



UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS  
PROGRAMA ACADÉMICO DE MATEMÁTICAS  
FORMATO SYLLABUS  
PLAN DE ESTUDIOS 298  
VERSIÓN: 2022

RESOLUCIÓN ACREDITACIÓN DE ALTA CALIDAD No. 007575 DE JULIO DE 2019

**FACULTAD:** Ciencias Matemáticas y Naturales

**NOMBRE DEL DOCENTE:**

**ÁREA DE FORMACIÓN:** Electiva

**NOMBRE DEL ESPACIO ACADÉMICO:** Analítica de Datos

**TIPO DE ESPACIO:** Teórico () Práctico () Teo-prac () Obligatorio () Electivo ()

**CÓDIGO:** SIN CÓDIGO

**NÚMERO DE CRÉDITOS:** 2

**HORARIO:** Total Horas Semanales Lectivas: \_\_\_\_\_

DÍA: \_\_\_\_\_ HORA: \_\_\_\_\_ SALÓN: \_\_\_\_\_

DÍA: \_\_\_\_\_ HORA: \_\_\_\_\_ SALÓN: \_\_\_\_\_

DÍA: \_\_\_\_\_ HORA: \_\_\_\_\_ SALÓN: \_\_\_\_\_

## 1. JUSTIFICACIÓN DEL ESPACIO ACADÉMICO

El auge y crecimiento masificado del nuevo mundo de los datos, ha provocado que hayan surgido, en los últimos años, diversas herramientas, basadas en inteligencia artificial, para manipular estos datos y abstraer conocimiento a partir de allí, y a su vez ha crecido la demanda de capacidades para este tipo de especialidad. Muy bien lo indican Azevedo y Santos (2008) en donde proponen metodologías ampliamente reconocidas y utilizadas como SEMMA y CRISP-DM. Multinacionales como Microsoft con su plataforma AZURE e IBM con Watson Analytics entre otras confirman este hecho.

## 2. PRERREQUISITOS (Contenidos)

Se recomienda poseer conocimientos básicos en algún lenguaje de programación en particular (estructuras y funciones).

## 3. DESCRIPCIÓN DEL ESPACIO ACADÉMICO

Inducir al estudiante en el entorno de la analítica de datos, ampliamente aplicado en muchos campos y utilizado como base la Matemática y la Estadística. Además conocer



terminología y variantes posibles con el fin de generar interés para su profundización posterior.

## **4. OBJETIVOS**

### **4.1 GENERAL**

Presentar una metodología clara para la introducción a la analítica de datos, así como las herramientas básicas, basada en estándares reconocidos mediante la aplicación práctica con el fin de llevar a cabo los diversos pasos que se deben tener en cuenta en esta labor.

### **4.2 ESPECÍFICOS**

- Tener claros los pasos que se deben llevar a cabo en un proceso de analítica de datos.
- Manipular el lenguaje Python (en entorno Jupyter Notebook) y los paquetes más sobresalientes (Numpy, Matplotlib, Pandas, Sklearn, Scikit-Learn, Tensorflow, NLTK, Keras entre otros)
- Abarcar problemas de aprendizaje supervisado y no supervisado con algoritmos de Machine Learning.

## **5. UNIDADES TEMÁTICAS Y/O PROBLEMÁTICAS**

- Introducción (Campos de aplicación, metodologías, herramientas)
- Lenguaje Python (Entornos de desarrollo, preparación, instalación)
- Acceso a fuentes de datos
- Extracción, limpieza, preparación y visualización de datos
- Creación de modelos
- Prueba y validación
- Casos prácticos

## 6. ESTRATEGIAS METODOLÓGICAS:

Clases magistrales alternadas con sesiones de ejercicios y problemas en trabajo cooperativo. Cada tema se presenta de manera concisa con suficientes ejemplos ilustrativos. Es indispensable la generación y construcción de resultados fundamentales en la teoría y su demostración rigurosa.

Basado en el sistema de créditos, la distribución de la dedicación horaria del estudiante para este espacio académico es la siguiente:

HORAS			Horas profesor/s emana	Horas Estudiante/sem ana	Total Horas Estudiante/semestre	Créditos
TD	TC	TA	(TD + TC)	(TD + TC +TA)	X 16 semanas	

### Convenciones:

*TD: Trabajo Presencial Directo; trabajo de aula con plenaria de todos los estudiantes.*

*TC: Trabajo Mediado cooperativo; Trabajo de tutoría del docente a pequeños grupos o de forma individual a los estudiantes.*

*TA: Trabajo Autónomo; Trabajo del estudiante sin presencia del docente, que se puede realizar en distintas instancias: en grupos de trabajo o en forma individual, en casa o en biblioteca, laboratorio, etc.)*

## 7. RECURSOS

### 7.1 TEXTO GUÍA

- González, Raúl. Python para todos
- Gopi Subramanian. Python Data Science Cookbook. Packt Publishing Ltd, 2015

### 7.2 TEXTOS COMPLEMENTARIOS

- Hands-On Machine Learning with Scikit-Learn and TensorFlow.
- Deep learning. Principios y fundamentos. Anna Bosch Rué, Jordi Casas Roma, Toni Lozano Bagén. Editorial UOC, 2019

### 7.3 REVISTAS



## 7.4 DIRECCIONES DE INTERNET

- <https://www.python.org/>
- <https://pypi.org/>
- <https://jupyter.org/try>
- <https://www.anaconda.com/products/individual>
- <https://keras.io/examples/>
- <https://matplotlib.org/>

## 7.5 MULTIMEDIA

## 7.6 MOODLE O PLATAFORMA ACADÉMICA (Link o enlace web)

- Enlace moodle: <https://aulasciencias.udistrital.edu.co/login/index.php>

## 7.7 SOFTWARE ESPECIALIZADO (Geogebra, R, Matlab...)

## 8. ORGANIZACIÓN /TIEMPO (Organizar contenidos por semanas)

- **Semana 1:** Introducción (Campos de aplicación, metodologías, herramientas), Lenguaje Python (Entornos de desarrollo, preparación, instalación).
- **Semana 2,3,4:** Uso de estructuras Python, sintaxis, variables. Lectura y escritura de archivos.
- **Semana 5,6:** Manipulación de datos en Dataframe (Paquete Pandas), talleres prácticos  
**Semana 7,8:** Acceso a fuentes de datos, extracción, limpieza, preparación y visualización de datos. **Semana 9,10,11,12:** Aprendizaje supervisado y no supervisado con Machine Learning (Regresión, clasificación, clustering).
- **Semana 13, 14:** Creación de modelos, prueba y validación
- **Semana 15, 16:** Casos prácticos.

## 9. EVALUACIÓN (Especificar porcentajes y formas de evaluación)



La evaluación debe ser coherente con la metodología. Para incentivar el estudio permanente y cultivar la disciplina, se recomiendan pruebas escritas cortas y frecuentes. Los estudiantes deben iniciarse en la aplicación de las formas de lenguaje, expresión y argumentación. La Universidad tiene reglamentado tres cortes:

1er corte: 35%	Fecha:
2do corte: 35%	Fecha:
3er corte: 30%	Fecha: