

PROGRAMA DE CURSO CIENCIA DE LOS DATOS

A. Antecedentes generales del curso:

Departamento	Ingeniería Industrial					
Nombre del curso	Ciencia de los datos	Código	IN5244	Créditos	6	
Nombre del curso en inglés	<i>Data Science</i>					
Horas semanales	Docencia	3	Auxiliares	1,5	Trabajo personal	5,5
Carácter del curso	Obligatorio			Electivo	X	
Requisitos	IN4151: Ingeniería de la información, IN4143: Análisis de datos e inferencia causal					

B. Propósito del curso:

Este curso tiene como propósito que el/la estudiante identifique métodos de análisis centrales de la ciencia de datos incluida la regresión, la inferencia causal, aprendizaje estadístico y aprendizaje de máquina.

Para ello, comprenden cómo acceder, tratar, y analizar bases de datos utilizando software estadístico.

El curso tributa a las siguientes competencias específicas (CE) y genéricas (CG):

CE1: Identificar, analizar y diagnosticar los diferentes elementos de los problemas complejos que surgen en las organizaciones, y que son claves para resolverlos.

CE2: Concebir y diseñar soluciones que crean valor para resolver problemas de las organizaciones, utilizando los conocimientos provenientes de la gestión de operaciones, tecnologías de información y comunicaciones, finanzas, economía y marketing.

CE3: Modelar, simular y evaluar problemas de gestión, para encontrar soluciones óptimas, a necesidades de la ingeniería industrial.

CE4: Emplear y aplicar los conocimientos de las distintas disciplinas constitutivas de la ingeniería industrial: gestión de operaciones, tecnologías de información y comunicaciones, finanzas, economía y marketing, en las respectivas áreas funcionales de las organizaciones.

CE5: Desarrollar habilidades para liderar equipos de trabajo, manejando las relaciones interpersonales.

CG1: Comunicación académica y profesional

Comunicar en español de forma estratégica, clara y eficaz, tanto en modalidad oral como escrita, puntos de vista, propuestas de proyectos y resultados de investigación fundamentados, en situaciones de comunicación compleja, en ambientes sociales,

académicos y profesionales.

CG3: Compromiso ético

Actuar de manera responsable y honesta, dando cuenta en forma crítica de sus propias acciones y sus consecuencias, en el marco del respeto hacia la dignidad de las personas y el cuidado del medio social, cultural y natural.

CG4: Trabajo en equipo

Trabajar en equipo, de forma estratégica y colaborativa, en diversas actividades formativas, a partir de la autogestión de sí mismo y de la relación con el otro, interactuando con los demás en diversos roles: de líder, colaborador u otros, según requerimientos u objetivos del trabajo, sin discriminar por género u otra razón.

C. Resultados de aprendizaje:

Competencias específicas	Resultados de aprendizaje
CE2, CE3	RA1: Usa métodos de análisis centrales de la ciencia de datos (regresión, inferencia causal, aprendizaje estadístico y aprendizaje de máquinas), para resolver problemas con datos reales.
CE3, CE4	RA2: Usa bases de datos, utilizando softwares estadísticos, así como evidencia observacional y evidencia experimental o cuasiexperimental, para resolver problemas reales de ciencias de los datos.
CE2, CE5	RA3: Resuelve un problema de ciencia de los datos en una empresa u organismo público, a fin de proponer una solución, considerando métodos de análisis de datos (estadísticos y/o de aprendizaje de máquina), en el contexto del desarrollo de un proyecto.
Competencias genéricas	Resultados de aprendizaje
CG1	RA4: Expone, en una presentación concisa y un informe escrito, la solución propuesta a un problema de ciencia de datos, comunicando de manera efectiva los resultados en base a análisis estadístico.
CG3	RA5: Analiza situaciones profesionales que conllevan dilemas éticos, con relación a acciones reñidas con el uso, privacidad y manejo de los datos e información.
CG4	RA6: Resuelve con sus pares problemas de ciencia de los datos, organizando su quehacer en base a una planificación, definición de tareas.

