



## PROCESO DE GESTIÓN DE FORMACIÓN PROFESIONAL INTEGRAL

### GUÍA DE APRENDIZAJE (PYTHON (MACHINE LEARNING + BIG DATA))

#### 1. IDENTIFICACIÓN DE LA GUIA DE APRENDIZAJE

- **Denominación del Programa de Formación:** Tecnología en Análisis y Desarrollo de Software
- **Código del Programa de Formación:** 228118
- **Nombre del Proyecto Formativo** (si aplica): Desarrollo de Software Empresarial Innovador Alineado con las Políticas y Estrategias Gubernamentales Locales y Nacionales.
- **Fase del Proyecto** (si aplica): Codificar
- **Actividad de Proyecto Formativo** (si aplica): Desarrollar la estructura de datos y la interfaz de usuario del software
- **Competencia:** Implementar la solución de software de acuerdo con los requisitos de operación y modelos de referencia.
- **Resultados de Aprendizaje:** Planear actividades de implantación del software de acuerdo con las condiciones del sistema.
- **Duración de la Guía de Aprendizaje** (horas): 120 horas

#### 2. PRESENTACIÓN

¡Bienvenidos a esta emocionante experiencia de aprendizaje con Big Data y Machine Learning! A través de esta guía de aprendizaje, los participantes tendrán la oportunidad de adentrarse en el fascinante mundo del análisis de datos masivos y la inteligencia artificial, comprendiendo cómo la información se convierte en conocimiento útil para la toma de decisiones estratégicas en distintos sectores. Este proceso formativo permitirá a los aprendices desarrollar competencias fundamentales en la recolección, procesamiento, análisis e interpretación de grandes volúmenes de datos, utilizando herramientas y lenguajes especializados como Python, R y plataformas de Big Data.

Durante este recorrido, se promoverá un aprendizaje integral que trasciende lo técnico, estimulando el pensamiento crítico, la capacidad analítica y la resolución de problemas complejos mediante la aplicación de modelos de aprendizaje automático (Machine Learning). La guía invita a construir sobre los conocimientos previos en programación, matemáticas y estadística, fortaleciendo la comprensión de los nuevos conceptos y técnicas que conforman el ecosistema de



la Ciencia de Datos, desde la limpieza y transformación de datos hasta la creación y evaluación de modelos predictivos.

El trabajo autónomo y colaborativo será un eje central de esta experiencia. Los aprendices deberán gestionar sus tareas de manera organizada y metódica, explorando diferentes fuentes de información, experimentando con datos reales y compartiendo sus hallazgos con el equipo. A través del trabajo en equipo, se fomentará el crecimiento tanto individual como colectivo, potenciando la comunicación efectiva, la co-creación de soluciones innovadoras y el aprendizaje a partir de la diversidad de perspectivas.

Esta guía está diseñada para inspirar a los aprendices a involucrarse activamente en su proceso de formación, entendiendo que la Ciencia de Datos no solo implica dominar herramientas tecnológicas, sino también desarrollar una mentalidad analítica, ética y orientada al descubrimiento. ¡Es momento de comenzar esta aventura con el desarrollo de las actividades!

### 3. FORMULACIÓN DE LAS ACTIVIDADES DE APRENDIZAJE

- **Descripción de la(s) Actividad(es)**

#### 3.1 Actividades de reflexión inicial: Datos que cuentan historias

**Descripción de la actividad:** Comprender los fundamentos de la ciencia de datos y su impacto en la toma de decisiones.

