

CURSO ANÁLISIS DE DATOS CON HERRAMIENTAS DE BIG DATA

Presencial **36** Horas

Fundación Universidad de América | Vigilada MinEducación



ISO 9001:2015
ISO 45001:2018
ISO 14001:2015
BUREAU VERITAS
Certification





PRESENTACIÓN

Este curso proporciona a los estudiantes herramientas clave para el análisis de grandes volúmenes de datos, superando las limitaciones de las hojas de cálculo tradicionales. Con un enfoque 75% práctico y 25% teórico, se trabajará con Python, Pandas y Plotly en un entorno 100% en la nube mediante Google Colaboratory, lo que facilita el acceso desde cualquier dispositivo.

A lo largo del programa, los participantes aprenderán a automatizar tareas, manipular grandes conjuntos de datos y generar visualizaciones interactivas, optimizando el análisis y la toma de decisiones en proyectos de Big Data.

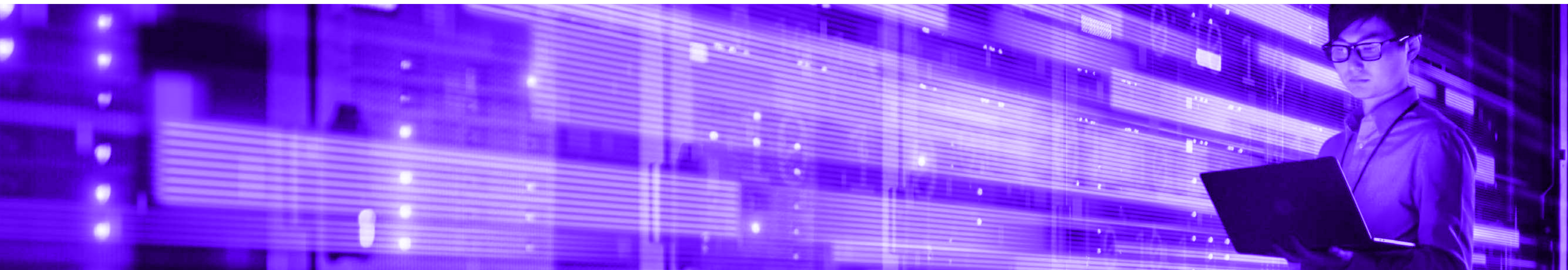
Además, se explorará la evolución del Big Data y su impacto global, brindando fundamentos sólidos para su aplicación en investigación y desarrollo profesional.

OBJETIVO GENERAL

Brindar herramientas a los estudiantes para el manejo, análisis y visualización de estructuras de datos con herramientas del lenguaje de programación Python.

OBJETIVOS ESPECÍFICOS

- Hacer una introducción a los criterios del panorama del Big Data y cómo estas herramientas están al alcance de cualquier persona dado su contexto open source.
- Hacer un repaso general del lenguaje de programación Python en el entorno de trabajo de Google Colaboratory.
- Enseñar el uso de la librería pandas, para cargar datos de diferentes orígenes, filtrar, transformar y guardar la información.
- Enseñar el uso de la librería plotly para la presentación interactiva de resultados.





