

Introducción a la ciencia de datos en salud con Python

Los notebooks y bases de datos actualizados del curso están disponibles en el [Github del curso](#), un repositorio muy usado en el mundo de la programación y ciencia de datos para intercambiar información.

Módulo I: Introducción a la programación en Python

- Cómo instalar Python
- Archivos de programa en Python
- Jupyter Notebooks
- Variables y tipos
- Operadores y comparaciones
- Tipos compuestos (cadenas, listas, tuplas)
- Control de flujo
- Ciclos
- Funciones
- Clases
- Módulos
- Excepciones

Módulo II: Exploración y procesamiento de datos

- Librerías principales
- Importación de datos: txt, csv, xlsx, y zip.
- Inspección inicial de la información.
- Manejo básico pandas
 - Selección de filas y/o columnas, selección condicionada.
 - Funciones apply, mapping and merge.
 - Funciones de manejo de texto.
 - Creación de muestras y subconjuntos de data frames: subsets and samples.

Módulo III: Paquetes Gráficos

- Introducción a plotly y matplotlib y demás librerías gráficas.
- Gráficos base: histogramas, scatter plot, bar plot, box plot.
- Otras herramientas para visualización de datos: manejo de mapas.
- Personalización de gráficos: ejes, colores, títulos.
- Exportación de gráficos.

Módulo IV: Georreferenciación

- Georreferenciación de información con el módulo Geopandas
- Introducción a las proyecciones y formatos de capas con información espacial (shape, GeoJson)
- Puntos y mapas coropleticos
- Operaciones de análisis espacial
- Isocronas

Bibliografía y sitios recomendados

- [Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython](#), Wes McKinney. Del creador de la librería Pandas, un libro detallado sobre su uso
- <http://www.python.org> - The official web page of the Python programming language.
- <http://www.python.org/dev/peps/pep-0008> - Guía de estilo para la programación en Python.
- <http://www.greenteapress.com/thinkpython/> - Libro gratuito sobre Python.
- [Python Essential Reference](#) - Un buen libro de referencia sobre programación en Python.

Recursos adicionales:

- Kaggle: Sitio con datagrames y notebooks preparados para practicar EDA y machine learning. Bases de datos de todo tipo, con código comentado. Recomendado para practicar con datos reales.
- Google Colab: Editor y ejecutor de notebooks en la nube. No hace falta instalar nada local. Gratis.
- [Corey Schafer](#): Excelente fuente de tutoriales sobre Python.
- [Keith Galli](#): Completo tutorial de Pandas