- Mediante trabajo colaborativo se conformarán grupos pequeños de trabajo de (3 o 4 aprendices por grupo).
- Cada grupo recibe una tarjeta temática con un contexto real en el que se usan datos. Ejemplos:
  - **Salud:** Predicción de enfermedades con IA.
  - **Educación:** Personalización del aprendizaje con analítica de datos.
- Cada grupo reflexiona y responde:
  - ¿Qué tipo de datos se generan en este contexto?
  - ¿Para qué se usan?
  - ¿Qué beneficios y riesgos identifican?
  - ¿Qué profesiones o tecnologías intervienen en este proceso?
  - ¿Cómo impacta en la sociedad o en la vida de las personas?
- Cada grupo elabora una infografía o poster digital (o en papel) con sus conclusiones.
  - El cartel debe incluir: título del contexto, tipos de datos, beneficios, riesgos y conclusiones.
- Posteriormente, cada grupo expone la infografía o poster digital generado.



**Ambiente requerido:** Ambiente de pluritecnológico con ventilación, mobiliario ergonómico, iluminación, acceso a internet, computadores de escritorio y/o portátil, pantalla y/o televisor para proyectar y acceso a la plataforma LMS.

**Estrategias o técnicas didácticas activas:**

- Aprendizaje Basado en Problemas (ABP) y discusión en pequeños grupos.

**Materiales de formación:** Papel periódico, marcadores, colores, lápices y/o computador con acceso a internet, pantalla y/o televisor para proyectar, plataforma LMS.

**Material de apoyo:**

- El Mundo de los Datos. (s.f.). El Mundo de los Datos. Recuperado de <https://elmundodelosdatos.com/>
- Domo. (2025). Data Never Sleeps: AI Edition 2025 [Infografía]. Domo. <https://www.domo.com/es/learn/infographic/data-never-sleeps-ai-edition-2025>

**Duración de la actividad:** 02 horas.

### 3.2 Actividades de contextualización e identificación de conocimientos necesarios para el aprendizaje: Data Sprint — Quiz de Ciencia de Datos”.

**Descripción de la actividad:** Diagnosticar los conocimientos previos sobre Ciencia de Datos (Big Data, Machine Learning, Estadística, Ética de datos y Fundamentos de Programación en Python).

- El aprendiz ingresa a la URL del quiz
  - En la pantalla de inicio registra su nombre, inicia el quiz y dispone de 15 minutos para responder 15 preguntas seleccionadas aleatoriamente.
  - Al finalizar, el sistema muestra nombre, fecha de realización, puntaje total, tiempo empleado, exactitud (%) y desglose por temática.
  - El aprendiz pulsa “Descargar resultados (.json)” para obtener el archivo de desempeño y realiza una captura de pantalla del puntaje mostrado.
  - Entrega en Classroom/LMS:
    - Archivo resultados\_[nombre].json.
    - Captura de pantalla del puntaje.
  - El instructor realizará una retroalimentación de conformidad a los resultados de cada aprendiz.

**Ambiente requerido:** Ambiente de pluritecnológico con ventilación, mobiliario ergonómico, iluminación, acceso a internet, computadores de escritorio y/o portátil, pantalla y/o televisor para proyectar y acceso a la plataforma LMS.



**Estrategias o técnicas didácticas activas:**

- Diagnóstico inicial individual con instrumento digital y Aprendizaje autónomo guiado (autoexploración y autorregulación del tiempo).

**Materiales de formación:** computador con acceso a internet, pantalla y/o televisor para proyectar, plataforma LMS.

**Material de apoyo:**

- IBM. (2023). ¿Qué es la ciencia de datos? IBM. Recuperado de <https://www.ibm.com/co-es/topics/data-science>
- Microsoft Learn. (2024). Introducción al aprendizaje automático con Python y Azure Machine Learning. Microsoft. Recuperado de <https://learn.microsoft.com/es-es/training/paths/intro-machine-learning-python-azure/>
- Python Software Foundation. (2024). The Python Tutorial. Python.org. Recuperado de <https://docs.python.org/es/3/tutorial/>
- Wes McKinney. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2.ª ed.). O'Reilly Media.

