

PYTHON FOR DATA ENGINEERING



BROCHURE 2025

Presentación

La creciente demanda de datos en proyectos de inteligencia de negocios, ciencia de datos e inteligencia artificial exige soluciones tecnológicas robustas para transportar y procesar información desde diversas fuentes hasta los puntos donde será explotada. Por ello, se requieren profesionales capaces de planear e implementar ETL's y data pipelines que garanticen datos oportunos, precisos y confiables.

El Curso Python for Data Engineering te permitirá emplear este demandado y versátil lenguaje para la implementación de soluciones ETL, destinadas al transporte y procesamiento de datos entre diversas fuentes, empleando en el proceso sus principales librerías, método y componentes.



Duración: 24 horas académicas





Objetivo General

Aprende a emplear el lenguaje Python para la implementación de soluciones ETL, destinadas al transporte y procesamiento de datos entre diversas fuentes, empleando en el proceso la librería Pandas, los DataFrames y sus diversos métodos.



Objetivos específicos

- Aprende sobre los fundamentos de los ETL.
- Aprende los principales comandos de este lenguaje de programación, con énfasis en aquellos aplicables a la implementación de ETL.
- Aprende a emplear el lenguaje Python para mapear bases de datos relacionales a través de las herramientas denominadas ORM.
- Aprende a emplear la famosa librería Pandas de Python, para la manipulación de conjuntos de datos en elementos denominados DataFrames.

Malla curricular



Malla curricular

1. Fundamentos de ETL con Python

- ETL. Definición y herramientas.
- Herramientas Python para ETL.
- Python y sus entornos de ejecución.

2. Introducción a Python

- Manejo de excepciones e instrucciones.
- Tipos de datos en Python.
- Creación de Programas en Python.
- Interactuando con el OS.
- API. Definición y librerías para extraer datos.
- Taller: Consulta de datos desde un API.

3. Object Relational Mapper

- ORM. Definición, ventajas de su uso.
- Tipos de ORM en Python.
- SQLAlchemy. Definición y características.
- Taller: Creación de un Engine con SQLAlchemy.
- Taller: Conexión a base de datos con SQLAlchemy.

4. Pandas, Series y DataFrames

- Pandas. Definición, carga en Python.
- Pandas Series. Características y uso de vectores.
- Operaciones con Series. Búsquedas, Slicing, operaciones aritméticas, tipos de datos.
- Pandas DataFrames. Características y uso de DataFrames. Diferencias respecto a Series.
- Operación con DataFrames. Creación, descripción, visualización.
- Operaciones de agrupación.
- Agrupaciones directas y por Agregación simple y múltiple (varios campos).
- Guardar DataFrames en archivos planos (Json y CSV) y base de datos (MySQL).
- Taller: Carga de datos de un API, procesamiento, y descarga en una base de datos.
- Taller: Carga de datos desde un archivo plano, procesamiento y descarga en una base de datos.

Certifícate con nosotros



- El participante podrá obtener un certificado a nombre de DMC INSTITUTE, una vez aprobado con la nota mínima establecida y asistir a clases con el mínimo de horas requeridas
- Una vez cumplidos los requerimientos el certificado será emitido de forma virtual



¿Por qué elegirnos?



Somos los primeros en Perú en apostar por el desarrollo de profesionales y empresas en data & analytics con más de 15 años de experiencia.



Las empresas worldclass de Latam confían en nosotros para acompañarlas en su transformación hacia el enfoque data driven.



Nuestros docentes son destacados expertos en data & analytics que lideran equipos de alto rendimiento en las empresas más grandes de Latam.



Nuestra metodología "Aprende haciendo" ha logrado que nuestra comunidad de +25K profesionales en todo Latam mejoren su situación laboral.



Tenemos el portafolio más completo con +150 capacitaciones sincrónicas y asincrónicas que se ajustan a diferentes perfiles y niveles de conocimiento.





Visita nuestra web

www.dmc.pe

Síguenos en:

