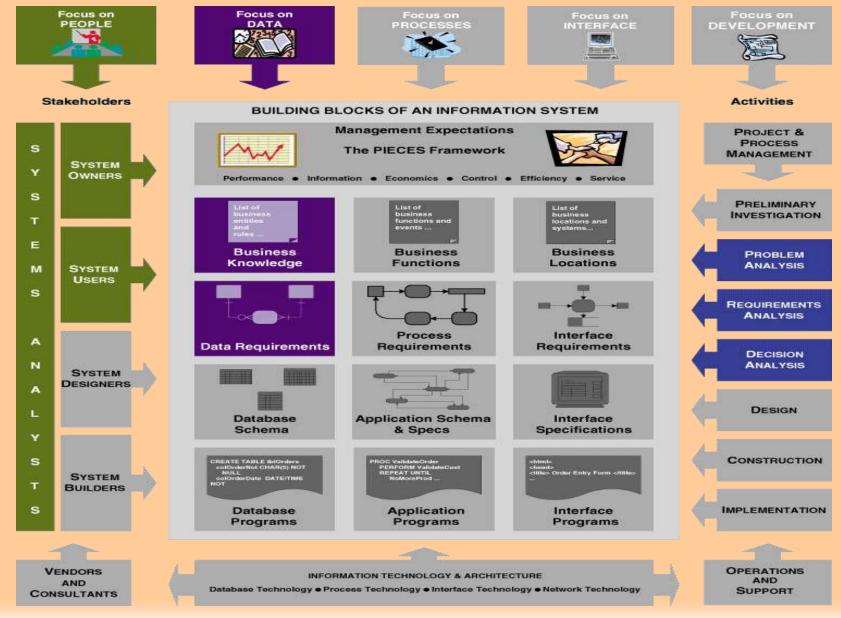**C H A P T E R**

7

DATA MODELING AND ANALYSIS

Chapter Seven   Data Modeling and Analysis

- Define systems modeling and differentiate between logical and physical system models.

- Define data modeling and explain its benefits.

- Recognize and understand the basic concepts and constructs of a data model.

- Read and interpret an entity relationship data model.

- Explain when data models are constructed during a project and where the models are stored.

- Discover entities and relationships.

- Construct an entity-relationship context diagram.

- Discover or invent keys for entities and construct a key-based diagram.

- Construct a fully attributed entity relationship diagram and describe all data structures and attributes to the repository or encyclopedia.

- Normalize a logical data model to remove impurities that can make a database unstable, inflexible, and nonscalable.

- Describe a useful tool for mapping data requirements to business operating locations.

Chapter Map

System Models

- A model is a representation of reality.
  - Logical models
  - Physical model

- Logical models show what a system is or does. They are implementation independent; that is, they depict the system independent of any technical implementation.

- Physical models show not only what a system is or does, but also how the system is physically and technically implemented.

Data Modeling
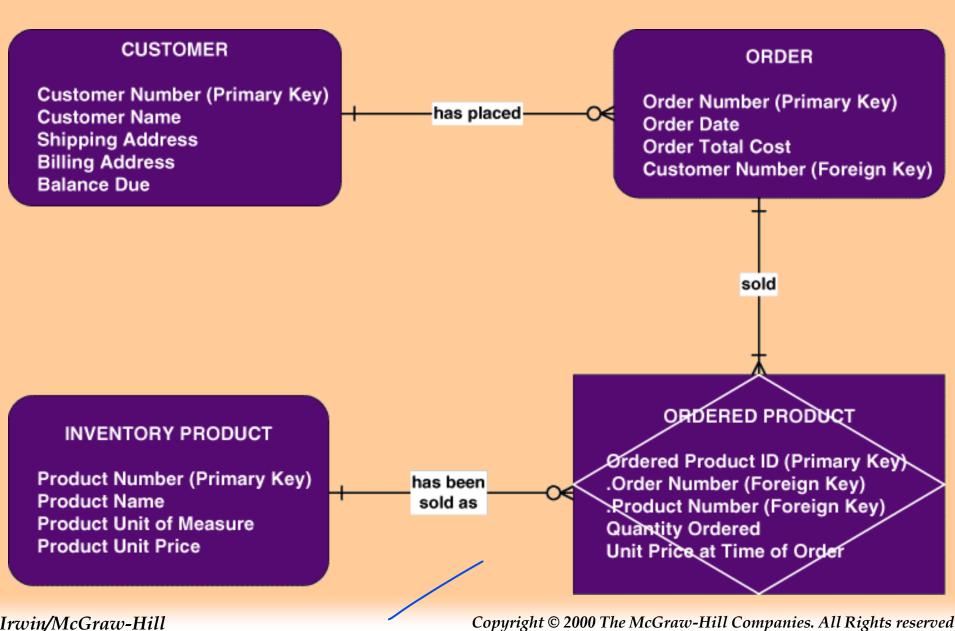
Data modeling is a technique for organizing and documenting a system's data.