**Duración de la actividad:** 02 horas.

### 3.3 Actividades de apropiación:

**Actividad No.01 – Programación en Python para Ciencia de Datos**

**Descripción de la actividad:** Aplicar los conceptos fundamentales de Python en el contexto de la Ciencia de Datos, fortaleciendo el uso de estructuras, funciones y librerías básicas para la manipulación de conjuntos de datos.

- Escuche con atención la explicación de su instructor(a) sobre los temas de fundamentos de programación en Python, incluyendo estructuras de control, tipos de datos, funciones y uso básico de librerías para análisis de datos (como pandas y numpy), con apoyo de presentación electrónica.
- De manera individual, desarrolle el “Taller Fundamentos de Programación en Python para Ciencia de Datos” entregado por el instructor, aplicando los conceptos vistos para resolver ejercicios prácticos de manipulación y análisis de conjuntos de datos.
- Una vez finalizado el taller, genere el archivo con las soluciones en formato .ipynb (Jupyter Notebook) o .py, y exporte una copia en formato PDF.
- Suba ambos archivos (código fuente y PDF) en el espacio dispuesto para la entrega en el LMS (Classroom o Zajuna), siguiendo las indicaciones del instructor.



**Ambiente requerido:** Ambiente de pluritecnológico con ventilación, mobiliario ergonómico, iluminación, acceso a internet, computadores de escritorio y/o portátil, pantalla y/o televisor para proyectar y acceso a la plataforma LMS.

**Estrategias o técnicas didácticas activas:**

- Aprendizaje Basado en Problemas (ABP) y simulaciones.

**Materiales de formación:** computador con acceso a internet, pantalla y/o televisor para proyectar, plataforma LMS.

**Material de apoyo:**

- Python Software Foundation. (2024). The Python Tutorial. Python.org. Recuperado de <https://docs.python.org/es/3/tutorial/>
- Wes McKinney. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2.ª ed.). O'Reilly Media.

**Duración de la actividad:** 20 horas.

**Actividad No.02 – Analisis de Grandes Volúmenes de Datos**

**Descripción de la actividad:** Apropiar los conceptos relacionados con el Analisis de grandes volúmenes de datos provenientes de fuentes abiertas o institucionales, aplicando el proceso de ciencia de datos (ingesta, limpieza, transformación, exploración y visualización) con Python (*pandas, numpy, matplotlib*) y/o herramientas de BI.

- Escuche con atención la explicación de su instructor(a) sobre los conceptos de Big Data, análisis de grandes volúmenes de información, y las etapas del proceso de ciencia de datos, con apoyo de presentación electrónica y ejemplos prácticos.
- Observe la demostración del uso de herramientas y librerías de análisis de datos (por ejemplo, *pandas, numpy, matplotlib* o Power BI) para la exploración, limpieza, transformación y visualización de conjuntos de datos extensos.
- De manera individual o en pequeños grupos (según indicación del instructor), descargue el conjunto de datos asignado desde la URL o fuente indicada, y realice el análisis exploratorio aplicando filtros, agrupamientos y cálculos estadísticos básicos que permitan identificar patrones o tendencias.
- Desarrolle el “Taller: Análisis de Grandes Volúmenes de Datos” entregado por el instructor, integrando los resultados obtenidos en su análisis y respondiendo las preguntas orientadoras del taller.
- Una vez finalizado el ejercicio, genere el archivo de trabajo en formato .ipynb (Jupyter Notebook) o .py, junto con un informe en formato PDF que incluya:



- Descripción del conjunto de datos.
  - Pasos realizados en el análisis (limpieza, procesamiento, visualización).
  - Principales hallazgos y conclusiones.
- Suba los archivos generados (código y PDF) en el espacio dispuesto para la entrega dentro del LMS (Classroom, Zajuna o Sofía), siguiendo las instrucciones del instructor.

**Ambiente requerido:** Ambiente de pluritecnológico con ventilación, mobiliario ergonómico, iluminación, acceso a internet, computadores de escritorio y/o portátil, pantalla y/o televisor para proyectar y acceso a la plataforma LMS.

