



# Propedéutico para el ingreso al diplomado Ciencia de datos

## ÁREA: DIPLOMADOS Y LÍNEAS DE ESPECIALIZACIÓN

### PRESENTACIÓN

La ciencia de datos es una disciplina que se emplea en diversas tecnologías para solucionar problemas en negocios u organizaciones, a través de las matemáticas, la programación y el método científico, que implica el planteamiento de hipótesis, la realización de experimentos y pruebas, el análisis de datos y el desarrollo de modelos predictivos. Este curso se ha diseñado como apoyo a los interesados en capacitarse en la Ciencia de datos.

### OBJETIVO

El participante reconocerá el uso de modelos matemáticos para el estudio de fenómenos aleatorios, identificará los conceptos esenciales de las bases de datos y se introducirá al lenguaje de programación Python.

### BENEFICIOS

Reforzar conocimientos sobre temas elementales de estadística, bases de datos y programación con Python, para presentar el examen de admisión al diplomado Ciencia de datos

### DIRIGIDO A:

Este curso está dirigido a profesionales de las áreas de Matemáticas, Física, Actuaría, Ciencias de la computación, Ingeniería en Computación o afines, que requieran reforzar su perfil de ingreso para presentar el examen de admisión al diplomado Ciencia de Datos.

### CONTENIDO

#### ESTADÍSTICA Y PROBABILIDAD

1. INTRODUCCIÓN AL PENSAMIENTO ESTADÍSTICO
2. MUESTREO Y ESTADÍSTICA DESCRIPTIVA
  - 2.1 Población

- 2.2 Muestra
- 2.3 Recolección de datos
- 2.4 Medidas de localización
- 2.5 Medidas de variabilidad
- 2.6 Datos discretos y continuos
- 3. PROBABILIDAD
  - 3.1 Espacio muestral
  - 3.2 Eventos
  - 3.3 Métodos de conteo
  - 3.4 Probabilidad de un evento
  - 3.5 Probabilidad condicional, independencia y regla del producto
  - 3.6 Variables aleatorias
- 4. DISTRIBUCIONES DE PROBABILIDAD
  - 4.1 Distribuciones discretas de probabilidad
  - 4.2 Distribuciones de probabilidad continua
  - 4.3 Distribuciones de probabilidad conjunta
  - 4.4 Distribuciones comúnmente usadas
    - 4.4.1 Bernoulli
    - 4.4.2 Binomial
    - 4.4.3 Poisson
    - 4.4.4 Normal
    - 4.4.5 LogNormal
    - 4.4.6 Exponencial
  - 4.5 Teorema del límite central
- 5. INTERVALOS DE CONFIANZA
- 6. INTRODUCCIÓN A LAS PRUEBAS DE HIPÓTESIS

## **FUNDAMENTOS DE BASES DE DATOS**

- 1. MODELO RELACIONAL
  - 1.1 Álgebra relacional
  - 1.2 Introducción a SQL
  - 1.3 Integridad de dominio
  - 1.4 Integridad referencial
- 2. DISEÑO LÓGICO
  - 2.1 Generación del modelo relacional
  - 2.2 Dependencias funcionales
  - 2.3 Normalización

### 3. CONCEPTOS DE SISTEMA DE GESTIÓN DE BASES DE DATOS RELACIONALES

3.1 Arquitectura de dos y tres capas

3.2 Propiedades ACID

3.3 Concurrencia

3.4 Seguridad

### 4. SQL

4.1 Lenguaje de definición de datos (DDL)

4.2 Lenguaje de manipulación de datos (DML)

4.3 Lenguaje de control de datos (DCL)

### 5. BASES DE DATOS NOSQL Y CLOUD

5.1 Bases de datos NoSQL

5.2 Bases de datos en la Nube

5.3 Bases de datos Graph

5.4 Bases de datos XML

5.5 Almacenamiento Key-Value y Bases de datos de documentos

5.6 Almacenamiento columnar

5.7 Bases de datos orientadas a objetos

## **INTRODUCCIÓN A LA PROGRAMACIÓN EN PYTHON**

### 1. FUNDAMENTOS DE PROGRAMACIÓN

1.1 Constantes y variables

1.2 Tipos de datos

1.3 Comentarios

1.4 Estructuras de datos

1.5 Operaciones aritméticas

1.6 Operadores lógicos

1.7 Condicionales

1.8 La indentación en Python

### 2. FUNCIONES

2.1 ¿Qué es una función?

