



## 1. Datos Generales de la asignatura

<b>Nombre de la asignatura:</b>	Introducción a la Ingeniería en Ciencia de Datos
<b>Clave de la asignatura:</b>	CDI-2413
<b>SATCA<sup>1</sup>:</b>	4-0-4
<b>Carrera:</b>	Ingeniería en Ciencia de Datos

## 2. Presentación

<b>Caracterización de la asignatura</b>
<p>La Ciencia de Datos, en actualidad, se ha convertido en una disciplina fundamental para comprender y tomar decisiones informadas en una amplia variedad de campos. Esta disciplina se centra en el estudio y el análisis de datos, con el objetivo de extraer conocimiento valioso y patrones significativos a partir de conjuntos de información cada vez más grandes y complejos.</p> <p>La Ciencia de Datos combina elementos de la estadística, la programación, la matemática y el conocimiento de dominio específico para abordar problemas reales. Algunos de los aspectos clave de esta disciplina incluyen la recopilación y limpieza de datos, la exploración y visualización, la construcción de modelos predictivos y la toma de decisiones basada en datos.</p> <p>Su aplicación es amplia y abarca áreas como la medicina, el comercio electrónico, la inteligencia artificial, la investigación científica, el marketing y muchas otras. Es por estas razones que la Ciencia de Datos se ha convertido en una herramienta esencial para entender el mundo que nos rodea y utilizar esa comprensión para mejorar procesos, productos y servicios en la búsqueda de soluciones más eficientes y efectivas.</p> <p>Dicho lo anterior, esta materia contribuye al perfil del egresado en su aspecto profesional al proporcionar al estudiante el conocimiento de los conceptos fundamentales involucrados en la Ciencia de Datos, así como las metodologías y herramientas relacionadas con los mismos, facilitando su comprensión mediante la realización de actividades de investigaciones de problemas de ingeniería complejos relacionados con el diseño de experimentos, el análisis y la interpretación de los datos, así como la síntesis de la información para proporcionar conclusiones válidas.</p>
<b>Intención didáctica</b>
<p>El propósito de este curso es introducir al estudiante a este campo de rápido crecimiento, y equiparlos con el conocimiento de los principios básicos y herramientas que necesitarán para tratar con las diferentes facetas involucradas en la ciencia de los datos, que incluyen la colección e integración de datos, el análisis exploratorio de los datos, modelos predictivos y descriptivos, y la evaluación de estos, así como su relación con otras áreas del conocimiento.</p>

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



En la unidad uno se presenta la teoría general sobre la ciencia de datos.  
 En la unidad dos se analiza la importancia de la ética en relación con la ciencia de datos y sus aplicaciones.  
 En la unidad tres conocerán las metodologías más comunes para desarrollar sistemas informáticos con la ciencia de datos.  
 En la unidad cuatro se describen las técnicas más usuales de análisis preliminar de datos, así como las formas de visualización asociadas con los mismos.  
 En la unidad cinco se presentan algunas de las aplicaciones más comunes de la ciencia de datos en las diferentes áreas de la tecnología, así como su potencial impacto social y económico.  
 El proceso de adquisición del conocimiento antes descrito se llevará a cabo considerando los siguientes principios:

- Propiciar la búsqueda y selección de información.
- Inducir al alumno a los temas de clase.
- Realimentar de forma permanente los temas tratados en clase.
- Promover la presentación de exposiciones por parte de los alumnos.
- Promover la solución de problemas en forma individual y grupal.
- Análisis y discusión por equipos de los temas investigados.
- Exposición verbal auxiliada de materiales de apoyo.

### 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.



Tecnológico Nacional de México del 22 al 24 de abril del 2024	Representante del Instituto Tecnológico de Querétaro e Instituto Tecnológico Superior de Alvarado.	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
Tecnológico Nacional de México del 27 al 31 de mayo del 2024.	Representantes de los Institutos Tecnológicos de: Morelia, Querétaro. Instituto Tecnológico Superior de Alvarado. CENIDET.	Consolidación curricular de la carrera de Ingeniería Ciencia de Datos

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

Competencia(s) específica(s) de la asignatura
Conoce los mecanismos usados en la ciencia de datos para la obtención de conocimientos en el soporte de toma de decisiones, y su relación con otras áreas del saber humano.

