



## 1. Datos Generales de la asignatura

<b>Nombre de la asignatura:</b>	Programación avanzada para ciencia de datos.
<b>Clave de la asignatura:</b>	CDC-2420
<b>SATCA<sup>1</sup>:</b>	2-2-4
<b>Carrera:</b>	Ingeniería en Ciencia de Datos

## 2. Presentación

### Caracterización de la asignatura

Esta asignatura está diseñada para proporcionar a los estudiantes los conocimientos y habilidades necesarios para utilizar técnicas de programación avanzada en el contexto específico de la ciencia de datos.

La caracterización de esta asignatura se describe a través de los siguientes puntos clave:

- **Enfoque práctico y aplicado:** la asignatura se enfoca en desarrollar habilidades prácticas y aplicadas en el uso de técnicas de programación avanzada para resolver problemas reales en el ámbito de la ciencia de datos. Los estudiantes aprenderán a aplicar conceptos de programación avanzada en la manipulación, análisis, procesamiento y visualización de datos.
- **Interdisciplinariedad:** la asignatura integra conocimientos de diversas disciplinas, incluyendo la informática, la estadística y la ciencia de datos. Se explorarán métodos y herramientas de programación avanzada que son fundamentales para el análisis de datos y el desarrollo de modelos predictivos y descriptivos.
- **Profundización en conceptos de programación:** los estudiantes adquirirán una comprensión más profunda de conceptos de programación avanzada, como la programación orientada a objetos, la programación funcional, el manejo avanzado de datos y la implementación de algoritmos de machine learning. Se explorarán también técnicas de procesamiento de datos a gran escala y visualización avanzada.
- **Aplicabilidad en el mundo real:** la asignatura se centra en la aplicación práctica de los conocimientos adquiridos en contextos del mundo real. Se incluirán proyectos prácticos y casos de estudio que permitirán a los estudiantes enfrentarse a problemas reales de análisis de datos y encontrar soluciones utilizando herramientas y técnicas de programación avanzada.
- **Desarrollo de habilidades profesionales:** además de adquirir habilidades técnicas en programación avanzada para ciencia de datos, los estudiantes también desarrollarán habilidades profesionales importantes, como trabajo en equipo, resolución de problemas, comunicación efectiva y pensamiento crítico. Estas habilidades son fundamentales para el éxito en el campo de la ciencia de datos y la ingeniería.

<sup>1</sup> Sistema de Asignación y Transferencia de Créditos Académicos



<p><b>Intención didáctica</b></p> <p>La intención didáctica de esta materia es proporcionar a los estudiantes una formación sólida y práctica en el uso de técnicas de programación avanzada en el contexto específico de la ciencia de datos. Esta asignatura tiene como objetivo principal desarrollar en los estudiantes las habilidades necesarias para analizar, procesar y visualizar datos de manera efectiva utilizando herramientas y técnicas de programación avanzada, así como implementar y evaluar modelos de machine learning para la solución de problemas reales en diversos campos.</p> <p><b>Metodología:</b></p> <p>la asignatura se desarrollará mediante una combinación de clases teóricas, sesiones prácticas en laboratorio y proyectos aplicados. Se utilizarán ejemplos y casos de estudio reales para ilustrar los conceptos y técnicas presentadas en clase. Se fomentará la participación activa de los estudiantes a través de discusiones, debates y resolución de problemas en equipo. Además, se promoverá el aprendizaje autónomo mediante la realización de tareas y proyectos individuales y grupales.</p> <p><b>Evaluación:</b></p> <p>la evaluación de los estudiantes se llevará a cabo mediante la realización de tareas prácticas, proyectos aplicados, exámenes teóricos y/o prácticos, así como la participación en clase y en actividades grupales. Se valorará tanto el dominio de los conceptos teóricos como la capacidad de aplicarlos en la resolución de problemas prácticos y en la comunicación efectiva de los resultados obtenidos.</p>
--

### 3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Instituto Tecnológico Superior de Alvarado del 21 al 23 agosto de 2023.	Representante del Instituto Tecnológico Superior de Alvarado.	Propuesta inicial.
Tecnológico Nacional de México 30 octubre 2023	Representante del Instituto Tecnológico de: Querétaro y del Instituto Tecnológico Superior de Alvarado.	Presentación de la propuesta de la carrera de Ingeniería en Ciencia de Datos.
Instituto Tecnológico de Querétaro Campus Norte del 19 al 22 de marzo 2024.	Representantes de los Institutos Tecnológicos de: Morelia, Puebla, Querétaro, Tehuacán. Instituto Tecnológico Superior de Alvarado. CENIDET. Representante de Ciencias Básica de los Institutos de: Celaya, Morelia y CIIDET.	Diseño y/o desarrollo curricular de la carrera de Ingeniería en Ciencia de Datos.



