
 <b>UNIVERSIDAD DISTRITAL</b> <b>FRANCISCO JOSÉ DE CALDAS</b>	<b>FORMATO DE SYLLABUS</b>		Código: AA-FR-003	 <b>SIGUD</b> <small>Sistema Integrado de Gestión</small>
	Macroproceso: Direccionamiento Estratégico		Versión: 01	
	Proceso: Autoevaluación y Acreditación		Fecha de Aprobación: 27/07/2023	

<b>FACULTAD:</b>	<b>Ingeniería</b>		
<b>PROYECTO CURRICULAR:</b>	<b>Ingeniería de Sistemas</b>	<b>CÓDIGO PLAN DE ESTUDIOS:</b>	

#### I. IDENTIFICACIÓN DEL ESPACIO ACADÉMICO

#### NOMBRE DEL ESPACIO ACADÉMICO: Programación para análisis de datos

Código del espacio académico:		Número de créditos académicos:	3			
Distribución horas de trabajo:	HTD	4	HTC	2	HTA	3
Tipo de espacio académico:	Asignatura	x	Cátedra			

#### NATURALEZA DEL ESPACIO ACADÉMICO:

Obligatorio Básico		Obligatorio Complementario	x	Electivo Intrínseco		Electivo Extrínseco	
--------------------	--	----------------------------	---	---------------------	--	---------------------	--

#### CARÁCTER DEL ESPACIO ACADÉMICO:

Teórico		Práctico		Teórico-Práctico	x	Otros:		Cuál: _____
---------	--	----------	--	------------------	---	--------	--	-------------

#### MODALIDAD DE OFERTA DEL ESPACIO ACADÉMICO:

Presencial	x	Presencial con incorporación de TIC		Virtual		Otros:		Cuál: _____
------------	---	-------------------------------------	--	---------	--	--------	--	-------------

#### II. SUGERENCIAS DE SABERES Y CONOCIMIENTOS PREVIOS

Conocimientos en programación de computadores y bases de datos

#### III. JUSTIFICACIÓN DEL ESPACIO ACADÉMICO

Actualmente las técnicas asociadas al análisis de datos constituyen una habilidad esencial para los profesionales en ingeniería de sistemas, la cual pueden aplicar en todas las áreas de conocimiento. El uso de la programación para recopilar, procesar y analizar grandes volúmenes de datos resulta indispensable, dadas las necesidades de automatización intrínsecas en el manejo de grandes volúmenes de datos.

#### IV. OBJETIVOS DEL ESPACIO ACADÉMICO (GENERAL Y ESPECÍFICOS)

##### Objetivo General:

Conocer las características básicas de los lenguajes de programación en Python y R para aprender a manipular conjuntos de datos a través de las estructuras disponibles en los lenguajes y aplicar librerías estadísticas y de aprendizaje automático para el análisis de datos; desarrollar habilidades para representar visualmente los datos

##### Objetivos específicos:

- \* Manejar y transformar datos
- \* Aplicar el análisis estadístico en los datos
- \* Desarrollar el aprendizaje automático básico
- \* Aplicar la visualización de datos

#### V. PROPÓSITOS DE FORMACIÓN Y DE APRENDIZAJE (PFA) DEL ESPACIO ACADÉMICO

Competencias	Dominio-Nivel	RA	Resultados de Aprendizaje
Manejo y Transformación de Datos	Cognitivo-Intelectual	1	Manipular y limpiar datos usando Pandas en Python y dplyr en R.
	Cognitivo-Intelectual	2	Realizar agregaciones y fusiones de datos.
Análisis Estadístico	Cognitivo-Intelectual	3	Aplicar funcionalidades estadísticas descriptivas para resumir datos.
	Cognitivo-Intelectual	4	Conocer las librerías para manejar las diferentes distribuciones probabilísticas
	Cognitivo-Intelectual	5	Usar modelos de regresión y análisis de correlación.

Procesamiento de lenguaje natural	Cognitivo-Intelectual	6	Describir los conceptos básicos y la importancia del procesamiento de lenguaje natural .
	Cognitivo-Intelectual	7	Aplicar técnicas de preprocesamiento de texto, incluyendo tokenización, lematización y eliminación de stop words .
	Cognitivo-Intelectual	8	Realizar análisis de sentimientos en datos textuales e interpretar los resultados .
	Cognitivo-Intelectual	9	Implementar algoritmos de clasificación de texto utilizando bibliotecas especializadas.
Visualización de Datos	Cognitivo-Intelectual	10	Crear visualizaciones efectivas de datos utilizando Matplotlib y Seaborn en Python y ggplot2 en R.
	Cognitivo-Intelectual	11	Desarrollar dashboards interactivos utilizando Plotly en Python y Shiny en R.