#### 5. Competencias previas

Ninguna
---------

#### 6. Temario

No.	Temas	Subtemas
1	Introducción a ciencia de datos.	1.1. Definición y conceptos en ciencia de datos. 1.2. Historia de la ciencia de datos. 1.3. Habilidades de un científico de datos. 1.4. Técnicas y herramientas para la ciencia de datos. 1.5. Áreas de conocimiento relacionadas con la ciencia de datos. 1.5.1. Minería de datos. 1.5.2. Machine Learning. 1.5.3. Inteligencia de negocios. 1.5.4. Big data.
2	Ética y responsabilidad en ciencia de datos.	2.1. Consideraciones éticas en la recopilación y uso de datos. 2.2. Regulaciones y normativas de privacidad. 2.3. Conceptos fundamentales para una ética aplicada a Big data. 2.4. Sesgos y discriminación algorítmica.



3	Metodologías para la ciencia de datos.	<p>3.1. Fuentes de datos y métodos de recopilación.</p> <p>3.2. Técnicas de limpieza y procesamiento de datos.</p> <p>3.3. Tratamiento de valores faltantes y datos atípicos.</p> <p>3.4. Análisis exploratorio de datos.</p>
4	Fundamento de la exploración y visualización de datos.	<p>4.1. Visualización de datos.</p> <p>4.2. Estadística inferencial y descriptivas.</p> <p>4.3. Identificación de patrones y tendencias.</p>
5	Aplicaciones específicas de la ciencia de datos.	<p>5.1. Estudios de caso en:</p> <p>5.1.1. Redes sociales y motores de búsqueda.</p> <p>5.1.2. Finanzas.</p> <p>5.1.3. Energía.</p> <p>5.1.4. Medicina.</p> <p>5.1.5. Seguridad.</p> <p>5.1.6. Aplicación social y gubernamental.</p> <p>5.2. Áreas de aplicación.</p> <p>5.2.1. Big data.</p> <p>5.2.2. Cómputo en la nube.</p> <p>5.2.3. IOT.</p>

## 7. Actividades de aprendizaje de los temas

1. Introducción a ciencia de datos	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Comprende e identifica los conceptos básicos sobre ciencia de datos.</p> <p><i>Genéricas:</i> <i>Competencias Instrumentales</i></p> <ul style="list-style-type: none"> <li>● Planea y administra el tiempo.</li> <li>● Capacidad de organización.</li> <li>● Compara diferentes alternativas de solución.</li> <li>● Selecciona la solución más adecuada, satisfaciendo los requerimientos.</li> </ul> <p><i>Competencias interpersonales</i></p> <ul style="list-style-type: none"> <li>● Capacidad crítica y autocrítica.</li> <li>● Trabajo en equipo.</li> <li>● Habilidades interpersonales.</li> </ul>	<ul style="list-style-type: none"> <li>● Propiciar la búsqueda y selección de información relacionada con los conceptos fundamentales de la ciencia de datos.</li> <li>● Propiciar la presentación verbal de exposiciones por parte de los alumnos acerca de las áreas del conocimiento relacionadas con la ciencia de datos.</li> <li>● Análisis y discusión por escrito sobre las habilidades requeridas por los científicos de datos.</li> </ul>



