

Facultad de Ingeniería y Ciencias
Escuela de Informática y Telecomunicaciones

DESCRIPTOR DE ASIGNATURA

Data Science

1. Identificación de la asignatura:

Nombre de la Asignatura: Data Science	
Códigos: CIT3620	Créditos: 6
Duración: Semestral	Ubicación en el plan de estudios: Semestre 9
Requisitos: CIT2309 Big Data, CIT2313 Inteligencia Artificial	
Sesiones cátedras semanales: 2 cátedras	
Sesiones de ayudantía: 1	

2. Descripción de la asignatura:

Este curso es aplicado, orientado al desarrollo de proyectos interdisciplinarios en ciencia de datos, donde los/as estudiantes diseñan, ejecutan y comunican soluciones basadas en datos para problemas reales de distintas industrias. Se trabaja con estudios de caso de alto impacto, integrando habilidades técnicas con la comprensión del contexto de aplicación, la ética del uso de datos y la colaboración con perfiles no técnicos.

3. Resultados de Aprendizaje:

1. Desarrolla soluciones de ciencia de datos, contextualizadas a problemas reales e interdisciplinarios.
2. Participa en equipos, colaborando con especialistas de otras áreas para entender problemas, formular hipótesis y comunicar resultados.
3. Aplica herramientas de ciencia de datos de manera estratégica, desde la recolección, limpieza de datos hasta el modelado, su interpretación y su visualización.
4. Comunica hallazgos de forma efectiva a públicos técnicos y no técnicos mediante informes, dashboards y presentaciones.

4. Unidades Temáticas:

Unidad 1: Contexto y desafíos en ciencia de datos interdisciplinaria

- Rol del científico de datos en distintas industrias.
- Comunicación efectiva con perfiles no técnicos.
- Ética, sesgo y responsabilidad en la toma de decisiones basadas en datos.

Unidad 2: Estructura del proyecto de ciencia de datos

- Principales marcos metodológicos para proyectos de ciencia de datos.
- Formulación del problema y validación con stakeholders.
- Definición de métricas de éxito y restricciones prácticas.

- Diseño del pipeline de trabajo.

Unidad 3: Estudios de caso por industria

Se desarrollarán ejemplos concretos y datasets para trabajo práctico. Además, se incorporan técnicas de limpieza y procesamiento de datos, análisis exploratorio para la selección de características relevantes y transformaciones que aseguren el cumplimiento de supuestos. Se sugieren temas a analizar:

- **Salud:** predicción de diagnósticos, optimización de recursos hospitalarios.
- **Finanzas:** segmentación de clientes, detección de fraude.
- **Retail/Marketing:** análisis de comportamiento, recomendadores, predicción de compras.
- **Transporte y logística:** optimización de rutas, demanda y congestión.
- **Energía y medio ambiente:** consumo energético, análisis climático.
- **Educación:** predicción de deserción, personalización de contenidos.

Unidad 4: Desarrollo del proyecto final

- Selección de dominio e identificación de problema real
- Desarrollo de etapas que justifiquen la toma de decisiones, asegurando la coherencia metodológica y éxito del proyecto. Desarrollo iterativo en equipo (con entregas parciales)
- Validación, visualización y presentación de resultados

5. Descripción general del método de enseñanza:

El curso se desarrolla como un laboratorio práctico con foco en trabajo en equipo y resolución de proyectos reales. Cada grupo de estudiantes abordará un proyecto aplicado desde una perspectiva interdisciplinaria, en colaboración con expertos o usuarios de otras disciplinas (o mediante simulaciones guiadas).

Se trabajará con herramientas estándar de ciencia de datos (Python, Pandas, Scikit-learn, Streamlit, GitHub, etc.) y se incentivará el uso de recursos abiertos (datasets públicos, APIs gubernamentales, Kaggle, etc.), así como la posibilidad de generar datos propios mediante técnicas de recolección y web scraping, o bien colaborando con una empresa o bureau industrial o servicio público. Las sesiones combinan clases teóricas breves con el desarrollo práctico de código utilizando ejemplos concretos que permiten ilustrar los conceptos aprendidos.

6. Descripción general de la modalidad de evaluación:

La evaluación se centra en el desarrollo progresivo de un proyecto de ciencia de datos con entregas parciales durante el semestre, complementado con trabajos prácticos y dos pruebas solemnes. Se evalúan tanto las competencias técnicas como la capacidad de comunicar resultados a distintos públicos.

Entregas del proyecto final (el proyecto se evalúa en etapas a lo largo del semestre):

- Propuesta de proyecto y definición del problema (10%)
- Avance 1: exploración y procesamiento de datos + visualización (30%)
- Avance 2: modelado y validación (30%)
- Informe final + presentación oral (30%)

7. Planificación de Sesiones:

La planificación de las sesiones dependerá del calendario académico vigente al momento de dictar la asignatura. En cualquier caso, cada docente que la imparta deberá informar a la Escuela (antes del comienzo del semestre) la planificación de sus actividades, evaluaciones, y mecanismos de diversificación de las mismas.

Esta información será informada al estudiantado durante la primera semana de clases.

8. Bibliografía Básica Obligatoria:

1. Cathy O'Neil y Rachel Schutt. Doing Data Science: Straight Talk from the Frontline, O'Reilly Media, 2013.
2. Jake VanderPlas. Python Data Science Handbook: Essential Tools for Working with Data, O'Reilly Media, 2do edition, 2023.
3. Aurélien Géron. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly Media.
4. Eric Ma, Alice Zhao. Communicating Data with Python, O'Reilly Media, 2024.
5. Wes McKinney. Python for Data Analysis, data wrangling with pandas, numpy, and ipython O'Reilly Media.
6. Documentación oficial: Pandas, Scikit-learn, Keras, Streamlit, FastAPI.
7. Casos de estudio y datasets: kaggle.com, data.gov, ourworldindata.org, UCI ML Repository, datos.gob.cl, etc.

PAUTAS ÉTICAS BÁSICAS

El aula es un espacio donde los intercambios buscan generar un clima que potencie el aprendizaje, basado en el respeto y el buen trato. Las diferencias, tanto entre estudiantes, como entre estudiante y docentes, deben abordarse desde este marco de respeto.

La universidad cuenta con dos reglamentos importantes de conocer:

- Reglamento de Convivencia Estudiantil
- Normativa de Prevención y Sanción de Acciones de Discriminación, Violencia Sexual y/o de Género.

Puedes consultar los reglamentos aquí: <https://www.udp.cl/universidad/reglamentos-y-politicas/>

Así también, el plagio es el uso de las ideas o trabajo de otra persona sin el adecuado consentimiento. El plagio puede ser intencional o no. El plagio intencional es el claro intento de hacer pasar el trabajo o ideas ajenas como el suyo propio para su beneficio. El plagio no intencional puede ocurrir si Ud. no conoce el mecanismo adecuado de referenciar la fuente de sus ideas e información. Si no está seguro de los métodos aceptados para referenciar, debería consultar con su profesor, tutor o personal de biblioteca.

El plagio comprobado es una actitud que puede resultar en severas sanciones disciplinarias y/o en la exclusión de la Universidad (Artículo 44, Reglamento del Estudiante de Pregrado).

Elaborado por: Jonathan Frez

Revisado por: Miguel Carrasco

Fecha revisión: Noviembre 2025

Fecha vigencia: Marzo 2026