



Código	Nombre			
MDS7103	Bases de Datos			
Nombre en inglés				
Databases				
CT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	10	3	1,5	5,5
Requisitos			Carácter del Curso	
CC3101/FI2002			Obligatorio Magíster en Ciencia de Datos	
Resultados de Aprendizaje				
<p>El propósito del curso es que los estudiantes del curso aprenderán sobre la estructura y funcionamiento de las bases de datos relacionales y NoSQL, modelamiento de datos, funcionalidades esenciales presentes en los productos que permiten la creación de ambos tipos de bases de datos y cuando se debe usar cada una de ellas.</p> <p>Dentro de las competencias que se espera que el estudiante desarrolle están:</p> <ul style="list-style-type: none"> <li>- Diseñar e implementar sistemas de almacenamiento de datos de acuerdo a las condiciones de la industria o sector en donde se desarrolla el proyecto de Data Science.</li> <li>- Aprender a analizar las variables de un problema de administración de datos que devengan en el uso de una base de datos relacional o de una NoSql</li> <li>- Conocer la estructura y funcionamiento de los sistemas de administración de bases de datos relacionales y NoSql</li> <li>- Comprender las funcionalidades esenciales presentes en los productos de Software que permiten la creación de bases de datos relacionales y NoSql</li> </ul> <p>Al finalizar el curso, se espera que los estudiantes hayan:</p> <ul style="list-style-type: none"> <li>- Aprendido a modelar datos usando el enfoque Entidad Relación.</li> <li>- Comprendido el funcionamiento de un motor de bases de datos relacional y NoSQL</li> <li>- Aprendido a programar en lenguaje SQL.</li> <li>- Entendido las distinciones fundamentales de un problema de administración de datos que devenga en el uso de una bases de datos relacional o NoSql</li> </ul>				

Metodología Docente	Evaluación General
Este curso tiene una connotación teórico-práctica. Está compuesto por cátedras y trabajo práctico de análisis de datos utilizando las técnicas aprendidas a lo largo del curso.	<p>El curso se evalúa a partir de tareas que permiten a los estudiantes aplicar y ejercitar las diferentes técnicas aprendidas en el curso</p> <p>El cálculo de esas notas se efectúa de la siguiente forma:</p> $NT = \text{Promedio de las notas parciales } (\sum w_i * P_i) / n, \text{ donde } P_i \text{ son las notas de las tareas y } w_i \text{ la ponderación que tiene cada una de ellas.}$



	<p>En caso de que el alumno rinda el examen, la nota final se calcula de la siguiente forma.  <math>NT * 0,6 + EX * 0,4</math></p> <p>La condición para aprobar el curso es:  <math>NT \geq 4.0</math></p>
--	--

### Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	<b>Bases de datos relacionales</b>	5
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>Modelamiento entidad relación               <ul style="list-style-type: none"> <li>Conceptos básicos</li> <li>Representaciones gráficas</li> <li>Ejemplos</li> </ul> </li> <li>Modelo Relacional               <ul style="list-style-type: none"> <li>Conceptos básicos: relación, tabla, dominio, tupla, esquemas</li> <li>Integridad de datos.</li> <li></li> </ul> </li> <li>Diseño de bases de datos relacionales               <ul style="list-style-type: none"> <li>Definición del problema</li> <li>Normalización: 1NF, 2NF y 3NF</li> <li>Criterios para normalizar</li> </ul> </li> <li>Lenguaje SQL ACID (Atomicity, Consistency, Isolation, Durability)</li> </ul>	<p>Los estudiantes adquieren el conocimiento adecuado respecto del modelamiento de datos usando el enfoque entidad relación, así como los fundamentos de la programación en lenguaje SQL, con su correspondiente realización práctica en la creación de funcionalidades basadas en requerimientos.</p>	<p>Connolly, T. M., &amp; Begg, C. E. (2015). <i>Database systems: a practical approach to design, implementation, and management</i>. (6th edition) Pearson Education.</p> <p>Harrington, J. L. (2016). <i>Relational database design and implementation</i>. Morgan Kaufmann.</p>

Número	Nombre de la Unidad	Duración en Semanas
2	<b>Sistemas de administración de bases de datos relacionales</b>	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía



<p>Estructura de una base de datos relacional</p> <ul style="list-style-type: none"> <li>• Tablespace y archivos</li> <li>• Instancia de la base de datos</li> <li>• Tablas particionadas</li> <li>• Vistas materializadas</li> <li>• Índices</li> <li>• Respaldo y recuperación de la base de datos</li> </ul>	<p>Los estudiantes comprenden la estructura básica de un motor de bases de datos relacionales, su funcionamiento mínimo. Comprenden el concepto de “instancia de bases de datos” y aprenden a utilizar estructuras de aceleración de consultas. También comprenderán la importancia del respaldo adecuado para la recuperación eficaz de una base de datos</p>	<p>C. J. Date, A. Kannan and S. Swamynathan, <i>An Introduction to Database Systems</i>, Pearson Education, Eighth Edition, 2009.</p> <p>Abraham Silberschatz, Henry F. Korth and S. Sudarshan, <i>Database System Concepts</i>, McGraw-Hill Education (Asia), Fifth Edition, 2006.</p> <p>Shio Kumar Singh, <i>Database Systems Concepts, Designs and Application</i>, Pearson Education, Second Edition, 2011.</p>
---	--	--

