

Planificación de la Asignatura: Bases de Datos Multidimensionales

Fecha: 23/10/2024 13:02

Código: T1637

Carrera: Tecnicatura Universitaria en Procesamiento y Explotación de Datos

Departamento Académico: Informática

Docente a cargo:

Correo del docente a cargo: sin datos

Régimen de Dictado: Cuatrimestral 2º Cuatrimestre

Carga Horaria Semanal: 5 horas semanales

Carga Horaria Total: 70 horas

Contenidos Mínimos:

Entorno de Depósito de Datos (Data Warehouse). Diseño conceptual, lógico y físico. Técnicas para la extracción, transformación y carga de los datos. Técnicas de explotación de la información. Seguridad de acceso y manipulación de la información. Modelado multidimensional.

Correlativas Regulares para cursar:

Bases de Datos

Modelado Estadístico

Exploración de Datos Multivariados

Correlativas Aprobadas para cursar:

Comprensión Lectora y Producción Escrita,

Espacio Integrador I,

Probabilidad y Estadística;

Algoritmos y Estructura de Datos

Aspectos Legales del Uso de la Información

Correlativas Aprobadas para promocionar o rendir el examen final:

No posee

Objetivo General:

- Desarrollar los conceptos de cómo se estructura los datos en modelos multidimensionales.
- Introducir la importancia de las bases de datos para el desarrollo de Sistemas de Información orientados a Business Intelligence (BI)
- Enunciar y explicitar las distintas metodologías para el modelado de datos.
- Desarrollar los fundamentos y la arquitectura de los Sistemas para BI
- Describir los fundamentos y características de las bases de datos multidimensionales, las operaciones de definición y manipulación de datos.

Objetivos Particulares:

- Estudiar los diferentes enfoques para el modelado de datos.
- Explicar los conceptos de hechos, dimensiones y su relación con los fundamentos de bases de datos.
- Desarrollar los elementos y sentencias del lenguaje de consulta para la definición, manipulación de datos, administración de usuarios y seguridad.
- Describir los aspectos vinculados a la construcción de las etapas ETL (Extract, Transform and Load)
- Describir los aspectos que hacen a la consistencia e integridad de los datos
- Describir las acciones a tomar relacionados con la seguridad física y lógica de los datos, resaltando su importancia.

Programa Analítico:

Unidad 1

Introducción. Data Warehouse. Concepto. Bases de datos relacionales y el modelo multidimensional. Arquitectura. Diseño cúbico. Métricas. Operaciones OLAP. ETL. Tiers. Variaciones de la arquitectura. Herramientas de diseño. Herramientas de inteligencia de negocios. Pentaho.

Unidad 2

Diseño conceptual de Data Warehouses. Esquemas, dimensiones y niveles. Atributos e identificadores. Hechos, roles y cardinalidad. Jerarquías balanceadas y no balanceadas. Generalización. Jerarquías alternativas, paralelas, no estrictas. Aspectos de modelización avanzada.

Unidad 3

Diseño lógico de Data Warehouse. OLAP Relacional (ROLAP). Implementación relacional del modelo conceptual. Dimensión temporal. Representación lógica de jerarquías. Operaciones SQL/OLAP. OLAP multidimensional (MOLAP) y OLAP Híbrido (HOLAP). Esquema estrella. Esquema snowflake, starflake y constellation.

Unidad 4

Data Warehouse. Diseño físico. Vistas materializadas. Índices de mapas de bits. Conexión de índices. Fragmentación. Particionamiento vertical. Particionamiento horizontal. Manejo de bases de datos particionadas. Reescritura de consultas. Mantenimiento y estrategias computacionales de un cubo de datos. Estimación del tamaño de un cubo. Indexación.

Unidad 5

Extracción, transformación y carga (ETL). Modelado de procesos de negocios (BPMN). Flujo de objetos, actividades, subprocessos, pasarelas, eventos. Diseño conceptual. Ejemplo de ETL. Modelado ETL usando Kettle.

Unidad 6

Técnicas de explotación de la información. Minería de datos. Tareas. Modelado descriptivo y predictivo. Descubrimiento de patrones. Clasificación supervisada. Clustering. Reglas de asociación. Algoritmo de crecimiento de patrones. Patrones secuenciales. Indicadores clave de rendimiento (KPI). Clasificación. Definiciones. Tableros. Clasificación. Diseño.

Unidad 7

Diseño y modelado multidimensional. Generalidades. Especificación de requerimientos. Identificación de usuarios. Determinación de las necesidades de análisis. Requerimientos documentales. Especificación de requisitos basados en análisis y basados en fuentes. Caso de estudio.

Listado de Actividades de Formación Práctica:

Metodología de Evaluación Durante el cursado:

Durante el cursado se efectúa un seguimiento continuo a través de la resolución de los problemas en clases y el desarrollo de un trabajo práctico (TP) grupal (2 alumnos por grupo) que debe entregarse en tiempo y forma. El trabajo práctico se evalúa de dos formas, en primer lugar se los califica cuan bien se lograron las consignas y objetivos del TP, como su presentación, completitud, etc., además se realiza un coloquio de defensa de los mismos donde los alumnos deberán exponer los fundamentos, recursos y medios utilizados para la realización del TP. El TP consiste en el desarrollo de una aplicación de Bases de Datos: diseño, implementación, carga de datos y consultas.

Metodología de Evaluación en Exámenes Finales:

El examen final para el alumno regular consiste en un examen teórico-práctico escrito el cual integra todos los temas dictados en la asignatura. Previo a dicho examen defenderá oralmente el TP integrador que se le planteó durante el cursado

El examen final para el alumno libre consiste en la presentación y defensa del desarrollo de una aplicación de Bases de Datos: diseño, implementación, carga de datos y consultas, cuyo enunciado será entregado con una semana de anticipación. Superada esta instancia el alumno podrá rendir un examen escrito teórico-práctico similar al del alumno regular aunque con un ejercicio/problema de práctica adicional.

Condiciones de Regularidad :

.Para adquirir la regularidad en la asignatura, los alumnos deben asistir al 75% de las clases prácticas y de laboratorio. Deben entregar en tiempo y forma el Trabajo Práctico y aprobar la defensa del mismo. Además deberán aprobar con 50% o más, los cuestionarios ó sus recuperatorios (se pueden recuperar ámbos parciales).

El régimen de promoción para esta asignatura exige cumplir las condiciones de regularidad y aprobar con 70% o más la defensa oral del TP integrador

Bibliografía Principal:

Datawarehouse Systems. Design and Implementation. 2nd Edition. Alejandro Vaisman, Esteban Zimányi.

ISBN10 3662651661

ISBN13 9783662651667. Editorial Springer-Verlag Berlin and Heidelberg GmbH & Co. KG. 2022

Bibliografía Complementaria:

1- ELMASRI R. y NAVATHE S., "Sistemas de Bases de Datos", Addison-Wesley Iberoamericana, 3ra ó 5ta Edición, 2002 (7 ejemplares disponibles en Biblioteca FI-UNER)

2- KORTH H. y SILBERSCHATZ A., "Fundamentos de Bases de Datos 3/e", McGraw-Hill, 1998 (2 ejemplares disponibles en Biblioteca FI-UNER).

3- JOYANES AGUILAR, L.. Big data : análisis de grandes volúmenes de datos en organizaciones. México D.F. : Alfaomega - 2013 (1 ejemplar disponible en Biblioteca FI-UNER)