
DS-306 Data Warehousing and Business Intelligence

Lecture 1: Introduction

Dr. Khurram Shahzad

Agenda

- Introduction
- Course Material
- Course Evaluation
- Course Contents

Course Material (Books)

- A. Vaisman, E. Zimanyi. Data Warehouse Systems: Design and Implementation. Springer
- Data Warehousing Fundamentals. A Comprehensive Guide for IT Professionals. by Paulraj Ponniah.

Grading Scheme (most likely)

- Assignments
- Quizes
- Mid Term (35 marks)
- Final Term (40 marks)

Job Description Example

Data Warehouse / Business Intelligence (BI) Developer

Employer

Syracuse University

Location

New York, United States

Salary

Salary Not specified

Date posted


Feb 26, 2024



Apply on website 

View more 

 Save

 Send job

Job Description Example

Job Description:

The D'Aniello Institute for Veterans and Military Families' (IVMF) research and analytics (R&A) team is seeking an experienced data warehouse / business intelligence developer (DW/BI) interested in joining our rapidly growing team. The DW/BI developer will work collaboratively with subject matter experts across the IVMF to provide enterprise-wide expertise in data management and analytics. This position would establish data management processes, develop data integrations, implement data access and security protocols, and develop dashboards and reports. The DW/BI developer will design, build, and develop data integration solutions. Responsibilities would include supporting, maintaining, improving existing data solutions, coding new ETL interfaces, and developing and maintaining data visualizations. This position requires initiative, independence, drive, and the ability to work successfully in a fast-paced, dynamic environment with the ability to perform a wide variety of responsibilities. Strong communication skills are an asset that will be helpful while working closely with various teams across the institute.

Job Description Example

Job Specific Qualifications:

- Expertise in developing legible and well-structured scripts and queries in SQL.
- Experience working with large data repositories and/or relational databases.
- Experience as an ETL developer with strong data architecture knowledge around data warehousing concepts, SQL development and optimization, and operational support models.
- Knowledge of SAS Visual Analytics or Tableau would be a plus.
- Strong aptitude to learn and adapt to technology.
- Ability to work on multiple projects simultaneously with minimal supervision.
- Highly self-motivated and directed with keen attention to detail.
- Candidates must possess excellent written and verbal communication skills.
- Positive customer-service mindset.
- Ability to work collaboratively in dynamic team environment.

Operational Sources (OLTP's)

- Operational computer systems did provide information to run day-to-day operations, and answer's daily questions, but...
- Also, called **online transactional processing system** (OLTP)
- Data is read or manipulated with **each** transaction
- Transactions/queries are simple, and easy to write
- Usually for middle management
- Examples
 - Sales systems
 - Hotel reservation systems
 - Campus management system
 - HRM Applications etc.

Basics: Database concepts

- A database is a shared collection of logically related data, and a description of that data, designed to meet the information needs and support the activities of an organization.
- A database is deployed on a **database management system (DBMS)**, which is a software system used to define, create, manipulate, and administer a database.

Basics: Database concepts

- Typically, DB design is a complex undertaking having four phases
 1. Requirements specification
 2. Conceptual design (ERD)
 - Top-down design (difficult and expensive for large DBs and inexperienced developers)
 - Bottom-up design (schema for each group of users and integrated later)
 3. Logical design (relational model)
 4. Physical design

Conceptual DB design

- Entity-relationship (ER) model is one of the most often used conceptual design
- **Entity types** are used to represent a set of real-world objects of interest
 - Employee, order, customer
 - An object belonging to an entity type is called **entity** or **instance**
 - The set of instances of an entity type is called its population
- Entities have **attributes**

Conceptual DB design

- Entities have attributes
- There are several types of attributes
 - Mono-valued attribute
 - Multivalued attribute
 - Simple (non-decomposed)
 - Composite
 - Derived
 - Non-derived

