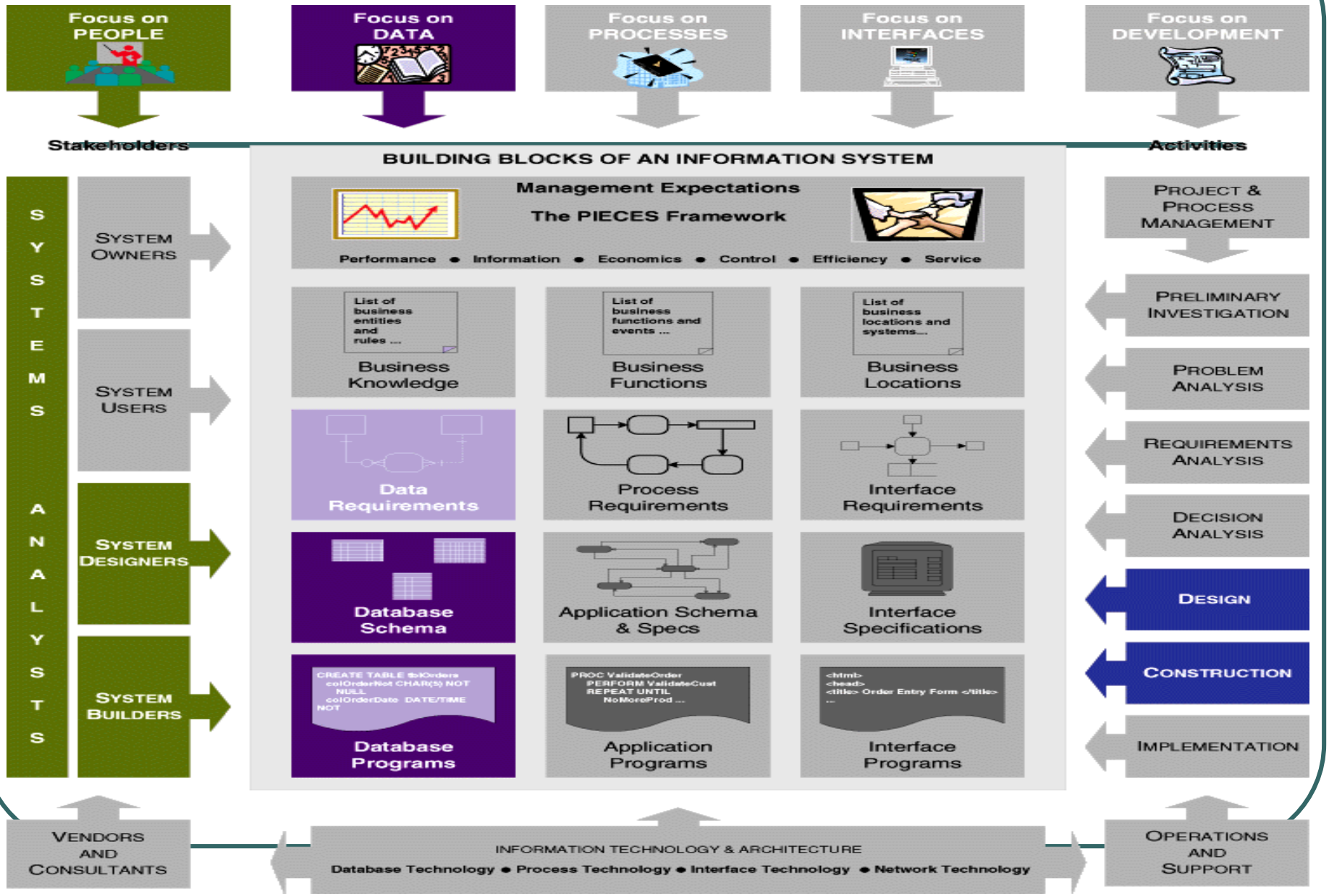
CHAPTER

12

DATABASE DESIGN

Chapter Map



Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Fields

A **field** is the physical implementation of a data attribute. They are the smallest unit of meaningful data.

A **primary key** is a field whose values identify one and only one record in a file.

A **secondary key** is an alternate identifier for a record.

A **foreign key** is a pointer to a record in a different file.

A **descriptive field** is any other (nonkey) field that stores business data.

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Records

A **record** is a collection of fields arranged in a predefined format.

- Fixed-length record structures
- Variable-length record structures

A **blocking factor** is the number of logical records included in a single read or write operation (from the computer's perspective).

