

# Planeación de Tópicos Selectos De Matemáticas Aplicadas II: Análisis de Datos con Python



## I. Información general

**Nombre de la UEA:** Análisis de Datos con Python

**Nombre del profesor:** Dr. Juan Alberto Martínez Cadena (jamc88@xanum.uam.mx)

## II. Información sobre el programa

**Contenido:**

### 1. Introducción a Python para Ciencia de Datos

- Repaso ágil de estructuras básicas (listas, diccionarios, etc.)
- Entorno Jupyter + buenas prácticas en notebooks científicos.
- Introducción a NumPy con enfoque en álgebra de matrices.

### 2. Análisis Exploratorio con Pandas

- Carga, limpieza y resumen de datasets reales.
- Estadística descriptiva (media, mediana, percentiles, correlaciones).
- Gráficas rápidas con `pandas.plot()` y `pairplots`.

### 3. Limpieza y Transformación de Datos

- Valores nulos, duplicados y codificación de variables.
- Agrupaciones (`groupby`), joins y creación de variables derivadas.
- Manejo de outliers con criterios estadísticos (`boxplot`, `z-score`).

### 4. Visualización de Datos

- `matplotlib` y `seaborn` para gráficos profesionales.
- Histogramas, diagramas de dispersión y relaciones multivariadas.
- Introducción a visualizaciones interactivas (`plotly` opcional).

## 5. Introducción a Series Temporales

- Fechas y tiempos en pandas, remuestreo y ventanas móviles.
- Detección de tendencias y estacionalidad.
- Caso práctico: datos ambientales o económicos.

## 6. Introducción a Machine Learning

- Flujo básico de un modelo en scikit-learn.
- División de datos, regresión lineal y clasificación básica.
- Validación cruzada y métricas (RMSE, accuracy, etc.).

### Calendarización de los temas:

Tema 1: Semana 1

Tema 2: Semanas 2 y 3

Tema 3: Semanas 4 y 5

Tema 4: Semanas 6 y 7

Tema 5: Semanas 8 y 9

Tema 6: Semanas 10 y 11

### Objetivos del curso:

Al finalizar el curso, el alumno será capaz de:

- Comprender y utilizar Python para análisis de datos
- Realizar análisis y limpieza de datos de manera eficiente
- Visualizar datos de manera efectiva
- Analizar datos en series temporales
- Aplicar técnicas de modelado y aprendizaje automático en Python
- Desarrollar un proyecto práctico de análisis de datos

## III. Evaluación

### Modalidades de evaluación y ponderación de exámenes:

- Dos exámenes parciales (70%) y un proyecto (30%)

### Fechas de exámenes parciales:

Primer parcial: viernes de la semana 4

Segundo parcial: viernes de la semana 8

Presentación de Proyecto: viernes de la semana 11

**Criterios y escalas para la asignación de la calificación definitiva.**

- Menos de 6, es **NA**
- De 6 a menos de 7.6, es **S**
- De 7.6 a menos de 8.6, es **B**
- De 8.6 a 10, es **MB**

**Bibliografía:**

1. Wes McKinney, *"Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython"*, 2nd Edition, O'Really Media, 2017.
2. Joel Grus, *"Data Science from Scratch: First Principles with Python"*, 2nd Edition, O'Really Media, 2019.
3. Jake VanderPlas, *"Python Data Science Handbook: Essential Tools for Working with Data"*, 2nd Edition, O'Really Media, 2023.
4. Andreas C. Müller and Sarah Guido, *"Introduction to machine learning with Python a guide for data scientists"*, O'Really Media, 2018.
5. Mendenhall, W. *Introducción a la Probabilidad y Estadística*, 14ª ed, 2015.