

El Colegio de México
Centro de Estudios Sociológicos
Maestría y Doctorado en Ciencia Social con especialidad en Sociología

Análisis multivariado y métodos de clasificación de datos

Segundo semestre agosto-diciembre 2024

Curso Obligatorio

Miércoles 8 a 11 horas.

Salón por definir

Isaac Cisneros Yescas

icisneros@colmex.mx

Presentación

El curso tiene como objetivo contribuir a una formación sólida y rigurosa sobre los conceptos y herramientas del análisis multivariado y a los métodos de clasificación de datos; se busca la aplicación e interpretación a problemas concretos de las Ciencias Sociales con el uso del software R.

Se parte de la premisa de que para realizar análisis cuantitativo es necesario la construcción de datos que permitan evaluar las hipótesis que plantean las teorías en la disciplina. Para lo anterior, el curso está constituido a partir de cinco temáticas en las que se presentarán, por parte del profesor, las nociones teóricas fundamentales para comprender cada tema y, posteriormente, se expondrán y replicarán de manera aplicada en R los análisis realizados en artículos académicos que utilicen las técnicas que sustantivamente se muestran los temas que se presentan.

Objetivos del curso

- 1) Dominar los conceptos, nociones y herramientas esenciales para el análisis multivariado y clasificación de datos.
- 2) Conocer y aplicar el análisis de componentes principales, análisis factorial, análisis de conglomerados, de correspondencias y de clusters, así como su fundamento, interpretación y supuestos.
- 3) Conocer los alcances y límites del análisis de componentes principales, análisis factorial, análisis de conglomerados, de correspondencias y de clusters como herramienta de análisis de los fenómenos sociales.
- 4) Aprender a realizar análisis multivariado y clasificación de datos con el paquete de procesamiento estadístico R.

Acerca del curso

Las y los estudiantes evaluarán las distintas implicaciones teóricas de diferentes investigaciones en Ciencias Sociales que incorporen el uso de análisis multivariado y clasificación de datos. Lo fundamental es que incluso en sus tesis puedan desarrollar el uso de las técnicas de las técnicas aprendidas en clase.

En ese sentido, el curso tiene por un lado un aspecto fundamentalmente teórico y, por el otro, un componente muy aplicado. La propuesta consiste en equilibrar tanto la exposición teórica, formal, de la teoría estadística --junto con nociones básicas de lógica de la investigación en ciencias sociales-- y, complementar la parte aplicada con el manejo del paquete de análisis estadístico. En consecuencia, la primera parte de cada clase se destinará a la exposición conceptual por parte del profesor, mientras que la segunda parte tendrá una dinámica tipo taller, orientada a la aplicación de los conceptos y técnicas mediante el uso del software de procesamiento de datos.

Lo anterior exige que, durante todo el curso, las y los estudiantes tengan una participación activa, vinculando los contenidos expuestos, con la orientación práctica aplicada; esto con el objetivo de explotar el potencial de las herramientas estadísticas para el análisis multivariado y clasificación de datos. Finalmente, las bases de datos que se empleen en el curso serán proporcionadas por el profesor.

Evaluación

La evaluación tendrá tres componentes diferentes:

1. Tareas durante todo el curso (15%)
2. Examen intermedio (25%)
3. Examen final (50%)
4. Asistencia (10%)¹

Las *tareas* tienen el objetivo de servir de repaso sobre lo visto en clase para que las y los estudiantes se puedan familiarizar aún más con los contenidos del curso. En caso de que sea necesario se pueden realizar más sesiones, independientemente del número previsto para repasar los contenidos. El objetivo fundamental es que lo expuesto en clase sea apprehendido teórica y de manera aplicada.

El *examen intermedio* tiene dos propósitos fundamentales. Primero, dar seguimiento a lo aprendido en el curso. Segundo, brindar una orientación a los estudiantes y al profesor acerca del avance en el aprendizaje del curso impartido. El examen intermedio consistirá en un

¹ "En el acápite I del artículo 52 del Reglamento General de Estudiantes de El Colegio de México se establece como obligación de los(as) estudiantes regulares la asistencia al menos al 90% de clases durante el semestre. El cumplimiento de esta norma es requisito para ser evaluado(o) en el curso. Se considerará como faltas justificadas únicamente las causadas por enfermedad, trámites burocráticos urgentes y/o fallecimiento de algún familiar cercano, las cuales, deberán ser documentadas y comunicadas oportunamente".

examen sistemático, presencial, sobre los contenidos revisados hasta la primera mitad del curso: Tema 1, Tema 2 y Tema 3 (**fecha prevista 25 de septiembre**).