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Files

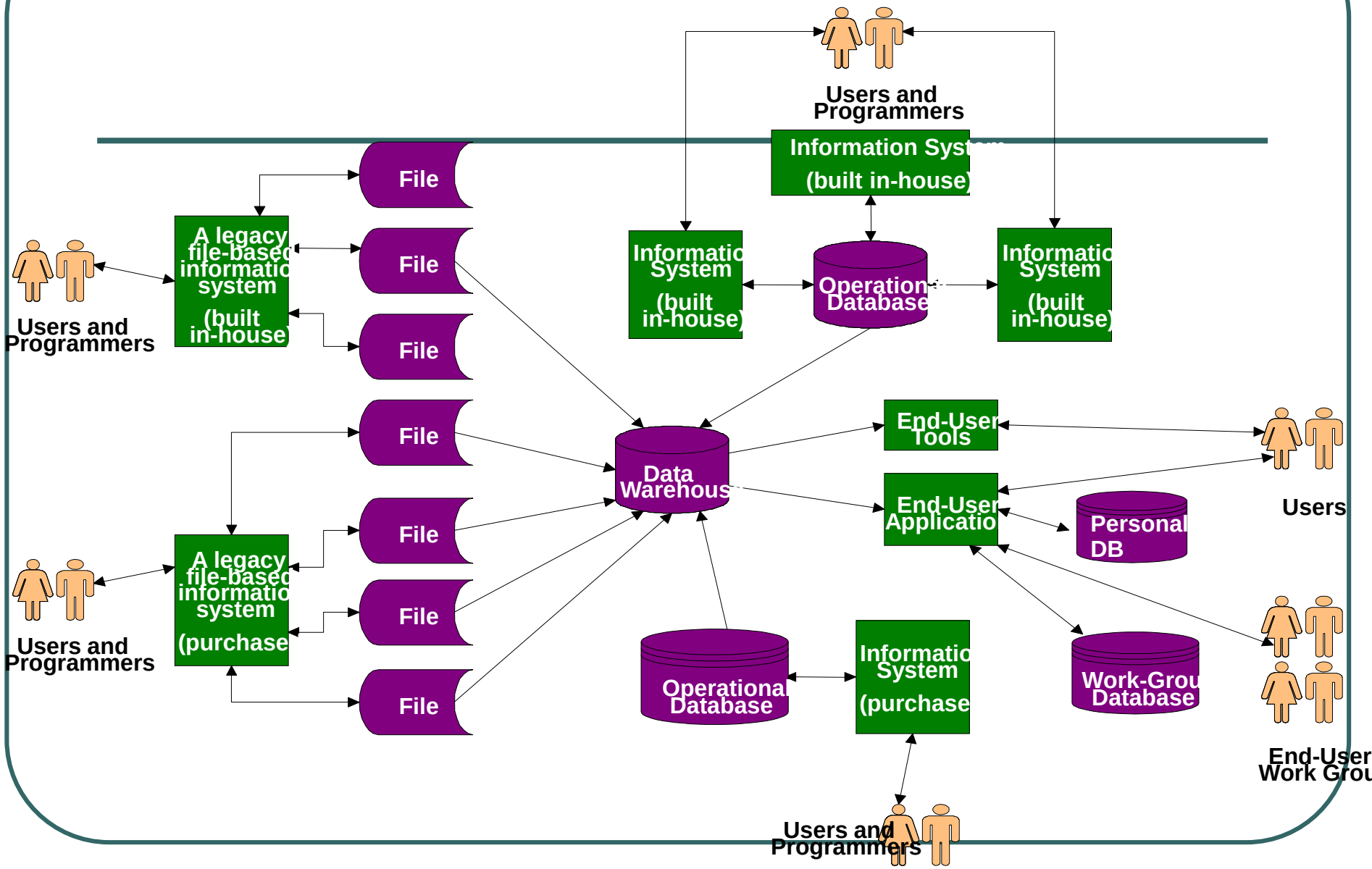
A **file** is the set of all occurrences of a given record structure.

A **table** is the relational database equivalent of a file.

- Types
  - Master files
  - Transaction files
  - Document files
  - Archival files
  - Table lookup files
  - Audit files
- File organization
- File access

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

A Modern Data Architecture



Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Administrators

A **data administrator** is responsible for the data planning, definition, architecture, and management.

One or more **database administrators** are responsible for the database technology, database design and construction, security, backup and recovery, and performance tuning.

