



Universidad del
Rosario

A VANCE
Educación Continua

CURSO | REMOTO

ANÁLISIS DE DATOS CON PYTHON

CONTENIDO TEMÁTICO

Módulo 1. **CONCEPTOS BÁSICOS DE PYTHON**

Python es una de las herramientas de software libre más utilizadas en el mundo por su versatilidad, poder gráfico y gran poder de cálculo. En esta primera parte se presentarán diferentes intérpretes del lenguaje y se trabajará con de sus algunos objetos.

- Identificación del lenguaje Python.
- Objetos básicos.
- Cargue de archivos y manipulación de objetos.

Módulo 2. **DEPURACIÓN Y ANÁLISIS EXPLORATORIO DE DATOS. (SCIPY, NUMPY, PANDAS, MATPLOTLIB, SEABORN)**

Una de las primeras dificultades encontradas cuando se analiza una gran cantidad de información es el determinar cuál de ella es verdaderamente importante, representativa y significativa. En esta primera parte se proporcionarán las herramientas para depurar, seleccionar y caracterizar la información que debe ser utilizada en la construcción de modelos.

- Depuración de datos.
- Análisis exploratorio de datos, gráfico y formal.
- Limpieza de datos.
- Análisis estadístico de la información.
- Análisis de información no normal to atípica.

Módulo 3. **ESTIMACIÓN Y ANÁLISIS DE MODELOS SUPERVISADOS. (SCIKIT-LEARN, SKLEARN, STATSMODELS)**

Después de haber seleccionado el conjunto de datos relevante, significativo y representativo, se presentan diversos modelos que permiten la generación de nuevo conocimiento de la situación, aportando de esta forma indicadores e índices que caracterizan cada una de las partes del problema a analizar.

- Modelos de regresión múltiple.
- Modelos Logit, Probit.
- Modelos multinomiales.

Módulo 4. **CONSTRUCCIÓN Y ANÁLISIS DE MODELOS DE REDUCCIÓN DE DIMENSIONES. (SKLEARN.DECOMPOSITION)**

En muchas ocasiones el fenómeno que desea analizarse está descrito por varias variables. El objetivo de la reducción de dimensiones es tratar de obtener un conjunto de variables menor al anterior, máximo de dos o tres variables, que resuman la mayor cantidad de información o variabilidad de las variables originales y faciliten el acercamiento al problema en cuestión.

- Construcción y análisis de Componentes Principales.
- Construcción y análisis Factorial de Correspondencias.
- Construcción y análisis de Correspondencias Múltiples.
- Construcción y análisis de indicadores e índices.

Módulo 5. **MODELOS DE TOMA DE DECISIONES. (KMODES. KPROTOTYPES, KMEANS)**

En muchas situaciones es necesaria la descripción detallada de un conjunto de individuos, o la agrupación de estos, o la generación de diferentes escenarios que faciliten la toma de decisiones. Aquí se presentan algunos modelos que facilitan este tipo de decisiones.

- Construcción y análisis de Árboles de decisión.
- Construcción y análisis de Clúster y segmentación.
- Redes Neuronales.

- Todos nuestros programas incluyen un certificado de asistencia. Los participantes que cumplan con la asistencia mínima requerida y los requisitos académicos, recibirán su certificado.
- La Universidad del Rosario se reserva el derecho de modificar el equipo académico de los programas de Educación Continua que estén anunciados en la programación, así como modificar el orden temático presentado en este programa, garantizado que se abordará la totalidad de temas propuestos.
- La Universidad del Rosario podrá modificar las fechas de desarrollo de los programas o de las sesiones de un programa en ejecución, en situaciones que así lo amerite. La notificación a inscritos o a los participantes de los programas se podrá hacer sin un plazo de tiempo previo estipulado, pero siempre intentando mantener un mínimo de horas previas para ello.
- Los programas de Educación Continua son educación informal, no conducen a título profesional.