Número	Nombre de la Unidad	Duración en Semanas
3	<b>Bases de datos No relacionales</b>	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>• Datos masivos: Orígenes de NoSQL</li> <li>• El Teorema CAP: Consistencia, Accesibilidad, Particionamiento (Conjetura de Brewer),</li> <li>• MapReduce</li> <li>• Sistemas NoSql:               <ul style="list-style-type: none"> <li>○ Clave valor(Redis, Riak KV)</li> <li>○ Bases de datos documentales (Mongo Db)</li> <li>○ Familia de columnas                   <ul style="list-style-type: none"> <li>▪ Google Big Table</li> <li>▪ Casandra</li> </ul> </li> <li>○ Basadas en Grafos                   <ul style="list-style-type: none"> <li>▪ Neo4j</li> </ul> </li> </ul> </li> </ul>	<p>Los estudiantes comprenden la estructura y funcionamiento a nivel conceptual de las bases de datos NoSql. Comprenden la necesidad de este nuevo enfoque cuando existen datos masivos en formatos complejos como lo son los textos, imágenes etc.</p>	<p>Harrison, G. (2015). <i>Next Generation Databases: NoSQL and Big Data</i>. Apress.</p> <p>Fox, A., &amp; Brewer, E. A. (1999). Harvest, yield, and scalable tolerant systems. In <i>Hot Topics in Operating Systems, 1999. Proceedings of the Seventh Workshop on</i> (pp. 174-178). IEEE.</p> <p>Gilbert, S., &amp; Lynch, N. (2002). Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web</p>



**fcfm**

ESCUELA DE POSTGRADO  
Y EDUCACIÓN CONTINUA  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
UNIVERSIDAD DE CHILE

		<p>services. <i>Acm Sigact News</i>, 33(2), 51-59.</p> <p>Dean, J., &amp; Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. <i>Communications of the ACM</i>, 51(1), 107-113.</p> <p>Hills, T. (2016). <i>NoSQL and SQL Data Modeling: Bringing Together Data, Semantics, and Software</i>. Technics Publications.</p>
--	--	--

Número	Nombre de la Unidad	Duración en Semanas
4	<b>Productos para la creación de bases de datos relacionales y NoSql</b>	4
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>● MySql</li> <li>● PostgreSQL</li> <li>● Motores de bases de datos comerciales</li> <li>● Hadoop</li> <li>● HDFS</li> <li>● Spark</li> <li>● Pig</li> <li>● Hive</li> <li>● Impala</li> <li>● Streaming</li> <li>● Scala</li> <li>● Mongo DB</li> </ul>	<p>Los estudiantes entienden la estructura y funcionamiento de varios productos de software utilizados en la creación de bases de datos relacionales y NoSql</p>	<p>White, T. (2012). <i>Hadoop-The Definitive Guide: Storage and Analysis at Internet Scale</i> (revised and updated).</p> <p>Wampler, Dean, and Alex Payne. <i>Programming Scala: Scalability=Functional Programming+ Objects</i>. " O'Reilly Media, Inc.", 2014.</p> <p>Carpenter, Jeff, and Eben Hewitt. <i>Cassandra: The Definitive Guide: Distributed Data at Web</i></p>



fcfm

ESCUELA DE POSTGRADO  
Y EDUCACIÓN CONTINUA  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
UNIVERSIDAD DE CHILE

		<p>Scale. " O'Reilly Media, Inc.", 2016.</p> <p>Chodorow, Kristina. <i>MongoDB: The Definitive Guide: Powerful and Scalable Data Storage.</i> " O'Reilly Media, Inc.", 2013.</p>
--	--	--

#### Bibliografía General

- Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia. Learning Spark. O'Reilly 2015
- Mahmoud Parsian. Data Algorithms: Recipes for Scaling Up with Hadoop and Spark. O'Reilly 2015
- Vladimir Bacvanski. Introduction to Big Data An Overview of Fundamental Big Data Concepts, Tools, Techniques and Practices. O'Reilly Media, 2015.
- Salvador García, Julián Luengo, and Francisco Herrera. Data preprocessing in data mining. New York: Springer, 2015.
- Elmasri, Ramez. *Fundamentals of database systems.* Pearson Education India, 2008.
- Elmasri, Ramez, and Sham Navathe. *Fundamentals of database systems.* London: Pearson, 2016.
- Tiwari, Shashank. *Professional NoSQL.* John Wiley & Sons, 2011.

Vigencia desde:	Otoño 2019
Elaborado por:	Juan Velásquez S y Claudio Gutierrez
Validado por:	Comité Académico Magíster de Ciencia de Datos