



## 1. Datos Generales de la asignatura

<b>Nombre de la asignatura:</b>	Metodologías para proyectos en ciencia de datos
<b>Clave de la asignatura:</b>	CDH-2417
<b>SATCA<sup>1</sup>:</b>	1-3-4
<b>Carrera:</b>	Ingeniería en Ciencia de Datos

## 2. Presentación

### Caracterización de la asignatura

La asignatura de Metodología en Proyectos de Ciencia de Datos contribuye significativamente al perfil de egreso al proporcionar a los estudiantes las habilidades y conocimientos necesarios para planificar, ejecutar y gestionar proyectos de ciencia de datos de manera efectiva. Esto incluye la comprensión de los enfoques metodológicos utilizados en el ciclo de vida de un proyecto de ciencia de datos, así como la capacidad de aplicar técnicas y herramientas para la recolección, limpieza, análisis y visualización de datos.

La asignatura es fundamental en la formación de profesionales en ciencia de datos, ya que proporciona las bases metodológicas necesarias para llevar a cabo proyectos de manera eficiente y efectiva. Sin una comprensión adecuada de la metodología de proyectos, los esfuerzos en ciencia de datos pueden ser poco sistemáticos, mal enfocados o incluso fracasar en alcanzar los objetivos deseados. Por lo tanto, entender los principios y prácticas de la metodología en proyectos de ciencia de datos es crucial para el éxito en este campo.

La asignatura de Metodología en Proyectos de Ciencia de Datos abarca una variedad de temas relacionados con la planificación, ejecución y gestión de proyectos en el contexto de la ciencia de datos. Esto incluye la comprensión de los marcos metodológicos comunes, como CRISP-DM (Cross-Industry Standard Process for Data Mining) y otros enfoques ágiles. También implica el estudio de técnicas para la identificación y definición de problemas, la recopilación y preparación de datos, el modelado y análisis de datos, la interpretación de resultados y la comunicación efectiva de hallazgos.

La asignatura de Metodología en Proyectos de Ciencia de Datos se relaciona estrechamente con otras asignaturas dentro del campo de la ciencia de datos, así como con disciplinas relacionadas como la estadística, la informática y la ingeniería. Se relaciona con temas como análisis exploratorio de datos, minería de datos, aprendizaje automático, visualización de datos y ética en la ciencia de datos. En términos de competencias específicas, se enfoca en habilidades de planificación, gestión de proyectos, resolución de problemas, pensamiento crítico, comunicación efectiva y trabajo en equipo, que son cruciales para llevar a cabo proyectos integradores en ciencia de datos.

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



### **Intención didáctica**

Los contenidos de la asignatura deben abordarse de manera práctica y teórica. Esto implica combinar la exposición de conceptos metodológicos clave con ejemplos de aplicación práctica en proyectos reales de ciencia de datos. Se pueden utilizar estudios de caso, ejercicios prácticos, análisis de proyectos previos y discusiones en clase para ilustrar los principios y técnicas de la metodología en proyectos de ciencia de datos.

Los contenidos deben ser tratados desde una perspectiva interdisciplinaria que integre aspectos de estadística, informática, matemáticas y gestión de proyectos. Se debe enfatizar la importancia de comprender las necesidades del negocio o del problema que se está abordando, así como la aplicación de enfoques metodológicos adecuados para resolverlos de manera eficiente y efectiva.

Los contenidos deben cubrirse de manera exhaustiva, abarcando desde los fundamentos de la metodología en proyectos de ciencia de datos hasta técnicas avanzadas de planificación, ejecución y gestión de proyectos. Se debe proporcionar una profundidad adecuada en cada área, asegurándose de que los estudiantes adquieran un conocimiento sólido que puedan aplicar en situaciones reales.

Se deben resaltar actividades como la resolución de problemas, la toma de decisiones, la comunicación efectiva, el trabajo en equipo y el pensamiento crítico. Estas actividades son fundamentales para el desarrollo de competencias genéricas que son aplicables en una variedad de contextos profesionales.

Las competencias genéricas que se desarrollan incluyen habilidades de análisis, síntesis y evaluación de información, capacidad para trabajar en equipo y liderar proyectos, habilidades de comunicación oral y escrita, capacidad para adaptarse a diferentes contextos y culturas, y ética profesional en el manejo de datos y resultados.

El docente debe desempeñar un papel activo y facilitador en el desarrollo de la asignatura. Esto implica proporcionar orientación y dirección a los estudiantes, fomentar la participación activa en clase, brindar retroalimentación constructiva sobre el trabajo realizado y crear un entorno de aprendizaje colaborativo. Además, el docente debe estar al tanto de las últimas tendencias y prácticas en metodología de proyectos de ciencia de datos para asegurar que los contenidos sean relevantes y actualizados.



