



**INSTITUTO**  
Data Science

# DIPLOMATURA EN CIENCIA DE DATOS CON R.



```
<p> Avalado por la Universidad  
Tecnológica Nacional. </p>
```



**INSTITUTO**  
Data Science

# Hola!

`<p> ¿Estás buscando potenciar tu  
carrera profesional? </p>`

Comienza hoy mismo tu camino como  
Científico de Datos con clases online  
en vivo.



LA DIPLOMATURA ES AVALADA  
ACADÉMICAMENTE POR LA  
UNIVERSIDAD  
TECNOLÓGICA NACIONAL  
SEDE RECONQUISTA





# 18 DURACIÓN SEMANAS

<p> La Diplomatura en Ciencia de Datos incluye además un curso de nivelación en Bases de Datos, Programación Lógica y Estadística para aquellos estudiantes sin conocimientos previos con clases en vivo. </p>



# APRENDE HACIENDO.

`<p>` Un fuerte componente práctico orientado a proyectos para que aprendas haciendo. Al final de la cursada presentas tu propio proyecto. `</p>`

# CLASES EN VIVO.

Todas las clases en vivo quedan grabadas a disposición de los estudiantes y no se exige asistencia.

También dispones de clases de ediciones anteriores.

Participa en el grupo de Telegram para realizar tus consultas fuera del horario de clases.

Se coordinan videollamadas para atender consultas específicas.

# PROGRAMA ACADÉMICO.



01

Introducción  
Conceptual.



02

Introducción a R.



03

Tests básicos.



04

Árboles de decisión.



05

Reglas de Asociación.



06

Algoritmos genéticos.

# PROGRAMA ACADÉMICO.



07

Método de Simulación  
de Montecarlo.



08

Regresiones.



09

Clusters.



10

Redes Neuronales.



11

Series temporales.



12

Minería de textos.



# PROGRAMA ACADÉMICO.



13

Bayes Ingenuo.



14

Random Forest.



15

Métodos bayesianos  
avanzados.



16

Máquina de soporte  
vectorial.



17

Diseño y construcción  
de un DW.



18

Big Data.



# 01 INTRODUCCIÓN CONCEPTUAL

`<p>` Introducción a la Ciencia de Datos. Niveles a los que opera la ciencia de datos. 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. `</p>`



# 02 INTRODUCCIÓN A R

<p> Variables, vectores y matrices. Dataframes. Paquetes y librerías. Manejo de archivos csv. Conversiones de tipos. Conexión a bases de datos. Ejecución condicional. Bucles. Funciones en R. Medidas estadísticas. Correlaciones. Funciones estadísticas. </p>



# 03 TEST BÁSICOS

<p> Repaso de conceptos de probabilidad y estadística. Test de Hipótesis. Correlaciones AB Test. Cálculo de correlaciones en R con cor. Funciones de distribución en R. Histogramas en R. Gráficos de líneas en R. Gráficos de áreas en R </p>



# 04 ÁRBOLES DE DECISIÓN

`<p>` Algoritmo básico en Excel. División en entrenamiento y prueba. Uso de Rpart y cp. Predicción y valoración de la solución. Uso de Party. Aplicación al problema del call center. `</p>`



# 05 REGLAS DE ASOCIACIÓN

<p> Algoritmo básico en Access. División en entrenamiento y prueba. Uso de arules en R. Ejemplo de una aplicación real votación de reglas. Ejemplo de una aplicación real a datos de ventas en supermercados. </p>



# 06 ALGORITMOS GENÉTICOS

<p> Algoritmo básico en Excel. Discusión del tipo de problemas en los que se aplica. Discusión de otros mecanismos de optimización. Implementación en R del uso de algoritmos genéticos como valores reales. Binarios Permutaciones. </p>



# 07 MÉTODO DE SIMULACIÓN DE MONTECARLOS

`<p>` Algoritmo básico en Excel. Discusión de la utilidad del método. Comparación con el análisis de escenarios. Ejemplo de una aplicación real. Ejemplo simple implementado en R. `</p>`





# 08 REGRESIONES

`<p>` Regresión Lineal. Regresión polinómica. Regresión exponencial y logarítmica. Regresión de dos variables. Cálculo de regresiones en R con `lm`. Cálculo de regresiones logísticas en R con `glm`. `</p>`



# 09 CLUSTERS

<p> Algoritmo básico en Excel. Uso de kmeans. Ejemplo de aplicación real. Otros algoritmos de agrupamiento en R. Ejercicio de aplicación de agrupamiento de mascotas. </p>



# 10 REDES NEURONALES

<p> Algoritmo básico en Excel. División en entrenamiento y prueba. Uso de neuralnet. Predicción y valoración de la solución. Discusión del problema de legibilidad de los resultados desde el negocio. Competencia entre árboles y redes en un caso concreto. </p>



# 1 1 SERIES TEMPORALES

<p> Taxonomía. Separación de componentes. Predicciones. ARIMA implementado en R. Predicción en series con un único período. Predicción en series con múltiples períodos. Predicción en series con períodos variables: renormalización </p>



# 1 2 MINERÍA DE TEXTOS

<p> Clasificación supervisada de piezas de texto. Construcción de una red semántica. Discusión de las posibilidades de reconocimiento de voz. Ejemplos de aplicación real del algoritmo de clasificación. Aplicación a la detección de sentimientos </p>



# 13 BAYES INGENUO

`<p>` Implementación en Excel. Ejemplo de aplicación en R.  
Comparación de 4 algoritmos predictivos (Bayes ingenuo, árboles,  
redes neuronales y regresión logística). `</p>`



# 14 RANDOM FOREST

`<p>` Descripción conceptual del método. Paquete randomForest.  
Ejemplo de aplicación. Comparación con otras técnicas. `</p>`



# 15 MÉTODOS BAYESIANOS AVANZADOS

`<p>` Descripción conceptual del método: Regresión lineal bayesiana. Regresión logística bayesiana. Inferencia bayesiana. Red Bayesiana. Ejemplos de aplicación. Paquete BAS. Paquete brms. Paquete arm. Paquete bnlearn. Comparación con otras técnicas.  
`</p>`





# 16 MÁQUINA DE SOPORTE VECTORIAL

`<p>` Descripción conceptual del método. Paquete e1071. Ejemplo de aplicación. Comparación con otras técnicas. `</p>`



# 17 DISEÑO Y CONSTRUCCIÓN DE UN DW

<p> Diferencias entre los DW y los OLTP. Tipos de datos y soportes. Dimensiones y jerarquías. Estimación de recursos y tiempos según tipos de datos. Ejemplos de staging. Interacción entre equipos (DB-ETL-Soporte-FrontEnd-Usuarios). Ejercicios de diseño de DW. </p>



# 18 BIG DATA

`<p>` ¿Qué es Big Data? ¿Cuándo usar Big Data? ¿Cómo armar una infraestructura para Hadoop? Ejemplo de una aplicación de Map Reduce: wordcount. Cómo evitar Hadoop. Pig Hive. Cassandra. `</p>`

# NUESTROS **DOCENTES.**



## **IGNACIO URTEAGA**

MBA en Dirección de  
Sistemas de  
Información.

## **CESAR PROCOPIO**

Maestría en Finanzas.



**INSTITUTO**  
Data Science



# MUCHAS GRACIAS!

[info@institutodatascience.org](mailto:info@institutodatascience.org)

[www.institutodatascience.org](http://www.institutodatascience.org)



WhatsApp:

+54 9 11 39148021



**INSTITUTO**  
Data Science