

**Facultad de Economía y Negocios / Ingeniería en Control de Gestión,
Mención Ciencia de Datos**

Nombre de la actividad curricular:	Ciencia de Datos I
Código:	ADM0578
Créditos:	7
Carácter:	Obligatorio
Prerrequisitos:	Estadística Modelo de Regresión Múltiple- R y Métodos de Programación II (PI)
Tipo:	Laboratorio
Horas cronológicas de dedicación:	Docencia directa: 3 Trabajo autónomo: 4

I. DESCRIPCIÓN

El objetivo del presente curso es fortalecer y desarrollar las capacidades de los estudiantes en el diseño, implementación y validación de modelos predictivos utilizando **técnicas de aprendizaje de máquina supervisado** implementados en lenguaje **Python** para resolver problemas.

El curso se estructura en base a la aplicación de algoritmos clásicos en la Ciencia de Datos en diversos campos de aplicación, los cuales posibiliten al estudiante tener una mirada integral para la resolución de problemas, empleando su conocimiento previo en estadística y programación en Python.

El curso plantea la revisión de diversos algoritmos de clasificación, tales como regresiones (simple, múltiple y logística), clúster, árboles de decisión, clasificadores bayesianos, máquinas de vectores de soporte, entre otros que posibiliten realizar la clasificación y predicción.

Esta actividad curricular, ubicada en el octavo trimestre del plan de estudios, contribuye al logro de las siguientes competencias del Perfil de Egreso: “Crea soluciones computacionales basadas en algoritmos de procesamiento de grandes volúmenes de intangibles digitales, externa y vinculada con la organización, para reducirla, visualizarla y propiciar la comprensión de esta, al interior de la empresa” y “Desarrolla procedimientos computacionales basados en ciencia de datos, tecnologías de la información y modelos de compliance para la generación y uso ético de la información y conocimiento de la organización”.

II. PROPÓSITOS DE APRENDIZAJE

Al término de esta actividad curricular los estudiantes serán capaces de:

1. Comprender qué es la Ciencia de Datos, su ciclo de vida, los roles involucrados y los distintos tipos y fuentes de datos.

2. Aplicar técnicas de análisis exploratorio y transformación de datos para prepararlos correctamente antes del modelado.
3. Comprender y aplicar algoritmos de agrupamiento para descubrir patrones ocultos en los datos.
4. Aplicar técnicas de reducción de dimensionalidad para interpretar datos de alta dimensión y mejorar la visualización de clusters.

III. CONTENIDOS

Unidad 1. Fundamentos de Ciencia de Datos

- Definición y aplicaciones de la Ciencia de Datos
- Diferencias entre ingeniería, análisis, visualización y ciencia de datos
- Etapas de un proyecto de ciencia de datos
- Tipos de datos: estructurados, semiestructurados, no estructurados

Unidad 2. Preparación y análisis exploratorio de datos (EDA)

- Estadísticas descriptivas y exploración de relaciones
- Visualización multivariada con Seaborn
- Manejo de valores nulos y outliers
- Escalamiento, normalización y codificación de variables
- Generación de datasets aptos para modelos no supervisados

Unidad 3. Modelos de aprendizaje no supervisado

- Principios de clustering y diferencias con modelos supervisados
- Métricas de similitud/distancia
- Implementación de K-Means y técnicas para determinar K óptimo
- DBSCAN para detección de outliers
- Clustering jerárquico: aglomerativo y dendrogramas

Unidad 4. Reducción de dimensionalidad y visualización avanzada

- Motivación para reducir dimensiones
- Interpretación de componentes principales (PCA)
- Combinación con clustering para validar visualmente los grupos
- Casos en los que estas técnicas mejoran el entendimiento del modelo

IV. METODOLOGÍA

El curso se desarrollará en alto porcentaje con una metodología de aprendizaje basado en problemas o bien una metodología de proyectos, propiciando de esta forma, el “aprender haciendo”. Asimismo, realizarán actividades en donde se propicien prácticas como la “programación de pares”, donde se produzca interacción y discusión entre los estudiantes, así como el desarrollo de trabajos en equipos para fortalecer el aprendizaje colaborativo.

El docente generará simulaciones basadas en casos reales para desafiar a los estudiantes a encontrar una solución a la problemática planteada, en dicho contexto tanto talleres como actividades grupales tienen como objeto poner en práctica el aprendizaje del estudiante en base sus conocimientos adquiridos empleando las herramientas y técnicas aprendidas a lo largo de su proceso de aprendizaje.

De igual forma se empleará el enfoque de clases expositivas, las cuales se reforzarán con prácticas de aprendizaje inverso, las cuales esperan generar un efecto de aprendizaje deductivo, lógico y construido de forma interactiva y participativa.

V. EVALUACIÓN DE APRENDIZAJES

El curso tendrá las siguientes evaluaciones:

- Prueba de contenidos (20%)
- 2 talleres prácticos (25% cada una)
- Examen (30%)

VI. RECURSOS PEDAGÓGICOS

Bibliografía Básica:

- Python Data Science Handbook. 2016. Jake Vander Plas.
ISBN: 978-1491912058. Editorial: O'Reilly Media.
- Data Science from Scratch. 2015. Joel Grus.
ISBN: 978-1492041139. Editorial: O'Reilly Media.
- Python for Data Analysis. 2011. Wes McKinney.
ISBN: 978-1491957660. Editorial: O'Reilly Media

Bibliografía Complementaria:

- Python Cookbook, 2013. Brian K. Jones, David Beazley
ISBN: 978-1449340377. Editorial: O'Reilly Media.
- Dive into Python 3, 2009. Pilgrim, Mark.
ISBN: 978-1430224150. Editorial: Apress.

Sesión	Contenido	Fecha	Sala	Tipo
1	- Revisión del programa - Unidad 1. Fundamentos de Ciencias de Datos	23 / 06 / 2025	Lab.	Formativa
2	- Unidad 1. Fundamentos de Ciencias de Datos	30 / 06 / 2025	Lab.	Formativa
3	- Unidad 2. Preparación y análisis exploratorio de datos (EDA)	04 / 07 / 2025	Teams	Formativa
4	- Unidad 2. Preparación y análisis exploratorio de datos (EDA)	07 / 07 / 2025	Teams	Formativa
5	- Prueba	14 / 07 / 2025	Lab.	Evaluativa (20%)
6	- Taller 1	28 / 07 / 2025	Lab.	Evaluativa (25%)
7	- Unidad 3. Modelos de aprendizaje no supervisado	04 / 08 / 2025	Teams	Formativa
8	- Unidad 3. Modelos de aprendizaje no supervisado	08 / 08 / 2025	Teams	Formativa
9	- Unidad 4. Reducción de dimensionalidad y visualización avanzada	11 / 08 / 2025	Lab.	Formativa
10	- Unidad 4. Reducción de dimensionalidad y visualización avanzada	18 / 08 / 2025	Teams	Formativa
11	- Taller 2	25 / 08 / 2025	Lab.	Evaluativa (25%)
12	- Preparación Examen	01 / 09 / 2025	Teams	Formativa
EX	- Examen	08 / 09 / 2025	Lab.	Evaluativa (30%)