

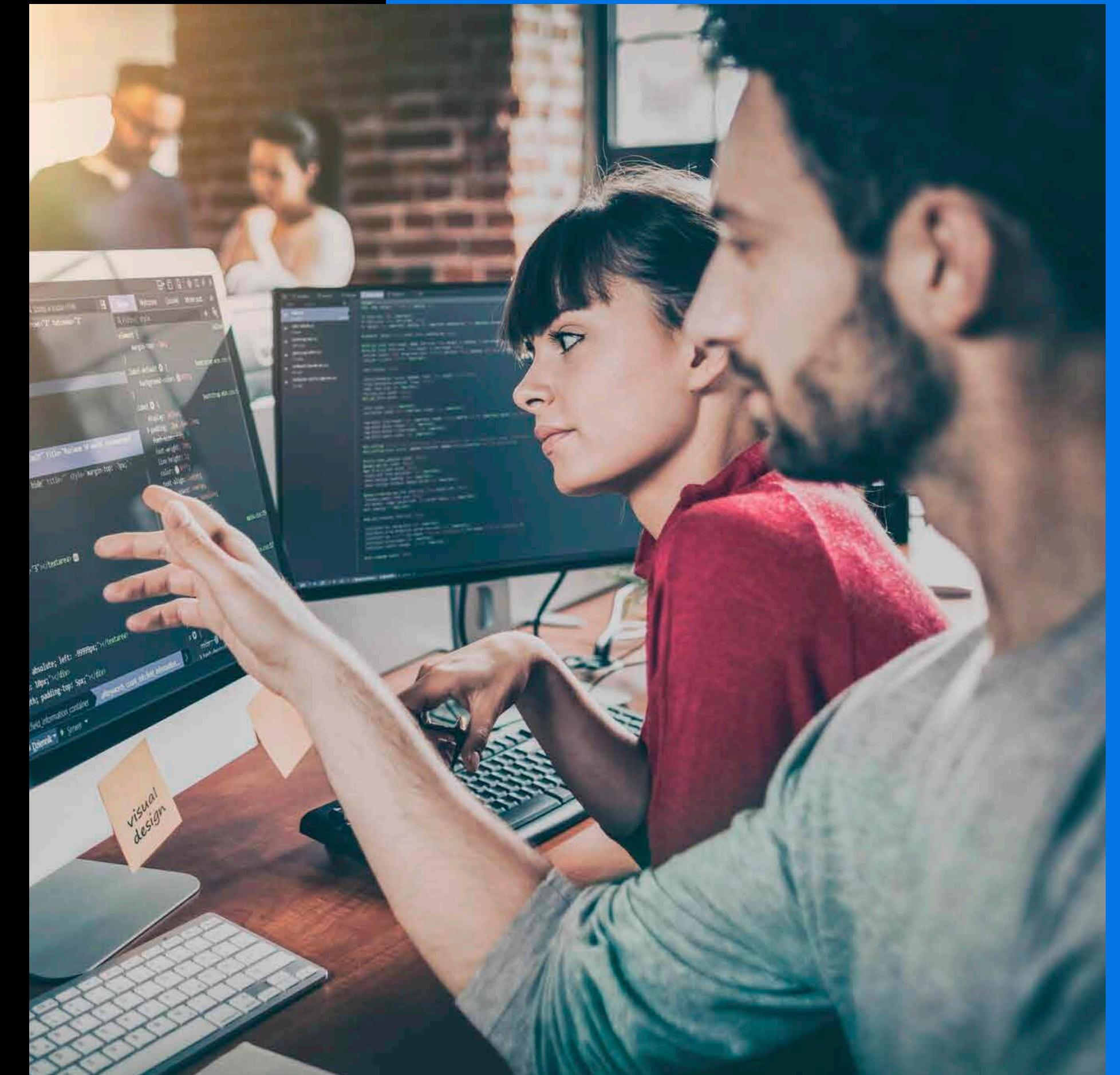
CURSO INTERNACIONAL:

Python para la Ciencia de Datos Aplicada

48 hrs. académicas / 3 créditos



CAD PERÚ
Corporación Americana de Desarrollo





. Presentación

En un mundo cada vez más impulsado por los datos, el análisis y la interpretación de información se han convertido en habilidades esenciales para la toma de decisiones estratégicas. Este curso está diseñado para brindar a los participantes los conocimientos y herramientas fundamentales para trabajar con datos utilizando Python, uno de los lenguajes de programación más versátiles y demandados en el ámbito de la Ciencia de Datos.

