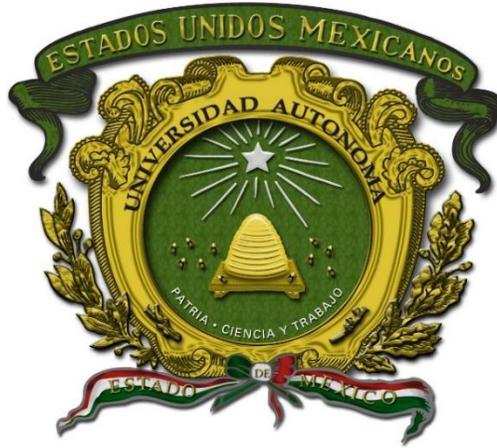


**UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO**  
**LICENCIATURA DE INGENIERÍA EN COMPUTACIÓN**



**PROGRAMA DE ESTUDIOS**

**CIENCIA DE LOS DATOS**

<b>Elaboró:</b>	Dra. Rosa María Valdovinos Rosas	Facultad de Ingeniería
	M. en I. Patricia Romero Ramírez	Facultad de Ingeniería
	Dra. Nely Plata Cesar	U.A.P. Tianguistenco
	Dr. Juan Fernando García Mejía	C.U. UAEM Atlacomulco
	Dr. en C.C. Asdrúbal López Chau	C.U. UAEM Zumpango

**Fecha de  
aprobación:**

**H. Consejo Académico**

**H. Consejo de Gobierno**

**Facultad de Ingeniería**



## Índice

	<b>Pág.</b>
<b>I. Datos de identificación.</b>	<b>3</b>
<b>II. Presentación del programa de estudios.</b>	<b>4</b>
<b>III. Ubicación de la unidad de aprendizaje en el mapa curricular.</b>	<b>5</b>
<b>IV. Objetivos de la formación profesional.</b>	<b>7</b>
<b>V. Objetivos de la unidad de aprendizaje.</b>	<b>8</b>
<b>VI. Contenidos de la unidad de aprendizaje y su organización.</b>	<b>9</b>
<b>VII. Acervo bibliográfico.</b>	<b>11</b>



### I. Datos de identificación.

Espacio académico donde se imparte	<b>Facultad de Ingeniería</b> <b>Centro Universitario UAEM Atlacomulco</b> <b>Centro Universitario UAEM Ecatepec</b> <b>Centro Universitario UAEM Texcoco</b> <b>Centro Universitario UAEM Valle de Chalco</b> <b>Centro Universitario UAEM Valle de México</b> <b>Centro Universitario UAEM Valle de Teotihuacán</b> <b>Centro Universitario UAEM Zumpango</b> <b>Unidad Académica Profesional Tianguistenco</b>
------------------------------------	---

Estudios profesionales	<b>Licenciatura de Ingeniería en Computación, 2019</b>
------------------------	--

Unidad de aprendizaje	<b>Ciencia de los datos</b>	Clave	<b>LINC16</b>
-----------------------	-----------------------------	-------	---------------

Carga académica	<b>1</b>	<b>3</b>	<b>4</b>	<b>5</b>
	Horas teóricas	Horas prácticas	Total de horas	Créditos

Carácter	<b>Obligatoria</b>	Tipo	<b>Taller</b>	Periodo escolar	<b>Séptimo</b>
----------	--------------------	------	---------------	-----------------	----------------

Área curricular	<b>Ingeniería Aplicada y Diseño en Ingeniería</b>	Núcleo de formación	<b>Integral</b>
-----------------	---	---------------------	-----------------

Seriación	<b>Bases de datos II</b>	<b>Ninguna</b>
-----------	--------------------------	----------------

UA Antecedente

UA Consecuente

Formación común

No presenta	<b>X</b>
-------------	----------



## II. Presentación del programa de estudios.

Las grandes cantidades de datos que actualmente se generan en aspectos cotidianos, se derivan de la utilización cada vez más generalizada de plataformas tecnológicas, por ejemplo, compras electrónicas, navegación por internet, uso de celulares o las redes sociales. También, se obtienen a partir de transacciones de grandes sistemas de información, llamadas telefónicas, acceso a wifi públicas, o bien, a partir de interacciones máquina a máquina (M2M) desde dispositivos como sensores de temperatura, luz, presión, ritmo cardíaco o ubicación (GPS); así mismo, provienen de la recolección de datos biométricos como escaneo de retina, de huella digital, cadenas ADN, etc.