**Estrategias o técnicas didácticas activas:**

- Aprendizaje Basado en Problemas (ABP) y simulaciones.

**Materiales de formación:** computador con acceso a internet, pantalla y/o televisor para proyectar, plataforma LMS.

**Material de apoyo:**

- Python Software Foundation. (2024). The Python Tutorial. Python.org. Recuperado de <https://docs.python.org/es/3/tutorial/>
- Wes McKinney. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2.ª ed.). O'Reilly Media.
- **McKinney, W. (2018).** *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython* (2.ª ed.). O'Reilly Media.

**Duración de la actividad:** 20 horas.

**Actividad No.03 – Modelos de Inteligencia Artificial**

**Descripción de la actividad:** Apropiar los conceptos relacionados con los **modelos de inteligencia artificial, específicamente el aprendizaje automático (Machine Learning)**, aplicando técnicas de entrenamiento, validación y evaluación de modelos predictivos mediante el uso de **Python y librerías especializadas** como *scikit-learn*, *pandas*, *numpy* y *matplotlib*.

- Escuche con atención la explicación de su instructor(a) sobre los principios del aprendizaje automático, los tipos de modelos (supervisado, no supervisado y por refuerzo) y las etapas del ciclo de Machine Learning, con apoyo de presentación electrónica y ejemplos aplicados.
- Observe la demostración del uso de herramientas y librerías de Python (por ejemplo, scikit-learn, pandas, numpy, matplotlib, seaborn) para la creación, entrenamiento y evaluación de modelos predictivos.



- De manera individual o en pequeños grupos (según indicación del instructor), descargue el conjunto de datos asignado desde la fuente indicada y realice el preprocesamiento necesario (limpieza, codificación de variables, normalización, división de datos en entrenamiento y prueba).
- Desarrolle el “Taller: Modelos de Inteligencia Artificial con Machine Learning” entregado por el instructor, implementando al menos dos algoritmos distintos (por ejemplo, Regresión Lineal, Árboles de Decisión, K-Vecinos más Cercanos o Random Forest), comparando sus resultados y métricas de rendimiento.
- Una vez finalizado el ejercicio, genere el archivo de trabajo en formato .ipynb (Jupyter Notebook) o .py, junto con un informe en formato PDF que incluya:
  - Descripción del conjunto de datos.
  - Explicación de los modelos implementados.
  - Comparación de resultados y métricas (accuracy, precision, recall, F1-score, etc.).
  - Conclusiones sobre el modelo con mejor desempeño y posibles mejoras futuras.
- Suba los archivos generados (código y PDF) en el espacio dispuesto para la entrega dentro del LMS (Classroom, Zajuna o Sofía Plus), siguiendo las instrucciones del instructor(a).

**Ambiente requerido:** Ambiente de pluritecnológico con ventilación, mobiliario ergonómico, iluminación, acceso a internet, computadores de escritorio y/o portátil, pantalla y/o televisor para proyectar y acceso a la plataforma LMS.

**Estrategias o técnicas didácticas activas:**

- Aprendizaje Basado en Problemas (ABP) y simulaciones.

**Materiales de formación:** computador con acceso a internet, pantalla y/o televisor para proyectar, plataforma LMS.

**Material de apoyo:**

- Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2.ª ed.). O'Reilly Media.
- IBM. (2023). Introducción al aprendizaje automático (Machine Learning). IBM. Recuperado de <https://www.ibm.com/co-es/topics/machine-learning>
- Python Software Foundation. (2024). The Python Tutorial. Python.org. Recuperado de <https://docs.python.org/es/3/tutorial/>

**Duración de la actividad:** 30 horas.



#### **Actividad No.04 – Despliegue de una Aplicación de Inteligencia Artificial**

**Descripción de la actividad:** Implementar y desplegar una aplicación de Inteligencia Artificial que integre un modelo de Machine Learning previamente entrenado, utilizando herramientas de desarrollo y servicios en la nube o en entornos locales.