2.2 Usar funciones (built-in functions)

2.3 Definir funciones

2.4 Scope

### 3. COLECCIONES

3.1 ¿Qué es un método?

3.2 Listas

3.3 Tuplas

- 3.4 Diccionarios
- 3.5 Conjuntos
- 3.6 Colección de colecciones
- 4. MANIPULACIÓN DE CADENAS DE TEXTO
  - 4.1 Indexación
  - 4.2 Concatenación
  - 4.3 Formateo
  - 4.4 Splitting
  - 4.5 Join
  - 4.6 Replace
  - 4.7 Slicing
- 5. BUCLES
  - 5.1 For Loop
  - 5.2 While Loop
  - 5.3 Break
  - 5.4 Continue
  - 5.5 Bucles anidados
- 6. MANIPULACIÓN DE COLECCIONES
  - 6.1 Indexado
  - 6.2 List-comprehensions y tuple-comprehensions
  - 6.3 Iterables y for loops
  - 6.4 Colecciones mutables e inmutables
- 7. PARADIGMAS DE PROGRAMACIÓN
  - 7.1 Principales paradigmas
    - 7.1.1 Procedural Programming
    - 7.1.2 Functional Programming
    - 7.1.3 Object-oriented programming
  - 7.2 Clases y objetos
  - 7.3 Encapsulación
  - 7.4 Abstracción
  - 7.5 Polimorfismo
  - 7.6 Herencia
- 8. TÓPICOS AVANZADOS
  - 8.1 Recursión
  - 8.2 Sorting y Searching
  - 8.3 Big-O notation y complejidad
  - 8.4 Pilas y colas
  - 8.5 Listas enlazadas

## **DURACIÓN**

40 horas

## MODALIDAD

A distancia

## FORMA DE TRABAJO

- El curso y la interacción grupal se llevarán a cabo mediante sesiones síncronas a través de una herramienta de videoconferencia, en el horario y el periodo programados.
- Las temáticas serán expuestas a través de presentaciones y diversos medios.
- Se dispondrá de materiales y actividades diseñadas especialmente para el curso.
- Las sesiones no serán grabadas, por lo que el usuario que no participe en alguna es responsable de ponerse al corriente a través de los contenidos dispuestos para tal fin.

## EVALUACIÓN

- La calificación mínima aprobatoria es de 8.0
- Se requiere el 80% de asistencia
- Los elementos de evaluación serán indicados por el instructor

## RECURSOS INFORMÁTICOS NECESARIOS

- Cuenta de correo electrónico
- Conexión a Internet

### HARDWARE:

- Procesador Intel Core i3 o superior
- Memoria RAM instalada de 4GB como mínimo recomendable
- Disco duro: 8 GB
- Pantalla con una resolución de 1024x768 o superior
- Micrófono
- Cámara
- Bocinas o audífonos multimedia

### SOFTWARE:

- Microsoft Windows 10 o superior
- Un navegador (Chrome, Microsoft Edge, Mozilla Firefox, etc.), es recomendable utilizar las versiones más actuales
- [Adobe Acrobat Reader](#) u otro software libre para abrir los archivos PDF
- Instalar la aplicación de Zoom, la cual se puede descargar de: <https://zoom.us/support/download/>
- Microsoft SQL Server 2019 Developer Edition

Mayo 2022

101530/ANC