Todos estos datos permiten obtener conocimiento valioso para cualquier persona u organización, por lo que con el interés de analizar e interpretar dichos datos surge una nueva área profesional denominada Ciencia de los Datos.

La ciencia de datos o análisis de datos, descrita con diferentes términos (data science, big data o data análisis) es actualmente una de las áreas con mayor demanda de profesionales calificados en el mundo, y es descrita como un campo interdisciplinar de la inteligencia artificial que hace uso de algoritmos como aprendizaje automático, estadística, matemáticas, informática, analítica, programación y el método científico para extraer conocimiento no explícito y útil a partir de bancos de datos.

Dada la importancia que tiene hoy en día el manejo de información, es fundamental que el estudiante de Ingeniería en Computación a su egreso cuente con las competencias profesionales para afrontar exitosamente proyectos basados en grandes volúmenes de datos, provenientes de múltiples orígenes que apoyen la toma de decisiones en las organizaciones en general.

En este curso, se introducen los fundamentos a la Ciencia de los Datos, estudiando los tipos de almacenamiento de datos, fundamentos de Big Data, las bases algorítmicas y matemáticas que permiten modelización y análisis correcto de datos, incluyendo aspectos como técnicas de aprendizaje automático y aprendizaje profundo (Deep Learning), sistemas de recomendación, procesamiento del lenguaje natural, extracción de información en imágenes, etc.

Para este cometido, es fundamental que el alumno tenga acceso a recursos de software actualizado tanto libre y/o propietario, que le permita la realización de proyectos relacionados a la actividad profesional.



### III. Ubicación de la unidad de aprendizaje en el mapa curricular

MAPA CURRICULAR DE LA LICENCIATURA DE INGENIERÍA EN COMPUTACIÓN, 2019

	PERIODO 1	PERIODO 2	PERIODO 3	PERIODO 4	PERIODO 5	PERIODO 6	PERIODO 7	PERIODO 8	PERIODO 9	PERIODO 10	
O B L I G A T O R I A S		Epistemología 3 1 4 7	Comunicación y relaciones humanas 3 1 4 7	Arquitectura de computadoras 3 1 4 7	Circuitos eléctricos y electrónicos 4 2 6 10	Sistemas analógicos 3 1 4 7	Sistemas digitales 3 1 4 7	Sistemas embebidos 2 2 4 6			
	Física 3 1 4 7	Ecuaciones diferenciales 3 1 4 7	Probabilidad y estadística 3 1 4 7	Métodos estadísticos 3 1 4 7	Transmisión de datos 3 1 4 7	Protocolos de comunicación de datos 3 1 4 7	Arquitectura de redes 1 1 4 5	Seguridad de la información 3 1 4 7	Proyecto integral de comunicación de datos 1 3 4 5		
	Álgebra superior 3 1 4 7	Álgebra lineal 3 1 4 7	Matemáticas discretas 3 1 4 7	Métodos numéricos 1 3 4 5	Investigación de operaciones 3 1 4 7	Administración de recursos informáticos 3 1 4 7	Administración de proyectos informáticos 3 1 4 7	Gestión de proyectos de investigación 0 4 4 4			
	Programación I 3 1 4 7	Programación II 3 1 4 7	Paradigmas de programación I 1 3 4 5	Paradigmas de programación II 1 3 4 5	Ingeniería de software I 3 1 4 7	Ingeniería de software II 3 1 4 7	<b>Ciencia de los datos</b> 1 3 4 5		Proyecto integral de ingeniería de software 1 3 4 5		
	Geometría analítica 3 1 4 7	Química 3 1 4 7	Bases de datos I 3 1 4 7	Bases de datos II 1 3 4 5	Ensambladores 3 1 4 7	Compiladores 3 1 4 7	Sistemas operativos 3 1 4 7	Tecnologías computacionales I 1 3 4 5	Tecnologías computacionales II 1 3 4 5		
	Cálculo I 3 1 4 7	Cálculo II 3 1 4 7	Cálculo III 3 1 4 7	Electromagnetismo 3 1 4 7	Inteligencia artificial 3 1 4 7	Procesamiento de imágenes digitales 3 1 4 7	Robótica 3 1 4 7	Integrativa profesional -- ** ** 8			
	El ingeniero y su entorno socioeconómico 3 1 4 7	Inglés 5 2 2 4 6	Inglés 6 2 2 4 6	Inglés 7 2 2 4 6	Inglés 8 2 2 4 6		Graficación computacional 1 3 4 5	Ética profesional y sustentabilidad 2 2 4 6			
											Práctica profesional -- -- -- 30
	O P T A T I V A S								Optativa 1 1 3 4 5	Optativa 2 1 3 4 5	
									Optativa 3 1 3 4 5		