<p>Tecnológico Nacional de México del 22 al 24 de abril del 2024</p>	<p>Representante del Instituto Tecnológico de Querétaro e Instituto Tecnológico Superior de Alvarado.</p>	<p>Contraste y ajuste de las asignaturas de Ingeniería en Ciencia de Datos con respecto a las de Ing. en Inteligencia Artificial, Ing. en Desarrollo WEB e Ing. en Ciberseguridad</p>
<p>Tecnológico Nacional de México del 27 al 31 de mayo del 2024.</p>	<p>Representantes de los Institutos Tecnológicos de: Morelia, Querétaro. Instituto Tecnológico Superior de Alvarado. CENIDET.</p>	<p>Consolidación curricular de la carrera de Ingeniería Ciencia de Datos</p>

#### 4. Competencia(s) a desarrollar

<p align="center"><b>Competencia(s) específica(s) de la asignatura</b></p>
<ul style="list-style-type: none"> <li>● Comprende los fundamentos de la programación avanzada y su aplicación en el análisis de datos y la ciencia de datos.</li> <li>● Domina los conceptos y técnicas de programación orientada a objetos, programación funcional y manejo avanzado de datos para la manipulación y procesamiento eficiente de grandes conjuntos de datos.</li> <li>● Adquiere habilidades avanzadas en el uso de bibliotecas y herramientas de programación específicas para la ciencia de datos, como librerías para manejo de datos tabulares, visualización y entrenamiento de algoritmos de aprendizaje automático básicos.</li> <li>● Desarrolla la capacidad de implementar y evaluar algoritmos de machine learning para la clasificación, regresión y clustering de datos, así como comprender los principios subyacentes de estos algoritmos.</li> <li>● Aprende a realizar análisis exploratorio de datos, limpieza de datos, preparación de datos y visualización avanzada para comunicar de manera efectiva información derivada de los datos.</li> <li>● Aplica los conocimientos y habilidades adquiridos en proyectos prácticos y casos de estudio que representen problemas reales en diferentes dominios, como negocios, salud, finanzas, medio ambiente, entre otros.</li> <li>● Fomenta el trabajo en equipo, la colaboración y la comunicación efectiva a través de proyectos grupales y discusiones en clase.</li> <li>● Desarrolla habilidades críticas y analíticas para la resolución de problemas complejos en el campo de la ciencia de datos, incluyendo la capacidad de evaluar y seleccionar las mejores herramientas y técnicas para abordar un problema específico.</li> </ul>



## 5. Competencias previas

- Fundamentos de programación orientada a objetos.
- Operaciones matriciales básicas.
- Compresión de descriptores estadísticos básicos.

## 6. Temario

No.	Temas	Subtemas
1	Herramientas de programación, desarrollo y análisis para ciencia de datos.	1.1. Lenguajes de programación para ciencia de datos. 1.2. Entornos de desarrollo de software. 1.3. Intérpretes en línea para desarrollo en ciencia de datos. 1.4. Análisis de complejidad algorítmica 1.5. Notación Big-O.
2	Manejo preliminar de datos.	2.1. Acceso a archivos de formato especializado para ciencia de datos. 2.2. Manipulación de datos en forma tabular. 2.3. Limpieza y preparación de datos. 2.4. Fusiones y agrupaciones de datos. 2.5. Estrategias para el manejo de datos faltantes y atípicos.
3	Programación funcional.	3.1. Funciones lambda y comprensión de listas. 3.2. Mapeos, filtrados y reducciones. 3.3. Listas, tuplas y diccionarios avanzados. 3.4. Uso de bibliotecas especializadas para operaciones matemáticas y científicas.
4	Estructuras de datos avanzadas para ciencia de datos.	4.1. Heaps y Colas de Prioridad. 4.2. Conjuntos y Mapas. 4.3. Tablas Hash. 4.4. Estructuras de Árboles para ciencia de datos. 4.4.1. Árbol de prefijos. 4.4.2. Árbol B+. 4.4.3. Árbol de regresión. 4.5. Data Frames. 4.6. Series. 4.7. Programación concurrente.