<p><i>Competencias sistémicas</i></p> <ul style="list-style-type: none"> <li>● Capacidad de generar nuevas ideas (creatividad).</li> <li>● Habilidad para trabajar en forma autónoma.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de aprender.</li> <li>● Búsqueda del logro.</li> </ul>	
<b>2. Ética y responsabilidad en ciencia de datos.</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Conoce y explica la importancia de la privacidad de los datos y reconoce sesgos en conjuntos de datos</p> <p><i>Genéricas:</i></p> <p><i>Competencias Instrumentales</i></p> <ul style="list-style-type: none"> <li>● Planea y administra el tiempo.</li> <li>● Capacidad de organización.</li> <li>● Compara diferentes alternativas de solución.</li> <li>● Selecciona la solución más adecuada, satisfaciendo los requerimientos.</li> </ul> <p><i>Competencias interpersonales</i></p> <ul style="list-style-type: none"> <li>● Capacidad crítica y autocrítica.</li> <li>● Trabajo en equipo.</li> <li>● Habilidades interpersonales.</li> </ul> <p><i>Competencias sistémicas</i></p> <ul style="list-style-type: none"> <li>● Capacidad de generar nuevas ideas (creatividad).</li> <li>● Habilidad para trabajar en forma autónoma.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de aprender.</li> <li>● Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>● Redactar informes derivados de la búsqueda y selección de información relacionada con la importancia de la ética en la ciencia de datos.</li> <li>● Promover la discusión y análisis en grupo sobre la normatividad existente para regular el impacto social del uso de la ciencia de datos en la toma de decisiones que afectan a los usuarios.</li> </ul>
<b>3. Metodologías para la ciencia de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Conoce y describe metodologías para la depuración de datos</p> <p><i>Genéricas:</i></p> <p><i>Competencias Instrumentales</i></p>	<ul style="list-style-type: none"> <li>● Investigar las características de las metodologías para la ciencia de datos.</li> <li>● Propiciar la presentación verbal de exposiciones por parte de los alumnos acerca de ejemplos de la aplicación de las metodologías descritas en esta unidad.</li> </ul>



<ul style="list-style-type: none"> <li>● Planea y administra el tiempo.</li> <li>● Capacidad de organización.</li> <li>● Compara diferentes alternativas de solución.</li> <li>● Selecciona la solución más adecuada, satisfaciendo los requerimientos.</li> </ul> <p><i>Competencias interpersonales</i></p> <ul style="list-style-type: none"> <li>● Capacidad crítica y autocrítica.</li> <li>● Trabajo en equipo.</li> <li>● Habilidades interpersonales.</li> </ul> <p><i>Competencias sistémicas</i></p> <ul style="list-style-type: none"> <li>● Capacidad de generar nuevas ideas (creatividad).</li> <li>● Habilidad para trabajar en forma autónoma.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de aprender.</li> <li>● Búsqueda del logro.</li> </ul>	
---	--

<b>4. Fundamento de la exploración y visualización de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Investiga y describe conjuntos de datos con apoyo de elementos gráficos</p> <p><i>Genéricas:</i> <i>Competencias Instrumentales</i></p> <ul style="list-style-type: none"> <li>● Planea y administra el tiempo.</li> <li>● Capacidad de organización.</li> <li>● Compara diferentes alternativas de solución.</li> <li>● Selecciona la solución más adecuada, satisfaciendo los requerimientos.</li> </ul> <p><i>Competencias interpersonales</i></p> <ul style="list-style-type: none"> <li>● Capacidad crítica y autocrítica.</li> <li>● Trabajo en equipo.</li> <li>● Habilidades interpersonales.</li> </ul> <p><i>Competencias sistémicas</i></p> <ul style="list-style-type: none"> <li>● Capacidad de generar nuevas ideas (creatividad).</li> </ul>	<ul style="list-style-type: none"> <li>● Propiciar la búsqueda y selección de información relacionada con el proceso de análisis exploratorio de los datos</li> <li>● Hacer uso de software de simulación para ejemplificar el uso de la visualización de datos en la toma de decisiones.</li> <li>● Elaborar un mapa conceptual del proceso de análisis exploratorio y visualización de los datos.</li> </ul>



<ul style="list-style-type: none"> <li>● Habilidad para trabajar en forma autónoma.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de aprender.</li> <li>● Búsqueda del logro.</li> </ul>	
--	--