Los participantes podrán aprender desde los conceptos básicos de Python y la instalación del entorno de desarrollo, hasta técnicas de manipulación, análisis y visualización de datos. Además, se abordarán fundamentos del Machine Learning para introducir a los asistentes en el desarrollo de modelos predictivos.

. Objetivo

Desarrollar las habilidades necesarias para utilizar Python como herramienta en la Ciencia de Datos, permitiendo a los participantes manipular, analizar y visualizar datos, así como aplicar técnicas básicas de Machine Learning para la toma de decisiones basadas en evidencia.

• **Dirigido a**

Programadores y analistas de datos que apoyan en las diversas gerencias de línea de las empresas o en áreas adjuntas a la gerencia general, y a los profesionales y alumnos de los últimos ciclos de las carreras diversas afines a la ciencia de datos, con especialidad e interés en la programación y el análisis de datos.

• **Modalidad**

Virtual asincrónico.



. Sesión 01

Introducción a Python y Ciencia de Datos

- ¿Qué es la Ciencia de Datos y su importancia?
- Introducción a Python como herramienta para Ciencia de Datos.
- Instalación y configuración del IDE Visual Studio Code
- Pruebas de funcionamiento
- Principales bibliotecas para Ciencia de Datos:
 - NumPy, Pandas, Matplotlib, Seaborn.

. Sesión 02

Manipulación y Análisis de Datos con Pandas

- Introducción a estructuras de datos en Pandas:
 - Series y DataFrames.
- Carga y exploración de datos:
 - Archivos CSV, Excel, y bases de datos.
- Limpieza de datos:
 - Manejo de valores nulos y duplicados.
 - Normalización de datos.

. Sesión 03

Operaciones Avanzadas

- Agrupaciones con groupby
- Tablas Dinámicas (pivot_table)
- Ordenamiento y Combinaciones de Datos
- Transformaciones y Creación de Nuevas Columnas

