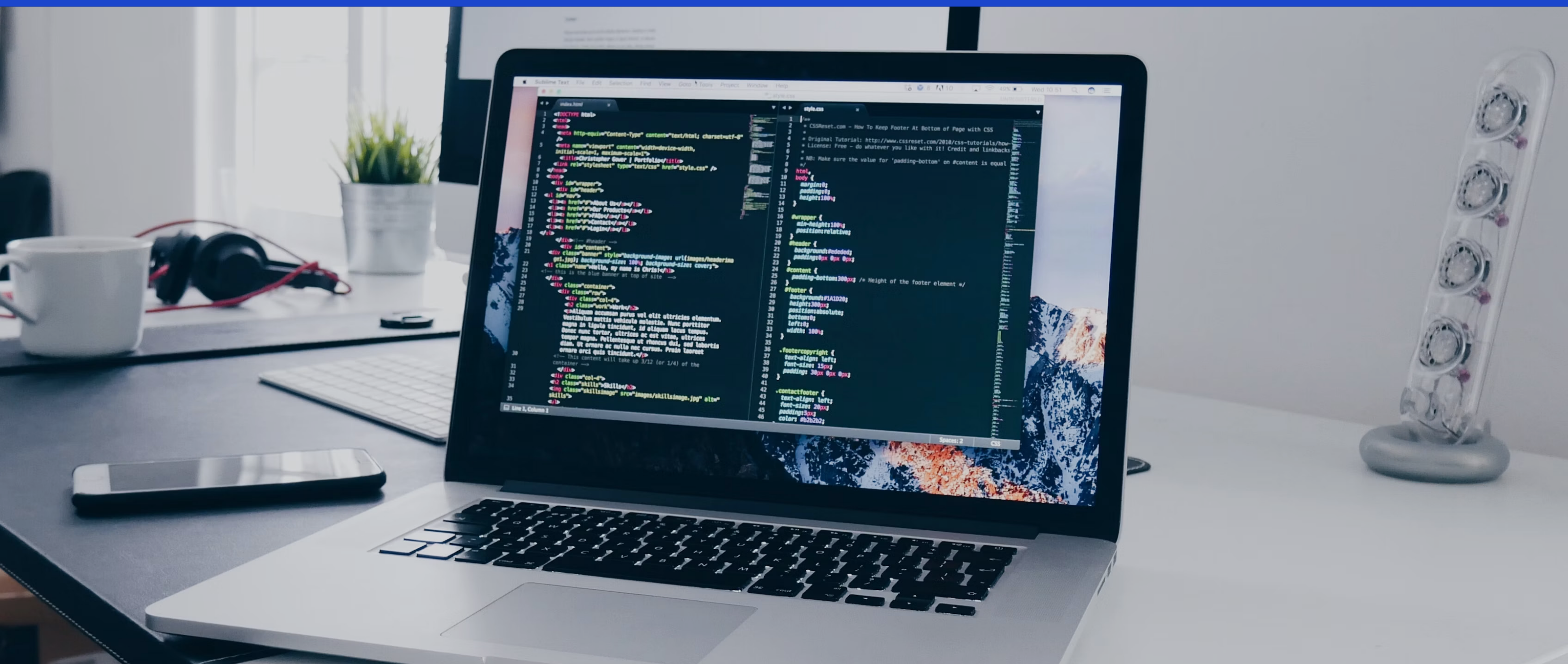




Python for Data Scientist

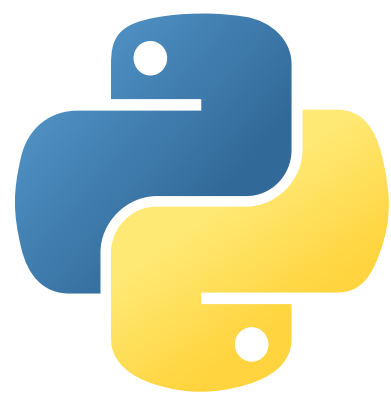
Curso de Especialización





Presentación

Aprende Python desde cero y desarrolla competencias en Data Science, análisis de datos y Machine Learning mediante un enfoque práctico y progresivo. A lo largo del curso trabajarás en un caso real de negocio, aplicando técnicas de exploración, modelado e interpretación de datos para generar insights y tomar decisiones basadas en evidencia. Además, integrarás herramientas de IA generativa para acelerar la creación, comprensión y depuración de código fortaleciendo tu aprendizaje.