El tercer elemento de la evaluación será un examen final sobre los temas de todo el curso (***se darán las indicaciones del examen final con anticipada anterioridad a las últimas sesiones del curso***). El objetivo será evaluar alguna hipótesis; realizar las pruebas necesarias y, posteriormente, llevar a cabo el análisis necesario para su evaluación.

Contenido del curso

Presentación del curso (1 sesión)

La primera parte de la sesión 1 estará dedicada a presentar el contenido, la forma de trabajo y, la modalidad de evaluación del curso durante el semestre. También se hará una breve presentación de los(as) integrantes del curso, puntualizando en los intereses temáticos, teóricos y/o en métodos.

Tema 1 - (2 sesiones) - **Introducción al análisis de datos componentes principales**

Objetivo: Dominar los conceptos, nociones y herramientas esenciales para el análisis de de componentes principales.

- a) Tipos de índices: sumatorio simple
- b) Componentes principales
- c) Rotación de componentes: discusión sobre la rotación

Tema 2 – (3 sesiones) – **Introducción al análisis factorial**

Objetivo: Identificar los elementos técnicos y metodológicos del análisis factorial

- a) Extracción del análisis factorial
- b) Tipos de métodos de extracción
- c) Matriz de factores

Tema 3 – (3 sesiones) – **Análisis de conglomerados**

Objetivo: Conocer y aplicar las nociones básicas del análisis de conglomerados

- a) Nociones sobre el análisis de conglomerados
- b) Métodos para el análisis de conglomerados

- c) Agrupamiento de variables

Tema 4- (3 sesiones) – Escalamiento multidimensional y análisis de correspondencias

Objetivo: Profundizar en el análisis de escalamiento multidimensional y en el análisis de correspondencias

- a) Escalamiento multidimensional
- b) Análisis de cluster
- c) Nociones básicas del análisis de correspondencias

Tema 5- (3 sesiones) – Clases latentes

- a) Nociones generales de clases latentes
- b) Ejemplos de clases latentes
- c) Resumen general del curso

Bibliografía

Libros

1. **Anderson, T. W. (2003).** *An Introduction to Multivariate Statistical Analysis* (3rd ed.). Wiley.
2. **Gnanadesikan, R. (1997).** *Methods for Statistical Data Analysis of Multivariate Observations* (2nd ed.). Wiley.
3. **Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010).** *Multivariate Data Analysis* (7th ed.). Pearson.
4. **Johnson, R. A., & Wichern, D. W. (2007).** *Applied Multivariate Statistical Analysis* (6th ed.). Pearson.
5. **Rencher, A. C., & Christensen, W. F. (2012).** *Methods of Multivariate Analysis* (3rd ed.). Wiley.
6. **Tabachnick, B. G., & Fidell, L. S. (2013).** *Using Multivariate Statistics* (6th ed.). Pearson.

7. Urbina, G., & Bárcena, S. (2019). Herramientas de análisis análisis multivariado para la investigación social. Una guía práctica en STATA. Tecnológico de Monterrey

Artículos

1. Krzanowski, W. J. (1988). *Principles of Multivariate Analysis: A User's Perspective*. Oxford University Press.
2. Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press.
3. Seber, G. A. F. (1984). *Multivariate Observations*. Wiley.

Políticas de la clase: La clase funciona bien cuando existen fuertes normas de reciprocidad. Para ello se requiere de empatía, respeto, comprensión, confianza y compromiso mutuos. Por lo tanto, a continuación, se establecen las reglas básicas para el éxito de este curso.

Obligaciones: Las **obligaciones del profesor** con las y los estudiantes son: (1) empezar y terminar puntualmente cada clase; (2) tratar a las y los estudiantes y a sus ideas con respeto; (3) evaluar de manera justa, cuidadosa y puntual; (4) estar disponible para responder preguntas y facilitar el aprendizaje y; (5) fomentar la creación de capital intelectual y social. Las **obligaciones de las y los estudiantes** entre sí y hacia el profesor son: (1) presentarse puntualmente a cada clase; (2) tener un trato respetuoso; (3) completar concienzudamente las tareas asignadas; (4) contribuir de manera dinámica y constructiva a las sesiones y; (5) hacer su propio trabajo.

Asistencia y puntualidad: Quien tenga menos de 90% de asistencia sin justificación no tendrá derecho a calificación final.

Trabajos: Todos los trabajos deberán ser redactados con cuidado. Los errores ortográficos, gramaticales, de puntuación y sintaxis tendrán una penalización en la calificación.

Calendario

Sesión	Fechas
1	7 de agosto
2	14 de agosto
3	21 de agosto
4	28 de agosto
5	4 de septiembre
6	11 de septiembre
7	18 de septiembre
8	25 de septiembre
9	2 de octubre
10	9 de octubre

11	16 de octubre
12	23 de octubre
13	30 de octubre
14	6 de noviembre
15	13 de noviembre
16	20 noviembre