# Introducción a la ciencia de datos

## Módulo 1. Cómputo científico con Numpy

### Sesión 1: Introducción al Cómputo Científico y Python (1 hora)

* **Conceptos Básicos de Python** (20 min): Repaso rápido de Python. Variables, tipos de datos básicos, estructuras de control (if, for, while).
* **Entorno de Desarrollo** (20 min): Instalación de Python y configuración del entorno. Uso de Jupyter Notebooks o Google Colab para ejemplos interactivos.
* **Introducción al Cómputo Científico** (20 min): Conceptos básicos y aplicaciones. Importancia de NumPy en el ecosistema de Python para ciencia e ingeniería.

### Sesión 2: Fundamentos de NumPy (1 hora)

* **Creación y Manipulación de Arrays** (30 min): Creación de arrays, dimensiones, acceso a elementos, slicing y dicing.
* **Operaciones con Arrays** (30 min): Operaciones elemento a elemento, agregaciones (max, min, sum), y operaciones de broadcast.

### Sesión 3: Álgebra Lineal con NumPy (1 hora)

* **Operaciones Básicas de Álgebra Lineal** (30 min): Vectores, matrices, operaciones matriciales básicas (suma, resta, multiplicación).
* **Sistemas de Ecuaciones y Eigenvalores** (30 min): Uso de NumPy para resolver sistemas de ecuaciones lineales, cálculo de eigenvalores y eigenvectores.

### Sesión 4: Análisis de Datos con NumPy (1 hora)

* **Estadísticas Descriptivas** (30 min): Funciones para cálculo de medias, medianas, desviaciones estándar, etc.
* **Manipulación y Limpieza de Datos** (30 min): Técnicas para manejar datos faltantes, filtrado de datos, y transformaciones.

### Sesión 5: Visualización de Datos (1 hora)

* **Introducción a Matplotlib** (30 min): Creación de gráficos básicos: líneas, barras, histogramas.
* **Integración de NumPy con Matplotlib** (30 min): Uso de arrays de NumPy para crear visualizaciones complejas.

### Sesión 6: Aplicaciones Científicas de NumPy (1 hora)

* **Simulaciones y Modelos** (30 min): Uso de NumPy para simulaciones estocásticas y modelado matemático.
* **Procesamiento de Imágenes** (30 min): Manipulación básica de imágenes con NumPy: filtros, transformaciones.

### Sesión 7: Optimización y Eficiencia (1 hora)

* **Mejorando el Rendimiento** (30 min): Consejos para escribir código NumPy eficiente. Uso de operaciones in-place, evitar bucles cuando sea posible.
* **Perfiles y Depuración** (30 min): Herramientas para medir y mejorar el rendimiento del código NumPy.

### Sesión 8: Proyecto Práctico (1 hora)

* **Desarrollo de un Proyecto** (1 hora): Aplicación de los conceptos aprendidos en un proyecto práctico, como análisis de un conjunto de datos o simulación de un problema científico.

## Módulo 2: Estadística Matemática

### Sesión 1: Fundamentos de Probabilidad (1 hora)

* **Conceptos Básicos de Probabilidad** (20 min): Introducción a la probabilidad, eventos, espacio muestral, probabilidad condicional.
* **Operaciones con Eventos** (20 min): Independencia, uniones, intersecciones, complementos.
* **Teorema de Bayes** (20 min): Introducción y aplicaciones prácticas.

### Sesión 2: Variables Aleatorias y Distribuciones (1 hora)

* **Variables Aleatorias** (30 min): Definición, variables aleatorias discretas y continuas, función de masa de probabilidad (PMF), función de densidad de probabilidad (PDF).
* **Distribuciones Importantes** (30 min): Distribución binomial, distribución normal, distribución uniforme. Uso de NumPy para generar datos aleatorios.

### Sesión 3: Uso de NumPy para Probabilidad y Estadística (1 hora)

* **Generación de Números Aleatorios** (30 min): Uso de numpy.random para generar muestras aleatorias.
* **Funciones Estadísticas Básicas con NumPy** (30 min): Media, mediana, desviación estándar, correlación. Ejemplos prácticos.

### Sesión 4: Introducción a SciPy para Probabilidad (1 hora)

* **Distribuciones con SciPy** (30 min): Uso de scipy.stats para trabajar con diferentes distribuciones de probabilidad.
* **Aplicaciones de las Distribuciones** (30 min): Cálculos de probabilidad, PDFs, CDFs, y percentiles con SciPy.

### Sesión 5: Estimación de Parámetros y Ajuste de Modelos (1 hora)

* **Métodos de Estimación** (30 min): Estimadores de máxima verosimilitud (MLE), mínimos cuadrados.
* **Ajuste de Distribuciones** (30 min): Utilizar SciPy para ajustar distribuciones a datos. Ejemplos con datos reales.

### Sesión 6: Pruebas de Hipótesis (1 hora)

* **Conceptos Básicos de Pruebas de Hipótesis** (30 min): Hipótesis nula, hipótesis alternativa, nivel de significancia, p-valor.
* **Pruebas de Hipótesis con SciPy** (30 min): Realizar pruebas Z, t-tests, y pruebas chi-cuadrado utilizando SciPy.

### Sesión 7: Regresión y Correlación (1 hora)

* **Correlación** (30 min): Pearson, Spearman. Cálculo de coeficientes de correlación con SciPy.
* **Regresión Lineal** (30 min): Uso de SciPy para ajuste de regresión lineal. Interpretación de resultados.

