# *Python Data Science, manipulación y visualización de datos*
## *Certificación TOSA® opcional*

*Curso práctico de 4 días - 28h*
*Ref.: IYT - Precio 2025: 1 700€ sin IVA*

La ciencia de datos es un campo multidisciplinar en constante expansión. Se basa en métodos, algoritmos y procesos científicos que Python domina gracias a un ecosistema especialmente rico. Se ha convertido en el lenguaje de referencia para el análisis de datos, sea cual sea su formato. Nuestra formación le permitirá familiarizarse con las herramientas, las bibliotecas y los módulos de Python para adquirir rápidamente competencias en ciencia de datos con este lenguaje.

### OBJETIVOS PEDAGÓGICOS

Al término de la formación, el alumno podrá:

Tener una visión general del ecosistema científico de Python

Conozca las bibliotecas científicas esenciales para la ciencia de datos

Ser capaz de manipular grandes conjuntos de datos con Python

Comprender las ventajas de la visualización de datos

Visualización de datos con Python

### MÉTODOS PEDAGÓGICOS

El formador proporciona enseñanza activa, retroalimentación y demostraciones para que los participantes puedan ponerlas en práctica más rápidamente.

Los temas se ilustran con numerosos ejercicios.

## PROGRAMA

última actualización: 08/2024

### 1) Presentación del ecosistema científico Python

- Visión general del ecosistema científico de Python: las bibliotecas esenciales.
- Saber dónde encontrar nuevas librerías y evaluar su sostenibilidad.
- Las principales herramientas y software de código abierto para la ciencia de datos.
- Por qué usar la distribución científica, Anaconda.
- Comprender las ventajas de un entorno virtual y saber cómo utilizarlo.
- El intérprete IPython y el servidor Jupyter.
- Prácticas recomendadas para empezar con buen pie un proyecto de ciencia de datos con Python.
- Formatos de archivos científicos y bibliotecas para manipularlos.

*Trabajo práctico : Configuración del entorno de desarrollo: instalación de Anaconda, creación de un entorno virtual, exportación y duplicación de un entorno, uso de cuadernos Jupyter.*

### 2) La pila SciPy

- La base de las bibliotecas científicas esenciales en las que se basan todas las demás: la pila SciPy.
- NumPy: cálculo numérico y álgebra lineal (vectores, matrices, imágenes).
- SciPy, basado en NumPy para estadística, análisis funcional, análisis geoespacial, procesamiento de señales, etc.

### PARTICIPANTES

Ingenieros, desarrolladores, investigadores, científicos de datos, analistas de datos y cualquier persona que desee conocer el mundo científico de Python.

### REQUISITOS PREVIOS

Práctica del lenguaje Python.

### COMPETENCIAS DEL FORMADOR

Los expertos que imparten la formación son especialistas en las materias tratadas. Han sido validados por nuestros equipos pedagógicos, tanto en el plano de los conocimientos profesionales como en el de la pedagogía, para cada curso que imparten. Cuentan al menos con entre cinco y diez años de experiencia en su área y ocupan o han ocupado puestos de responsabilidad en empresas.

### MODALIDADES DE EVALUACIÓN

El formador evalúa los progresos pedagógicos del participante a lo largo de toda la formación mediante preguntas de opción múltiple, escenificaciones de situaciones, trabajos prácticos, etc.
El participante también completará una prueba de posicionamiento previo y posterior para validar las competencias adquiridas.

### MEDIOS PEDAGÓGICOS Y TÉCNICOS

• Los medios pedagógicos y los métodos de enseñanza utilizados son principalmente: ayudas audiovisuales, documentación y soporte de cursos, ejercicios prácticos de aplicación y ejercicios corregidos para los cursillos prácticos, estudios de casos o presentación de casos reales para los seminarios de formación.
• Al final de cada cursillo o seminario, ORSYS facilita a los participantes un cuestionario de evaluación del curso que analizarán luego nuestros equipos pedagógicos.
• Al final de la formación se entrega una hoja de presencia por cada media jornada de presencia, así como un certificado de fin de formación si el alumno ha asistido a la totalidad de la sesión.

### MODALIDADES Y PLAZOS DE ACCESO

La inscripción debe estar finalizada 24 horas antes del inicio de la formación.

### ACCESIBILIDAD DE LAS PERSONAS CON DISCAPACIDAD

¿Tiene alguna necesidad específica de accesibilidad? Póngase en contacto con la Sra. FOSSE, interlocutora sobre discapacidad, en la siguiente dirección psh-accueil@orsys.fr para estudiar de la mejor forma posible su solicitud y su viabilidad.

- Pandas: análisis de datos tabulares (CSV, Excel, etc.), estadísticas, pivotes, filtros, búsquedas, etc.
- Matplotlib: la biblioteca esencial para la visualización de datos.

*Trabajo práctico : Mida el rendimiento del NumPy instalado por su Linux y el de Anaconda. Procesamiento de imágenes con NumPy. Primeros gráficos. Análisis estadístico de ficheros CSV. Primeros elementos de mapeo. Transformadas de Fourier.*

### 3) Mostrar bibliotecas

- Visión general de las bibliotecas de visualización de Python: 2D/3D, escritorio/web, estadística, cartografía, big data, etc.
- Bibliotecas orientadas al escritorio: Matplotlib, Pandas, Seaborn.
- Bibliotecas orientadas a la web: Bokeh, Altair, Plotly...
- Librerías 3D: Plotly, pythreejs, ipyvolume, etc.
- Bibliotecas cartográficas: Cartopy, folium, ipyleaflet, Bokeh, cesiumpy, etc.
- Bibliotecas de big data: datashader, Vaex...

*Trabajo práctico : Una serie de ejercicios utilizando algunas de las bibliotecas presentadas. Big data, visualización cartográfica, 2D y 3D.*

### 4) Visualización de datos

- Ventajas de la visualización de datos
- Uso de PyViz y el ecosistema HoloViz.
- Presentación de las herramientas SuperSet, Mayavi, Paraview y VisIt.

*Trabajo práctico : Seguir utilizando bibliotecas de visualización y herramientas de manipulación.*

### 5) Formatos de archivos científicos y tratamiento de datos voluminosos

- Panorama de los principales formatos de archivos científicos: NetCDF, HDF5, GRIB, JSON, PARQUET, MATLAB, CGNS, etc.
- Manejo de big data con Dask, Vaex y Xarray.

*Trabajo práctico : Manejo de datos de más de GB, lectura y escritura de archivos NetCDF/HDF5. Visualización de datos climáticos, imágenes de satélite, creación de vídeos/animaciones gráficas.*

## FECHAS

Contacto