

## Ciencia De Datos Conductual

| ENGIN656   |  | 3 Créditos |
|------------|--|------------|
| Profesores |  |            |
| Ayudantes  |  |            |

### DESCRIPCIÓN DEL CURSO

El auge del Big Data ha llevado a organizaciones y negocios a enfocarse en la toma de decisiones basadas en datos. Sin embargo, la disponibilidad de datos y el desarrollo tecnológico no han ido suficientemente de la mano con las habilidades necesarias para utilizar la información y tomar mejores decisiones. Este curso presenta marcos conceptuales, herramientas y aplicaciones para abordar los sesgos y errores sistemáticos que surgen en los procesos de generación y análisis de datos, con el fin de mejorar la toma de decisiones analíticas en las organizaciones.

### I.- OBJETIVOS DE APRENDIZAJE DEL CURSO

|                                                                                                                                             |  |
|---------------------------------------------------------------------------------------------------------------------------------------------|--|
|                                                                                                                                             |  |
| 1. Conocer sesgos derivados de los procesos de generación y análisis de datos en la toma de decisiones.                                     |  |
| 2. Identificar heurísticas y sesgos cognitivos comunes para optimizar la toma de decisiones basada en datos.                                |  |
| 3. Comprender el rol de los métodos de análisis causal en las organizaciones y su integración con modelos de aprendizaje automático causal. |  |

## II.- CONTENIDOS

| Tema | Contenido                                                                                                                                                         | Lectura Obligatoria |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| 1    | Racionalidad limitada y sesgos en el análisis de datos (datos faltantes, contrafactuales, poblaciones vs. subpoblaciones).                                        |                     |
| 2    | Decisiones basadas en buena/mala evidencia (tipos de validez, factores de confusión, etapas de pruebas A/B, sesgos en causalidad, ética en el análisis de datos). |                     |
| 3    | Introducción al análisis con modelos de machine learning causal.                                                                                                  |                     |
| 4    | Sesgos cognitivos y heurísticas relacionadas en la generación y el análisis de datos.                                                                             |                     |

## III.- METODOLOGÍA, EVALUACIÓN Y NORMATIVA BÁSICA

### 3.1.- Metodología:

- El curso está diseñado para que exista interacción en clases como parte del proceso de aprendizaje.
- Clases expositivas y participativas con discusión de contenidos, lecturas y aplicaciones a través de actividades.
- El curso revisa varios tópicos, se darán las referencias apropiadas para aquellos que deseen profundizar en algún tema en particular.
- Las cátedras serán en español. El material del curso será en inglés.

### 3.2.- Evaluación:

- Quizzes (preguntas cortas sobre conceptos y lecturas) (30%)
- Ejercicios prácticos y casos en clases (40%).
- Examen (30%).

\*Las evaluaciones del curso son presenciales

### 3.3.- Normativa Básica

1. Los/las estudiantes deberán tener al menos un 75% de asistencia en el curso para poder aprobarlo.
2. Las clases serán los días y horas indicados por la dirección del programa.
3. Para cada clase los/las estudiantes deberán haber leído y estudiado anticipadamente la bibliografía correspondiente.
4. La calificación de todas las evaluaciones se hará con nota de 1 a 7.
5. El/la profesor/a se reserva el derecho de agregar, eliminar o reemplazar bibliografía durante el transcurso del programa si así lo estimara conveniente para la buena marcha de la asignatura.

6. La ausencia injustificada de un/a estudiante a una exigencia será calificada con nota 1.
7. Es importante enfatizar que cada estudiante debe asumir su propia responsabilidad en el cumplimiento del programa, especialmente en lo relativo a:
  - a. Estar al día en el desarrollo de la materia y de las diversas indicaciones que entregue tanto el/la profesor/a como la coordinación del curso. Por ejemplo, la ausencia a una sesión de clases no lo exime de las obligaciones académicas señaladas ese día.
  - b. Velar por el fiel cumplimiento de las fechas y plazos establecidos para las distintas actividades de evaluación. Una vez fijadas y conocidas no se procederá a modificarlas.
  - c. Obtener el material de apoyo indicado para la cátedra cuando corresponda.
8. Todos los trabajos que se presenten en el transcurso del programa solo tendrán valor en la medida en que su autor sea capaz de explicar y respaldarlos personalmente. No se aceptan entregas que contradigan lo anterior. Toda justificación médica correspondiente a la inasistencia a una exigencia debe ser presentada a través de los canales regulares establecidos por la Universidad.
9. Toda forma de **copia y/o plagio** está penalizada y en caso de identificarse esta situación, se seguirá el [procedimiento disciplinario respectivo](#).

#### IV.- BIBLIOGRAFÍA

##### Lecturas Obligatorias

- (serán entregadas por el profesor)

##### Lecturas Recomendadas

- The Book of Why: The New Science of Cause and Effect (Judea Pearl y Dana Mackenzie)
- Research Methods in Social Relations (Rick H. Hoyle)
- A / B Testing: The Most Powerful Way to Turn Clicks Into Customers (Dan Siroker y Pete Koomen)
- Field Experiments: Design, Analysis, and Interpretation (Alan S. Gerber y Donald P. Green)
- Decisive: How to Make Better Choices in Life and Work (Chip Heath y Dan Heath)
- Thinking, Fast and Slow (Daniel Kahneman)
- Misbehaving: The Making of Behavioral Economics (Richard H. Thaler)
- Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 113(523), 1228-1242.

\*Syllabus sujeto a cambios