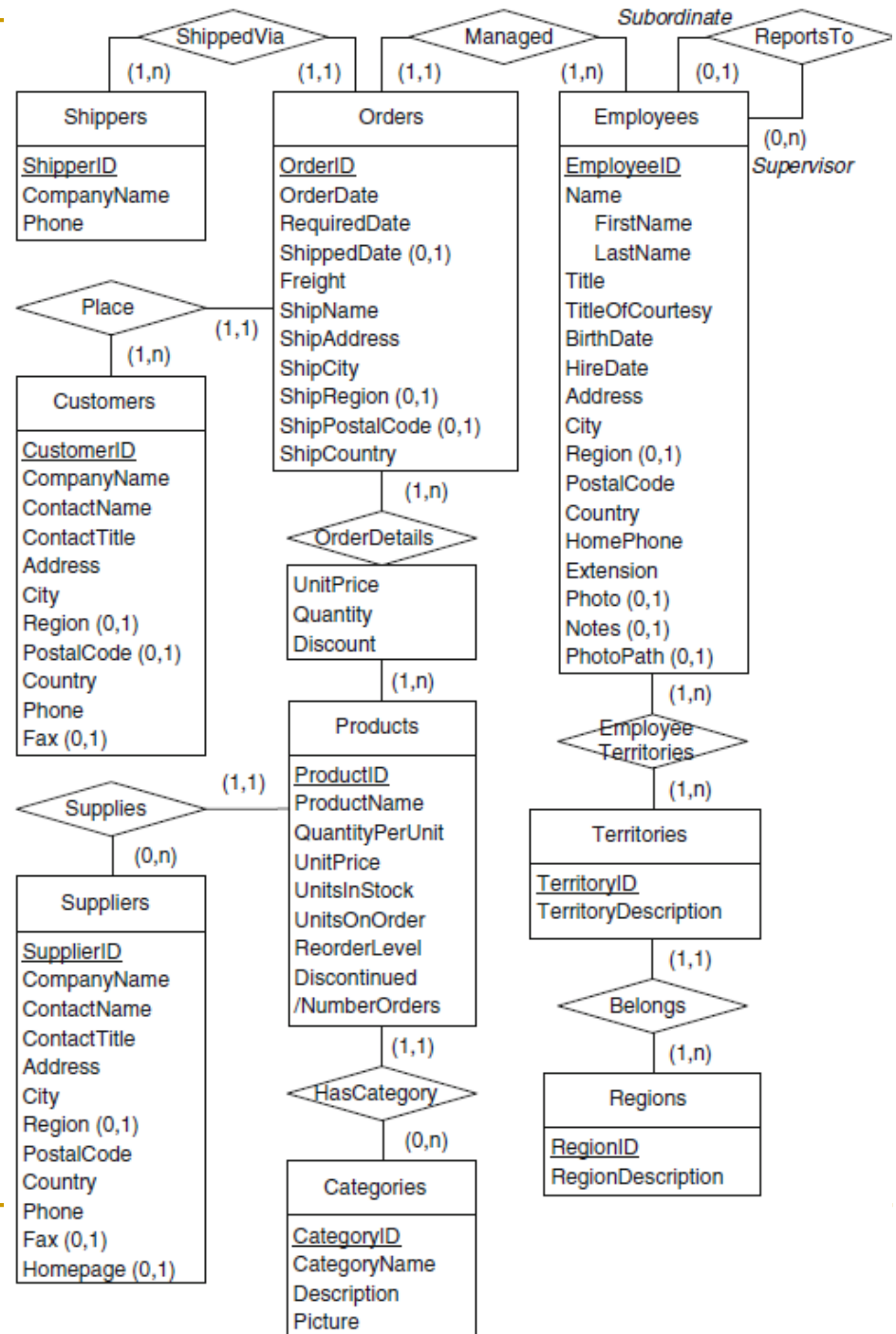
Conceptual DB design

- Relationship types are used to represent these associations between objects
- An association between objects of a relationship type is called a **relationship** or an **instance**.
- Each entity has associated with it a pair of **cardinalities** that describe minimum and maximum number of items
- An entity role is **Optional** or **Mandatory** depending on the minimum cardinality is 0 or 1

