



**SILABO**  
**Módulo: Python for Analytics**

<b>Del (los) programa(s)</b>	:	<b>Diploma Data Scientist</b>
<b>Edición</b>	:	<b>19</b>
<b>Duración</b>	:	<b>9 horas cronológicas   3 sesiones</b>
<b>Horario</b>	:	<b>07:30pm a 10:30pm</b>
<b>Docente</b>	:	<b>Geider D. Nuñuvero Angeles</b>
<b>Correo electrónico</b>	:	<b>jeider1624@gmail.com</b>

## I. SUMILLA

El módulo es de naturaleza teórico y práctico. Le permitirá al participante utilizar el lenguaje Python para cargar y manipular DataFrames como imput de datos para futuros proyectos de Data Science, combinándolos con variables, otras estructuras de datos y estructuras de control. Las temáticas que comprenden este módulo son:

- a) Introducción a Python.
- b) Variables y estructuras de datos en Python
- c) Estructuras de control en Python
- d) Manejo de Dataframes en Python (Pandas)

## II. METODOLOGÍA

### Metodología DMC: Learning Agile

La metodología DMC garantiza un aprendizaje significativo mediante la práctica intensiva, combinando tres enfoques:

#### Secuencia Didáctica Descubre-Explora-Aplica

- **Descubre** (Introducción al tema) → Videos, diapositivas y ejemplos.
- **Explora** (Práctica guiada) → Ejercicios en clase con apoyo del docente.
- **Aplica** (Práctica autónoma) → Ejercicios en casa con solucionarios (Puede ser parte de la Evaluación Formativa).

## III. HERRAMIENTAS

- Python, Google Colab.  
Provisto por DMC: No.  
Costo asociado por alumno: Ninguno.  
El alumno debe tener una cuenta Gmail – Google Drive con espacio disponible.

## IV. CONTENIDO

### Sesión 1

#### Introducción a Python

- Python. Definición, proceso de instalación (solo descriptivo).
- Software y editores de código para ciencia de datos.
- Taller: Exploración del entorno de trabajo de Google Colab.

**Caso de Uso:** diferenciación entre los distintos tipos de software para ciencia de datos y machine learning, para el desarrollo y despliegue de modelos.

**Aplicación:** dependiendo del equipo que integres en el trabajo, utilizarás distinto software y componentes.

**Ejemplo:** si en tu organización utilizan Databricks y dependiendo de las directrices de la misma, los desarrolladores pueden utilizar los workspace (Pyspark, Python) y los ML engineers (despliegue), MLFlow.

### Sesión 2

#### Variables y estructuras de datos en Python

- Variables. Definición y tipos de datos.
- Taller: Declaración, escritura y lectura de variables de diferentes tipos. Operaciones aritméticas básicas.
- Taller: Declaración, escritura y lectura de listas. Casos de uso
- Taller: Declaración, escritura y lectura de matrices. Casos de uso.
- Taller: Declaración, escritura y lectura de diccionarios. Casos de uso.

**Caso de Uso:** almacenamiento de las variables (features) de un modelo de ML en una variable colección.

**Aplicación:** llamado a esta variable colección cuando se está desarrollando-entrenando el modelo.

**Ejemplo:** creación de un data frame contenido solo algunas variables (indicadas en la variable colección) a partir del original, para entrenar un RandomForestClassifier().

### Sesión 3

#### Estructuras de control en Python

- Taller: Uso de la condicional (IF) y sus variantes en Python.
- Taller: Uso de bucles (WHILE, FOR) en Python.

#### Manejo de Dataframes en Python (Pandas)

- Taller: Declaración, escritura y lectura de dataframes.
- Taller: Agregar, cambiar y eliminar elementos de dataframe.
- Taller: Técnicas de selección de filas y columnas en un dataframe.

**Caso de Uso:** bucle FOR sobre elementos de una lista que contenga las columnas de un data frame.

**Aplicación:** función para obtener el tipo de dato de cada una de estas columnas y almacenarlos en otra lista.

**Ejemplo:** aplicación de la función type() utilizando el bucle FOR, así posteriormente aplicar One Hot Encoding a las variables string y poder utilizarlas en un modelo de ML.

## V. SISTEMA DE EVALUACIÓN

La calificación de este MÓDULO se realizará considerando lo siguiente:

CRITERIO	PESO	FECHA DE EVALUACIÓN
Evaluación Final 100%  1. Evaluación Final : Trabajo práctico.	100%	SESIÓN 2 : Inicio [04/12/2025] Fin [12/12/2025]

Leyenda:

Inicio: Fecha que inicia la evaluación

Fin: Fecha que finaliza la evaluación

Calificación de Diploma Total:

$$\text{Nota de Diploma} = \text{Promedio de Módulos } (N_1, N_2, \dots, N_{10}) * 40\% + N \text{ Proyecto} * 60\%$$

Leyenda:

N : Nota

### EVALUACIÓN FINAL

#### CORRESPONDIENTE A:

Programa	Diploma Data Scientist, Edición 19
Módulo	Python for Analytics
Objetivo	El módulo permitirá al participante comprender y aplicar Python en la manipulación y análisis de datos, utilizando estructuras y tipos de datos clave en ciencia de datos. Se abordarán técnicas de exploración y transformación de información, con un enfoque en la evaluación teórica y práctica a través de un cuestionario de opción múltiple.

#### CONSIDERACIONES GENERALES:

Modalidad	Individual
Medio y plazo de entrega	Aula virtual DMC, hasta el 12/12/2025 las 23:59.
Archivos y herramientas	• Es un archivo en formato de notebook
Producto académico (entregables esperados).	Se les dejará una serie de ejercicios aprendidas de las sesiones 1 y 2 y sus talleres, en total serán 20 preguntas cada uno de 1 punto. La nota mínima de aprobación será 14.

## VI. FUENTES DE INFORMACIÓN COMPLEMENTARIA

Las siguientes son fuentes de información que el docente le sugiere revisar para que pueda complementar los contenidos revisados en clase:

- (Raúl González Duque). (Python para todos).  
(<https://persoal.citius.usc.es/eva.cernadas/informaticaparacientificos/material/libros/Python%20para%20todos.pdf>).
- (Valentina Alvarez Gálvez & Adrián Soto Suárez). (Introducción a la programación en Python).  
(<https://adriansoto.cl/pdf/pythonbook.pdf>).
- (O'Reilly). (Python for Data Analysis). Recuperado de  
(<https://fasttacks.com/wpcontent/uploads/2024/07/Wes-McKinney-Python-for-Data-Analysis -Data-Wrangling-withpandas-NumPy-and-Jupyter-OReilly-Media-2022.pdf>).