- Escuche con atención la explicación de su instructor(a) sobre los conceptos de despliegue de modelos de Inteligencia Artificial, entornos de producción, y diferencias entre entornos locales y cloud, con apoyo de presentación electrónica y ejemplos demostrativos.
- Observe la demostración práctica del instructor sobre cómo empaquetar un modelo de Machine Learning entrenado en Python (por ejemplo, con joblib o pickle) y cómo integrarlo en una aplicación web sencilla con Flask o Streamlit.
- De manera individual o en pequeños grupos (según indicación del instructor), utilice el modelo entrenado en la actividad anterior o uno nuevo asignado por el instructor, y desarrolle una aplicación funcional que permita realizar predicciones a partir de datos ingresados por el usuario.
- Configure la aplicación para su ejecución local o despliegue en la nube (por ejemplo, Render, Hugging Face Spaces, Vercel o Railway), asegurando que el modelo cargue correctamente y que la interfaz muestre los resultados de manera clara y comprensible.
- Desarrolle el “Taller: Despliegue de una Aplicación de Inteligencia Artificial” entregado por el instructor, documentando los pasos de configuración, ejecución y publicación del proyecto.
- Una vez finalizada la práctica, genere los siguientes entregables:
  - Carpeta del proyecto (código fuente y archivos del modelo entrenado).
  - Capturas de pantalla de la aplicación en funcionamiento.
  - Documento en formato PDF con los siguientes apartados:
    - Descripción del modelo implementado.
    - Arquitectura y componentes del despliegue.
    - Capturas o enlace de la aplicación publicada.
    - Conclusiones sobre el proceso de implementación y retos encontrados.
- Suba los archivos generados (código y PDF) en el espacio dispuesto para la entrega dentro del LMS (Classroom, Zajuna o Sofía Plus), siguiendo las instrucciones del instructor(a).

**Ambiente requerido:** Ambiente de pluritecnológico con ventilación, mobiliario ergonómico, iluminación, acceso a internet, computadores de escritorio y/o portátil, pantalla y/o televisor para proyectar y acceso a la plataforma LMS.

#### **Estrategias o técnicas didácticas activas:**

- Aprendizaje Basado en Problemas (ABP) y simulaciones.

**Materiales de formación:** computador con acceso a internet, pantalla y/o televisor para proyectar, plataforma LMS.

#### **Material de apoyo:**





- Chollet, F. (2021). Deep Learning with Python (2.<sup>a</sup> ed.). Manning Publications.
- IBM. (2023). Despliegue de modelos de Machine Learning. IBM Cloud. Recuperado de <https://www.ibm.com/docs/es/cloud-paks/cp-data/4.7.x?topic=models-deploying-machine-learning>
- Streamlit Inc. (2024). Getting Started with Streamlit. Recuperado de <https://docs.streamlit.io/get-started>

**Duración de la actividad:** 30 horas.

### **3.4 Actividades de Transferencia el Conocimiento:** Proyecto Final - Aplicación del Proceso de Ciencia de Datos

**Descripción de la actividad:** Aplicar de manera integral los conocimientos y habilidades adquiridas en el curso mediante el desarrollo de un **proyecto completo de Ciencia de Datos o Inteligencia Artificial**, abordando todas las fases del proceso: **definición del problema, recolección, limpieza, análisis exploratorio, modelado, evaluación y despliegue de resultados.**

