

# Introducción a Data Science: Visualización y análisis de datos científicos utilizando Python

Rubén Bartolo Morales

## Contenido

### 0. Introducción

Python es un lenguaje de programación multiparadigma, dinámico e interpretado fácil de aprender, también conocido como un lenguaje de scripts.

#### ¿Por qué Python?

#### Aplicaciones de Python

- Desarrollo web
  - Django
  - Flask
- Aplicaciones de escritorio
- Scripting
  - Integración con otros lenguajes

#### Python en la ciencia de datos

**1995: *Numeric***, Jim Hugunin, et. al. Objeto Array básico, Python para calculo científico.

**2001: *Travis E. Oliphant (Inicios de SciPy)***

Fernando Pérez crea IPython

John Hunter crea matplotlib

#### ***SciPy, IPython y Matplotlib***

**Matplotlib** considerado el estándar de facto para visualización con Python

Basado en la API de Matlab.

Gráficos de alta calidad

**SciPy** colección de algoritmos para tareas comunes:

Integración y EDOs (scipy.integrate)

Procesamiento de señales (scipy.signal)

Optimización (scipy.optimize)

Entre otras...

### **SymPy** para calculo simbólico

Permite la creación de sistemas de algebra computacional (CAS) estilo Maple, Mathematica...

Escrito en Python

Soporte para LaTeX

### **IPython** Interprete de Python mejorado

Entorno interactivo ideal para el aprendizaje

Ha evolucionado a Jupyter permitiendo la creación de notebooks en la web sin instalación del interprete.

Multiplataforma gracias al proyecto Anaconda.

### **2006: NumPy**

Arrays multidimensionales, funciones rápidas y eficientes para operar con ellos.

Vectorización de bucles lentos de Python.

### ***Python VS R / Matlab***

## **1. Introducción a la sintaxis de Python**

**Tipos numéricos**

**Operaciones**

**Operaciones con números complejos**

**Valor absoluto**

**Conversiones entre tipos de datos**

**Asignación y operadores de comparación**

**Secuencias: listas y tuplas**

**Estructuras de control**

**Condicionales**

**Bucles**

**Funciones**

## **2.- Introducción a NumPy (Arrays)**

**¿Qué es NumPy?**

**Constantes y funciones matemáticas**

**¿Qué es un Array?**

**Eficiencia**

**Indexación de Arrays**

**Creación de Arrays**

**Operaciones con Arrays**

**Funciones de comparación**

**Ejemplos y Ejercicios**

## **3.- Representación gráfica con Matplotlib y Pandas**

**¿Qué es Matplotlib?**

**Interfaz PyPlot**

**Función plot**

**Personalización**

**Otros tipos de gráficas**

**Interfaz orientada a objetos de Matplotlib**

**Pandas**

**¿Qué problemas resuelve Pandas?**

**Creación de un DataFrame**

**Principales funciones para manipular un Dataframe**

**Widgets de Matplotlib**

## **4.- SciPy**

**El ecosistema de SciPy**

**Resolver ecuaciones matemáticas con SciPy**

**Integración Numérica**

**Ecuaciones diferenciales ordinarias (EDOs)**

**Sistemas de ecuaciones no lineales**

**Optimización**

**Interpolación**

**Ajuste de curvas**

## **5.- SymPy**

**Creación de símbolos**

**Expresiones**

**Simplificación**

**Derivadas e Integrales**

**Limites**

**Series**

**Resolución de ecuaciones**

**Ecuaciones diferenciales**