Conceptual DB design

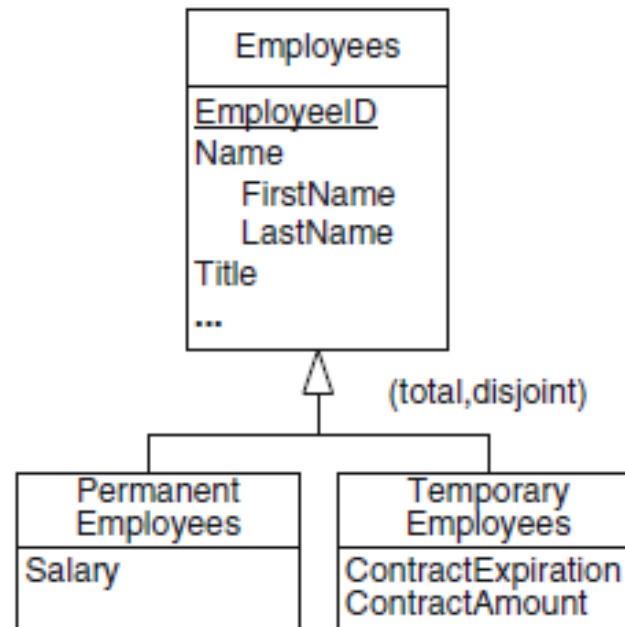
- Based on optional or mandatory, a role is said to be monovalued or multivalued
- In case same entity type is related more than once in a relationship type, the relationship type is called **recursive**

