

PROGRAMA DE ESTUDIOS Análisis Avanzado de Datos II

I. Identificación.

Código:	SOC9035
Créditos:	5
Período Académico:	1º semestre 2025
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:	Fernanda Hurtado fernanda.hurtado@mail.udp.cl Francisca Hernández francisca.hernandez_c@mail.udp.cl

II. Presentación.

Este curso profundiza en las técnicas de análisis multivariante, permitiendo examinar de manera integrada la interacción de múltiples factores en el estudio de problemas sociales. Se enfatiza la comprensión de los procedimientos y la interpretación rigurosa de los resultados, con un enfoque aplicado mediante el uso de herramientas computacionales, principalmente R y RStudio.

El curso prioriza la aplicación práctica de los métodos estadísticos sin requerir una profundización en sus fundamentos matemáticos. Al finalizar, el estudiante será capaz de analizar datos primarios y secundarios, identificar la técnica estadística adecuada según el contexto y generar informes que comuniquen de manera efectiva los hallazgos obtenidos.

III. Resultados de Aprendizaje

III.1. Resultado general:

Desarrollar la capacidad de aplicar técnicas de estadística multivariante descriptiva e inferencial para analizar datos, formular hipótesis y construir modelos explicativos en investigaciones sociales, asegurando la correcta interpretación y comunicación de los resultados.

III.2. Resultados específicos:

1. Gestionar bases de datos complejas, asegurando su preparación y estructuración para el análisis.
2. Definir y estructurar problemas de análisis multivariante, formulando preguntas de investigación con sustento metodológico.
3. Seleccionar técnicas estadísticas adecuadas, justificando su aplicación en función del tipo de datos y los objetivos del análisis.
4. Utilizar herramientas computacionales especializadas, aplicando modelos en R y RStudio de manera efectiva.
5. Interpretar de forma rigurosa los resultados del análisis, identificando sus alcances y limitaciones.
6. Redactar informes técnicos y académicos, comunicando los hallazgos de manera clara, estructurada y fundamentada.

IV. Contenidos.

1. Gestión de datos con R

- Introducción a la manipulación de datos con **tidyverse**.
- Cálculo e interpretación de estadísticos descriptivos.
- Visualización de información con **ggplot2**.
- Elementos introductorios de investigación reproducible.

2. Introducción al uso de muestras complejas en R

- Conceptos fundamentales de muestreo complejo.
- Diseño de encuestas y aplicación de ponderaciones.
- Inferencia estadística en muestras complejas.
- Manejo de encuestas con los paquetes **survey** y **svyr**.

3. Introducción a los modelos multivariados

- Rol de los modelos en las ciencias sociales.
- Diferencias entre enfoques exploratorios y confirmatorios.
- Repaso de conceptos estadísticos clave: covarianza, correlación e inferencia.
- Supuestos del análisis multivariante y su importancia.

4. Repaso de modelos de regresión

- Relevancia del control estadístico.
- Regresión lineal múltiple.
- Regresión logística binaria.

5. Análisis Factorial Exploratorio

- Aplicación en la investigación sociológica.
- Comparación entre análisis de componentes principales y factor común; supuestos del método.
- Métodos de extracción de factores, criterios de selección y técnicas de rotación.
- Interpretación de la matriz factorial y evaluación del ajuste. Cálculo de puntuaciones factoriales.
- Introducción al paquete **psych**.

6. Análisis Factorial Confirmatorio

- Diferencias principales con el análisis factorial exploratorio.
- Especificación e identificación del modelo, estimación de parámetros.
- Evaluación del ajuste, reespecificación y validez (convergente y discriminante).
- Ejemplo práctico de análisis confirmatorio.
- Introducción al paquete **lavaan**.

7. Análisis de Sendero

- Fundamentos y aplicación en ciencias sociales.
- Especificación del modelo de sendero y verificación de supuestos.
- Evaluación de la medición y capacidad confirmatoria.
- Ejemplo aplicado en investigación sociológica.