Conocimientos previos

- Manejo básico de Excel (funciones, tablas y gráficos).
- Ayuda tener nociones básicas de programación y/o estadística, aunque no son indispensables: **se aprende Python desde cero.**

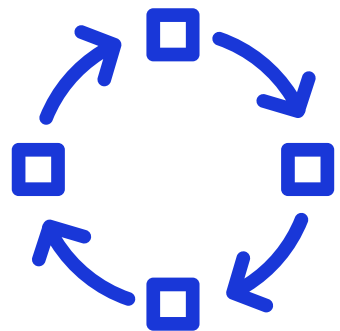
Objetivos

- Comprender los fundamentos de la programación en Python: variables, tipos de datos, estructuras de datos, control de flujo y funciones.
- Aplicar estadística descriptiva y análisis exploratorio de datos (EDA) para la toma de decisiones.
- Comprender el procedimiento y las técnicas del Machine Learning, y saber diferenciar algoritmos supervisados y no supervisados.





Metodología



Sesiones sincrónicas
basadas en casuísticas
(casos reales).

Duración



El curso tiene una
duración de **24 horas**
cronológicas.

Beneficios



Certificación con **validez internacional**.



Instructores y Consultores
especialistas en Data y
Analytics, reconocidos a
nivel latinoamericano.



Acompañamiento,
asesoramiento y desarrollo a
los alumnos.



Cuenta educativa de
Microsoft Learning Partner
(acceso a Office 365 online).



Acceso a la comunidad y
Bolsa de Trabajo EXCLUSIVA
para nuestros estudiantes.



Temario

Nivel Básico - Fundamentos de Python

Tema 1

Fundamentos de Python y entorno de trabajo

- **Python en la ciencia de datos:** ¿qué es Data Science?, tipos de análisis y fases de un proyecto analítico.
- **Entorno de trabajo:** Google Colab y notebooks. Instalación e importación de librerías.
- **Variables y tipos de datos:** string, int, float, boolean. Operadores básicos.
- **Operaciones de entrada/salida y control de flujo:** condicionales (if/elif/else) y bucles (for/while).
- **Taller:** Primer notebook en Colab con captura de datos, operaciones y estructuras de control.

Tema 2

Estructuras de datos, funciones y NumPy

- Listas, tuplas, diccionarios y conjuntos: indexing y slicing.
- **Funciones personalizadas:** parámetros, retorno, lambda y manejo de errores (try/except).
- **Librería NumPy:** arrays, vectores y matrices.
- Operaciones matemáticas y algebraicas con arrays.
- **Lab:** Funciones reutilizables y operaciones con listas y diccionarios.





Nivel Intermedio – Datos, estadística y visualización

Tema 3

Manipulación de datos con Pandas

- **Series y DataFrames:** definición y principales diferencias.
- Lectura y escritura desde archivos externos (CSV, Excel, TXT).
- Selección, filtrado y agrupación de datos (**GROUPBY, AGG, PIVOT_TABLE**).
- Combinación de DataFrames (merge, concat) y extracción de muestras.
- **Taller:** Exploración y agrupación de un dataset real.



Tema 4

Estadística descriptiva y visualización de datos

- **Tipos de datos y medidas de tendencia central:** media, mediana y moda.
- **Medidas de dispersión y posición:** desviación estándar, varianza, cuartiles. Asimetría y curtosis.
- Correlación y covarianza. Diagrama de cajas.
- **Visualización con Matplotlib y Seaborn:** gráficos para variables cuantitativas y cualitativas, subplots.
- **Taller:** Funciones reutilizables y operaciones con listas y diccionarios.



