

# Diplomatura en Ciencia de Datos y Análisis Avanzado

## Información General

Fecha de inicio: 18/08/2026

Frecuencia: Semanal

Carga horaria por clase: 120 min

## Presentación

La Diplomatura en Ciencia de Datos y Análisis Avanzado es un programa interdisciplinario diseñado para capacitar a los participantes en la extracción de conocimiento y percepciones valiosas a partir de datos estructurados y no estructurados. El objetivo principal es formar profesionales capaces de analizar tendencias, identificar patrones y resolver problemas complejos mediante el uso de métodos científicos, algoritmos y sistemas avanzados. Este enfoque permite tomar decisiones estratégicas y abordar desafíos en una amplia gama de sectores.

## Objetivos Generales

Desarrollar competencias en ciencia de datos aplicando métodos avanzados de análisis de datos, visualización, machine learning y modelado estadístico para transformar datos en información útil para la toma de decisiones.

## Objetivos Específicos

- Proporcionar habilidades prácticas a través de proyectos de análisis de datos aplicados.
- Desarrollar los conocimientos fundamentales en estadística, programación y visualización de datos.
- Fomentar el análisis crítico de resultados y la comunicación efectiva de hallazgos.
- Introducir a los participantes en técnicas de modelado predictivo, machine learning y deep learning.
- Capacitar en el uso de herramientas de análisis y procesamiento de grandes volúmenes de datos (Big Data).

## Destinatarios

- Profesionales de diversas áreas que estén interesados en el campo de la Ciencia de Datos.
- Estudiantes con alguna experiencia en programación que deseen identificar de qué se trata este rol en la industria de Tecnologías de la Información (IT).
- Estudiantes que quieran iniciar una carrera en Ciencia de Datos.

## Requisitos Previos

- Conocimientos básicos de estadística (nivel secundario) y programación (Python).

### **Módulo 1: Módulo 1 - Introducción a la Ciencia de Datos y Estadística Descriptiva.**

- Tema 1 - ¿Qué es la ciencia de datos? Ciclo de vida y aplicaciones.
- Tema 2 - Fundamentos de estadística descriptiva: medidas de centralización y dispersión.
- Tema 3 - Distribuciones de frecuencia y representaciones gráficas.
- Tema 4 - Herramientas y recursos disponibles en ciencia de datos.
- Tema 5 - Prácticas: Ejercicios de análisis descriptivo en datasets simples, utilizando herramientas estadísticas básicas.

### **Módulo 2: Módulo 2 - Programación para Ciencia de Datos (Python/R, manejo de datos, librerías básicas).**

- Tema 1 - Introducción a Python/R: sintaxis básica y estructura de datos.
- Tema 2 - Manipulación de datos con pandas (Python) o dplyr (R).
- Tema 3 - Librerías de visualización (matplotlib, seaborn en Python; ggplot2 en R).
- Tema 4 - Introducción a Numpy y Scipy para operaciones matemáticas.
- Tema 5: Prácticas: Ejercicios de manipulación de datos y gráficos básicos utilizando conjuntos de datos predefinidos.

### **Módulo 3: Módulo 3 - Análisis Exploratorio de Datos (AED/EDA).**

- Tema 1 - Técnicas de limpieza y preparación de datos: manejo de valores faltantes y outliers.
- Tema 2 - Transformación de variables y escalado. (Feature Engineering, Normalización y Estandarización).
- Tema 3 - Detección de patrones y relaciones entre variables.
- Tema 4 - Análisis de correlación y técnicas de resumen.
- Tema 5 - Prácticas: Aplicación de técnicas de limpieza y exploración en datasets reales para descubrir insights iniciales.

### **Módulo 4: Módulo 4 - Visualización de Datos.**

- Tema 1 - Principios de diseños y mejores prácticas en visualización de datos.
- Tema 2 - Gráficos básicos y avanzados: histogramas, gráficos de dispersión, mapas de calor.
- Tema 3 - Herramientas avanzadas (Plotly, Seaborn en Python; ggplot2 en R).
- Tema 4 - Introducción a Dashboards interactivos.
- Tema 5 - Prácticas: Creación de gráficos interpretativos y dashboards interactivos con herramientas visuales.

### **Módulo 5: Módulo 5 - Modelado Predictivo I: Regresión y Clasificación**

- Tema 1 - Introducción al modelado predictivo y machine learning.
- Tema 2 - Modelos de regresión lineal y logística.
- Tema 3 - Algoritmos de clasificación básicos: k-Nearest Neighbors, Naive Bayes.
- Tema 4 - Algoritmos de Clusterización.
- Tema 5 - Evaluación de modelos: métricas y validación cruzada.

- Tema 6 - Prácticas: Creación y validación de modelos de regresión y clasificación en problemas de datasets abiertos.

### **Módulo 6: Módulo 6 - Modelado Predictivo II: Técnicas Avanzadas de Machine Learning**

- Tema 1 - Árboles de decisión y Random Forest.
- Tema 2 - Máquinas de soporte vectorial (SVM).
- Tema 3 - Ensamblados de modelos y Gradient Boosting.
- Tema 4 - Optimización de hiperparámetros y técnicas de regularización.
- Tema 5 - Prácticas: Aplicación de técnicas avanzadas en conjuntos de datos complejos y ajustes de modelos para mejorar precisión.

### **Módulo 7: Módulo 7: Deep Learning y Redes Neuronales**

- Tema 1 - Fundamentos de redes neuronales artificiales.
- Tema 2 - Redes neuronales profundas (DNN) y aplicaciones prácticas.
- Tema 3 - Introducción a redes convolucionales (CNN) y redes recurrentes (RNN).
- Tema 4 - Frameworks de deep learning: TensorFlow y Keras.
- Tema 5 - Prácticas: Desarrollo de modelos de deep learning para clasificación de imágenes y series temporales usando TensorFlow / Keras.

### **Módulo 8: Módulo 8: Big Data y Herramientas de Procesamiento (Hadoop, Spark).**

- Tema 1 - Introducción a Big Data y su ecosistema.
- Tema 2 - Fundamentos de Hadoop y sistemas distribuidos.
- Tema 3 - Apache Spark para procesamiento en tiempo real.
- Tema 4 - Conexión con bases de datos y herramientas de consultas.
- Tema 5 - Prácticas: Procesamientos y análisis de datos masivos en Spark, usando ejemplos reales y simulados.

### **Módulo 9: Módulo 9: Proyecto Final Integrador.**

- Tema 1 - Elección de un conjunto de datos complejo y definición del problema a resolver.
- Tema 2 - Desarrollo de un plan de análisis que incluya etapas de limpieza, exploración y modelado.
- Tema 3 - Implementación de visualizaciones, modelos predictivos y conclusiones.
- Tema 4 - Presentación del proyecto final, mostrando hallazgos y recomendaciones.
- Tema 5 - Prácticas: Desarrollo completo del proyecto y defensa ante un panel de evaluación.

## **Bibliografía**

- Material generado por el equipo docente.