### 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

Aplica bases metodológicas de Ciencia de Datos para la resolución de proyectos reales en el contexto de salud, educación, finanzas, redes sociales, energía, seguridad, Big Data entre otras.

#### 5. Competencias previas

- Capacidad de análisis y pensamiento crítico.
- Conocimientos de la materia de introducción a la ingeniería de ciencia de datos.
- Conocimientos de la materia de programación avanzada para ciencia de datos.
- Conocimientos de la materia de aprendizaje automático.
- Conocimiento de la materia de visualización de datos.
- Conocimientos de la materia de tópicos selectos para ciencia de datos.

#### 6. Temario

No.	Temas	Subtemas
1	Ciclo de vida y metodologías de desarrollo de aplicaciones.	1.1. Ciclo de vida. 1.1.1. Formular y refinar la pregunta. 1.1.2. Recolección de datos. 1.1.3. Explorar los datos. 1.1.4. Construir modelos estadísticos. 1.1.5. Interpretar los resultados. 1.1.6. Comunicar los resultados. 1.2. Introducción a las metodologías. 1.2.1. Metodología fundacional para la ciencia de datos de IBM. 1.2.2. CRISP-DM. 1.2.3. Otras metodologías.
2	Ejemplos de aplicaciones de ciencia de datos.	2.1. Ciudades inteligentes. 2.2. Salud. 2.3. Educación. 2.4. Finanzas. 2.5. Redes sociales y motores de búsqueda. 2.6. Energía. 2.7. Seguridad. 2.8. Big data.
3	Tendencias de la ciencia de datos	3.1. En salud. 3.2. En educación. 3.3. En inteligencia computacional. 3.4. En otras áreas.



## 7. Actividades de aprendizaje de los temas

<b>1. Ciclo de vida y metodologías de desarrollo de aplicaciones</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Aplica el uso de metodologías de Ciencia de datos para el desarrollo de proyectos.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Habilidad para manejo de equipo de cómputo.</li> <li>● Capacidad para trabajar en equipo.</li> <li>● Solución de problemas.</li> </ul>	<ul style="list-style-type: none"> <li>● Leer estudios de casos o ejemplos donde se describa cómo se formuló una pregunta específica para un proyecto de ciencia de datos.</li> <li>● Practicar escribir preguntas claras y específicas sobre conjuntos de datos de tu interés.</li> <li>● Realizar ejercicios de recolección de datos simulados, donde tengas que recopilar información de diversas fuentes.</li> <li>● Aprender sobre técnicas de muestreo y cómo aplicarlas en diferentes contextos.</li> <li>● Utilizar herramientas como Pandas en Python o funciones de manipulación de datos en R para explorar conjuntos de datos.</li> <li>● Realizar visualizaciones de datos para identificar patrones, tendencias y posibles valores atípicos.</li> <li>● Participar en desafíos de modelado de datos donde tengas que aplicar algoritmos de aprendizaje automático para resolver problemas específicos.</li> <li>● Estudiar diferentes técnicas de modelado estadístico como regresión lineal, clasificación, agrupamiento, etc.</li> </ul>
<b>2. Ejemplos de aplicaciones de ciencia de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Aplica los conceptos de Ciencia de Datos para el desarrollo de aplicaciones en el contexto de algunos campos como son: ciudades inteligentes, salud, educación, finanzas, redes sociales, Big Data, entre otros.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> <li>● Habilidades de gestión de información (habilidad para buscar y analizar información proveniente de fuentes diversas).</li> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de comunicación oral y escrita.</li> </ul>	<ul style="list-style-type: none"> <li>● Buscar casos reales de aplicaciones de Ciencia de Datos en diferentes campos como la salud, finanzas, marketing, etc.</li> <li>● Analizar cómo se recopilaron los datos, qué técnicas de análisis se utilizaron y cuáles fueron los resultados obtenidos.</li> <li>● Diseñar proyectos prácticos donde los estudiantes puedan aplicar técnicas de Ciencia de Datos a conjuntos de datos reales o simulados. Por ejemplo, podrían trabajar en la predicción de ventas, análisis de sentimientos en redes sociales, o detección de fraudes.</li> <li>● Organizar debates sobre el uso ético de la Ciencia de Datos y sus implicaciones en la privacidad y la seguridad. También pueden</li> </ul>



