

1. Datos Generales de la asignatura

Nombre de la asignatura:	<i>Diseño y Construcción de Data Warehouse</i>
Clave de la asignatura:	<i>TBD-1802</i>
SATCA¹:	<i>2-3-5</i>
Carrera:	<i>Ingeniería en Sistemas Computacionales, Ingeniería en Tecnologías de la Información y Comunicaciones e Ingeniería en Informática.</i>

2. Presentación

<p>Caracterización de la asignatura</p> <p>Esta asignatura aporta al perfil del egresado los conocimientos teóricos y prácticos indispensables para el diseño y construcción de los sistemas. Data Warehouse, Los Data Warehouse también conocidos como almacenes, repositorios y en algunas ocasiones como Bodega de Datos han adquirido relevancia en la actualidad por las necesidades de las Empresas y Organizaciones de manejar grandes volúmenes de datos (Big Data) y en plena Era de las sociedades de la Información la obtención de la misma de manera suficiente y oportuna ha sido un elemento que reduce la incertidumbre a la hora de tomar decisiones en cualquier organización.</p> <p>La importancia de la asignatura en el contexto actual radica en que se dan a conocer herramientas utilizadas en la Inteligencia de negocios (BI), la cual es una estrategia empresarial que persigue incrementar el rendimiento de la empresa o la competitividad del negocio, a través de la organización inteligente de sus datos históricos (transacciones u operaciones diarias), usualmente residiendo en Data Warehouse corporativos o Data Marts departamentales. En las últimas décadas, la implantación y el uso de sistemas de información ha favorecido la acumulación de grandes volúmenes de datos, sobre todo en sistemas y bases de datos</p>

¹ Sistema de Asignación y Transferencia de Créditos Académicos

transaccionales, que en su mayoría requieren ser depuradas, organizadas y analizadas para su uso eficiente.

La asignatura consiste en proveer al alumno los conocimientos acerca de la Estructura de datos, flujo de datos y guardarlo en repositorio de datos para la identificación y la extracción de conocimiento o información relevantes. El propósito es dar soporte, de la mejor manera, a la toma de decisiones tácticas y estratégicas que contribuyan al desarrollo de ventajas competitivas en una organización.

La relación que tiene la asignatura con las otras que conforman el módulo de Especialidad de Base de datos, es que los Data Warehouse almacena los datos de las bases de datos de manera física o en la nube, utilizando minería y análisis de datos el usuario consulta y las respuestas que obtiene dependen de los volúmenes del Data warehouse de su estructura y organización de los datos para soportar funciones de administración,

Para esto se han generado diversas técnicas, herramientas y metodologías, tanto para la depuración y la selección de la información, como para el diseño e implantación orientada, en su conjunto, a proveer información útil para la toma de decisiones en una organización.

Intención didáctica

La asignatura está organizada en seis temas:

El primer tema, se centra en la evolución de los sistemas de información, haciéndose énfasis en la importancia de la información en la toma de decisiones y como han evolucionado las bases de datos.

El segundo, aborda la importancia del BI (del inglés business intelligence), también conocido como inteligencia de negocios o inteligencia empresarial, como el conjunto de estrategias, aplicaciones, datos, productos, tecnologías y arquitectura técnicas, están enfocados a la administración y creación de conocimiento sobre el medio, a través del análisis de los datos.

El tercer tema se centra en las características del Data Warehousing recientes, de la misma forma se aborda el tema de lenguaje de presentación CSS

El cuarto tema, se centra en el diseño conceptual del Repositorio

El quinto tema se avoca al estudio conceptual de los principales procesos de carga y actualización

El sexto tema nos lleva al estudio de aspectos tecnológicos y metodológicos que se requieren para analizar las estructuras de los datos para el análisis de negocios, arquitecturas de Data warehouse, Tecnologías de Base de Datos, hardware utilizado y las estrategias de seguridad.

3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Instituto Tecnológico de Villahermosa del 12 de Febrero al 01 de Junio de 2018.	Academia de Ingenierías en Sistemas computacionales, Ingeniería en TIC's e Ingeniería en Informática.	Reunión para la elaboración de la Especialidad de Tecnologías de Base de Datos.

4. Competencia(s) a desarrollar

Competencia(s) específica(s) de la asignatura
Aplica las tecnologías orientadas al análisis de datos, como es el Data Warehousing, para descubrir el conocimiento implícito en una base de datos y tener elementos que ayuden a la toma de decisiones.

