

1^a
Emisión

DATA SCIENCE

Módulo 05 Manipulación y visualización de datos con Python

Mtro. Ricardo Daniel Alanis Tamez



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
Dirección General de Cómputo y de Tecnologías de información y Comunicación
Dirección de Docencia en TIC



Educación
Continua
1971 - 2021

Módulos del diplomado

No.	Módulo	Duración
1	Introducción a la ciencia de datos	4
2	Matemática necesaria para la ciencia de datos	13
3	Estadística y probabilidad para ciencia de datos	15
4	Metodologías de ciencia de datos	8
5	Manipulación y visualización de datos con Python	14
6	Aprendizaje de máquina supervisado	25
7	Aprendizaje de máquina no supervisado	15
8	Introducción al Deep Learning	18
9	Procesamiento de lenguaje natural o minería de textos	18
10	Gobierno de datos	5
11	Arquitectura tecnológica para la práctica de la analítica	10
12	Datos masivos	14
13	Data operations (DataOps)	4
14	Storytelling en ciencia de datos	4
15	Casos de éxito en la industria	18

Presentación

En este módulo tendrás la oportunidad de conocer y fortalecer:

- Carga y Análisis de datos
- Manipulación y Exploración de Datos
- Visualización de la información y librerías

Agenda de hoy

1. Quienes estamos aquí
2. Expectativas
3. Objetivo del Módulo
4. Contenido
5. Acuerdos
6. Forma de evaluación
7. Importancia de las clases que vienen

¿Quiénes estamos aquí?

Perfil del Profesor



- Head Data Science @ Nowports
- Data Science Lead @ Civica Digital
- Data Scientist at Codeando México

Consultant, Advisor, Mentor,
Community Leader, Activist

Fellow in several futurist
organizations such as Global
Shapers and Singularity
University

Compartamos expectativas

Objetivo

El participante estructurará código en **Python** para **manipular, limpiar y visualizar** datos, haciendo uso de las **librerías básicas** y más importantes del lenguaje enfocadas en datos y su análisis.

Contenido

- Encuadre (esta clase)
- Introducción a Pandas
- Lectura de datos
- El concepto de Dataframe y su API
- Agrupación y agregación de datos
- Filtrado y selección de datos
- Transformación de datos
- Librerías de visualización
- Gráficos estadísticos básicos
- Estrategias de visualización

Acuerdos

- Sobre la diferencia del aprendizaje adulto

Implicaciones

- Si estás aquí, estás aquí. Si no sabes algo, no pasa nada
- Si quieres preguntar algo, pregunta - aunque ya lo hayamos dicho
- Si quieres participar, participa; Si no quieres participar, participa.
- Si te quedaste con la duda, mándame un mensaje y buscaré atenderte
- **Regla Google Brain:** Si te atorras media hora con algo, mándame un mensaje.

Por qué este módulo es clave

Por qué **crees** que este módulo es clave

Por qué **creo** que este módulo es clave

Sondeo de Grupo

Respuestas

1. Import pandas as pd
2. Crear modelos de aprendizaje automático
3. `dataframe.column.plot()`
4. Descargarlo, cargarlo, encontrar la columna, `dataframe.column.mean()`

Es un sondeo, no tiene repercusiones

Evaluación

- 50% Prácticas en clase
- 20% Ejercicios de práctica en casa
- 30% Caso final (2 horas finales)

Contacto

Mtro. Ricardo Daniel Alanis Tamez

ricardo@codeandomexico.org

LinkedIn: Ricardo Alanís