<b>5. Aplicaciones específicas de la ciencia de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Interpreta y explica casos de uso a partir de conjuntos de datos de diverso origen</p> <p><i>Genéricas:</i> <i>Competencias Instrumentales</i></p> <ul style="list-style-type: none"> <li>● Planea y administra el tiempo.</li> <li>● Capacidad de organización.</li> <li>● Compara diferentes alternativas de solución.</li> <li>● Selecciona la solución más adecuada, satisfaciendo los requerimientos.</li> </ul> <p><i>Competencias interpersonales</i></p> <ul style="list-style-type: none"> <li>● Capacidad crítica y autocrítica.</li> <li>● Trabajo en equipo.</li> <li>● Habilidades interpersonales.</li> </ul> <p><i>Competencias sistémicas</i></p> <ul style="list-style-type: none"> <li>● Capacidad de generar nuevas ideas (creatividad).</li> <li>● Habilidad para trabajar en forma autónoma.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de aprender.</li> <li>● Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>● Propiciar la búsqueda y selección de información relacionada con las aplicaciones más comunes de la ciencia de datos.</li> <li>● Propiciar la presentación verbal de exposiciones por parte de los alumnos acerca de las áreas aplicación de la ciencia de datos</li> <li>● Promover la asistencia virtual a conferencias con especialistas del área del conocimiento.</li> </ul>

## 8. Práctica(s)

<ul style="list-style-type: none"> <li>● Hacer uso de herramientas de software para el manejo de referencias bibliográficas (tales como Mendeley) en las actividades que requieran la entrega de reportes escritos como resultado de una investigación documental.</li> <li>● Redactar un informe escrito sobre las principales áreas de aplicación de la ciencia de datos.</li> <li>● Redactar un informe sobre tres metodologías para el desarrollo de aplicaciones de ciencia de datos (se recomienda hacer referencia a los procedimientos desarrollados y estandarizados de</li> </ul>
---



empresas de alto impacto en la industria del software, tal como *Team Data Science* de Microsoft).

- Codificar ejemplos simples del uso de software para ciencia de datos (como *R* y *Python*) que hagan llamada a funciones de librerías de uso común para el análisis exploratorio y la visualización preliminar de datos (se recomienda mencionar de una manera simplificada un ejemplo de regresión lineal y otro de *clustering*).
- Realizar exposiciones verbales en público donde se exponga, y se discuta grupalmente, la importancia de la ética en la ciencia de datos.

## 9. Proyecto de asignatura

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:

- **Fundamentación:** 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.
- **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 deberá realizarse de forma 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 para este curso se sugiere el uso de estrategias metacognitivas como: mapas conceptuales, reportes de prácticas, exposiciones en clase, ensayos, resúmenes, reportes de conferencias (tanto presenciales como virtuales), trípticos, guías de entrevista, observación y cuestionarios. Mientras que para verificar el nivel del logro de las



competencias del estudiante se recomienda utilizar no solo exámenes por escrito, sino incluir también un portafolio de evidencias que haga uso de listas de cotejo, rúbricas, matrices de valoración, guías de observación, además de estrategias en las que se logren las competencias blandas.

### 11. Fuentes de Información

1. Stephen Klosterman. Data Science Projects with Python (2nd Edition). Packt Publishing, 2021.
2. Joel Grus. Data Science from Scratch: First Principles with Python (2nd Edition). O'Reilly, 2019.
3. Ian Witten, Eibe Frank, Mark Hall, Christopher Pal. Data Mining: Practical Machine Learning Tools and Techniques (4th Edition). Elsevier, 2017.
4. Harrington. Machine Learning in Action. Manning, 2012.
5. Pérez, M. Business Intelligence: Técnicas, herramientas y aplicaciones. Alfaomega, 2015.
6. Joyanes, L. Big Data. Marcombo SA, 2013.
7. Provost F., Fawcett T. Data science for business. O'Reilly 2013.
8. Pérez, M. Minería de datos a través de ejemplos. Alfaomega, 2015.
9. Sharda, R, Delen, D, Turban, E. Analytics, Data Science & Artificial Intelligence. Pearson, 2021.
10. Starmer, J. The StatQuest Illustrated Guide to Machine Learning. 2022.