ggplot.

2. Introducción a los modelos multivariados

- Uso de modelos en ciencias sociales
- Modelos exploratorios y modelos confirmatorios
- Repaso de conceptos estadísticos: covarianza, correlación e inferencia.
- Supuestos de análisis.

3. Formas funcionales de los Modelos de Regresión

- Introducción
- Modelo Logarítmico bivariado
- Modelo Logarítmico Multivariado
- Interpretación de los coeficientes
- Ejemplos de aplicación en R.

4. Análisis Factorial Exploratorio

- Análisis factorial exploratorio y su aplicación en la investigación sociológica. El concepto de varianza y covarianza.
- Diferencias entre el análisis de componente principal y de factor común. Supuestos de la técnica.
- La extracción de factores comunes. Métodos de extracción de factores, criterios de selección del número de factores y rotación.
- La matriz factorial y su interpretación. Evaluación del modelo factorial y cálculo de puntuaciones factoriales.

5. Análisis Factorial Confirmatorio

- Análisis factorial confirmatorio y su aplicación en la investigación sociológica. El análisis confirmatorio frente al análisis exploratorio. Supuestos de la técnica.
- Especificación e identificación del modelo. Estimación de parámetros, evaluación del ajuste y reespecificación.
- Fiabilidad y validez de los modelos de medida. Validez convergente y validez discriminante. Evaluación de la capacidad confirmatoria del modelo.
- Ejemplo de análisis factorial confirmatorio.

6. Análisis de Sendero

- Análisis de sendero y su aplicación en la investigación sociológica. Supuestos de la técnica.
- Especificación del modelo.
- Evaluación del modelo de medición y de la capacidad confirmatoria del modelo.
- Ejemplo de análisis de sendero.

7. Modelos de Ecuaciones estructurales

- Análisis de ecuaciones estructurales y su aplicación en la investigación sociológica.
- La obtención de un modelo de ecuaciones estructurales: la estructura del modelo. Supuestos de la técnica.
- Estimación del modelo empírico.
- Evaluación del modelo.
- Introducción de modificaciones y/o conclusión del análisis. Evaluación de la capacidad confirmatoria del modelo.
- Ejemplos sobre modelamiento de ecuaciones estructurales.

V. Metodología.

Este curso se desarrolla sobre la base de clases expositivas y talleres de aplicación con uso intensivo de software computacionales a partir del sitio web del curso, lo cual permite trabajar cada uno de los distintos tópicos detallados en el contenido.

Se realizarán 2 sesiones semanales, 1 de clases y otra en forma de taller. Durante el curso cada alumno desarrollará trabajos de aplicación.

Adicionalmente se realizarán ayudantías cada 2 semanas aproximadamente, centradas especialmente en la aplicación en R de lo aprendido en la cátedra.

VI. Evaluación.

Se utilizarán distintas formas de evaluación: Tareas, pruebas y trabajos de investigación

- Tareas de uso de R
- 1 prueba solemne (35 %) (Formas funcionales de los Modelos de Regresión y AFE).
- Trabajo final (AFC, Senderos o SEM).

	Ponderación	Fecha (semana)
Tareas uso de R	30%	25 de marzo 15 de abril 27 de mayo 17 de junio
Prueba	35%	29 de abril
Trabajo final	35%	1 de julio

La nota de eximición es 5,5 siempre y cuando no se tengan notas bajo 4,0 en la prueba.
La nota mínima en el examen para aprobar el curso es un 3,5.

Notas:

1. Si un estudiante falta a una evaluación, puede presentar ante la Secretaría de Estudios un certificado médico dentro de los cinco días hábiles posteriores a dicha evaluación, para tener derecho a rendir una prueba recuperativa. Las evaluaciones no rendidas tendrán nota 1.0. Los estudiantes tendrán derecho a un máximo de una prueba recuperativa, dicha prueba se realizará la última semana de clases e incluirá toda la materia del semestre.

Recorrección

Según reglamento, los estudiantes tienen derecho a conocer las notas y la pauta de corrección de todas las evaluaciones en un plazo que no podrá exceder de 10 días hábiles contados desde la fecha desde que éstas fueran rendidas, como también a conocer el resultado de sus evaluaciones solemnes o principales, antes de rendir las siguientes de igual categoría en una misma asignatura.

Los estudiantes pueden imputar la corrección de cualquier evaluación escrita, en un plazo no superior a 5 días hábiles desde la entrega de la nota. Ello, a través de una carta dirigida al secretario de estudios, que debe ir acompañada por la prueba, trabajo o informe evaluado y además, por las correcciones y comentarios que eventualmente haya entregado el/la docente. El académico deberá comunicar al estudiante el resultado de la corrección, fundamentando por escrito sus argumentos, en un plazo no superior a los 5 días hábiles.

VII. Bibliografía

VII.1. Bibliografía obligatoria.