8. Modelos de Ecuaciones Estructurales

- Definición y aplicación de los SEM en ciencias sociales.
- Estructura del modelo y supuestos de la técnica.
- Estimación y evaluación del modelo, introducción de modificaciones.
- Ejemplos de modelamiento de ecuaciones estructurales.

V. Metodología.

El curso se desarrolla a través de clases expositivas y talleres prácticos, con uso intensivo de herramientas computacionales, principalmente R, apoyado en los recursos disponibles en el sitio web del curso. Esta metodología permite trabajar de manera estructurada cada uno de los temas abordados.

Se llevarán a cabo dos sesiones semanales: una clase teórica y un taller aplicado, donde los estudiantes desarrollarán ejercicios y análisis basados en datos reales.

Además, se realizarán ayudantías cada dos semanas, enfocadas en reforzar la aplicación en R de los conceptos abordados en la cátedra, proporcionando un espacio para resolver dudas y fortalecer habilidades prácticas.

VI. Evaluación.

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

- Tareas de uso de R
- 1 prueba solemne
- Trabajo grupal de investigación.

	Ponderación	Fecha (semana)
Tareas uso de R	30%	24 de marzo
		28 de abril
		2 de junio
		16 de junio
Prueba	35%	19 de mayo
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.

Sobre inasistencia a evaluaciones:

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.

Re corrección

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 recorreción, fundamentando por escrito sus argumentos, en un plazo no superior a los 5 días hábiles.

VII. Bibliografía

VII.1. Bibliografía obligatoria.

- **Batista-Foguet, J. M., Coenders, G., & Alonso, J.** (2004). Análisis factorial confirmatorio: su utilidad en la validación de cuestionarios relacionados con la salud. *Medicina Clínica*, 122(Supl. 1).
- **Boccardo, G. & Ruiz, F.** (2019). *RStudio para estadística descriptiva en ciencias sociales: Manual de apoyo docente para la asignatura Estadística Descriptiva*. Departamento de Sociología, Facultad de Ciencias Sociales, Universidad de Chile.
- **Cea D'Ancona, M. A.** (2002). *Análisis multivariable: teoría y práctica de la investigación social*. Síntesis Sociológica. ISBN: 84-7738-943-8
- **Darlington, R. B. & Hayes, A. F.** (2017). *Regression analysis and linear models: concepts, applications, and implementation*. Guilford Press.
- **De la Garza García, J.** (2013). *Análisis estadístico multivariante: un enfoque teórico y práctico*. McGraw-Hill. ISBN: 978-607-15-0817-1
- **Ferrando, P. J. & Anguiano-Carrasco, C.** (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31, 18–33.
- **Gana, K. & Broc, G.** (2019). *Structural equation modeling with lavaan*. Wiley. ISBN: 978-1786303691
- **Levi, J. P. & Varela, J.** (2001). *Análisis multivariante para las ciencias sociales*. Prentice Hall. ISBN: 84-205-3727-6
- **Ministerio de Desarrollo Social y Familia.** (2023). *Metodología de diseño muestral: Encuesta de Caracterización Socioeconómica Nacional (Casen 2022)* (Versión revisada en octubre 2023). Santiago, Chile: Ministerio de Desarrollo Social y Familia. Recuperado de <https://observatorio.ministeriodesarrollosocial.gob.cl/storage/docs/casen/2022/Diseño Muestral Casen 2022 v20oct23.pdf>.
- **Pérez, E., Medrano, L. A., & Rosas, J. S.** (2013). El path analysis: conceptos básicos y ejemplos de aplicación. *Revista Argentina de Ciencias del Comportamiento*, 5(1), 52–66.
- **Ruiz, M. A., Pardo, A., & San Martín, R.** (2010). Modelos de ecuaciones estructurales. *Papeles del Psicólogo*, 31(1), 34–45.
- **Uriel, E. & Aldás, J.** (2005). *Análisis multivariado aplicado*. Editorial Thomson. ISBN: 8448136101
- **Wickham, H.** (2019). *R para ciencia de datos*. [Versión en español]. Disponible en: <https://es.r4ds.hadley.nz/index.html>
- **Zimmer, S. A., Powell, R. J., & Velásquez, I. C.** (2024). *Exploring complex survey data analysis using R: A tidy introduction with {srvyr} and {survey}*. CRC Press.