Data modeling is sometimes called database modeling because a data model is eventually (finally) implemented as a database.

It is sometimes called information modeling.

The actual model is frequently called an entity relationship diagram (ERD) because it depicts data in terms of the entities and relationships described by the data.

Sample Entity Relationship Diagram (ERD)

CUSTOMER

Customer Number (Primary Key)
Customer Name
Shipping Address
Billing Address
Balance Due

─── has placed ───

ORDER

Order Number (Primary Key)
Order Date
Order Total Cost
Customer Number (Foreign Key)

sold

INVENTORY PRODUCT

Product Number (Primary Key)
Product Name
Product Unit of Measure
Product Unit Price

─── has been sold as ───

ORDERED PRODUCT

Ordered Product ID (Primary Key)
.Order Number (Foreign Key)
.Product Number (Foreign Key)
Quantity Ordered
Unit Price at Time of Order

Data Modeling Concepts: Entity

An entity is a class of persons, places, objects, events, or concepts about which we need to capture and store data.

Name of Entity

- Persons: agency, contractor, customer, department, division, employee, instructor, student, supplier.

- Places: sales region, building, room, branch office, campus.

- Objects: book, machine, part, product, raw material, software license, software package, tool, vehicle model, vehicle.

- Events: application, award, cancellation, class, flight, invoice, order, registration, renewal, requisition, reservation, sale, trip.

- Concepts: account, block of time, bond, course, fund, qualification, stock.

Data Modeling Concepts: Entity

An entity instance is a single occurrence of an entity.

Example: instances of the entity STUDENT may include

- Betty Arnold

- John Taylor

- Lisa Simmons

- Bill Macy

- Heather Leath

- Tim Wrench

Data Modeling Concepts: Attributes

An attribute is a descriptive property or characteristic of an entity. Synonyms include element, property, and field.

A compound attribute is one that actually consists of other attributes

STUDENT

Name
.Last Name
.First Name
.Middle Initial
Address
.Street Address
.City
.State or Province
.Country
.Postal Code
Phone Number
.Area Code
.Exchange Number
.Number Within Exchange
Date of Birth
Gender
Race
Major
Grade Point Average

Data Modeling Concepts: Domains

The data type for an attribute defines what type of data can be stored in that attribute.

The domain of an attribute defines what values an attribute can legitimately (legally) take on.

The default value for an attribute is the value that will be recorded if not specified by the user.

Data Modeling Concepts: Identification

A key is an attribute, or a group of attributes, that assumes a unique value for each entity instance.

A group of attributes that uniquely identifies an instance of an entity is called a concatenated key.

A candidate key is a "candidate to become the primary key" of instances of an entity.

A primary key is that candidate key that will most commonly be used to uniquely identify a single entity instance.

Any candidate key that is not selected to become the primary key is called an alternate key.

A subsetting criteria is an attribute (or concatenated attribute) whose finite values divide all entity instances into useful subsets.

Data Modeling Concepts: Identification Keys & Subsetting Criteria

STUDENT

Student Number (Primary Key)
Social Security Number (Alternate Key)
Name
.Last Name
.First Name
.Middle Initial
Address
.Street Address
.City
.State or Province
.Country
.Postal Code
Phone Number
.Area Code
.Exchange Number
.Number Within Exchange
Date of Birth
Gender (Subsetting Criteria 1)
Race (Subsetting Criteria 2)
Major (Subsetting Criteria 3)
Grade Point Average
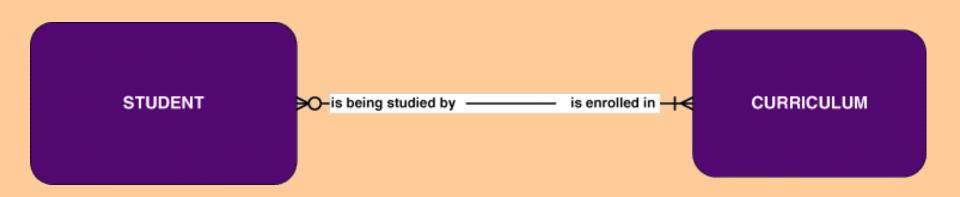
Race-jatigosthi

Data Modeling Concepts: Relationships

A relationship is a natural business association that exists between one or more entities. The relationship may represent an event that links the entities or merely a logical affinity (similarity) that exists between the entities.