. Sesión 04

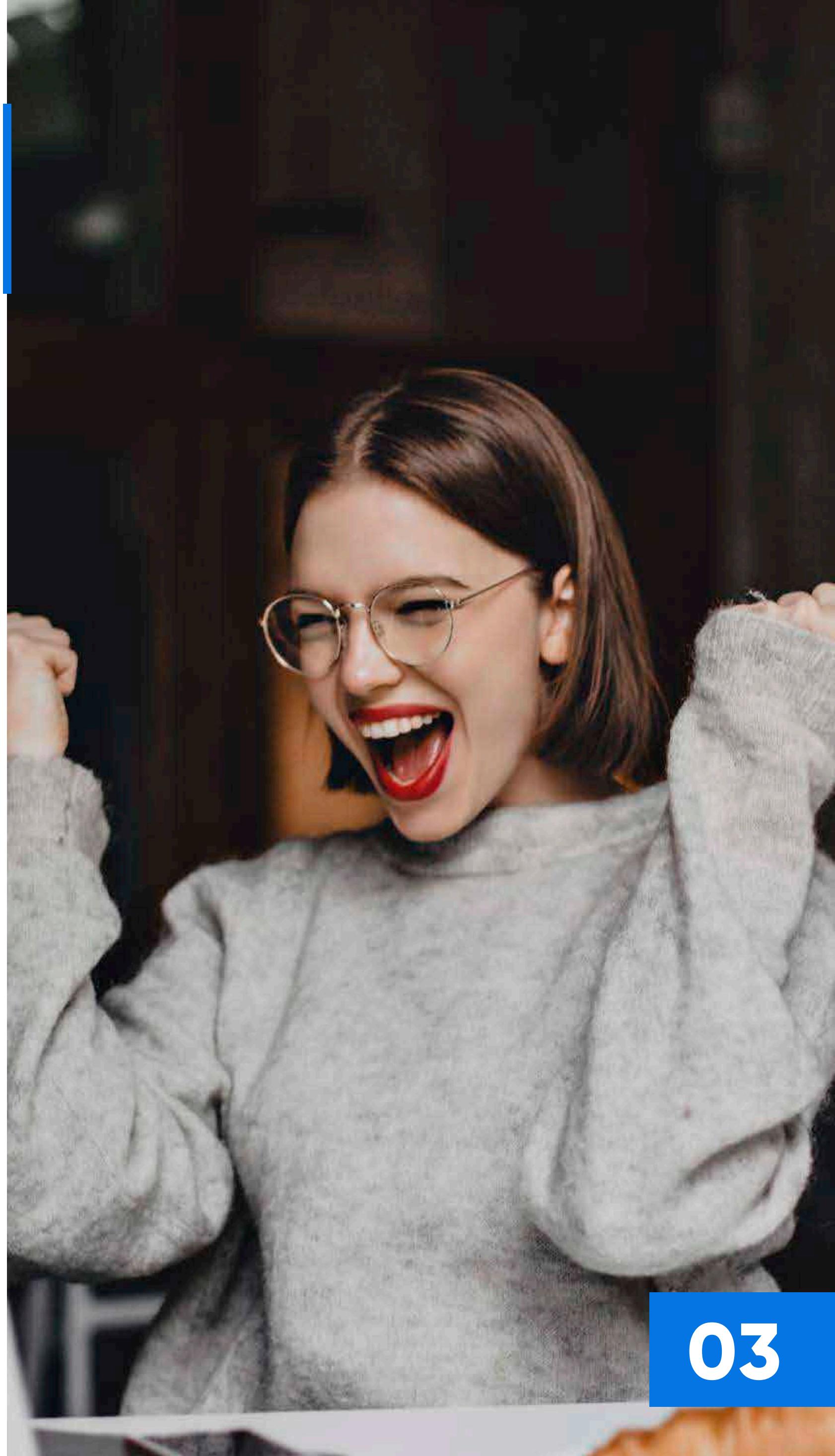
Introducción al Análisis Exploratorio de Datos

- ¿Qué es el EDA y su importancia?
- Identificación de patrones, tendencias y relaciones.
- Casos prácticos:
 - Exploración de un conjunto de datos de ejemplo.

. Sesión 05

Introducción a Machine Learning con Scikit-learn

- Conceptos básicos de Machine Learning:
 - Supervisado vs no supervisado.
- Introducción a Scikit-learn:
 - Preparación de datos: División en conjuntos de entrenamiento y prueba.
 - Algoritmos básicos:
- Regresión lineal.
- Clasificación con k-Nearest Neighbors.
 - Evaluación de modelos:
- Métricas como precisión, recall y F1-score.



. Docente



Dr. Jesús Fernando Cruz Álvarez

- Doctor en Educación, especialista en Ciencia de Datos y Programación en Python, con más de 18 años en educación superior. Capacitador internacional en Inteligencia Artificial y análisis de datos, y experto en desarrollo de aplicaciones y bases de datos.

. Metodología de Enseñanza

- **Video:** Medio audiovisual que contempla el desarrollo panorámico del tema, con ejemplos que ayuden a entender lo sustancial de cada unidad.
- **Recursos Complementarios:** Lecturas seleccionadas y/o material audiovisual, a fin de profundizar la comprensión y/o análisis del tema, facilitando la apropiación del contenido y los referentes bibliográficos que complementan la capacitación y enriquecen el conocimiento.
- **Evaluación:** Test que desarrolla el participante al finalizar el curso, permitiéndole evaluar su proceso de aprendizaje.





CAD PERÚ
Corporación Americana de Desarrollo

www.cadperu.com