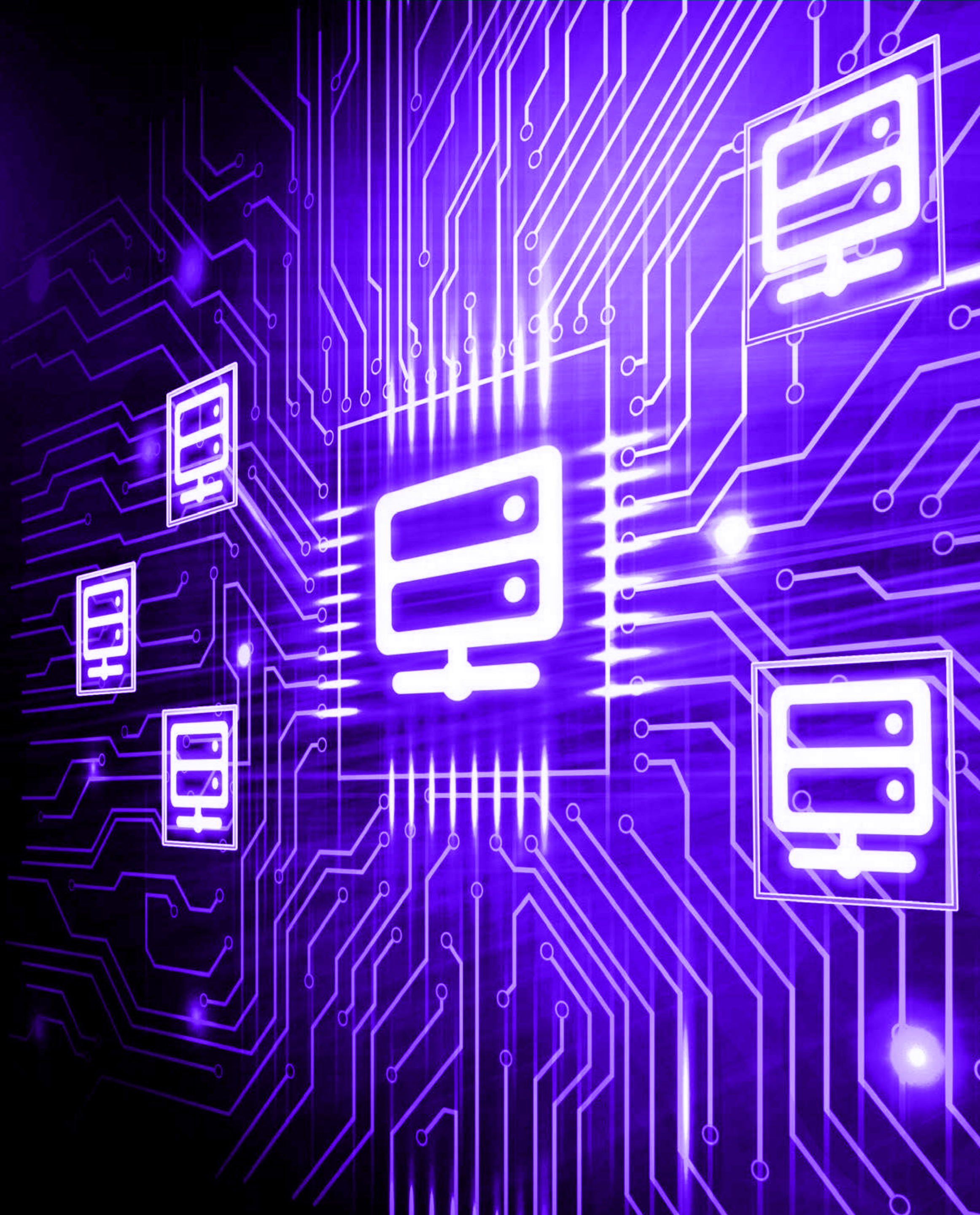
PERFIL DEL ASPIRANTE

Este curso está dirigido a todas aquellas personas con conocimientos básicos de programación que deseen manejar grandes volúmenes de datos de manera eficiente, superando las limitaciones de herramientas tradicionales como las hojas de cálculo.

Es ideal para profesionales y estudiantes de áreas como ingeniería, economía, administración, ciencias sociales y cualquier disciplina en la que el análisis de datos sea clave para la toma de decisiones. También, es una excelente oportunidad para quienes buscan adentrarse en el mundo del Big Data y la ciencia de datos, fortaleciendo sus habilidades en el procesamiento y visualización de información.

COMPETENCIAS PREVIAS:

Conocimientos básicos en programación en cualquier lenguaje, tener experiencia en bases de datos en Excel u otras herramientas, tener una cuenta de Google.



METODOLOGÍA

Se propone una metodología virtual de seis horas a la semana, dividida en tres clases de dos horas donde los primeros 30 minutos son teóricos y se explica el tema a tratar, una hora donde se desarrollará de manera simultánea con los estudiantes el ejercicio de la aplicación con la herramienta Colaboratory, para que de esta manera realicen las preguntas y expresen las dificultades que pudieran tener en la resolución del problema; en los últimos 30 minutos se hace una introducción al tema de la siguiente clase y se propone un ejercicio para reforzar los conocimientos adquiridos.

