

CERTIFICACIÓN PROFESIONAL

+ POWER BI

# DIPLOMATURA EN INGENIERÍA DE DATOS

MODALIDAD	DURACIÓN
ONLINE, EN VIVO	8 MESES

1 CLASE EN VIVO POR SEMANA



# INSTITUTO

## Data Science

Aval académico: Universidad Tecnológica Nacional Sede Reconquista  
Resolución Nro 290/20.



# Objetivos

- **Entenderás qué es la Ingeniería de Datos** y las responsabilidades diarias de un Ingeniero de Datos. No solo aprenderás teorías; vivirás la emoción de descubrir cómo la ingeniería de datos está cambiando industrias y mejorando vidas.
- **Conocerás el ecosistema de la Ingeniería de Datos** y las herramientas más empleadas. Desde Big Data hasta Machine Learning, te sumergirás en las últimas tecnologías que están revolucionando el campo.
- **Dominarás técnicas de diseño de pipelines de datos** y la extracción, consolidación y carga de datos. Con proyectos prácticos y casos de estudio reales, podrás aplicar lo aprendido y ver los resultados tangibles de tu trabajo.
- **Serás capaz de manejar datos** y fuentes diversas, diseñando arquitecturas robustas para su consumo en distintas áreas de una organización. Imagina ser capaz de analizar grandes conjuntos de datos para predecir tendencias y resolver problemas complejos.

## ¿A quién esta dirigido?

- Entusiastas de la Ciencia de Datos y Emprendedores que buscan habilidades prácticas.
- Ingenieros de Datos y Analistas de Datos que buscan actualizar sus habilidades.
- Estudiantes de Ingeniería, Matemáticas y Estadística interesados en la ciencia de datos.

- Profesionales de Negocios y Finanzas que necesitan trabajar con grandes volúmenes de datos.

**NO SE REQUIERE EXPERIENCIA PREVIA**

# Estructura de la cursada

La estructura del curso ha sido cuidadosamente diseñada para maximizar el aprendizaje y garantizar que todos los estudiantes, independientemente de su nivel previo de conocimientos, puedan avanzar de manera efectiva.

La **Etapa de Nivelación** permite a los estudiantes reforzar sus fundamentos en programación lógica, bases de datos y estadísticas, asegurando una base sólida para el aprendizaje avanzado.

La **Etapa de Introducción** proporciona un marco conceptual integral sobre la ciencia de datos y las herramientas esenciales, estableciendo un contexto claro y relevante.

Finalmente, la **Etapa de Especialización** se enfoca en técnicas avanzadas, librerías y aplicaciones prácticas, permitiendo a los estudiantes aplicar sus conocimientos en proyectos del mundo real y prepararse para oportunidades laborales de alto nivel.

<b>Etapa de Nivelación (optativa)</b>	<b>Programación Lógica</b>	<b>Bases de Datos</b>	<b>Probabilidad y Estadística</b>
<b>Etapa de Introducción</b>	<b>Introducción Conceptual</b>	<b>Introducción a Python</b>	
<b>Etapa de Especialización (circular)</b>	<b>Especialización en clase teórica-práctica</b>		

# Programa académico

## **Nivelación (Optativa)**

- Programación lógica.
- Bases de datos relacionales y no relacionales.
- Probabilidad y Estadística.
- GitHub.

## **Introducción Conceptual (Optativa)**

- Introducción a Ingeniería de Datos
- Introducción a Data Analytics.
- Introducción a Data Warehouse.
- Introducción a Data Mining.
- Introducción a Knowledge Discovery.
- Introducción a Herramientas OLAP y Tableros de comando.
- Repaso de herramientas disponibles.
- Taxonomía de las competencias de un científico de datos.

## **Introducción a Python (Optativa)**

- Descarga e instalación.
- Principales librerías.
- Variables y tipos de datos.
- Listas, tuplas y diccionarios.
- Ejecución condicional.
- Ciclos definidos e indefinidos.
- Manejo de Excepciones.
- Funciones y Generadores.
- Clases y objetos.
- Manejo de archivos y directorios.
- Conexión a archivos planos.
- Conexión a Excel.
- Conexión a Bases de Datos.

## **Etapa de Especialización**

## Introducción a la Ingeniería de Datos (1 clase)

- ¿Qué es la Ingeniería de Datos?
- El ecosistema de datos actual
- Evolución de la Ingeniería de Datos
- Responsabilidades y competencias de un Ingeniero de Datos
- Pantallazo de algunas de las aplicaciones típicas y del software más usado

## El Ecosistema de Datos (2 clases)

- Tipos de datos
- Fuentes de datos
- Trabajando con datos y fuentes diversos
- Introducción a RDBMS y NoSQL
- Data Warehouses, Data Marts, y Data Lakes
- ETL, ELT, y el Data Pipeline

## El Ciclo de Vida en Ingeniería de Datos (2 clases)

- Diseñando repositorios de datos
- Consideraciones de seguridad
- Recolección y almacenamiento de datos
- Introducción a “Data Wrangling”
- Análisis y visualización de datos
- Gobernanza y Compliance

## Bases de Datos Relacionales (2 clases)

- Introducción al formato de datos y arquitectura
- Big Query: Setup y queries más usadas
- Ejemplos de aplicación en Ingeniería de Datos

## NoSQL (2 clases)

- Introducción al formato de datos y arquitectura
- MongoDB: Setup y queries más usadas
- Ejemplos de aplicación en Ingeniería de Datos

## Shell Scripting (1 clase)

- Comandos básicos de Linux y su uso en Ingeniería de Datos
- Data Wrangling: limpieza y consolidación de los datos

## Big Data en Ingeniería de Datos (2 clases)

- Introducción a Apache Spark
- Introducción a Hadoop
- Ejemplos de aplicación

## Introducción a ETL (4 clases)

- ETL y ELT: Fuentes de datos, consolidación, y criterios importantes
- Data Wrangling
- Diseñando data pipelines: consideraciones importantes
- Data pipelines con Shell, Kafka y Airflow

## Visualización de Datos (2 clases)

- Introducción a Tableau y Looker
- Tipos de visualización adecuados para cada tipo de dato, contexto, y audiencia
- Cómo contar historias con datos

## Introducción a Machine Learning (2 clases)

- Definiciones y conceptos básicos
- La relación entre Ingeniería de Datos y Machine Learning
- Preparación y análisis de datos teniendo en cuenta los objetivos buscados
- Algunos algoritmos importantes y cuál usar en cada caso
- Cómo explicar los resultados de Machine Learning a distintas audiencias

# Evaluaciones

- **Proyecto Final:** Presentación y defensa.
- **Examen Final:** Mínimo 60% de respuestas correctas.

# Soporte y Recursos Adicionales

**Chatbot Híbrido:** Podes consultarle sobre contenidos, horarios, grabaciones y links de acceso a las clases en vivo.

**Consultas con docentes:** En la clase en vivo o través del grupo de Telegram que comparten docentes y estudiantes. También se pueden coordinar videollamadas.

**Clases Grabadas:** Disponibles para repaso. No se exige asistencia a las clases en vivo.

**Grupos de Estudio:** Coordinados por tutores.

**Más información en [www.institutodatascience.org](http://www.institutodatascience.org)**

**Consultas: [info@institutodatascience.org](mailto:info@institutodatascience.org) - +549 11 3914 8021**