

## PYTHON PARA LA CIENCIA DE DATOS

Duración: 24 hrs.

Código:

### Curso:

### Descripción del curso

El curso abordará el procesamiento y análisis de datos con Python. Para ello se realizará un estudio profundo de las principales librerías para el manejo avanzado de las estructuras de datos. Finalmente se desarrolla un pipeline de analítica predictiva, usando un algoritmo básico de clasificación y la librería más popular de Python (Sklearn)

### Dirigido a:

- Estudiantes.
- Profesionales.

### Objetivos:

El Participante al finalizar el curso será capaz de:

- \* Utilizar el lenguaje PYTHON para ciencia de datos.

### REQUISITOS MÍNIMOS

Conocimientos previos de programación (deseable).



## CONTENIDO

### Sesión 1

- Manejo avanzado de Estructuras de Datos
- Definición y operaciones con objetos Numpy
- Lectura de datos con Pandas
- El concepto de Dataframe y su API
- Operaciones con DataFrames en pandas
- Cláusula GroupBy
- El método agg
- Agregadores básicos: SUM (), avg(), median(), min (), Max()

### Sesión 2

- Operaciones Avanzadas con DataFrames
- Filtrado, clasificación y selección de dataframes
- Ensamblado de dataframes
- Transformaciones Avanzadas (stack, unstack, pivot, Melt)
- Dataframes "tidy"

### Sesión 3

- Visualización de Datos
- Introducción a matplotlib
- Gráficos básicos de Pandas con Matplotlib
- Seaborn para gráficos estadísticos
- Creación de gráficos interactivos con Chartify
- Reportes con Markdown



## CONTENIDO

### Sesión 4

- Machine Learning
- Introducción a scikit-learn
- Tuberías de scikit-learn