Tema 5

Análisis exploratorio (EDA), feature engineering y preparación de datos

- **Objetivos del análisis exploratorio:** relaciones entre variables y detección de patrones.
- Tratamiento de valores perdidos (missings) y valores atípicos (outliers).
- Generación de nuevas variables adecuadas al negocio (feature engineering).
- Imputación de datos mediante técnicas univariadas y multivariadas.
- **Taller:** Limpieza, creación y selección de variables sobre un dataset con datos faltantes y atípicos.

Nivel Avanzado – Machine Learning aplicado

Tema 6

Fundamentos de ML con Scikit-Learn: preprocesamiento y aprendizaje no supervisado

- **Introducción al Machine Learning:** tipos de aprendizaje (supervisado y no supervisado).
- Scikit-Learn y Train/Test Split.
- **Preprocesamiento:** StandardScaler y One Hot Encoding.
- Construcción de Pipelines de preprocesamiento y modelado.
- **Aprendizaje no supervisado:** clustering con K-Means.
- **Taller:** Pipeline de preprocesamiento + K-Means sobre el dataset del caso.

Tema 7

Machine Learning supervisado: modelado, validación y ajuste

- **Clasificación y Regresión:** Árbol de decisión, Random Forest, Regresión Lineal y XGBoost/LightGBM.
- **Overfitting vs. underfitting:** el equilibrio sesgo-varianza.
- Validación cruzada (cross-validation).
- **Optimización de hiperparámetros:** Grid Search y Random Search.
- **Taller:** Entrenamiento, validación y tuning de un modelo del caso.

Tema 8

Evaluación, interpretabilidad, IA para Data Science y MLOps – Presentación final

- **Métricas de evaluación:** matriz de confusión, accuracy, precisión, recall y F1-score (RMSE/MSE en regresión).
- **Interpretabilidad de modelos:** importancia de variables e introducción a SHAP.
- **IA para Data Science:** ChatGPT, Claude y prompt engineering aplicados al flujo analítico.
- **Introducción a MLOps:** del modelo a producción.
- Presentación final del caso real de negocio de extremo a extremo.
- **Taller de presentaciones efectivas de resultados.**





Instructor



Geanfranco Palomino
Head COE Data & Analytics

Ejecutivo en Data & Analytics, con más de 9 años liderando iniciativas de transformación analítica y posicionando el dato como un activo estratégico para la toma de decisiones y la generación de valor sostenible. Lidera equipos multidisciplinarios en entornos regionales, diseñando e implementando estrategias de Data & AI que impulsan la competitividad, habilitan eficiencia operativa y aceleran la adopción de una cultura data-driven en las organizaciones.

Con un enfoque centrado en conectar negocio, datos y tecnología para transformar capacidades analíticas en impacto real, asegurando la construcción adopción y monetización de soluciones.

Certificación

Al finalizar el curso, el alumno obtendrá doble certificación a nombre de la institución.





Empresas líderes del mercado que confían en nosotros





Métodos de pago



BCP - Banco de Credito
Nombre: Levo Learning SAC
RUC: 20606130661
Cuenta Corriente: 19171497185076
CCI: 00219117149718507658



BBVA - Banco Continental
Nombre: Levo Learning SAC
RUC: 20606130661
Cuenta Corriente: 00107330200144657
CCI: 01173300020014465730



LEVO LEARNING CENTER SAC
@levolearning



Información Adicional

Importante

Las fechas previstas para el inicio del curso ***están sujetas a cambios basados en la cantidad de personas registradas.***

Nuestra prioridad es asegurar que todos los inscritos puedan participar y aprovechar al máximo esta experiencia educativa. Se notificará con antelación cualquier ajuste necesario.

Sé disruptivo en el mundo tecnológico

LLC

Levo Learning Center

Educación Continua
Av. Pershing 890
Lima - Perú

www.levolearning.edu.pe
Cel: +51 986 619 728
informes@levolearning.edu.pe

Avalados por:



Microsoft
Learning Partner



Microsoft
Partner Network