



## Syllabus

### *Introducción Lenguaje Python*

Programa IA&DL República  
Dominicana

Guillermo Sánchez González

([guillermo.gonzalez@strategybigdata.com](mailto:guillermo.gonzalez@strategybigdata.com))

## 1. Presentación de la materia

El lenguaje de programación Python es uno de los que tiene mayor crecimiento recientemente. Su simplicidad, obligatoria estructuración y parecido con el pseudocódigo lo convierten en una herramienta versátil y muy útil para prototipar.

Junto a R, actualmente, es uno de los dos lenguajes más usados en ciencia de datos, constando de una amplia colección de librerías aplicables en el ámbito.

En este módulo se da una introducción a Python para rápidamente pasar a ver las herramientas fundamentales en la analítica de dato: tratamiento y analítica de tablas y representación gráfica.

## 2. Objetivos de aprendizaje

El objetivo general del módulo es el de ser capaz de montar un entorno de uso de Python así como aplicar las herramientas fundamentales de ciencia de datos.

Los objetivos específicos son los siguientes:

1. Instalación y uso de Jupyter Notebooks.
2. Carga de datos en tablas, análisis y transformaciones.
3. Representación gráfica de información en tablas.

### 3. Programa de la materia: estructura y contenido

TEMA. Conceptos generales y Python científico

Usaremos Anaconda para la versión de Python 3.6.

- Python en ciencia de datos:
  - ❖ Numpy
  - ❖ Pandas
  - ❖ Visualización (matplotlib, seaborn)
  - ❖ Representación de datos.
- Actividad 1. Análisis de un Dataset

### 4. Metodología y Actividades

El programa se desarrollará metodológicamente planteando los conceptos fundamentales de la actividad del data scientist, y fomentando la discusión de los principales aspectos de esa actividad.

El elemento central de la metodología docente es el trabajo práctico con un entorno de análisis de datos, para entender el trabajo del data scientist y adquirir habilidades básicas para seleccionar y aplicar algoritmos de aprendizaje automático. El entorno de análisis de datos es el IPython Notebook (ahora denominado Jupyter). Las clases presenciales se plantean con una primera parte teórica y una segunda parte de prácticas guiadas por el profesor.

Los alumnos que tengan un ordenador razonablemente potente ( 4 hilos y 8 GB de RAM o más), pueden instalar el *stack* de Anaconda que contiene todas las librerías incluido el Jupyter Notebook, en la versión Python 3.5+ entrando en <https://www.anaconda.com/download> y en la pestaña correspondiente a su sistema operativo. Se recomienda instalarlo si se puede antes de comenzar el módulo.

## CRONOGRAMA

Itinerario	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Presentación del módulo	Videoconferencia apertura					
Introducción al Machine Learning	Lectura Tema1					
Python en Data Science Tema 1		Python genérico y Numpy	Pandas	Representación visual		
Actividad 1. Análisis de dataset				Realizar actividad		
Cierre y conclusiones semana online						Videoconferencia cierre

Durante la realización de todo el módulo permanecerá abierto un canal para contactar con el profesor.

## Descripción de las actividades:

### Actividad 1. Análisis de un Dataset

**Carácter:** Individual

**Herramientas:** Foros, materiales Tema 1

**Desarrollo y plazo de ejecución:** Después de trabajar los materiales del tema 1, la primera actividad a realizar será la realización del análisis de un DataSet usando Pandas, se proporcionará un notebook con los enunciados y un csv.

Esta actividad es evaluable a través de un cuestionario en la plataforma.

## 5. Evaluación

El modelo de evaluación se basará en las actividades de curso desarrolladas (trabajos, casos, ejercicios, etc.). Las actividades se valorarán como sigue:

Actividad 1: Análisis de un DataSet :	50%
Participación en Foros y Videoconferencia:	50%

## 6. Bibliografía y materiales de consulta

### Materiales obligatorios

Instalación de Anaconda 3 descargando en <https://www.anaconda.com/distribution/> .

### Referencias recomendadas

- Python Data Science Handbook: Essential Tools for Working with Data

- Pandas Cookbook

## 7. CV del Profesor

Guillermo Gonzalez, Licenciado en Matemáticas, científico de datos en Strategy Big Data donde desarrolla modelos avanzados de machine Learning usando redes neuronales e implementando publicaciones de MIT o Stanford para tener las mejores algoritmos.