5. Competencias previas

<ul style="list-style-type: none"> Comprende y aplica los conceptos básicos de lógica matemática, conjuntos y relaciones para aplicarlos en modelos que resuelvan problemas computacionales Analiza requerimientos y diseña bases de datos para generar soluciones al tratamiento de información basándose en modelos y estándares. Implementa bases de datos para apoyar la toma de decisiones considerando las reglas de negocios.

6. Temario

No.	Temas	Subtemas
1.	Introducción	1.1 Cronología de los sistemas de información 1.2 Información para la toma de decisiones 1.3 Evolución de las bases de datos
2.	Inteligencia De Negocios	2.1 Conocimiento del negocio 2.2 Estrategias para la definición de indicadores de desempeño 2.2.1 Top-Down 2.2.2 Button-Up

	3. Data Warehousing	<p>3.1 Características</p> <p>3.2 Estructura de datos y flujo de datos</p> <p>3.2.1 Dos enfoques: Inmon y Kimball</p> <p>3.2.2 Conocimiento del negocio</p>
	4. Diseño Conceptual	<p>4.1 Conceptos Generales y Proceso de Diseño.</p> <p>4.2 Diseño Conceptual a partir de requerimientos.</p> <p>4.3 Aspectos Prácticos</p>
	5. Procesos De Carga Y Actualización	<p>5.1 Conceptos Generales</p> <p>5.2 Técnicas de limpieza de datos</p>
	6. Aspectos Tecnológicos Y Metodológicos	<p>6.1 Estructuras de datos para el análisis de negocio</p> <p>6.1.1 Non-Architected</p> <p>6.1.2 Data Marts dependientes</p> <p>6.1.3 Bus</p> <p>6.2 Arquitecturas de Sistemas de DW.</p> <p>6.3 Tecnologías de DBMS.</p> <p>6.4 Selección de Hardware.</p> <p>6.5 Seguridad</p>

7. Actividades de aprendizaje de los temas

Introducción	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> • Visualiza la evolución de los sistemas de Información • Comprende la importancia de los sistemas de Información para la toma de decisiones. <p>Genéricas:</p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis • Habilidad para manejo de equipo de cómputo • Habilidad para trabajar de forma autónoma 	<ul style="list-style-type: none"> • Investigar en diversas fuentes la evolución de los sistemas de Información, elaborando una línea del tiempo • Elabora un ensayo acerca de la importancia de la información en la toma de decisiones • Investiga la Evolución de Las Bases de Datos sobre el Paradigma SQL comparándolas con el Paradigma No Sql., para elaborar una tabla comparativa
Inteligencia De Negocios	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> • Comprende la trascendencia de la inteligencia de negocios (B.I) como una estrategia empresarial que permite optimizar recursos, monitorear el cumplimiento de los objetivos y la capacidad de tomar buenas decisiones para así obtener mejores resultados. <p>Genéricas:</p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Capacidad de organizar y planificar. • Comunicación oral y escrita. • habilidad para buscar y analizar información • proveniente de fuentes diversas. 	<ul style="list-style-type: none"> • Investigar la importancia de un sistema de medición de desempeño y la importancia de los indicadores para cuantificar los procesos, elaborando un resumen. • Elaborar un diagrama de integración de Indicadores Top Down o Vertical Descendente aplicándolo a un caso real • Elaborar un diagrama de integración de Indicadores Bottom Up o aplicándolo a un caso real

Data Warehousing	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> • Comprende el concepto de Data warehousing, como la construcción de un almacén de datos, integrado no volátil y variable en el tiempo. • Identifica las aportaciones Kimball e Inmon en el diseño de Data warehouses. <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidad para manejo de equipo de cómputo • Habilidad para trabajar de forma autónoma • Solución de problemas • Habilidades de investigación. • Capacidad de aprender. • Capacidad de adaptarse a nuevas situaciones. 	<ul style="list-style-type: none"> • Elaborar un diagrama que Describa un Data Warehouse. • Mediante un algoritmo representa el proceso de carga utilizando estructura 3FN, en la que los datos son depurados antes de pasar a la estructura normalizada del datawarehouse según la propuesta Top Down de Inmon. • Representar gráficamente el Modelo de un data warehouse de acuerdo a La visión de Kimball.
Diseño Conceptual	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> • Identifica los Conceptos Generales empleados en el diseño de almacenes de datos. • Conoce las fases del diseño Conceptual a partir de requerimientos. • Realiza aplicaciones prácticas 	<ul style="list-style-type: none"> • Realiza un listado de los principales términos empleados para el diseño de almacenes de datos. • Desarrollar una metodología para el diseño conceptual de Almacenes de Datos a partir de la