STUDENT ⊃O— is being studied by ———— is enrolled in —⊢< CURRICULUM

Data Modeling Concepts: Cardinality

Cardinality defines the minimum and maximum number of occurrences of one entity that may be related to a single occurrence of the other entity.

Because all relationships are bidirectional, cardinality must be defined in both directions for every relationship.
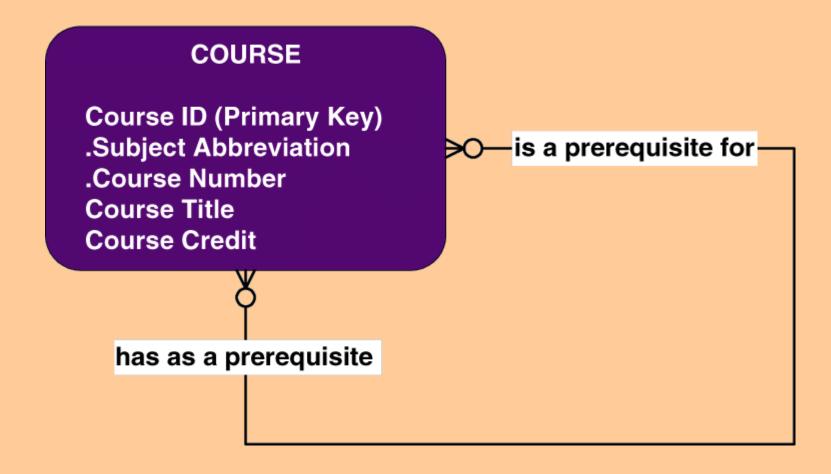
Data Modeling Concepts: Degree

The degree of a relationship is the number of entities that participate in the relationship.

Data Modeling Concepts: Degree

A recursive relationship is a relationship that exists between different instances of the same entity

COURSE

Course ID (Primary Key)
.Subject Abbreviation
.Course Number
Course Title
Course Credit

is a prerequisite for

has as a prerequisite

Data Modeling Concepts: Degree

PROJECT

Project Number (Primary Key)
Description
Projected Start Date
Projected End Date

EMPLOYEE

Employee ID (Primary Key)
Employee Name
.Last Name
.First Name
.Middle Initial

requires

is given

ASSIGNMENT

Assignment ID (Primary Key)
.Project Number
.Employee ID
.Location Number
Begin Date
End Date

offers

LOCATION

Location Number (Primary Key)
Address
.Street
.City
.State
.Zipcode

Relationships may exist between more than two entities and are called N-ary relationships.
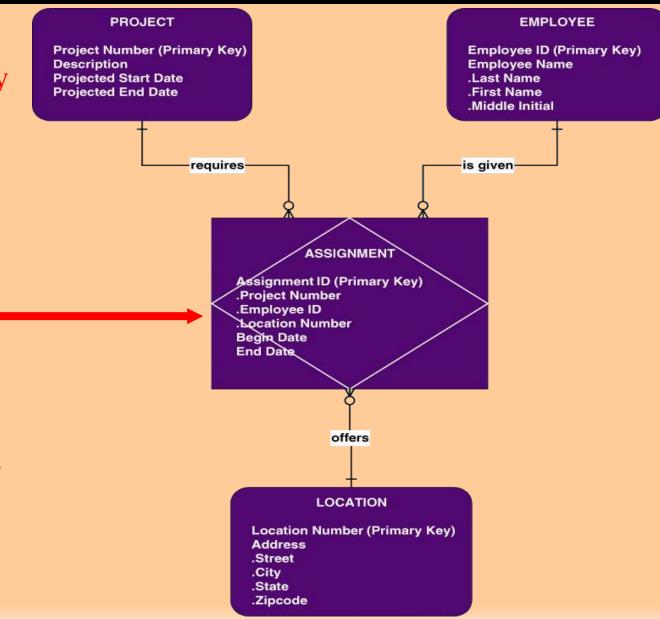
The example ERD depicts a ternary relationship.