Los set de datos y los notebooks con los que se trabajará en la clase estarán preparados con los enunciados para que el estudiante solo proponga la solución y los tenga como referencia para resolver problemas con las herramientas adquiridas en el curso.

CONTENIDO TEMÁTICO

MÓDULO 1

4 horas

Big Data

- ¿Qué es Big Data y por qué ha influido tanto en la última década?
- 5 V del Big Data.
- Herramientas para el procesamiento de grandes volúmenes de datos
- Importancia de las visualizaciones en la presentación de conclusiones y hallazgos de análisis de datos
- Computación en la nube como factor de éxito en el Big Data

MÓDULO 2

6 horas

Repaso de conceptos Python

- Uso de Colaboratory de Google con Python
- Tipos de datos
- Entrada y salida de información por medio de consola
- Bloques condicionales
- Bloques de iteración
- Instalación de librerías externas

MÓDULO 3

16 horas

Librería Pandas

- Generalidades, origen de la librería y principales usos
- Tipos de datos soportados por Pandas
- Transformación de tipos de datos
- Carga de datos de diferentes orígenes
- Selección de filas y columnas
- Filtrado de datos
- Filtrado condicional de datos
- Tratamiento de datos faltantes
- Agrupaciones condicionales de datos
- Herramientas básicas de visualización de Pandas
- Combinación lógica de diferentes fuentes de datos en una sola fuente
- Manejo de campos de tipo fecha
- Manejo básico de series temporales
- Repaso estadística básica con Pandas

La Universidad de América se reserva el derecho de ajustar, cambiar o modificar el contenido del programa.



CONTENIDO TEMÁTICO

MÓDULO 4

6 horas

Librería Plotly

- Generalidades, origen de la librería y principales usos
- Visualizaciones de puntos
- Visualizaciones de líneas
- Visualizaciones de barras e histogramas
- Visualizaciones circulares
- Visualizaciones especiales en 3D y mapas de calor

MÓDULO 5

4 horas

Proyectos propuestos

- Análisis exploratorio de datos
- Preparación de datos para algoritmos de Machine Learning
- Detección de anomalías

La Universidad de América se reserva el derecho de ajustar, cambiar o modificar el contenido del programa.



DOCENTE

ÓSCAR RIOJAS



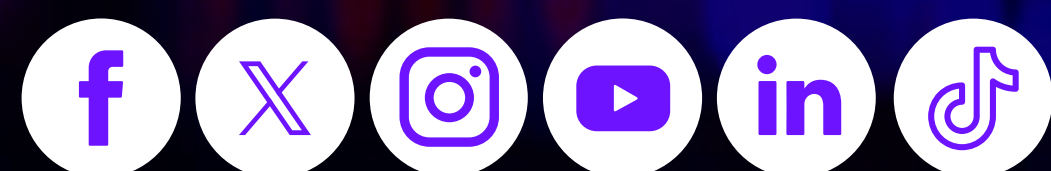
Científico de datos con experiencia en la construcción de modelos de análisis de datos y tecnología blockchain, así como en la creación de algoritmos para extraer valor de los datos, comprensión de los lenguajes de programación para procesar los datos, como Python, R y SQL.

CERTIFICADO

Se hará entrega de un Certificado expedido por la Universidad de América correspondiente al programa, a los participantes que asistan a por lo menos el 80% del tiempo total en horas del curso.

De lo contrario, se expedirá una constancia con el número de horas lectivas.





Más información:

Dirección de Mercadeo y Admisiones
educacion.continua@uamerica.edu.co

 **310 560 1538**

EcoCampus de Los Cerros - Avenida Circunvalar No 20 -53
Tel: (60 1) 3376680 opción 1

Sede Norte - Calle 106 No. 19 - 18
Tel: (60 1) 6580658 opción 1

Bogotá D.C., Colombia.

www.uamerica.edu.co

Fundación Universidad de América | Vigilada MinEducación



ISO 9001:2015
ISO 45001:2018
ISO 14001:2015
BUREAU VERITAS
Certification