<p>Genéricas:</p> <ul style="list-style-type: none"> Habilidad para manejo de equipo de cómputo Habilidad para trabajar de forma autónoma Solución de problemas Habilidades de investigación. Capacidad de aprender. Capacidad de adaptarse a nuevas situaciones. 	<p>descripción conceptual de las bases de datos operacionales y de los requisitos de usuario, definiendo sus tres fases:</p> <p>Fase 1 Derivación de esquemas multidimensionales: Fase 2. Especificación de requisitos de usuario. Fase 3. Integración.</p> <ul style="list-style-type: none"> Realizar aplicaciones prácticas
<p>Procesos De Carga Y Actualización</p>	
<p>Competencias</p>	<p>Actividades de aprendizaje</p>
<p>Específica(s):</p> <ul style="list-style-type: none"> Identifica los principales conceptos empleados en los procesos de carga y Descarga. Conoce las principales técnicas de limpieza de datos. <p>Genéricas:</p> <ul style="list-style-type: none"> Capacidad de análisis y síntesis Habilidad para manejo de equipo de cómputo Habilidad para trabajar de forma autónoma 	<ul style="list-style-type: none"> Investigar en diversas fuentes la los principales conceptos empleados en las técnicas de limpieza de datos. Conoce el proceso de Limpieza de Datos compuesta por las siguientes etapas: <ol style="list-style-type: none"> Auditoría de Datos Definición de Workflow (Flujo de Trabajo):. Ejecución de Workflow Post-Proceso y Control:
<p>Aspectos Tecnológicos Y Metodológicos</p>	
<p>Competencias</p>	<p>Actividades de aprendizaje</p>
<p>Específica(s):</p> <ul style="list-style-type: none"> Identifica los Conceptos Generales empleados en el diseño de almacenes de datos. 	<ul style="list-style-type: none"> Identifica las estructuras de datos para análisis de negocios Reconoce la arquitectura básica de un almacén de datos

<ul style="list-style-type: none"> • Conoce las fases del diseño Conceptual a partir de requerimientos. • Realiza aplicaciones prácticas <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidad para manejo de equipo de cómputo • Habilidad para trabajar de forma autónoma • Solución de problemas • Habilidades de investigación. • Capacidad de aprender. • Capacidad de adaptarse a nuevas situaciones. 	<ul style="list-style-type: none"> • Selecciona el hardware en base al diseño conceptual del Data Warehouse. • Identifica las principales medidas de seguridad para asegurar el proceso <p>Selecciona Herramientas de software Open Source, que permitan importar los datos desde un Data Source y obtener información</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

8. Práctica(s)

<ol style="list-style-type: none"> 1. Diseñar un DWH mediante la metodología Kimball 2. Elaborar un esquema Estrella para evaluarse mediante un Sistema OLAP 3. Elaborar un Esquema Copo de Nieve 4. Elaborar un esquema de Constelación de Hechos. 5. Creación de DWH a partir de diferentes orígenes. Access, Excel, Mysql utilizando la Herramienta Pentaho.

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

Son las técnicas, instrumentos y herramientas sugeridas para constatar los desempeños académicos de las actividades de aprendizaje.

11. Fuentes de información

1. Ralph Kimball & Margy Ross The Data warehouse Toolkit. The definitive Guide To Dimensional Modeling Third Edition. Kimball Group 2013.
2. Sid Adelman, Larissa Terpeluk Moss, Data Warehouse Project Management, Third Edition. Addison Wesley , primera 2010.
3. Ballard, Ch.; Rollins, J.; Ramos, J.; Perkins, A.; Hale, R.; Dorneich, A.; Cas Milner, E. & Chodagam, J. (2012). Dynamic Warehousing: Data Mining Made Easy. IBM International Technical Support Organization. IBM Press. USA.
4. Arabí, U., 2013. Ethical data mining and social Science data exploration and description: scope and limitations in social Science research. In: Ethical data mining applications for socio-economic development. United States of America: Idea Group Inc (IGI), pp. 22-39.
5. Barros, R. C., Basgalupp, M. P., Carvalho, A. C. R L. F. d. & Freitas, A. A., 2012. A Survey of Evolutionary Algorithms for Decision-Tree Induction. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on V42 Issue 3 , pp. 291-312.
6. **Hajian, S. & Domingo, F. J.**, 2013. A Methodology for Direct and Indirect Discrimination Prevention in Data Mining. IEEE Transactions on Knowledge and Data Engineering Vol25No. 7,pp. 1445-1459.