<ul style="list-style-type: none"> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Habilidades en el uso de las tecnologías de la información y de la comunicación.</li> <li>● Habilidad para manejo de equipo de cómputo.</li> <li>● Capacidad para trabajar en equipo.</li> <li>● Solución de problemas.</li> </ul>	<p>discutir sobre casos controvertidos donde se haya utilizado la Ciencia de Datos.</p> <ul style="list-style-type: none"> <li>● Solicitar a los estudiantes que investiguen sobre un tema específico relacionado con aplicaciones de Ciencia de Datos y que presenten sus hallazgos ante el resto de la clase. Esto fomentará la investigación independiente y el desarrollo de habilidades de comunicación. os de manera cuestionable.</li> <li>● Ofrecer recursos y tutoriales para que los estudiantes puedan desarrollar habilidades técnicas en herramientas y lenguajes de programación utilizados en Ciencia de Datos, como Python, R, SQL, etc.</li> </ul>
<b>3. Tendencias de la ciencia de datos</b>	
<b>Competencias</b>	<b>Actividades de aprendizaje</b>
<p><i>Específica(s):</i> Aplica el uso de tendencias de la Ciencia de Datos en el contexto de la salud, educación, inteligencia computacional y otras áreas de interés.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> <li>● Habilidades de gestión de información (habilidad para buscar y analizar información proveniente de fuentes diversas).</li> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de comunicación oral y escrita.</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Habilidades en el uso de las tecnologías de la información y de la comunicación.</li> <li>● Habilidad para manejo de equipo de cómputo.</li> <li>● Capacidad para trabajar en equipo.</li> <li>● Solución de problemas.</li> </ul>	<ul style="list-style-type: none"> <li>● Proporcionar a los estudiantes una lista de artículos y pappers recientes sobre tendencias en Ciencia de Datos y pídeles que los analicen. Deben identificar las principales tendencias, tecnologías emergentes y posibles aplicaciones en diferentes campos.</li> <li>● Organizar seminarios o charlas con profesionales de la industria o investigadores que estén trabajando en áreas de vanguardia en Ciencia de Datos. Esto permitirá a los estudiantes conocer de primera mano las últimas tendencias y tener la oportunidad de hacer preguntas y discutir sobre el tema.</li> <li>● Dividir a los estudiantes en grupos y asignar a cada grupo un tema de tendencia en Ciencia de Datos para que investiguen. Pueden explorar temas como inteligencia artificial explicativa, aprendizaje federado, ética en el aprendizaje automático, entre otros. Al finalizar, cada grupo puede presentar sus hallazgos a la clase.</li> <li>● Analizar estudios de casos reales que ejemplifiquen cómo se están aplicando las tendencias de Ciencia de Datos en diferentes industrias y sectores. Los estudiantes pueden examinar casos de éxito y entender cómo se están abordando los desafíos y oportunidades en la actualidad.</li> </ul>



## 8. Práctica(s)

- Solicitar a los alumnos que elijan un tema o problema específico para trabajar y que definan claramente los objetivos del proyecto de Ciencia de Datos. Deben identificar el alcance del proyecto y establecer metas alcanzables y medibles.
- Guiar a los alumnos en el proceso de recolección de datos relevantes para su proyecto. Pueden utilizar conjuntos de datos públicos o recopilar datos propios si es posible. Luego, anímalos a explorar los datos utilizando técnicas como visualizaciones, estadísticas descriptivas y análisis preliminares.
- Enseñar a los alumnos la importancia de limpiar y preprocesar los datos antes de comenzar el análisis. Pueden realizar tareas como eliminar valores atípicos, manejar datos faltantes, normalizar variables y convertir datos no estructurados en formatos utilizables.
- Introducir a los alumnos en diferentes técnicas de análisis de datos, como regresión, clasificación, clustering, entre otras. Pídeles que seleccionen las técnicas adecuadas para su proyecto y que las apliquen de manera apropiada, justificando sus decisiones.
- Enseñar a los alumnos a evaluar la efectividad y el rendimiento de los modelos de Ciencia de Datos que han desarrollado. Deben utilizar métricas adecuadas y técnicas de validación cruzada para asegurarse de que los resultados sean confiables y generalizables.
- Solicitar a los alumnos que documenten todo el proceso de su proyecto de Ciencia de Datos, desde la definición de los objetivos hasta la evaluación de los resultados. Anímalos a crear informes claros y concisos que comuniquen los hallazgos de manera efectiva. También pueden realizar presentaciones orales para compartir sus resultados con la clase.
- Fomentar la reflexión y el análisis crítico sobre el proceso de desarrollo del proyecto. Pide a los alumnos que identifiquen los desafíos encontrados, las lecciones aprendidas y las áreas de mejora para futuros proyectos de 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 de la(s) competencia(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 las competencias genéricas y específicas 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”, la metacognición, 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 conceptuales, reportes de prácticas, exposiciones en clase, ensayos, resúmenes, reportes de visitas industriales, trípticos, guías de entrevista, observación y cuestionarios. 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, guías de observación, además de estrategias en las que se logren las competencias blandas.

## 11. Fuentes de información

1. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.
2. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An Introduction to Statistical Learning: with Applications in R*. Springer.
3. Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media.
4. Wickham, H., & Grolemund, G. (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media.
5. VanderPlas, J. (2016). *Python Data Science Handbook: Essential Tools for Working with Data*. O'Reilly Media.