

Nombre de la actividad formativa

Python – Introducción al lenguaje Python para ciencia de datos.

Modalidad	Estrategia metodológica	Tipo de capacitación	Duración
Por participación	Capacitación no presencial-virtual *	Curso	16.0 horas efectivas **

* Sesiones telepresenciales y simultáneas a través de la Plataforma Zoom.

** Cuatro sesiones semanales y simultaneas de 3 horas y 1 hora semanal de trabajo a distancia.

Descripción:

Actualmente Python es uno de los lenguajes de programación más utilizados por los Analistas y Científicos de Datos. Nos permite realizar análisis exploratorios de datos, visualizaciones, modelos de machine learning, deep learning y mucho más.

No importa si no tienes una gran experiencia en programación ya que el curso inicia desde el nivel básico y te llevará por los fundamentos de Python, luego aprenderás manipulación de datos y por último la construcción de visualizaciones de datos.

Dirigido a:

Todos aquellos profesionales del área de las ciencias económicas (administradores, financistas, recursos humanos, mercadólogos, entre otros) que quieran aprender las bases de la analítica de datos y ciencia de datos para llevar a sus empresas a otro nivel y convertirlas en lo que hoy como conocemos como "**organizaciones data driven**".

Requisito:

Es un curso introductorio a Python por lo que no se tienen requisitos previos.

Objetivo general

- Conocer el entorno de programación Python
- Conocer los diferentes IDE más utilizados en la analítica de datos sobre el lenguaje Python.
- Realizar procesos de ETL conectándose desde bases de datos, archivos planos, csv, Excel y otros.
- Conocer los fundamentos para realizar procesos de análisis EDA
- Conocer las principales librerías de Python para la generación de plots.

Objetivos específicos	Contenido temático	Actividades a realizar (tarea semanal)
<p>Semana 1</p> <p>Introducción a Python: Conociendo el entorno de programación.</p>	<ul style="list-style-type: none"> • Conceptos básicos de Python. • Conocer los diferentes IDE para programación en Python • Introducción al lenguaje de programación Python. • Definición de variables y números • Conversión de tipos de datos. • Manejo de cadenas de texto • Manejo de operadores: operadores aritméticos, de asignación, de comparación, lógicos, de identidad y de pertenencia. • Colecciones de datos, listas, tuplas, conjuntos, diccionarios. • Condiciones if. • Bucles for, while 	<p>Actividad 1. Tiempo: 1 hora</p> <p>El estudiante deberá desde un archivo plano extraer las palabras claves para realizar un análisis de sentimientos.</p>
<p>Semana 2</p> <p>Manejo de archivos y base datos</p>	<ul style="list-style-type: none"> • Manejo de archivos tipo json • Expresiones regulares • Conexión a bases de datos con Python 	<p>Actividad 2. Tiempo: 1 hora</p> <p>El estudiante deberá crear un modelo de datos relacional en algunas de las siguientes herramientas:</p> <ol style="list-style-type: none"> 1. SQLite. 2. MariaDB.
<p>Semana 3</p> <p>Bases de datos y fundamentos para el análisis de datos</p>	<ul style="list-style-type: none"> • Introducción al lenguaje de consultas SQL. • Generación de consultas SQL desde Python • Modulo Numpy. • Módulo Pandas • Usos de los Dataframe 	<p>Actividad 3. Tiempo: 1 hora</p> <p>El estudiante deberá a partir del modelo generado en la semana 2 realizar una serie de extracciones de datos desde la base de datos y que se visualicen en Python mediante estructuras tipo DataFrame</p>
<p>Semana 4</p> <p>Análisis estadísticos en Python.</p>	<ul style="list-style-type: none"> • Limpieza y transformación de datos. • Generación de un análisis EDA. • Generación de plots. 	<p>Actividad 4. Tiempo: 1 hora</p> <p>Proyecto final, el estudiante deberá entregar un proyecto con un set de datos seleccionado donde aplique todo lo aprendido durante el curso.</p>

Objetivos específicos	Contenido temático	Actividades a realizar (tarea semanal)
		Se califica: <ol style="list-style-type: none"> 1. Set de datos utilizado. 2. Modelado de datos. 3. Limpieza de datos. 4. Selección de tipos de gráficos adecuado según el tipo de datos.

Facilitador: Ing. David Quirós Tencio

Ingeniero en Sistemas, con una licenciatura en gestión de la calidad del software y con especializaciones en las áreas de Big Data, Business Intelligence, y Machine Learning. Cuento con más de 12 años de experiencia en sistemas bancarios participando en importantes proyectos para empresas como lo son Banco Popular, ICE, JASEC, entre otros, desempeñando funciones como analista de sistemas, consultor de bases de datos, líder de proyectos y actualmente como gerente de proyectos para la empresa Novacom.

Posee experiencia en base de datos Oracle, SQL Server, MongoDB, RedisDB, Cassandra, lenguajes de programación Python, R, .net, PL/SQL, transact-sql, frameworks y herramientas para análisis de datos como pyspark, databricks, herramientas de visualización como Power Bi y Tableau, experiencia en servicios en la nube principalmente en Azure y conocimientos en el manejo de tecnologías para BI, Big data, Machine Learning

Educación:

- Ingeniería en Informática. Universidad Internacional de las Américas, 2004
- Scrum Master, International Scrum Institute, 2017
- Licenciatura en Gestión de la Calidad del Software. Universidad Fidelitas, 2019 (Graduado con honores)
- Especialización en Big Data, Instituto Tecnológico de Costa Rica, 2019(Graduado con honores)
- Especialización en Inteligencia de Negocio, Escuela de inteligencia de Negocios, Colombia,2020
- Programación C# en Visual Studio 2010. Grupo Asesor en informática.
- Programación MVC con HTML5. NOVACOMP
- General Product Support Assessment (V3.1). Oracle
- Oracle Database 12c Presales Specialist Assessment. Oracle
- Oracle Databas 12c Presales topics. Oracle
- Oracle Database 12c Sales Requirements Gathering and Discovery. Oracle
- Oracle Database 12c Objection Handling and competition. Oracle
- Oracle Database 12c Product Area and Features and Functions. Oracle
- Oracle Database 12c Target Customer Profile. Oracle
- Oracle Database 12c Market trends and Business Drivers. Oracle