ERD Example



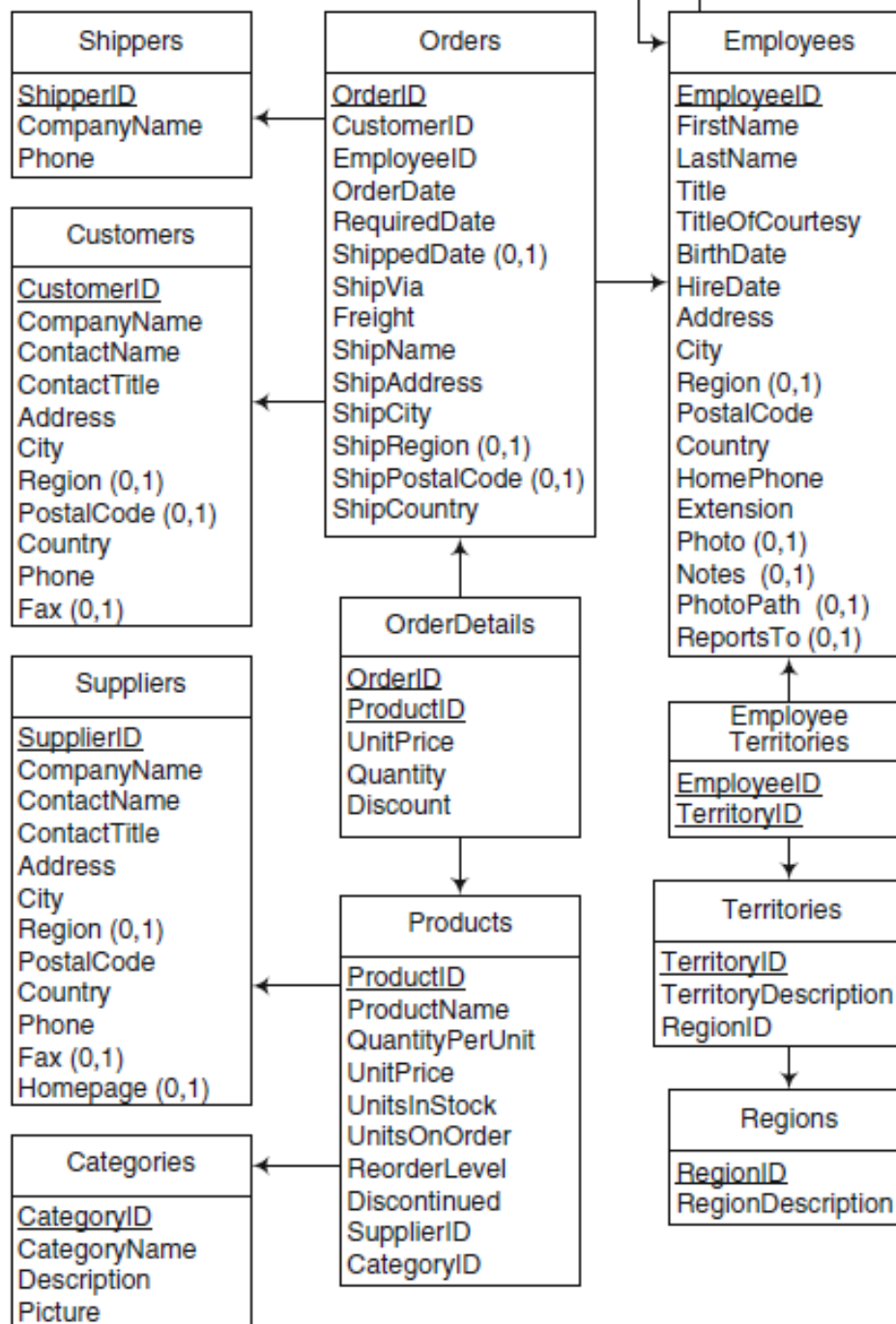
Conceptual DB design

- The **generalization** (or **is-a**) relationship captures the complexity of conceptualizing real world
 - **Supertype** and **subtype** means both types represent the same concept at different level of detail

