

### Plan del curso

<b>Materia:</b>	Fundamentos de Ciencia de Datos / Minería de Datos		
<b>Clave:</b>	DAT501 / ITI563		
<b>Número de horas:</b>	48 horas		
<b>Catedrático titular:</b>	Dra. Rosa María Cantón Croda		
<b>Datos de contacto:</b>	<a href="mailto:rosamaría.canton@upaep.mx">rosamaría.canton@upaep.mx</a>		
<b>Período:</b>	Primavera 2021		
	<b>Día:</b>	<b>Hora:</b>	<b>Lugar:</b>
<b>Clases:</b>	<b>Martes</b>	<b>19:00 – 22:00</b>	<b>BB-Ultra</b>
<b>Asesorías:</b>	<b>Jueves</b>	<b>17:00 – 18:00</b>	<b>Google Meet</b>
<b>Descripción</b>	<p>La información digital ha redefinido la forma en la que las organizaciones se enfrentan a la explotación de los datos para mejorar su posición competitiva y tomar decisiones.</p> <p>Con un adecuado análisis del Big Data se pueden:</p> <ul style="list-style-type: none"> <li>▪ detectar tendencias</li> <li>▪ realizar predicción de sucesos futuros</li> <li>▪ extraer patrones del comportamiento de los usuarios adaptando mejor los servicios a sus necesidades</li> </ul>		
<b>Los estudiantes en el curso serán capaces de:</b>	<ul style="list-style-type: none"> <li>✓ <i>Saber qué es Ciencia de Datos, cuáles son sus orígenes, sus principales fuentes, los retos a los que se enfrenta y los principales tipos de aplicaciones que se pueden llevar a cabo.</i></li> <li>✓ <i>Determinar cómo y cuándo surgen éstos retos en diferentes campos de actividad.</i></li> <li>✓ <i>Utilizar herramientas de Ciencia de Datos</i></li> <li>✓ <i>Identificar las consecuencias que la generalización del uso de Ciencia de Datos está teniendo en las organizaciones.</i></li> </ul>		

Parámetros de evaluación	Ponderación	Descripción
Reportes de aprendizaje	20%	Realizar los reportes (ClickStream) asignados para reforzar los temas de los módulos de aprendizaje del curso.
Proyecto final	40%	Seleccionar un área del conocimiento y replicar los reportes de aprendizaje, como seguimiento de un caso práctico (ClickStream).
Exámenes rápidos, control de lecturas, actividades	10%	En BB encontrarán la descripción de las actividades de reforzamiento de aprendizaje.
Examen final	20%	Realizar un examen final escrito del contenido visto durante el cuatrimestre.
Participación en clase	10%	Se llevará un registro de las participaciones en clase a través de diversas herramientas.
<b>Total</b>	<b>100</b>	

Normas y políticas internas	
Reportes de aprendizaje.	<ul style="list-style-type: none"> <li>✓ Elabora una plantilla propia para la entrega de todos los reportes.</li> <li>✓ Se desarrollan de manera individual.</li> <li>✓ Se deben cubrir las especificaciones de cada entregable.</li> <li>✓ Calidad en la redacción y estructura del reporte.</li> <li>✓ Subirlos a BB en tiempo y forma. No se aceptan reportes por correo electrónico ni fuera de tiempo.</li> <li>✓ Aportación obtenida del reporte.</li> </ul>
Proyecto final	<ul style="list-style-type: none"> <li>✓ El proyecto final consistirá en la réplica de los reportes (ClickStream) en un área de aplicación específica seleccionada por los estudiantes (salud, gobierno, mercadotecnia, finanzas, educación, etc.)</li> <li>✓ Se desarrollará en equipos de tres personas.</li> <li>✓ Lo ideal es contar con una empresa, o seleccionar un giro de organización que sea posible importar o bajar los archivos con datos de Internet.</li> <li>✓ Debe incluir todas las actividades elaboradas en los reportes.</li> <li>✓ Asegurar la calidad en la redacción y estructura de los entregables.</li> </ul>
Exámenes rápidos, control de lecturas, actividades.	<ul style="list-style-type: none"> <li>✓ Se deben cubrir las especificaciones de cada entregable.</li> <li>✓ Las actividades deben ser entregadas en tiempo y forma en BB.</li> <li>✓ Cualquier trabajo de investigación o lectura podrá ser motivo de examen rápido.</li> <li>✓ Calidad en la redacción y estructura de las actividades.</li> <li>✓ Las investigaciones deben incluir mínimo seis referencias bibliográficas.</li> </ul>
Examen final	✓ Examen escrito que incluye todo el material visto durante el curso.
Asistencia:	✓ 80% de asistencia según reglamento de UPAEP
Políticas generales	<ul style="list-style-type: none"> <li>✓ Todos los estudiantes deben cumplir y hacer cumplir los reglamentos de la Institución.</li> <li>✓ Todo trabajo sometido por un estudiante en este curso debe ser fruto de su propio trabajo. Para este curso el trabajo colaborativo es permitido en las actividades que así lo indiquen.</li> <li>✓ Los estudiantes deben estudiar y discutir la información y los conceptos que el curso cubre de manera individual y en equipo. Debe existir la disponibilidad de colaborar en equipo.</li> <li>✓ La colaboración permitida en el curso nunca debe conducir a que un estudiante presente una copia o parte del trabajo realizado por otro estudiante.</li> <li>✓ En el caso de que un estudiante incurra en el error de copiar el trabajo de otro o cometer plagio en alguna entrega se les aplicará a ambos las sanciones establecidas en los reglamentos de la Institución.</li> </ul>

Clase	Fecha	Temas	Actividades de aprendizaje
1	19-Ene	Introducción al curso M1: Evolución en los Sistemas de Gestión y Análisis de Datos.	<b>Act1.</b> Del bit... al Big Data
2	26-Ene	M1: Introducción a Ciencia de Datos: definiciones: datos, información, conocimiento, El concepto de Big Data	<b>Act2.</b> Rasgos de Big Data ClickStream01
3	02-Feb	M1: EL rol que Big Data desempeña en las organizaciones. M2: Obtención de datos	<b>Act3.</b> Empresas impulsadas por los datos/roles de ciencia de datos ClickStream02
4	09-Feb	M2: Trabajo con datos	<b>PF1:</b> Propuesta Ejecutiva ClickStream03
5	16-Feb	Presentación propuesta ejecutiva del proyecto final M2: Trabajo con datos: transformación de datos originales a un modelo multidimensional	ClickStream04
6	23-Feb	M2: Trabajo con datos: Consulta de información del modelo multidimensional	ClickStream05 <b>PF2:</b> Modelado y transformación de datos
7	02-Mar	M3: Gestión de datos: Almacenamiento de datos M3: Gestión de datos: Ecosistema Hadoop	Act4. MapReduce ClickStream06
8	09-Mar	Presentación de avance de Proyecto Final	
9	16-Mar	M3: Gestión de datos: Uso de HDFS y Hive en el ecosistema Hadoop.	ClickStream07
10	23-Mar	M3: Gestión de datos: Fundamentos de estadística	ClickStream08
11	06-Abr	M3: Gestión de datos: Análisis de datos	
12	13-Abr	M3: Gestión de datos: Análisis de datos	ClickStream09 <b>PF3:</b> Entrega final
13	20-Abr	M4: Visualización y consumo de datos	ClickStream10
14	27-Abr	Examen final escrito	
16	04-May	Presentación de Proyectos Finales	