D. Unidades temáticas:

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
1	RA1	Correlación y causalidad	1 semana
Contenidos		Indicador de logro	
1.1. El auge del uso de datos y la evidencia empírica en la academia y en la industria.		El/la estudiante: <ol style="list-style-type: none"> 1. Reconoce la importancia del uso de los datos en el mundo contemporáneo. 2. Compara entre evidencia observacional y evidencia experimental o cuasiexperimental, estableciendo sus diferencias. 	
Bibliografía de la unidad		[Wooldridge: "Econometric Analysis of Cross Section and Panel Data"]	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
2	RA2, RA3	Causalidad con datos de panel	2 semanas
Contenidos		Indicador de logro	
2.1. Uso de datos de panel.		El/la estudiante: <ol style="list-style-type: none"> 1. Usa datos de panel, incluyendo método de diferencias en diferencias y estudio de eventos en la resolución de problemas de análisis de causalidad. 	
Bibliografía de la unidad		[Wooldridge: "Econometric Analysis of Cross Section and Panel Data"]	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
3	RA2, RA3, RA4, RAA5, RA6	Causalidad con datos de corte transversal	2 semanas
Contenidos		Indicador de logro	
3.1. Uso de datos de corte transversal.		El/la estudiante: <ol style="list-style-type: none"> 1. Resuelve problemas con datos de corte transversal, incluyendo métodos de variables instrumentales y regresión discontinua. 2. Evalúa la calidad de los datos, los que estandariza para la aplicación de modelos de ciencia de datos. 	
Bibliografía de la unidad		[Angrist and Pischke, "Mostly Harmless Econometrics"]	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
4	RA1, RA2, RA3	Aprendizaje estadístico	2 semanas
Contenidos		Indicador de logro	
4.1. Aprendizaje estadístico.		El/la estudiante: 1. Utiliza métodos básicos de aprendizaje estadístico, incluyendo validación, Ridge/Lasso, y métodos de clasificación en problemas de ciencia de los datos.	
Bibliografía de la unidad		[James, Witten, Hastie and Tibshirani, "An Introduction to Statistical Learning with Applications in R"]	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
5	RA1, RA2, RA3, RA4, RA5, RA6	Proyecto sobre un problema de una empresa u organismo público.	7 semanas
Contenidos		Indicador de logro	
5.1. Comprensión del problema. 5.2. Preparación de los datos. 5.3. Aplicación de métodos de análisis de datos (estadísticos y/o de aprendizaje de máquina).		El/la estudiante: 1. Resuelve con sus pares resuelven un problema de una empresa u organismo público, utilizando métodos de análisis de datos (estadísticos y/o de aprendizaje de máquina). 2. Analiza un problema de una empresa pública o privada. 3. Expone con su equipo sobre una solución a un problema de ciencia de los datos, explicando su propuesta, los pasos a seguir, la toma de datos y resultados.	
Bibliografía de la unidad		[9], [10], [11]	

E. Estrategias de enseñanza – aprendizaje:

El curso considera estrategias de enseñanza – aprendizaje activo participativas:

- **Resolución de problemas:** en la primera mitad del semestre los y las estudiantes **resolverán problemas** con datos reales en forma grupal.
- Durante la segunda mitad **resolverán un problema de una empresa u organismo público.**
- Charlas: habrá además charlas de invitados sobre ciencia de los datos. Luego de cada charla habrá una evaluación individual (CTP) sobre sus contenidos.

F. Estrategias de evaluación:

La evaluación es de proceso y contempla:

- Nota promedio casos
- Nota proyecto
- Nota promedio evaluación individual (CTP)

Es importante señalar que, al inicio de cada semestre, el cuerpo académico informará sobre la cantidad y tipo de evaluaciones, así como las ponderaciones correspondientes.

G. Recursos bibliográficos:

Bibliografía obligatoria será publicada en UCursos.

Bibliografía complementaria:

- [1] Wooldridge: "Econometric Analysis of Cross Section and Panel Data".
- [2] Angrist and Pischke: "Mostly Harmless Econometrics".
- [3] James, Witten, Hastie and Tibshirani: "An Introduction to Statistical Learning with Applications in R".
- [4] Hansen: "Econometrics".

Referencias sobre Habilidades de Presentación (presentation skills):

<https://www.skillsyouneed.com/presentation-skills.html>

<https://www.skillsyouneed.com/present/presentation-tips.html>

<https://hbr.org/2013/06/how-to-give-a-killer-presentation>

<https://www.wordstream.com/blog/ws/2014/11/19/how-to-improve-presentation-skills>.

H. Datos generales sobre elaboración y vigencia del programa de curso:

Vigencia desde:	Otoño, 2023
Elaborado por:	Pablo Muñoz, Richard Weber
Validado por:	COMDOC
Revisado por:	Área de Gestión Curricular