



CENTRO MICRODATOS
Departamento de Economía

CURSO PARA EL ANÁLISIS DE DATOS CON PYTHON (NIVEL BÁSICO)

1. Descripción del Curso

Este curso tiene como objetivo entregar las herramientas necesarias para manipular y analizar datos de forma efectiva utilizando Python como lenguaje de programación. El curso se separa en cinco partes. La primera contempla una introducción detallada a este lenguaje junto con la revisión de las librerías básicas de manipulación de datos. La segunda sección revisa algunas herramientas útiles para el análisis tales como el uso de condicionales, funciones, iteradores, entre otros. En la tercera parte se revisará detalladamente como importar, unir, manipular, limpiar y explorar bases de datos. La cuarta sección revisa en las librerías de visualización de datos con el fin de confeccionar gráficos estáticos e interactivos. La última sección cubre: análisis estadístico utilizando modelos de regresión lineal y la obtención de datos desde internet utilizando herramientas de *web-scraping*.

Cada clase contará con una revisión del taller de la clase previa, exposición de los contenidos y múltiples ejemplos en clases con el fin de que se adquieran los conocimientos de forma práctica. Algunas clases contarán con un ejercicio práctico que se realizara en clases.

2. Público Destinatario

Profesionales interesados en la gestión y análisis de datos tanto del mundo público como del privado que estén interesados en complementar su formación profesional y aplicar los conocimientos obtenidos en un contexto de análisis de datos.

3. Modalidad

- Presencial o
- Distancia

4. Duración

30 horas, distribuidas en 10 clases.

5. Contenidos

Clase 1: Introducción a Python I: conceptos básicos

- Instalar Python.
- Interfaz de Python.
- Códigos en Python.
- Sintaxis básica.
- Semántica: Variables, objetos, operadores.
- Tipos de datos y estructuras.
- Importar paquetes e instalar librerías.
- Documentación de librerías
- Uso de foros.
- Taller 1.

Clase 2: Introducción a Python II: NumPy, Matplotlib y Pandas

- Revisión Taller 1.
- Arreglos.
- Operaciones matemáticas y estadísticas descriptivas básicas.
- Manipulación de arreglos.
- Gráficos básicos del tipo: barras, líneas, histogramas y de punto con Matplotlib.
- *Broadcasting*.
- Librer
- Taller 2.

Clase 3: Herramientas I: condicionales, loops, iteradores, *lists comprehension* y generadores.

- Revisión Taller 2.
- Condicionales.
- Loops.
- Utilización de iteradores.
- *Lists Comprehension*.
- Generadores.

- Taller 3.

Clase 4: Herramientas II: escribir funciones en Python

- Revisión Taller 3.
- Usar funciones.
- Definir funciones simples.
- Definir funciones con múltiples parámetros.
- Definir funciones anidadas.
- Definir funciones anónimas (lambda).
- Ejercicio de aplicación (en clases).
- Taller 4.

Clase 5: Análisis de datos I: importar datos de distintas fuentes y unir bases de datos

- Revisión Taller 4.
- Repaso librería Pandas.
- Importar datos de distinto tipo.
- Combinar bases de datos con distintos tipos de operadores.
- Unión de bases de datos en distintos formatos.
- Ejercicio de aplicación (en clases).
- Taller 5.

Clase 6: Análisis de datos II: manipulación de bases de datos

- Revisión Taller 5.
- Transformar datos: inspeccionar, ordenar, crear subconjuntos, agregar columnas.
- Agregación de datos: Estadísticas descriptivas, contar, agrupar y filtrar.
- Indexación avanzada: cambios, construcción, multi-indexación e indexación jerárquica.
- Ejercicio de aplicación (en clases).
- Taller 6.

Clase 7: Análisis de datos III: limpiar, preparar y explorar base de datos

- Revisión Taller 6.
- Problemas típicos al limpiar bases de datos (ej. duplicados, *missing values*).
- Texto y variables categóricas.
- Exploración de datos.
- Ejemplo de aplicación (en clases).
- Taller 7.

Clase 8: Visualización de datos: creación de gráficos estáticos e interactivos

- Revisión Taller 7.
- Visualización de datos con Matplotlib y Plotly.
- Confeccionar gráficos de distintos tipos, graficar series de tiempo, automatizar procesos de visualización, exportar y personalizar gráficos.
- Taller 8.

Clase 9: Web-Scrapping con Python

- Revisión Taller 9.
- Breve Introducción a HyperText Markup Language (HTML), XML Path Language (XPath) y Cascading Style Sheets (CSS).
- Spiders.
- Librería Selenium.
- Ejercicio de aplicación (1): extraer datos desde Wikipedia.
- Ejercicio de aplicación (2): extraer datos desde sitios web complejos con Selenium.
- Ejercicio de aplicación (3): extraer datos utilizando API de Google Street.
- Taller 10.

Clase 10: Análisis estadístico con Python

- Revisión Taller 10.
- Explorar datos de variables continuas y discretas analizando su estadística descriptiva.
- Introducción a testeo de hipótesis.
- Análisis de regresión lineal y logística.
- Taller de evaluación final.

6. Certificación del curso

El curso requiere de un 75% de asistencia y de la aprobación de la evaluación final.