Data Modeling Concepts: Degree
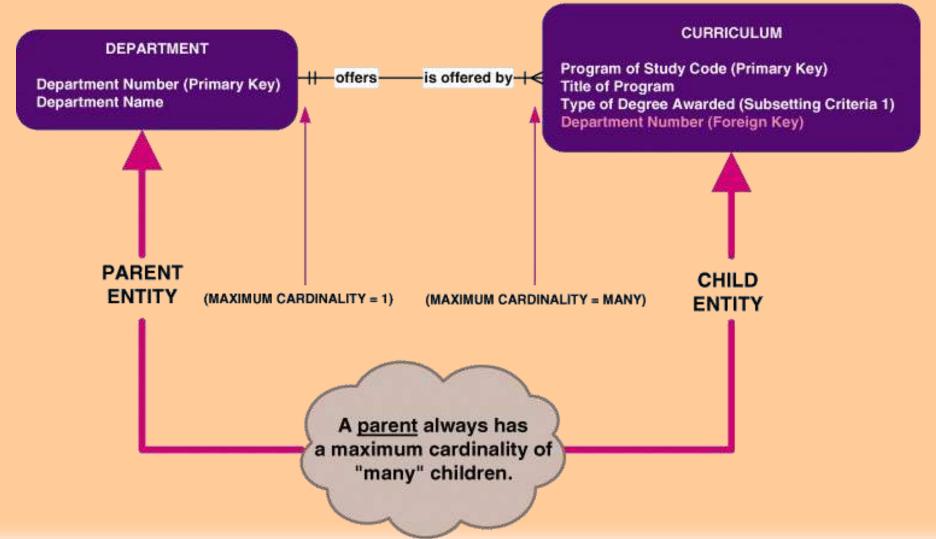
An associative entity is an entity that inherits its primary key from more than one other entity (called parents).

Each part of that concatenated key points to one and only one instance of each of the connecting entities.

PROJECT

Project Number (Primary Key)
Description
Projected Start Date
Projected End Date

EMPLOYEE

Employee ID (Primary Key)
Employee Name
.Last Name
.First Name
.Middle Initial

requires

is given

ASSIGNMENT

Assignment ID (Primary Key)
.Project Number
.Employee ID
.Location Number
Begin Date
End Date

offers

LOCATION

Location Number (Primary Key)
Address
.Street
.City
.State
.Zipcode

Data Modeling Concepts: Foreign Keys

(a) A foreign key is a primary key of one entity that is contributed to (duplicated in) another entity to identify instances of a relationship.

DEPARTMENT

Department Number (Primary Key)
Department Name

╫── offers ──── is offered by ──┤<

CURRICULUM

Program of Study Code (Primary Key)
Title of Program
Type of Degree Awarded (Subsetting Criteria 1)
Department Number (Foreign Key)

PARENT ENTITY

(MAXIMUM CARDINALITY = 1)

(MAXIMUM CARDINALITY = MANY)

CHILD ENTITY

A <u>parent</u> always has a maximum cardinality of "many" children.
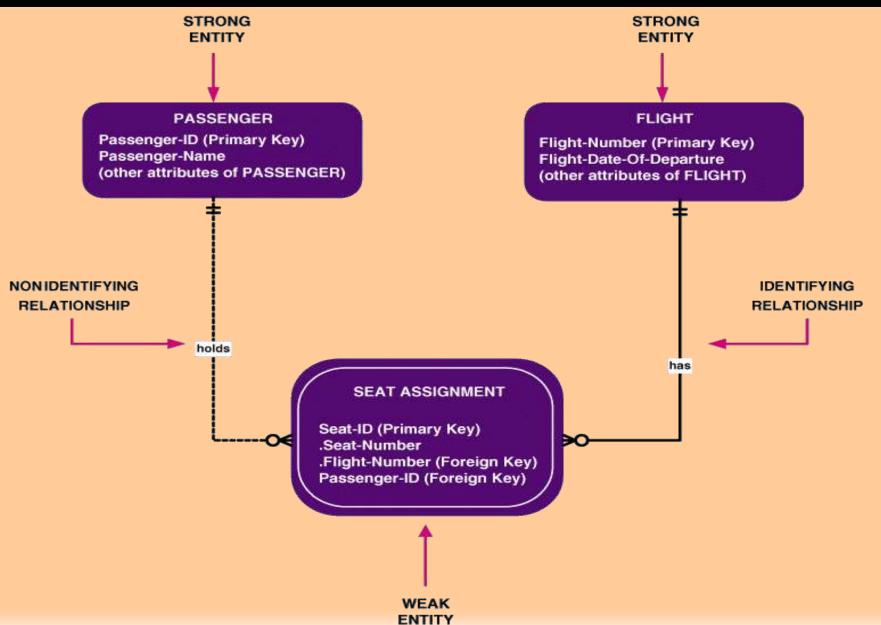
Data Modeling Concepts: Foreign Keys

Nonidentifying relationships are those in which each of the participating entities has its own independent primary key, In other words, none of the primary key attributes is shared.