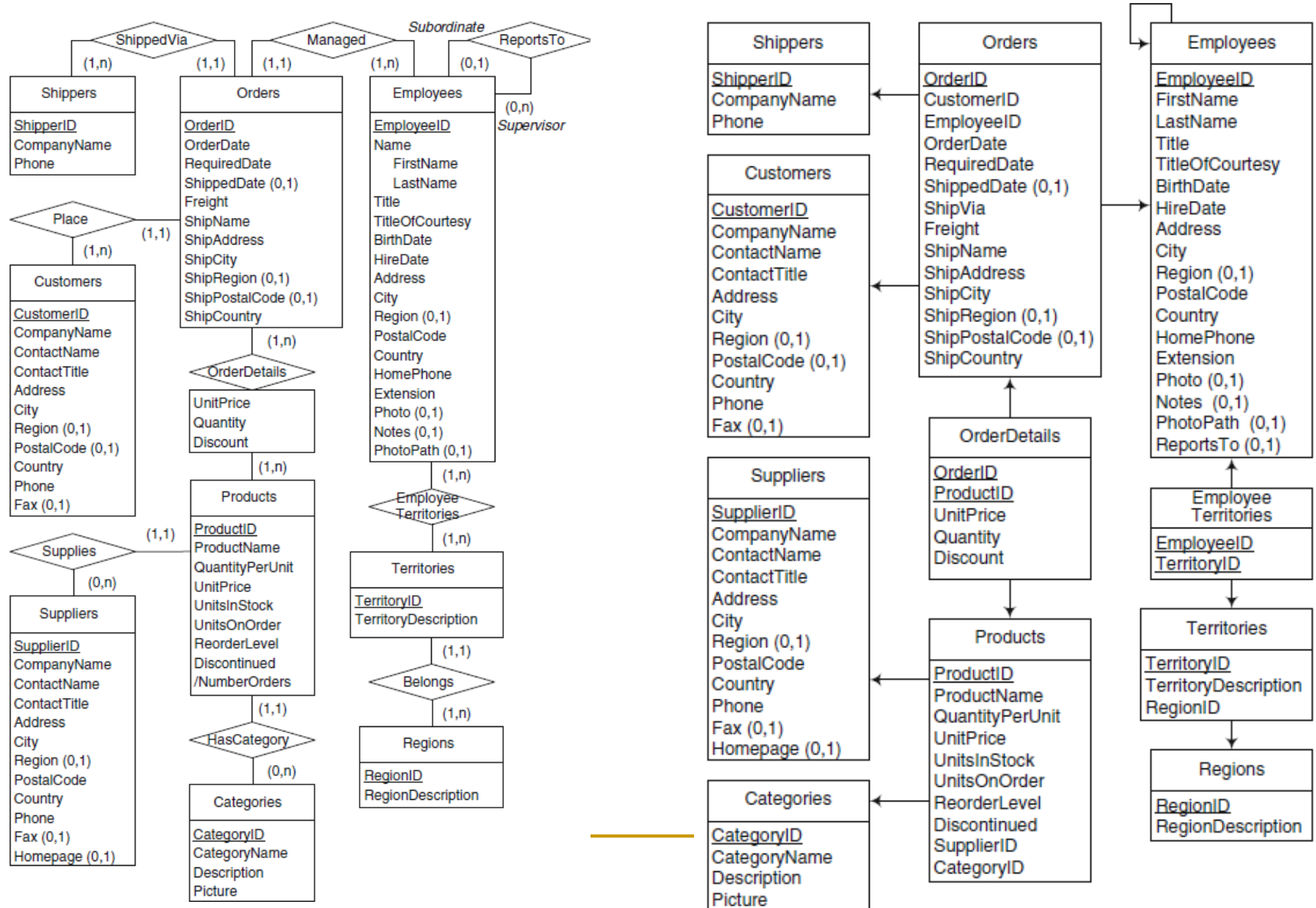


Logical DB design

- The **relational model** has a simple data structure, a **relation** (or table) composed of one or several **attributes** (or columns).
- Properties of a **relation**
- A **relational** schema describes the structure of a set of relations.



Conceptual vs Logical



Logical DB design

- Keys of
 - Primary key
 - Foreign key
- Referential integrity

DW and BI

Business Intelligence

- **Business intelligence** “a set of techniques and tools for the acquisition and transformation of raw data into meaningful and useful information for business analysis and purposes”
- **Business intelligence** comprises a collection of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information for decision making.

Business Intelligence

- **Business intelligence** and **decision-support systems** provide assistance to managers at various organizational levels for analyzing strategic information.
- **BI & DSS** systems collect vast amounts of data and reduce them to a form that can be used to analyze organizational behavior.

Business Intelligence

- This **data transformation** comprises a set of tasks that take the data from sources, through extraction, transformation, integration, and cleansing processes, store the data in a common repository called a **data warehouse**

Business Intelligence

- From the early days of data warehousing, the typical mechanism for those tasks has been **Online Analytical Processing (OLAP)**.
- Data mining tools have also been used in 1990s to infer and extract interesting knowledge hidden in data warehouse
- The new paradigm is generally called **data analytics**

Path to BI consultant

- Software Engineer (DW/BI Developer)
 - Analyze DW/decision support business requirements.
 - Develop scripts using PL/SQL to perform DML functions
 - Dashboard and reports development

Path to BI consultant

- Senior SE (DW/BI Administrator)
 - ❑ Database design, indexing, performance monitoring, and object creation such as cursors, triggers, stored procedures
 - ❑ DW Management, DW creation, taking backups and restoration.
 - ❑ Process of Data Migration from modules
 - ❑ Develop routines for loading data in DWH.
 - ❑ Schedule database jobs for the extraction of data.

Path to BI consultant

- Principal DW Architect
 - ETL Process Flow Design & Development
 - ETL Mappings Design & Development
 - Data Migrations

Path to BI consultant

- **BI Manager**
 - Develop, prepare and execute strategies for BI processes
 - Identify critical success factors and mitigate risk
 - Data warehouse methodology, project planning,
 - Modeling complex BI projects
 - Setting and enforcing BI stands and architecture
- **BI Consultant**