- Escuche con atención la orientación del instructor(a) sobre los lineamientos del proyecto final, las etapas a desarrollar y los criterios de evaluación, con apoyo de presentación electrónica y ejemplos de proyectos de Ciencia de Datos aplicados en distintos contextos.
- De manera individual o en equipos (según indicación del instructor), seleccione un tema o conjunto de datos provenientes de fuentes abiertas, institucionales o propias, y formule una pregunta o problema de análisis que pueda resolverse mediante técnicas de Ciencia de Datos o Machine Learning.
- Planifique el proyecto estructurando las siguientes fases:
  1. Definición del problema y objetivos.
  2. Recolección o selección de los datos.
  3. Limpieza, transformación y análisis exploratorio.
  4. Modelado predictivo o descriptivo (Machine Learning).
  5. Evaluación del modelo y visualización de resultados.
  6. Despliegue o presentación de la solución.
- Desarrolle el proyecto completo utilizando Python y librerías especializadas (pandas, numpy, matplotlib, seaborn, scikit-learn), o herramientas de Business Intelligence como Power BI, según el enfoque del proyecto.
- Elabore un informe final en formato PDF que incluya:
  1. Portada con nombre del proyecto, autor(es) y fecha.
  2. Descripción del problema y objetivos.
  3. Fuentes y características del conjunto de datos.
  4. Etapas del proceso de análisis y modelado.



5. Resultados, visualizaciones y métricas de evaluación.
  6. Conclusiones, recomendaciones y reflexiones éticas sobre el uso de los datos.
- Prepare una breve presentación o demostración (pitch) del proyecto, mostrando la funcionalidad, resultados o despliegue de la aplicación, según el alcance del trabajo desarrollado.
  - Suba los siguientes entregables al espacio dispuesto en el LMS (Classroom o Zajuna):
    1. Carpeta del proyecto (código fuente y datasets utilizados).
    2. Informe final en formato PDF.
    3. Presentación (PDF o diapositivas).
    4. (Opcional) Enlace a la aplicación desplegada o video de demostración.

**Ambiente requerido:** Ambiente de pluritecnológico con ventilación, mobiliario ergonómico, iluminación, acceso a internet, computadores de escritorio y/o portátil, pantalla y/o televisor para proyectar y acceso a la plataforma LMS.

**Estrategias o técnicas didácticas activas:**

- Aprendizaje Basado en Problemas (ABP) y simulaciones.

**Materiales de formación:** computador con acceso a internet, pantalla y/o televisor para proyectar, plataforma LMS.

**Material de apoyo:**

- Provost, F., & Fawcett, T. (2013). Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking. O'Reilly Media.
- Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2.ª ed.). O'Reilly Media.
- IBM. (2023). Ciclo de vida de la ciencia de datos. IBM. Recuperado de <https://www.ibm.com/cos/topics/data-science>
- 

**Duración de la actividad:** 16 horas.



#### 4. PLANTEAMIENTO DE EVIDENCIAS DE APRENDIZAJE PARA LA EVALUACIÓN EN EL PROCESO FORMATIVO.

Fase del proyecto formativo	Actividad del proyecto formativo	Actividad de Aprendizaje	Evidencias de Aprendizaje	Criterios de Evaluación	Técnicas e Instrumentos de Evaluación
Codificar	Desarrollar la estructura de datos y la interfaz de usuario del software	Actividad No.01 – Programación en Python para Ciencia de Datos	Archivo de programación y PDF con la solución del taller	Aplica estructuras y funciones de Python para manipular y analizar datos de forma correcta y eficiente.	Instrumento: Rubrica  Técnica: Taller ABP
		Actividad No.02 – Analisis de Grandes Volúmenes de Datos	Archivo de programación y PDF con la solución del taller	Analiza e interpreta grandes volúmenes de datos, aplicando técnicas de limpieza, transformación y visualización.	Instrumento: Rubrica  Técnica: Taller ABP
		Actividad No.03 – Modelos de Inteligencia Artificial	Archivo de programación y PDF con la solución del taller	Implementa y evalúa modelos de aprendizaje automático, seleccionando algoritmos y métricas adecuadas	Instrumento: Rubrica  Técnica: Taller ABP
		Actividad No.04 – Despliegue de una Aplicación de Inteligencia Artificial	Archivo de programación y PDF con la solución del taller	Integra y despliega un modelo de IA en una aplicación funcional y documentada.	Instrumento: Rubrica  Técnica: Taller ABP
		Proyecto Final - Aplicación del Proceso de Ciencia de Datos	Archivo de programación y PDF con la solución del proyecto	Elabora el plan de instalación de acuerdo con las características del software a implantar.	Instrumento: Rubrica  Técnica: Taller ABP



