



Fundamentos de Data Science con Python

Sesión 1: Introducción a Data Science y Python

Dr. Julio Lopez-Nunez

Diciembre, 2025



Agenda

02 al 04 dic.
09 al 11 dic.

18:00 a 21:30

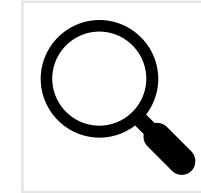


Sesión	Objetivos
Sesión 1	<ul style="list-style-type: none">• Comprender qué es Data Science y su relevancia en ingeniería y negocios.• Identificar el rol del Data Scientist.• Conocer notebooks interactivos (Google Colab, CoCalc).• Manipular datos básicos con pandas y numpy.
Sesión 2	<ul style="list-style-type: none">• Comprender qué es el Análisis Exploratorio de Datos (EDA).• Detectar problemas comunes en los datos: valores faltantes, duplicados, tipos incorrectos.• Aplicar técnicas de limpieza básica en Python con pandas.• Generar visualizaciones iniciales con matplotlib y seaborn.• Desarrollar criterios para interpretar patrones y tendencias.
Sesión 3	<ul style="list-style-type: none">• Comprender qué es Machine Learning y sus tipos.• Diferenciar entre regresión y clasificación.• Entrenar modelos simples con scikit-learn.• Evaluar la calidad de los modelos (métricas básicas).
Sesión 4	<ul style="list-style-type: none">• Comprender los fundamentos de las redes neuronales artificiales y su relación con el aprendizaje automático tradicional.• Conocer conceptos clave de Deep Learning: neuronas, capas, activaciones.• Comprender el perceptrón.• Comprender la red neuronal multicapa (MLP).
Sesión 5	<ul style="list-style-type: none">• Comprender conceptos avanzados de redes neuronales.• Conocer CNN y Transfer Learning de forma intuitiva.• Reflexionar sobre riesgos, sesgos y uso ético de la IA.
Sesión 6	<ul style="list-style-type: none">• Presentar y evaluar el proyecto final integrador del curso.• Integrar todas las herramientas aprendidas (EDA, ML, Deep Learning).• Reflexionar sobre desafíos reales al aplicar Ciencia de Datos.

Objetivos de la sesión.



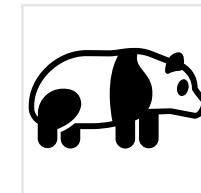
Comprender qué es Data Science y su relevancia en ingeniería y negocios.



Identificar el rol del Data Scientist.



Conocer notebooks interactivos (Google Colab, CoCalc).



Manipular datos básicos con pandas y numpy.

¿Qué es Data Science?



Definición: Ciencia interdisciplinaria (estadística + programación + conocimiento del dominio).



Ciclo de vida:



Obtención de datos



Limpieza/preparación



Análisis exploratorio



Modelado/predicción



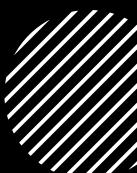
Comunicación de resultados



Aplicaciones en negocios, ingeniería, salud, ciencias sociales.



El rol del Data Scientist.



Perfil híbrido:

Estadística y
Programación
(Python, R).

**Habilidades
esenciales:** Análisis,
ML, Comunicación.

Conocimiento del
contexto.

Enfoque ético en el
uso de datos.

Introducción a Python.



About Google Colaboratory

(1) Historia: https://en.wikipedia.org/wiki/Google_Colab



Lenguaje abierto, versátil, con gran comunidad.



Entornos: Google Colab y CoCalc.



Bibliotecas clave: numpy, pandas, matplotlib, seaborn, scikit-learn.

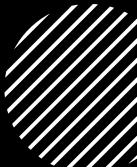


Ejemplo:

```
print("Hola, Ciencia de Datos!")
```



Primeros pasos con numpy.



(1) Historia: <https://es.wikipedia.org/wiki/NumPy>

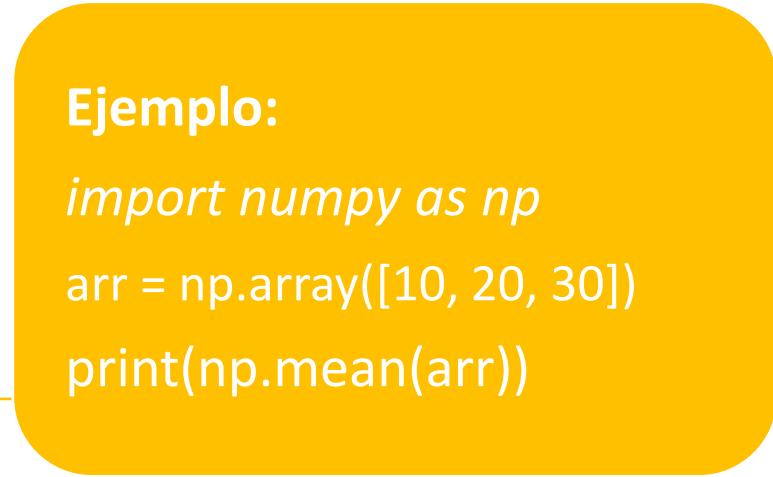
(2) Documentación: <https://numpy.org/>



Arreglos numéricos.



Operaciones vectorizadas.



Ejemplo:

```
import numpy as np  
arr = np.array([10, 20, 30])  
print(np.mean(arr))
```

Tareas comunes con Pandas.



(1) Historia: <https://es.wikipedia.org/wiki/Pandas>

(2) Documentación: <https://pandas.pydata.org/docs/>



Tablas de datos: DataFrames.



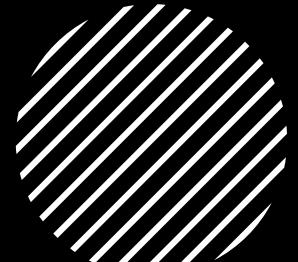
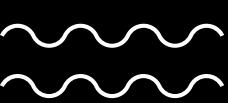
Importar, explorar, limpiar.



Métodos básicos:
.head(), .info(), .describe().



Ejemplo: cargar CSV en Colab.



Actividad práctica.

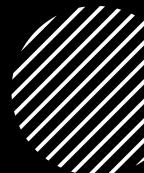
Importar un dataset (ejemplo: CSV con estudiantes o ventas).

Explorar filas y columnas.

Identificar valores faltantes.

Discusión de hallazgos iniciales.

Discusión grupal.



Pregunta guía:



¿Cómo aplicarían la Ciencia de Datos en sus áreas de enseñanza o investigación?

“If you torture the data long enough, it will confess.” – atribuido a Ronald Coase

¿Qué creen ustedes?...

¿Se trata de aprovechar al máximo la información o de manipularla para confirmar sesgos?

Cierre de la sesión 1.



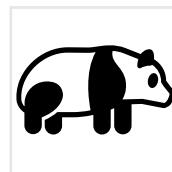
Definición y ciclo de Data Science.



Rol del Data Scientist.



Conocer notebooks interactivos (Google Colab, CoCalc).



Primer análisis de datos con numpy y pandas.