Identifying relationships are those in which the parent entity contributes its primary key to become part of the primary key of the child entity.

(b)

BUILDING

Building Name (Primary Key)

⊩ contains — is located in ⊫◁ **ROOM**

Room ID (Primary Key)
.Building Name (Foreign Key)
.Room Number

Identifying relationships have the foreign key participating as part of the child entity's *primary key*.

Data Modeling Concepts: Foreign Keys

STRONG ENTITY

STRONG ENTITY

PASSENGER

Passenger-ID (Primary Key)
Passenger-Name
(other attributes of PASSENGER)

FLIGHT

Flight-Number (Primary Key)
Flight-Date-Of-Departure
(other attributes of FLIGHT)

NONIDENTIFYING RELATIONSHIP

IDENTIFYING RELATIONSHIP

holds

has

SEAT ASSIGNMENT

Seat-ID (Primary Key)
.Seat-Number
.Flight-Number (Foreign Key)
Passenger-ID (Foreign Key)

WEAK ENTITY

Data Modeling Concepts: Generalization

Generalization is a technique wherein the attributes that are common to several types of an entity are grouped into their own entity, called a supertype.

An entity supertype is an entity whose instances store attributes that are common to one or more entity subtypes.

An entity subtype is an entity whose instances inherit some common attributes from an entity supertype and then add other attributes that are unique to an instance of the subtype.

The Key-based Data Model

MEMBER ORDER
- Primary Key
- Order-Number   [PK1]

sells

is a

MEMBER ORDERED PRODUCT
- Primary Key
- Order-Number   [PK1] [FK]
- Product-Number   [PK2] [FK]

responds to

places

MEMBER
- Primary Key
- Member-Number   [PK1]

has conducted

sold as

generates

PRODUCT
- Primary Key
- Product-Number   [PK1]

TRANSACTION
- Primary Key
- Transaction-Reference-Number   [PK1]

binds

is featured as

AGREEMENT
- Primary Key
- Agreement-Number   [PK1]

TITLE PROMOTION
- Primary Key
- Product-Number   [PK1] [FK]
- Promotion-Number   [PK2] [FK]

features

PROMOTION
- Primary Key
- Promotion-Number   [PK1]

The Key-based Data Model With Generalization

MEMBER ORDER
- Primary Key
- Order-Number [PK1]

sells

is a

responds to

MEMBER ORDERED PRODUCT
- Primary Key
- Order-Number [PK1] [FK]
- Product-Number [PK2] [FK]

places

has conducted

MEMBER
- Primary Key
- Member-Number [PK1]

TRANSACTION
- Primary Key
- Transaction-Reference-Number [PK1]

sold as

binds

AGREEMENT
- Primary Key
- Agreement-Number [PK1]

PRODUCT
- Primary Key
- Product-Number [PK1]

generates

PROMOTION
- Primary Key
- Promotion-Number [PK1]

is a

is a

MERCHANDISE
- Primary Key
- Product-Number [PK1] [FK]

TITLE
- Primary Key
- Product-Number [PK1] [FK]

is featured as

PROMOTION TITLE
- Primary Key
- Promotion-Number [PK1] [FK]
- Product-Number [PK2] [FK]

features

is a

is a

is a

AUDIO TITLE
- Primary Key
- Product-Number [PK1] [FK]

VIDEO TITLE
- Primary Key
- Product-Number [PK1] [FK]

GAME TITLE
- Primary Key
- Product-Number [PK1] [FK]

The Fully-Attributed Data Model

MEMBER ORDER
- Primary Key
  - Order-Number [PK1]
- Non-Key Attributes
  - Order-Creation-Date
  - Order-Fill-Date
  - Member-Name
  - Member-Address
  - Shipping-Address-Name
  - Shipping-Street-Address
  - Shipping-City
  - Shipping-State
  - Shipping-Zip
  - Shipping-Instructions
  - Order-Sub-Total
  - Order-Sales-Tax
  - Order-Shipping-Method
  - Order-Shipping-&-Handling-Cost
  - Order-Status
  - Order-Prepaid-Amount
  - Order-Prepayment-Method
  - Promotion-Number [FK]
  - Member-Number [FK]
  - Member-Number-1 - Member-Number [FK]
  - Ordered-Product-Number
  - Quantity-Shipped
  - Quantity-Backordered
  - Ordered-Product-Description
  - Ordered-Product-Title
  - Purchased-Unit-Price
  - Extended-Price