## 5. GLOSARIO DE TÉRMINOS

- Algoritmo: Conjunto de instrucciones lógicas y ordenadas que permiten resolver un problema o realizar un cálculo.
- Análisis de datos: Proceso de examinar, limpiar, transformar e interpretar datos para obtener conclusiones útiles.
- Big Data: Conjunto de datos tan grandes o complejos que requieren herramientas y técnicas especiales para su procesamiento.
- Ciencia de Datos: Disciplina que combina programación, estadística y conocimiento del dominio para extraer conocimiento y valor de los datos.
- Dataset: Conjunto estructurado de datos organizados generalmente en filas (registros) y columnas (atributos).
- Machine Learning: Rama de la inteligencia artificial que permite a los sistemas aprender automáticamente a partir de los datos sin ser programados explícitamente.
- Modelo predictivo: Representación matemática entrenada con datos históricos para realizar predicciones sobre nuevos casos.
- Limpieza de datos: Etapa del proceso donde se eliminan o corrigen valores erróneos, duplicados o incompletos en un conjunto de datos.
- Visualización de datos: Representación gráfica de la información que facilita la comprensión de patrones y tendencias.
- Python: Lenguaje de programación ampliamente utilizado en ciencia de datos por su simplicidad y sus librerías especializadas (pandas, numpy, scikit-learn, etc.).

## 6. REFERENTES BIBLIOGRÁFICOS

- Chollet, F. (2021). Deep Learning with Python (2.ª ed.). Manning Publications.
- De la Torre Díez, I., & García-Zapirain, B. (2019). Introducción a la Ciencia de Datos con Python. Editorial Universidad de Deusto. Recuperado de <https://repositorio.deusto.es/handle/10125/9291>



- Domo. (2025). Data Never Sleeps: AI Edition 2025 [Infografía]. Domo. Recuperado de <https://www.domo.com/es/learn/infographic/data-never-sleeps-ai-edition-2025>
- El Mundo de los Datos. (s.f.). El Mundo de los Datos. Recuperado de <https://elmundodelosdatos.com/>
- Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2.ª ed.). O'Reilly Media.
- IBM. (2023). Ciclo de vida de la ciencia de datos. IBM. Recuperado de <https://www.ibm.com/co-es/topics/data-science>
- IBM. (2023). Introducción al aprendizaje automático (Machine Learning). IBM. Recuperado de <https://www.ibm.com/co-es/topics/machine-learning>
- McKinney, W. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2.ª ed.). O'Reilly Media.
- Microsoft Learn. (2024). Introducción al aprendizaje automático con Python y Azure Machine Learning. Microsoft. Recuperado de <https://learn.microsoft.com/es-es/training/paths/intro-machine-learning-python-azure/>
- Provost, F., & Fawcett, T. (2013). Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking. O'Reilly Media.
- Python Software Foundation. (2024). The Python Tutorial. Python.org. Recuperado de <https://docs.python.org/es/3/tutorial/>
- Streamlit Inc. (2024). Getting Started with Streamlit. Recuperado de <https://docs.streamlit.io/get-started>

## 7. CONTROL DEL DOCUMENTO

	Nombre	Cargo	Dependencia	Fecha
<b>Autor (es)</b>	Alvaro Pérez Niño	Instructor	Centro de Servicios y Gestión Empresarial	Octubre 06 de 2025



**8. CONTROL DE CAMBIOS** (diligenciar únicamente si realiza ajustes a la guía)

	Nombre	Cargo	Dependencia	Fecha	Razón del Cambio
Autor (es)					