HT	18	HT	20	HT	18	HT	14	HT	21	HT	18	HT	15	HT	9	HT	5	HT	--
HP	6	HP	8	HP	10	HP	14	HP	9	HP	6	HP	13	HP	15+**	HP	15	HP	**
TH	24	TH	28	TH	28	TH	28	TH	30	TH	24	TH	28	TH	24+**	TH	20	TH	**
CR	42	CR	48	CR	46	CR	42	CR	51	CR	42	CR	43	CR	41	CR	25	CR	30





#### **IV. Objetivos de la formación profesional.**

##### **Objetivos del programa educativo:**

Son objetivos de la Licenciatura de Ingeniería en Computación, formar profesionistas que sean capaces de proveer soluciones computacionales innovadoras y sustentables a los problemas, requerimientos y necesidades específicas de la sociedad con responsabilidad ética y mediante la aplicación de metodologías y normas adecuadas en el desarrollo, implantación, optimización, administración y mantenimiento de sistemas de cómputo, que impliquen el uso o la integración de hardware, software y comunicación en diferentes plataformas y dispositivos y desarrollar los aprendizajes y competencias para:

##### **Generales**

- Ejercer el diálogo y el respeto como principios de la convivencia con sus semejantes, y de apertura al mundo.
- Reconocer la diversidad cultural y disfrutar de sus bienes y valores.
- Convivir con las reglas de comportamiento socialmente aceptables, y contribuir en su evolución.
- Adquirir los valores de cooperación y solidaridad.
- Cuidar su salud y desarrollar armoniosamente su cuerpo; ejercer responsablemente y de manera creativa el tiempo libre.
- Ampliar su universo cultural para mejorar la comprensión del mundo y del entorno en que vive, para cuidar de la naturaleza y potenciar sus expectativas.
- Participar activamente en su desarrollo académico para acrecentar su capacidad de aprendizaje y evolucionar como profesional con autonomía.
- Asumir los principios y valores universitarios, y actuar en consecuencia.
- Emplear habilidades lingüístico-comunicativas del idioma inglés.
- Evaluar el progreso, integración e incertidumbre de las ciencias, ante la creciente complejidad de las profesiones.

##### **Particulares**

- Crear proyectos de sistemas computacionales a través de la identificación de necesidades, metodologías ad hoc, teorías de la computación, empleo de sistemas de programación, sistemas electrónicos, comunicaciones y de sistemas, señales y control, para mejorar la cobertura y calidad de los servicios de cómputo de la sociedad y en sectores prioritarios como la educación, salud y seguridad social.



- Evaluar redes de cómputo a través del análisis, el diseño y la administración de la interconexión de dispositivos en redes de computadoras de área local y abierta, considerando estándares y modelos internacionales, para garantizar el rendimiento óptimo en la transmisión de datos.
- Crear nuevas tecnologías computacionales, empleando tecnologías emergentes tales como la inteligencia artificial, la visión computacional, el reconocimiento de patrones, la graficación por computadora, los sistemas embebidos y la ciencia de los datos; para resolver problemas específicos de la sociedad y en sectores prioritarios como la educación, salud y seguridad social.

#### **Objetivos del núcleo de formación:**

Proveer al alumno de escenarios educativos para la integración, aplicación y desarrollo de los conocimientos, habilidades y actitudes que le permitan el desempeño de funciones, tareas y resultados ligados a las dimensiones y ámbitos de intervención profesional o campos emergentes de la misma.

#### **Objetivos del área curricular o disciplinaria:**

Evaluar sistemas computacionales empleando paradigmas como la teoría de la computación, la teoría de juegos, la teoría de las bases de datos, las comunicaciones de datos, la ingeniería de software, tecnologías de vanguardia y arquitecturas de hardware, para optimizar la cobertura y calidad de los servicios de cómputo en las organizaciones de los sectores prioritarios como la educación, salud y seguridad social, así como la creación de aplicaciones específicas.

#### **V. Objetivos de la unidad de aprendizaje.**

Crear modelos de análisis de datos estructurados y no estructurados con la utilización de programación, matemáticas, informática, estadística, algoritmos de aprendizaje automático y la aplicación del método científico para identificar patrones, interpretar fenómenos implícitos en los datos y ofrecer una solución estructurada a problemas de la vida real en la que se involucren grandes volúmenes de datos.



