

Temario del Curso: Introducción a la Ciencia de Datos

Unidad 1. Introducción y motivación

- 1 Qué es la ciencia de datos y su relación con el método científico.
- 2 El ciclo de vida de un proyecto de datos.
- 3 Aplicaciones interdisciplinarias: física, deportes, biología.

Unidad 2. Primeros pasos en R y Python

- 1 Entornos de desarrollo: Jupyter, RStudio.
- 2 Tipos de datos, estructuras y librerías básicas.
- 3 Importación y visualización inicial de datos.

Unidad 3. Exploración de datos

- 1 Estadística descriptiva aplicada.
- 2 Distribuciones, medidas de centralidad y dispersión.
- 3 Visualización de datos categóricos y numéricos.

Unidad 4. Modelos básicos

- 1 Regresión lineal simple y múltiple.
- 2 Interpretación de parámetros.
- 3 Diagnóstico de modelos.

Unidad 5. Modelos generalizados

- 1 Regresión logística y de Poisson.
- 2 Modelos lineales generalizados (GLM).
- 3 Aplicaciones en física y deportes.

Unidad 6. Evaluación y validación de modelos

- 1 Métricas de desempeño: RMSE, precisión, recall, F1.
- 2 Validación cruzada.
- 3 Selección de modelos.

Unidad 7. Análisis no supervisado

- 1 PCA y reducción de dimensionalidad.
- 2 Clustering: k-means, jerárquico.
- 3 Aplicaciones experimentales.

Unidad 8. Series de tiempo

- 1 Conceptos básicos y componentes de series.
- 2 Modelos ARIMA.
- 3 Aplicaciones en física y deportes.

Unidad 9. Obtención y preparación de datos

- 1 Fuentes de datos abiertos.
- 2 Web scraping con Python y R.
- 3 Limpieza y depuración de datos reales.

Unidad 10. Comunicación de resultados

- 1 Narrativas visuales con datos.
- 2 RMarkdown, Jupyter y reportes reproducibles.
- 3 Buenas prácticas de comunicación científica.

Unidad 11. Ética y aplicaciones

- 1 Sesgos en los datos.
- 2 Privacidad, justicia algorítmica y responsabilidad.
- 3 Reflexiones interdisciplinarias.

Unidad 12. Proyecto final

- 1 Selección de dataset y pregunta de análisis.
- 2 Desarrollo del análisis completo.
- 3 Presentación de resultados (oral y escrita).