MEMBER ORDERED PRODUCT
- Primary Key
  - Order-Number [PK1] [FK]
  - Product-Number [PK2] [FK]

MEMBER
- Primary Key
  - Member-Number [PK1]
- Non-Key Attributes
  - Member-Name
  - Member-Status
  - Member-Street-Address
  - Member-Post-Office-Box
  - Member-City
  - Member-State
  - Member-Zip-Code
  - Member-Daytime-Phone-Number
  - Member-Date-of-Last-Order
  - Member-Balance-Due
  - Member-Credit-Card-Type
  - Member-Credit-Card-Number
  - Member-Credit-Card-Expire-Date
  - Member-Bonus-Balance-Available
  - Audio-Category-Preference
  - Audio-Media-Preference
  - Date-Enrolled
  - Expiration-Date
  - Game-Category-Preference
  - Game-Media-Preference
  - Number-of-Credits-Earned
  - Video-Category-Preference
  - Video-Media-Preference
  - Agreement-Number [FK]
  - Privacy-Code
  - Email-Address

TRANSACTION
- Primary Key
  - Transaction-Reference-Number [PK1]
- Non-Key Attributes
  - Transaction-Date
  - Transaction-Type
  - Transaction-Description
  - Transaction-Amount
  - Member-Number [FK]
  - Order-Number [FK]

PRODUCT
- Primary Key
  - Product-Number [PK1]
- Non-Key Attributes
  - "Universal-Product-Code (Alternate Key)"
  - Quantity-in-Stock
  - Product-Type
  - Suggested-Retail-Price
  - Default-Unit-Price
  - Current-Special-Unit-Price
  - Current-Month-Units-Sold
  - Current-Year-Units-Sold
  - Total-Lifetime-Units-Sold

AGREEMENT
- Primary Key
  - Agreement-Number [PK1]
- Non-Key Attributes
  - Agreement-Expire-Date
  - Agreement-Active-Date
  - Fulfillment-Period
  - Required-Number-of-Credits

MERCHANDISE
- Primary Key
  - Product-Number [PK1] [FK]
- Non-Key Attributes
  - Merchandise-Name
  - Merchandise-Description
  - Merchandise-Type
  - Unit-of-Measure

TITLE
- Primary Key
  - Product-Number [PK1] [FK]
- Non-Key Attributes
  - Title-of-Work
  - Title-Cover
  - Catalog-Description
  - Copyright-Date
  - Entertainment-Category
  - Credit-Value

PROMOTION
- Primary Key
  - Promotion-Number [PK1]
- Non-Key Attributes
  - Promotion-Release-Date
  - Promotion-Status
  - Promotion-Type
  - Product-Number
  - Title-of-Work

TITLE PROMOTION
- Primary Key
  - Promotion-Number [PK1] [FK]
  - Product-Number [PK2] [FK]

AUDIO TITLE
- Primary Key
  - Product-Number [PK1] [FK]
- Non-Key Attributes
  - Artist
  - Audio-Category
  - Audio-Sub-Category
  - Number-of-Units-in-Package
  - Audio-Media-Code
  - Content-Advisory-Code

VIDEO TITLE
- Primary Key
  - Product-Number [PK1] [FK]
- Non-Key Attributes
  - Producer
  - Director
  - Video-Category
  - Video-Sub-Category
  - Closed-Captioned
  - Language
  - Running-Time
  - Video-Media-Type
  - Video-Encoding
  - Screen-Aspect
  - MPA-Rating-Code

GAME TITLE
- Primary Key
  - Product-Number [PK1] [FK]
- Non-Key Attributes
  - Manufacturer
  - Game-Category
  - Game-Sub-Category
  - Game-Platform
  - Game-Media-Type
  - Number-of-Players
  - Parent-Advisory-Code

Relationships: sells, sold as, responds to, places, is a, has conducted, binds, generates, is featured as, features, is a (MERCHANDISE/TITLE), is a (AUDIO TITLE/VIDEO TITLE/GAME TITLE)