Centro de Investigación en Computación Subdirección Académica Departamento de Diplomados y Extensión Profesional

# "PROGRAMACIÓN PYTHON CON APLICACIONES AL ÁMBITO CIENTÍFICO" (35 HORAS)

#### **DIRIGIDO A.**

Estudiantes, profesionistas y público en general que se enfrenten día a día al manejo, análisis, visualización o automatización de datos y deseen aumentar su productividad gracias al lenguaje de programación de Python y sus librerías científicas.

#### **OBJETIVO**

Formar estudiantes que comprendan y dominen el flujo de trabajo del análisis de datos que consiste en Analizar, Visualizar y Automatizar datos desde su adquisición desde bases de datos, archivos de excel, servicios web, etc. Hasta su visualización y automatización mediante reportes gráficos, de texto y de tablas que concentren información descriptiva e indicativa basada en técnicas de estadística y machine learning. Preparar a los estudiantes para el uso de herramientas y librerías de Python para atacar problemas en áreas como Marketing Avanzado, Experimentación en Laboratorios, Ciencias Sociales, Geología y Cartografía, entre otras.

#### **PERFIL DE LOS PARTICIPANTES**

Aunque el curso puede ser tomado por el público general con conocimientos elementales de cualquier licenciatura, se recomienda tener nociones de programación, estadística y estar actualizado en temas de ciencia de los datos para un mejor aprovechamiento del curso. Se debe resaltar que el objetivo del curso no es enseñar a programar, por lo que el participante debe estar consciente que si desea aprender cómo programar en Python debería tomar en su lugar el curso de Python Básico.

# **PERFIL DEL INSTRUCTOR**

El instructor cuenta con un profundo conocimiento de los siguientes ejes sin excepción alguna:

- 1. Amplio dominio en el lenguaje de programación de Python: uso de colecciones (listas, tuplas, diccionarios), dominio de funciones y módulos, programación funcional, modulas y orientada a objetos, entre otros conceptos del lenguaje.
- **2.** Experiencia y dominio utilizando librerías científicas de Python: uso de numpy, matplotlib, pandas, sklearn, keras, entre otras.

Centro de Investigación en Computación
Subdirección Académica
Departamento de Diplomados y Extensión Profesional

**3.** Conocimientos en análisis de datos (estadística y machine learning), visualización de datos (graficación de datos categóricos, escalares, geoespaciales, etc.) y automatización de datos (generación de reportes automatizados en pdf, excel, bases de datos y reportes gráficos).

## **CONTENIDO:**

## 1. Introducción al Lenguaje de Programación de Python:

- 1.1 Conceptos fundamentales de la Programación
- 1.2 Variables y tipos de datos
- 1.3 Colecciones: Listas, tuplas y diccionarios
- 1.4 Funciones
- 1.5 Módulos
- 1.6 Clases y objetos\*

#### 2. Análisis de datos:

- 2.1 Introducción al análisis de datos
- 2.2 Flujo de análisis: Adquisición, Pre-Post Procesamiento,

Visualización y Automatización

- 2.3 Adquisición de datos: Vectores, Matrices, Series y DataFrames
- 2.4 Pre-Procesamiento de datos: Limpieza de valores inválidos,

reparación y reestructuración de datos

2.5 Procesamiento de datos: Ejes de datos,

tipos de ejes (categóricos, escalares, geoespaciales),

construcción de descriptores e indicadores, estadísticos elementales,

técnicas de aprendizaje

- 2.6 Post-Procesamiento: Construcción de tablas descriptoras y matrices indicadoras
- 2.7 Visualización de datos: Reportes de texto, reportes gráficos y reportes de tablas
- 2.8 Automatización de datos: Construcción de módulos reportadores

# 3. Visualización de datos:

Centro de Investigación en Computación Subdirección Académica
Departamento de Diplomados y Extensión Profesional

- 3.1 Introducción a la visualización de datos
- 3.2 Visualización en 1-eje: Categorías (conteos y porcentajes),

Escalares (Histogramas y estadísticos),

Geospaciales (Puntos en mapas)

3.3. Visualización en 2-ejes: Categoría vs Categoría,

Escalar vs Escalar,

Geoespacial vs Geospacial,

Otras combinaciones

3.4 Visualizaciones superiores: Combinaciones de 3-ejes en 2 dimensiones

# 4. Automatización de datos:

- 4.1 Introducción a la automatización de datos
- 4.2 Segmentaciones automáticas
- 4.3 Tablas estadísticas auto-generadas
- 4.4 Gráficas automatizadas
- 4.5 Reportes PDF con texto y gráficas
- 4.6 Reportes en Excel\*

# 5. Librerías científicas de Python

- 5.1 Introducción a las librerías científicas
- 5.2 Numpy y Pandas
- 5.3 Matplotlib y Seaborn
- 5.4 Folium y Basemap
- 5.5 Tweepy y Scrapy

#### 6. Prácticas del curso

- 6.1 Introducción a las prácticas del curso
- 6.2 Revisión y evaluación de las prácticas

## 7. Proyecto Final

- 7.1 Introducción al proyecto final
- 7.2 Revisión y evaluación del proyecto final