VI. CONTENIDOS TEMÁTICOS

<p>Introducción los lenguajes de programación usados para el análisis de datos</p> <ul style="list-style-type: none"> <li>- Sintaxis básica y estructuras de datos propias de los lenguajes de programación empleados al análisis de datos</li> <li>- Entornos de desarrollo integrados (IDE) y herramientas</li> <li>- Manejo de Datos en Python y R</li> </ul> <p>Introducción a Pandas (Python) y dplyr (R)</p> <ul style="list-style-type: none"> <li>- Limpieza y transformación de datos</li> <li>- Operaciones avanzadas con Pandas y dplyr</li> <li>- Análisis Estadístico</li> </ul> <p>Estadísticas descriptivas</p> <ul style="list-style-type: none"> <li>- Pruebas de hipótesis</li> <li>- Regresión y correlación</li> </ul> <p>Herramientas de visualización</p> <ul style="list-style-type: none"> <li>- Uso de Matplotlib y Seaborn (Python) y ggplot2 (R)</li> <li>- Dashboards interactivos con Plotly (Python) y Shiny (R)</li> </ul> <p>Procesamiento de lenguaje natural</p> <ul style="list-style-type: none"> <li>- Preprocesamiento de texto</li> <li>- Bolsa de palabras</li> <li>- Análisis de sentimientos</li> <li>- Clasificación de texto</li> </ul> <p>Análisis exploratorio de datos</p> <ul style="list-style-type: none"> <li>- Visualización inicial</li> <li>- Identificación de patrones y tendencias</li> <li>- Resumen estadístico</li> </ul> <p>Técnicas de validación y evaluación</p> <ul style="list-style-type: none"> <li>- Validación cruzada</li> <li>- Técnicas de muestreo</li> <li>- Overfitting y underfitting</li> <li>- Evaluación de rendimiento ROC-AUC y curvas PR</li> </ul>
--

VII. ESTRATEGIAS DE ENSEÑANZA QUE FAVORECEN EL APRENDIZAJE

Tradicional	X	Basado en Proyectos	X	Basado en Tecnología	X
Basado en Problemas	X	Colaborativo	X	Experimental	X
Aprendizaje Activo		Autodirigido		Centrado en el estudiante	

VIII. EVALUACIÓN

Resultados de aprendizaje (RA) a ser evaluados:	Resultados de aprendizaje asociados a las evaluaciones (T: Teórico / P: Práctico)					
	Actividades Entregables	Talleres	Parciales	Informes de proyecto final	Proyecto final	Exposiciones
RA1	X	X			X	
RA2		X			X	
RA3		X				
RA4			X			
RA5		X		X	X	
RA6			X			
RA7			X			
RA8			X			

RA9		X		X	X	
RA10		X		X	X	
RA11		X		X	X	
Tipo de evaluación						
Porcentaje de evaluación (%)	20	20	30	5	25	
Trabajo Individual (I) o Grupal (G)	I,G	I,G	I	G	G	
Tipo de nota	0-5	0-5	0-5	0-5	0-5	0-5
<b>IX. MEDIOS Y RECURSOS EDUCATIVOS</b>						
<ul style="list-style-type: none"> <li>• Salón normal con pizarrón para sesiones de cátedra y para sesiones de discusión.</li> <li>• Acceso a Videobeam.</li> <li>• Página web para publicar material didáctico, guías de trabajo, talleres, etc.</li> <li>• Videos didácticos alrededor de los temas de la asignatura.</li> <li>• Ttalleres investigativos y prácticos.</li> <li>• Acceso al material bibliográfico recomendado</li> </ul>						
<b>X. PRÁCTICAS ACADÉMICAS - SALIDAS DE CAMPO</b>						
No aplica						
<b>XI. BIBLIOGRAFÍA</b>						
<b>Básicas:</b> McKinney, W. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media. VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media. Grus, J. (2019). Data Science from Scratch: First Principles with Python. O'Reilly Media. Wickham, H., & Grolemund, G. (2016). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer. Kabacoff, R. (2015). R in Action: Data Analysis and Graphics with R. Manning Publications.						
<b>Complementarias:</b>						
<b>Páginas web</b>						
<b>XII. SEGUIMIENTO Y ACTUALIZACIÓN DEL SYLLABUS</b>						
Fecha revisión por Consejo Curricular:						
Fecha aprobación por Consejo Curricular:				Número de acta:		