5	Visualización de datos.	<p>5.1. Introducción a bibliotecas de visualización.</p> <p>5.2. Creación de gráficos personalizados y avanzados.</p> <p>5.3. Visualización interactiva.</p>
---	-------------------------	--

## 7. Actividades de aprendizaje de los temas

<b>1. Herramientas de programación, desarrollo y análisis para ciencia de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Diseña entornos de desarrollo a partir de un análisis para la selección adecuada de software.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>● Comprensión conceptual</li> <li>● Identificación de problemas</li> <li>● Capacidad para trabajar en equipo</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas</li> <li>● Resolución de problemas básicos.</li> </ul>	<ul style="list-style-type: none"> <li>● Investiga y evalúa diferentes lenguajes de programación populares para ciencia de datos, como Python, R o Julia.</li> <li>● Redacta un breve informe comparativo que destaque las fortalezas y debilidades de cada lenguaje, así como su idoneidad para diferentes tipos de proyectos de análisis de datos.</li> <li>● Investiga diferentes IDEs disponibles para el desarrollo de proyectos de ciencia de datos y elige uno para explorar en profundidad.</li> <li>● Realiza una lista de características clave de la IDE seleccionada y crea un breve tutorial que muestre cómo utilizar algunas de sus funcionalidades básicas.</li> </ul>
<b>2. Manejo preliminar de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Aplica métodos de limpieza de datos para la solución de problemas valorando su pertinencia en el uso eficaz de los recursos.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Habilidad en el manejo de equipo de cómputo.</li> <li>● Capacidad para trabajar en equipo.</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> </ul>	<ul style="list-style-type: none"> <li>● Investiga diferentes formatos de archivos utilizados comúnmente en ciencia de datos, como CSV, JSON, XML, HDF5, entre otros.</li> <li>● Elige uno de estos formatos y crea un pequeño conjunto de datos de muestra en ese formato para practicar la lectura y escritura de datos utilizando el lenguaje de programación de tu elección.</li> <li>● Utiliza conjuntos de datos tabulares para la manipulación de datos utilizando las operaciones básicas disponibles.</li> <li>● Propone estrategias para el manejo de datos faltantes, como la imputación de valores, la detección y eliminación de valores atípicos.</li> </ul>



<b>3. Programación funcional</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Aplica diferentes formatos de lista, mapeos y diccionarios en conjuntos de datos.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>• Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>• Capacidad de análisis y síntesis</li> <li>• Habilidad en el manejo de equipo de cómputo</li> <li>• Capacidad para trabajar en equipo.</li> <li>• Capacidad de investigación.</li> <li>• Capacidad de aplicar los conocimientos en la práctica.</li> </ul>	<ul style="list-style-type: none"> <li>• Realiza operaciones simples en listas, como mapear, filtrar u ordenar elementos.</li> <li>• Utiliza mapeos y filtrados, comentando la eficiencia y legibilidad del código resultante.</li> </ul>
<b>4. Estructuras de datos avanzadas para ciencia de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Aplica estructuras de datos avanzadas para el desarrollo y análisis.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>• Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>• La comprensión y manipulación de ideas y pensamientos.</li> <li>• Metodologías para solución de problemas, organización del tiempo y para el aprendizaje.</li> <li>• Habilidad en el manejo de equipo de cómputo.</li> <li>• Capacidad para trabajar en equipo.</li> <li>• Capacidad de aplicar los conocimientos en la práctica.</li> </ul>	<ul style="list-style-type: none"> <li>• Investigar y seleccionar información acerca de los</li> <li>Implementa el algoritmo de Dijkstra para encontrar el camino más corto en un grafo utilizando heaps.</li> <li>• Implementa una tabla hash para encontrar elementos similares en una colección de datos.</li> <li>• Analizar un conjunto de datos utilizando DataFrames y series.</li> </ul>