## VI. Contenidos de la unidad de aprendizaje, y su organización.

### Unidad temática 1. Fundamentos de la ciencia de datos

**Objetivo:** Investigar los conceptos y fundamentos de la ciencia de datos, mediante revisión bibliográfica, para lograr su diferenciación e interrelación con las disciplinas involucradas, así como identificar las regularidades, las características y los patrones que caractericen un problema de estudio.

**Temas:**

- 1.1 Conceptos básicos.
- 1.2 Relación entre análisis de datos, minería de datos, inteligencia de negocios y ciencia de datos.
- 1.3 Recuperación de la información, fuentes de datos.
  - 1.3.1 Datos estructurados (Bases de datos, Almacenes de datos).
  - 1.3.2 Datos no estructurados (e-mail, multimedia, texto).
  - 1.3.3 Datos semi-estructurados (HTML, XML, JSON, Binary JSON).
- 1.4 Metodología para integrar proyectos de ciencia de datos.
  - 1.4.1 Diseño de experimentos.
  - 1.4.2 Pruebas de hipótesis.

### Unidad temática 2. Modelos analíticos de datos

**Objetivo:** Analizar los diferentes modelos de análisis de datos, mediante la descripción de los métodos utilizados en cada uno de ellos, a fin de identificar el tipo de análisis más adecuado para un problema de estudio en particular.

**Temas:**

- 2.1 Análisis descriptivo de datos. Descubrimiento de características y complejidades en los datos.
  - 2.1.1 Visualización de datos.
  - 2.1.2 Tratamiento de complejidades de datos (dimensionalidad, ruido, desbalance y traslape de clases, datos faltantes, outliers).
  - 2.1.3 Técnicas de análisis descriptivo: agrupamiento, reglas de asociación.
- 2.2 Análisis predictivo de datos.
  - 2.2.1 Aprendizaje automático.
  - 2.2.2 Estadística aplicada.



### Unidad temática 3. Fundamentos de Big Data

**Objetivo:** Desmenuzar el ciclo de vida que sigue el Big Data, mediante su revisión conceptual, a fin de clarificar cómo métodos de diferentes disciplinas son utilizadas para la recolección, extracción, almacenamiento, integración y manejo de datos masivos y de fuentes diversas orientados a brindar una solución a problemas reales.

**Temas:**

- 3.1 Conceptos básicos de Big Data y analítica de datos.
- 3.3 Proceso o ciclo de vida de Big Data.
- 3.4 La estrategia MapReduce.
- 3.5 Relación con los algoritmos de aprendizaje automático y minería.
- 3.6 Seguridad y tendencias en Big Data.

### Unidad temática 4. Herramientas disponibles para ciencia de datos

**Objetivo:** Escoger diferentes herramientas disponibles para el análisis de grandes volúmenes de datos existentes en el mercado y de acceso abierto, mediante la ejecución de los algoritmos con los que se dispone, a fin de distinguir, asociar y clasificar un conjunto de datos, para dar solución a un problema real y apoyar la toma de decisiones.

**Temas:**

- 4.1 Herramientas de uso libre para análisis de cúmulos de datos: WEKA, RAPID MINER, KEEL, TANAGRA, AdaM, entre otros.
- 4.2 Herramientas de uso libre para Big Data: SPARK, Hadoop, Flink.
- 4.3 Otras herramientas para procesamiento ciencia y analítica de datos (comerciales y otras).



## VII. Acervo bibliográfico

### Básico:

Bruce, P. y Bruce A., (2017), *Practical Statistic for Data Scientists: 50 Essential Concepts*, O'Reilly Media, Inc.

EMC Education Services, (2015), *Data Science and Big Data Analytics*, 1ra. Edición, Wiley.

Grus, J., (2015), *Data Science from Scratch*, O'Reilly Media, Inc.

Mohanty, H., Bhuyan, P., Chenthati, D., (2015), *Big Data: A Primer*, 1ra. Edición, Springer India.

Provost, F. y Fawcett, T., (2013), *Data Science for Business*, O'Reilly Media, Inc.

### Complementario:

Inoubli, W., Aridhi, S., Mezni, H., Jung, A., (2016), *Big data frameworks: A comparative study*.

Ruili W., Wanting J., Mingzhe L., et al., (2018), Review on mining data from multiple data sources, *Pattern Recognition Letters*.