VII.2. Bibliografía complementaria.

- **Ato, M. & López, J. J.** (1996). *Análisis estadístico para datos categóricos*. Síntesis. ISBN: 8477383928
- **Díaz de Rada Iguzquiza, V.** (2002). *Técnicas de análisis multivariante para la investigación social y comercial*. Ra-Ma. ISBN: 8478975152
- **Escofier, B. & Pagès, J.** (1994). *Análisis factoriales simples y múltiples: objetivos, métodos e interpretación*. Servicio Editorial, Universidad del País. ISBN: 8475853838
- **Johnson, D. E.** (2000). *Métodos multivariados aplicados al análisis de datos*. International Thomson Editores. ISBN: 9687529903
- **Rossee, Y.** (2017). *The lavaan tutorial*. Department of Data Analysis, Ghent University. Retrieved from <https://faculty.washington.edu/matsueda/courses/529/Readings/Lavaan%20tutorial.pdf>

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 vician 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

Fecha	Tema	Lectura Obligatoria	Evaluaciones
10 de marzo	Gestión de datos con R	Boccardo, G. & Ruiz, F. (2019). <i>RStudio para estadística descriptiva en ciencias sociales</i> .	
17 de marzo	Gestión de datos con R	Wickham, H. (2019). <i>R para ciencia de datos</i> .	
24 de marzo	Introducción al uso de muestras complejas en R	Zimmer, S. A., Powell, R. J., & Velásquez, I. C. (2024). <i>Exploring complex survey data analysis using R: A tidy introduction with {srvyr} and {survey}</i> .	Pauta Tarea 1
31 de marzo	Introducción al uso de muestras complejas en R		Entrega Tarea 1

		Ministerio de Desarrollo Social y Familia. (2023). <i>Metodología de diseño muestral: Encuesta de Caracterización Socioeconómica Nacional (Casen 2022</i>	
7 de abril	Introducción a los modelos multivariados		
14 de abril	Repaso de modelos de regresión	Darlington, R. B. & Hayes, A. F. (2017). <i>Regression analysis and linear models</i>	
21 de abril	Repaso de modelos de regresión		
28 de abril	Análisis Factorial Exploratorio	Cea D'Ancona, M. A. (2002). <i>Análisis multivariable: teoría y práctica de la investigación social.</i> Síntesis Sociológica. ISBN: 84-7738-943-8	Pauta Tarea 2
5 de mayo	Receso		
12 de mayo	Análisis Factorial Exploratorio		Entrega Tarea 2
19 de mayo	Prueba 1		
26 de mayo	Análisis Factorial Confirmatorio	Ferrando, P. J. & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. <i>Papeles del Psicólogo</i> , 31, 18–33. Gana, K. & Broc, G. (2019). <i>Structural equation modeling with lavaan.</i> Wiley. ISBN: 978-1786303691	
2 de junio	Análisis Factorial Confirmatorio		Pauta Tarea 3
9 de junio	Análisis de sendero	Pérez, E., Medrano, L. A., & Rosas, J. S. (2013). El path analysis: conceptos básicos y ejemplos de aplicación. Gana, K. & Broc, G. (2019). <i>Structural equation modeling with lavaan.</i> Wiley. ISBN: 978-1786303691	Entrega Tarea 3
16 de junio	Análisis de sendero		Pauta Tarea 4
23 de junio	Modelos de ecuaciones estructurales	Ruiz, M. A., Pardo, A., & San Martín, R. (2010). Modelos de ecuaciones estructurales Gana, K. & Broc, G. (2019). <i>Structural equation modeling with lavaan.</i> Wiley. ISBN: 978-1786303691	Entrega Tarea 4
30 de junio	Modelos de ecuaciones estructurales		Entrega Trabajo Final

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