5. Visualización de datos	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Implementa gráficos a partir de conjuntos de datos.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>• Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>• La comprensión y manipulación de ideas y pensamientos.</li> <li>• Metodologías para solución de problemas, organización del tiempo y para el aprendizaje.</li> <li>• Habilidad en el manejo de equipo de cómputo.</li> <li>• Capacidad para trabajar en equipo.</li> <li>• Capacidad de aplicar los conocimientos en la práctica.</li> </ul>	<ul style="list-style-type: none"> <li>• El estudiante instalará y configurará la biblioteca Seaborn en su entorno de programación.</li> <li>• Utiliza las API de estilo de Seaborn para crear gráficos de violín, gráficos de distribución de probabilidad (KDE), gráficos de calor y gráficos de matriz.</li> </ul>

## 8. Práctica(s)

<ul style="list-style-type: none"> <li>• <b>Objetivo:</b> crear gráficos personalizados y avanzados utilizando la biblioteca Seaborn en Python.</li> <li>• El estudiante instalará y configurará la biblioteca Seaborn en su entorno de programación.</li> <li>• Cargará un conjunto de datos en un DataFrame de Pandas.</li> <li>• Utilizará las API de estilo de Seaborn para crear gráficos de violín, gráficos de distribución de probabilidad (KDE), gráficos de calor y gráficos de matriz.</li> <li>• Personaliza la apariencia de los gráficos con paletas de colores, estilos de marcadores y temas predefinidos.</li> <li>• Implementará estilos de cuadrícula personalizados y anotaciones para mejorar la claridad y la comunicación de los resultados.</li> </ul>
--

## 9. Proyecto de asignatura

<p>El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance del(los) logro(s) formativo(s) de la asignatura, considerando las siguientes fases:</p> <p><b>Fundamentación:</b> marco referencial (teórico, conceptual, contextual, legal) en el cual se fundamenta el proyecto de acuerdo con un diagnóstico realizado, mismo que permite a los estudiantes lograr la comprensión de la realidad o situación objeto de estudio para definir un proceso de intervención o hacer el diseño de un modelo.</p>
--



**Planeación:** con base en el diagnóstico en esta fase se realiza el diseño del proyecto por parte de los estudiantes con asesoría del docente; implica planificar un proceso: de intervención empresarial, social o comunitario, el diseño de un modelo, entre otros, según el tipo de proyecto, las actividades a realizar los recursos requeridos y el cronograma de trabajo.

**Ejecución:** consiste en el desarrollo de la planeación del proyecto realizada por parte de los estudiantes con asesoría del docente, es decir en la intervención (social, empresarial), o construcción del modelo propuesto según el tipo de proyecto, es la fase de mayor duración que implica el desempeño de los saberes, habilidades y destrezas a desarrollar.

**Evaluación:** es la fase final que aplica un juicio de valor en el contexto laboral-profesión, social e investigativo, ésta se debe realizar a través del reconocimiento de logros y aspectos a mejorar se estará promoviendo el concepto de "evaluación para la mejora continua", el desarrollo del pensamiento crítico y reflexivo en los estudiantes.

## 10. Evaluación por competencias

La evaluación debe hacerse diagnóstica, formativa y sumativa. De igual manera, para fortalecer la parte actitudinal, se recomienda guiar al estudiante hacia la introspección para utilizar la autoevaluación y la coevaluación.

En el caso de las actividades de aprendizaje se sugiere el uso de estrategias metacognitivas como: mapas mentales, mapas conceptuales, reportes de prácticas, exposiciones en clase, ensayos, resúmenes, observación y cuestionarios, cuadros comparativos, informes.

Mientras que para verificar el nivel del logro de las competencias del estudiante se recomienda utilizar: el portafolio de evidencias, listas de cotejo, rúbricas, matrices de valoración, exámenes, guías de observación, además de estrategias en las que se logren las competencias blandas.

## 11. Fuentes de Información

1. McKinney, W. (2022). Python for data analysis. " O'Reilly Media, Inc."
2. Hadley, W., & Garrett, G. (2016). R for data science: import, tidy, transform, visualize, and model data. O'Reilly Media, Inc.
3. Müller, A. C., & Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. " O'Reilly Media, Inc."
4. Matloff, N. (2011). The art of R programming: A tour of statistical software design. No Starch Press.
5. Kwon, C. (2019). Julia Programming for Operations Research. Changhyun Kwon.
6. Martinez, W. L., Martinez, A. R., & Solka, J. (2017). Exploratory data analysis with MATLAB. Chapman and Hall/CRC.