

Diplomatura en Big Data
Módulo 1

Data Warehousing y OLAP

Alejandro VAISMAN

Departamento de Ingeniería Informática
Instituto Tecnológico de Buenos Aires
avaisman@itba.edu.ar

Objetivo del Curso de DW & OLAP

- ◆ Introducir los principales conceptos de OLAP y Data Warehousing
 - Establecer las diferencias entre BD transaccionales y DW
- ◆ Manejar las técnicas de Modelado Multidimensional
- ◆ Dar un panorama general de las herramientas disponibles
- ◆ Introducir las tecnologías necesarias para DW y OLAP en Big Data
- ◆ Introducir las Bases de Datos de Grafos, y cómo se pueden utilizar las BD NoSQL en DW
- ◆ Presentar los desafíos futuros para DW
- ◆ Notar que:
 - No vamos a cubrir ETL
 - Es el curso que requiere mayores conocimientos previos de bases de datos

Esquema del Curso

- ◆ Presentación del Curso - SQL
- ◆ Introducción a Data Warehousing y OLAP
- ◆ Diseño Conceptual, Lógico y Físico
- ◆ Metodología de diseño
- ◆ Consultas OLAP
- ◆ OLAP en Big Data
- ◆ Evaluación:
 - Realización de un proyecto (diseño de un DW)
 - Evaluación en la última clase

Cronograma

- ◆ Clase 1 (12/4)
 - Modelo lógico: Presentación
 - Introducción a SQL - PostgreSQL
- ◆ Clase 2 (13/4)
 - Introducción a OLAP y DW
 - Diferencias entre OLAP y OLTP: consultas analíticas vs consultas transaccionales
 - Ubicación del campo en el contexto de Big Data
 - Arquitectura de un DW
 - Modelo multidimensional: el cubo de datos
- ◆ Clase 3 (26/4)
 - Diseño conceptual
 - Diseño lógico
- ◆ Clase 4 (27/4)
 - Modelo lógico: Slowly changing dimensions
 - Explotación del DW: SQL/OLAP
 - Definición de proyectos
- ◆ Clase 5 (3/5)

- Explotación del DW: SQL/OLAP
- Diseño físico del DW
- Revisión proyectos
- ◆ Clase 6 4/5)
 - OLAP en Big Data: Kylin - introducción
 - Revisión de proyectos
- ◆ Clase 7 (10/5)
 - OLAP sobre bases de datos no relacionales: grafos, Neo4J, Cypher.
 - Revisión proyectos
- ◆ Clase 8 (11/5)
 - Exposición y discusión de los proyectos.
 - Examen.

Bibliografía

- ◆ Slides y papers en IOL
- ◆ *The Data Warehouse Toolkit*. R. Kimball. 1st Ed. Wiley, 1996
- ◆ *The Data Warehouse Lifecycle Toolkit*. Ralph Kimball, Margy Ross, Warren Thornthwaite and Joy Mundy, 2008
- ◆ *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Wiley, 2013
- ◆ *Multidimensional Databases and Data Warehousing (Synthesis Lectures on Data Management)*. Christian S. Jensen, Torben Bach Pedersen, Christian Thomsen and M. Tamer Ozsu, 2013 (Research)
- ◆ *Data Warehouse Systems: Design and Implementation*. A. Vaisman and E. Zimányi, 2014 (Textbook) (En biblioteca)
- ◆ Multitud de libros prácticos: herramientas de SW, dashboards, KPIs, ETL, etc.