Bibliografía Obligatoria					
Nº	TÍTULO	AUTOR	ED. Y AÑO	EDITORIAL	ISBN
1	ANÁLISIS ESTADÍSTICO MULTIVARIANTE - UN ENFOQUE TEÓRICO Y PRÁCTICO	DE LA GARZA GARCIA, JORGE	2013	McGraw-Hill	978-607-15-0817-1
2	Análisis de datos multivariantes.	Peña, Daniel	2002	McGraw-Hill	8448136101
3	Análisis Multivariado Aplicado .	URIEL Y ALDAS	2005, 1ª Edición	Editorial Thomson	8448136101
4	Análisis Multivariante para las Ciencias Sociales	LEVI J.P. y VARELA J.;	2001	Prentice Hall.	84-205-3727-6
5	Modelos de Ecuaciones Estructurales Cuadernos de estadística	Batista, J.M. y Coenders G.,	2012, 2º ed.	La Muralla	
6	Introducción al Análisis de Regresión Lineal	Montgomery, Peck y Vining	2006, 3º edición.	Continental	970-24-0327-8
7	El análisis factorial como técnica de investigación en Psicología	Ferrando, P. J., & Anguiano-Carrasco, C.	2010	Papeles del Psicólogo, 31, 18-33	
8	Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud.	J. Batista-Foguet, J.M., Coenders, G., & Alonso, J.	2004	Medicina Clínica, 122(Supl.1).	
9	El Path Analysis: conceptos básicos y ejemplos de aplicación.	Pérez, E., Medrano, L. A., & Rosas, J. S	2013	Revista Argentina de Ciencias Del Comportamiento, 5(1), 52-66.	
10	Modelos de ecuaciones estructurales	Ruiz, M.A., Pardo, A., & San Martín, R.	2010	Papeles del Psicólogo, Vol. 31, Núm. 1, pp. 34-45	
11	RStudio para Estadística Descriptiva en Ciencias Sociales. Manual de apoyo docente para la asignatura Estadística Descriptiva	Boccardo, G. y Ruiz, F.	2019	Departamento de Sociología, Facultad de Ciencias Sociales, Universidad de Chile	
12	R para Ciencia de Datos https://es.r4ds.hadley.nz/index.html	Wickham, H	2019		

VII.2. Bibliografía complementaria.

Bibliografía secundaria / complementaria					
Nº	TÍTULO	AUTOR	ED. Y AÑO	EDITORIAL	ISBN
1	Técnicas de Análisis Multivariante para la Investigación Social y Comercial	Díaz de Rada Igúzquiza, Vidal	2002	Ra-Ma	8478975152

2	APPLIED LOGISTIC REGRESSION	HOSMER, DAVID - LEMESHOW,	2º edición	WILEY	0-471-35632-8
3	Métodos Multivariados Aplicados al análisis de Datos	D.E. Johnson	2000	International Thomson Editores	9687529903
4	Análisis multivariable teoría y práctica de la investigación social.	Cea D'Ancona, M. Ángeles	2002	Síntesis Sociológica	84-7738-943-8
5	Análisis factoriales simples y múltiples. Objetivos, métodos e interpretación	B. Escofier y J. Pagès	1994	Servicio Editorial, Universidad del País	8475853838
6	Análisis estadístico para datos categóricos	Ato, M. y López, J. J.	1996	Sintesis	8477383928

VIII. Cláusula Ética

La Escuela establece severas sanciones para las faltas graves que los estudiantes realicen, tales como plagio o falsificación de documentos, señalados en el Reglamento del Estudiante de Pre-grado en el Título XIII de las Conductas Contrarias a los Principios Universitarios. El Comité de Ética es la instancia responsable de revisar y resolver las situaciones que sean contrarias a los principios universitarios, en relación al desarrollo de conductas ilícitas que vicien los procesos de evaluación. Son definidas como conductas que vician una evaluación, todos aquellos actos que implican un rompimiento de los estándares éticos que un estudiante debe resguardar en una evaluación, distorsionando la apreciación de los conocimientos y habilidades reales que el estudiante posee (por ejemplo, copia y uso indebido de documento y/o medios electrónicos, modificación o falsificación de documentos, plagio en trabajos escritos).

Se consideran como fuentes de información que se deben citar, los textos e imágenes en formato electrónico o en papel, en todas sus modalidades: libros, diarios, documentos de trabajo, artículos académicos, etc.

IX. Cronograma de trabajo

11 de marzo	Análisis de datos y uso de R
18 de marzo	Análisis de datos y uso de R
25 de marzo	Introducción a los modelos multivariados
1 de abril	Formas funcionales de los Modelos de Regresión
8 de abril	Formas funcionales de los Modelos de Regresión
15 de abril	Análisis Factorial Exploratorio
22 de abril	Análisis Factorial Exploratorio
29 de abril	Prueba 1
6 de mayo	Receso
13 de mayo	Análisis Factorial Confirmatorio
20 de mayo	Análisis Factorial Confirmatorio
27 de mayo	Análisis de sendero
3 de junio	Análisis de sendero
10 de junio	Modelos de ecuaciones estructurales

17 de junio	Modelos de ecuaciones estructurales
24 de junio	Modelos de ecuaciones estructurales
1 de julio	Repaso

X. NORMATIVA DE PREVENCIÓN Y SANCIÓN DE ACCIONES DE DISCRIMINACIÓN, VIOLENCIA SEXUAL Y DE GÉNERO DE LA UNIVERSIDAD DIEGO PORTALES.

Tal como señala esta normativa:

“La Universidad Diego Portales (UDP) promueve el pluralismo y la independencia crítica de quienes la integran y tienen como uno de sus principales objetivos el desarrollo integral de las personas que forman parte de ella y de la sociedad. En ese marco, la UDP promueve una convivencia reflexiva basada en el reconocimiento a la dignidad de las personas, la libertad de expresión, la libertad sexual, la identidad y expresión de género, y la autonomía.”

Por lo tanto:

“Quienes integren la Universidad tienen la obligación de conocer esta normativa y los protocolos, sitios web y programas que de ella se deriven. Particularmente responsables de su implementación serán las autoridades de la Universidad, quienes tendrán la primera obligación de hacer frente a situaciones que deñen a sus integrantes. Para lograr estos objetivos, se debe realizar la correcta difusión y socialización de la normativa.”

Los detalles de la normativa pueden ser consultados en el siguiente enlace:

https://staticcl1.fidelizador.com/udportales/2022/enero/actualizacion-normativa/RESOLUCION_%20RECTORIA_%20N%3F03_%28APRUEBA_MODIF_NORMATIVA_VIOLENCIA%202022%29OK.pdf