### Sesión 8: Proyecto Práctico (1 hora)

* **Desarrollo de un Proyecto** (1 hora): Implementar un proyecto que incluya análisis estadístico, estimación de parámetros, pruebas de hipótesis y regresión. Se podría analizar un conjunto de datos y modelar utilizando distribuciones de probabilidad.

## Módulo 3: Álgebra lineal

### Sesión 1: Introducción y Configuración del Entorno (1 hora)

* **Introducción al Álgebra Lineal** (20 min): Panorama general y aplicaciones.
* **Configuración del Entorno de Desarrollo** (20 min): Instalación y configuración de Python, NumPy, SciPy, y SymPy.
* **Introducción a NumPy, SciPy, y SymPy** (20 min): Breve introducción a las bibliotecas y sus roles en el cálculo matemático y simbólico.

### Sesión 2: Álgebra Matricial Básica con NumPy (2 horas)

* **Creación y Operaciones con Matrices** (1 hora): Uso de NumPy para crear matrices, operaciones básicas (suma, resta, multiplicación).
* **Determinantes, Inversas y Rango** (1 hora): Cálculo de determinantes, inversas de matrices y rango utilizando NumPy y SciPy.

### Sesión 3: Sistemas de Ecuaciones Lineales (2 horas)

* **Resolución de Sistemas Lineales** (1 hora): Uso de NumPy y SciPy para resolver sistemas de ecuaciones lineales.
* **Aplicaciones Prácticas y Problemas** (1 hora): Resolución de problemas aplicados utilizando sistemas de ecuaciones lineales.

### Sesión 4: Espacios Vectoriales (1 hora)

* **Conceptos Básicos** (30 min): Introducción a espacios vectoriales, subespacios, bases y dimensión.
* **Implementación con Python** (30 min): Uso de NumPy para explorar espacios vectoriales y subespacios.

### Sesión 5: Espacios Nulos, Columnas y Filas (1 hora)

* **Cálculo de Espacios Nulos** (30 min): Uso de SciPy para calcular espacios nulos de matrices.
* **Espacios de Columnas y Filas** (30 min): Teoría y ejemplos prácticos utilizando NumPy y SciPy.

### Sesión 6: Valores y Vectores Propios (2 horas)

* **Conceptos y Aplicaciones** (1 hora): Introducción a eigenvalores y eigenvectores, importancia en el álgebra lineal.
* **Cálculo con SciPy y NumPy** (1 hora): Uso de bibliotecas para encontrar eigenvalores y eigenvectores, ejercicios prácticos.

### Sesión 7: Diagonalización y Descomposición de Matrices (1 hora)

* **Teoría de la Diagonalización** (30 min): Conceptos y condiciones para la diagonalización de matrices.
* **Descomposiciones** (30 min): Uso de SciPy para descomposición LU, QR y SVD.

### Sesión 8: Aplicaciones Avanzadas y Simbólicas (2 horas)

* **Cálculo Simbólico con SymPy** (1 hora): Uso de SymPy para operaciones de álgebra lineal simbólica, simplificación de expresiones.
* **Proyectos de Aplicación** (1 hora): Desarrollo de pequeños proyectos que integren los conceptos aprendidos, como sistemas dinámicos, optimización o análisis de redes.

## Módulo 4 Cálculo

### Sesión 1: Introducción a los Conceptos Básicos del Cálculo (1 hora)

#### Parte 1: Conceptos Básicos del Cálculo Diferencial (30 minutos)

* Breve historia y significado del cálculo diferencial.
* Introducción a las derivadas: definición y significado físico.

#### Parte 2: Primeros Pasos con SymPy (30 minutos)

* Instalación y configuración del entorno de trabajo con SymPy.
* Introducción a SymPy: Variables simbólicas y expresiones.
* Ejercicio guiado: Cálculo de derivadas simples con SymPy.

### Sesión 2: Derivadas y Aplicaciones (1 hora)

#### Parte 1: Reglas de Derivación y Aplicaciones (30 minutos)

* Reglas de derivación: producto, cociente y cadena.
* Interpretación geométrica de las derivadas y aplicaciones a problemas de tasas de cambio.

#### Parte 2: Explorando Derivadas con SymPy (30 minutos)

* Uso de SymPy para aplicar reglas de derivación.
* Visualización de funciones y sus derivadas: un enfoque práctico.

### Sesión 3: Fundamentos del Cálculo de Varias Variables (1 hora)

#### Parte 1: Derivadas Parciales y Gradiente (30 minutos)

* Introducción a funciones de varias variables: concepto de derivadas parciales.
* El gradiente: interpretación y aplicación.

#### Parte 2: Cálculo Multivariable con SymPy (30 minutos)

* Cálculo de derivadas parciales y gradientes usando SymPy.
* Ejercicios prácticos: Análisis de curvas de nivel y superficies.

### Sesión 4: Introducción a la Optimización (1 hora)

#### Parte 1: Conceptos de Optimización (30 minutos)

* Introducción a la optimización de funciones: buscar máximos y mínimos.
* Condiciones para extremos en funciones de una y varias variables.

#### Parte 2: Optimización con SymPy (30 minutos)

* Uso de SymPy para encontrar puntos críticos y determinar su naturaleza (máximos, mínimos, puntos de silla).
* Ejercicio de cierre: Optimización de una función de dos variables, interpretación de resultados.