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Database Architecture

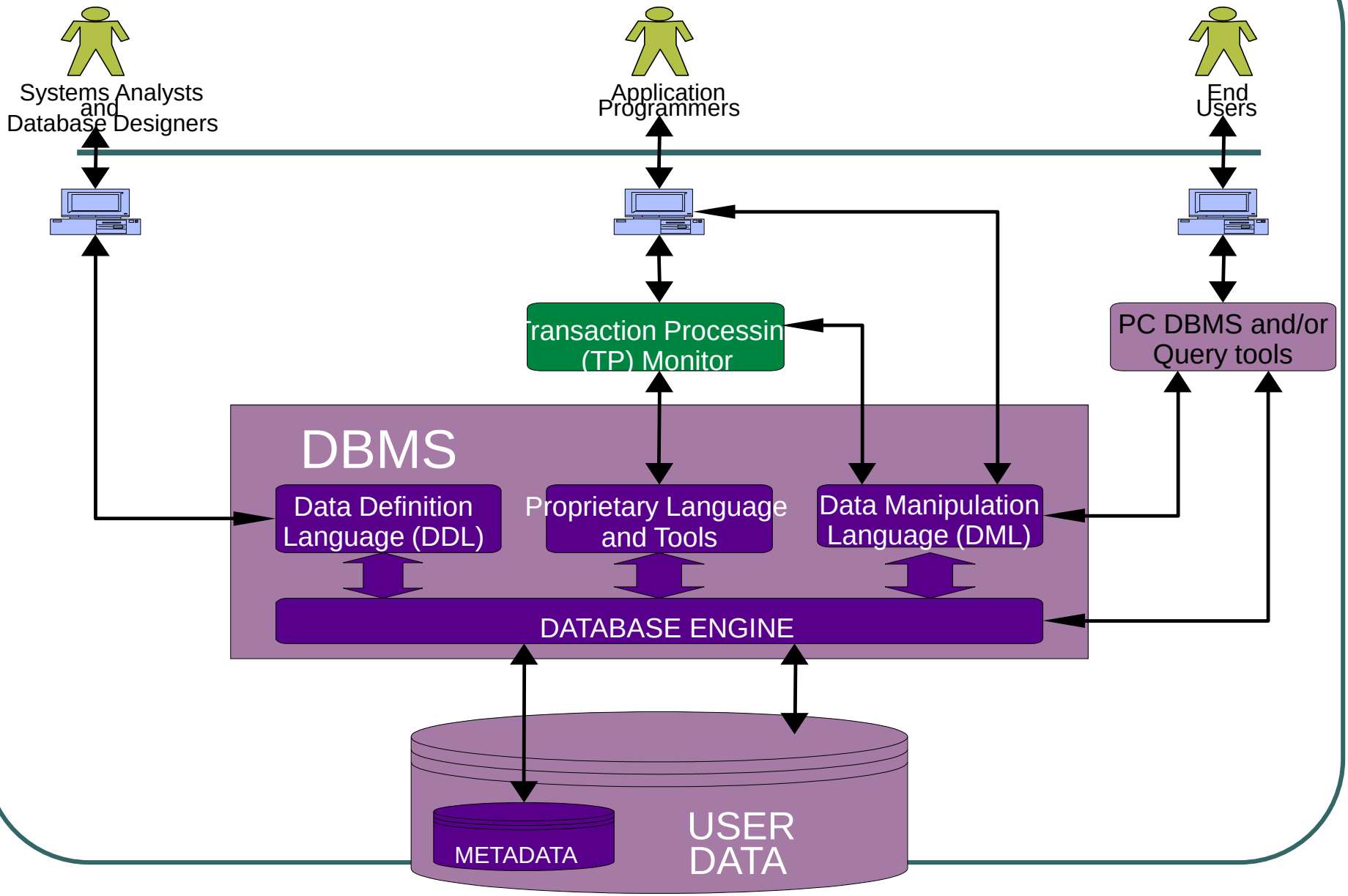
Database architecture refers to the database technology including the database engine, database utilities, CASE tools, and database development tools.

A **database management system (DBMS)** is specialized software that is used to create, access, control, and manage the database. The core of the DBMS is a **database engine**.

- A data definition language (DDL) is that part of the engine used to physically define tables, fields, and structural relationships.
- A data manipulation language (DML) is that part of the engine used to create, read, update, and delete records in the database, and navigate between different files (tables) in the database.

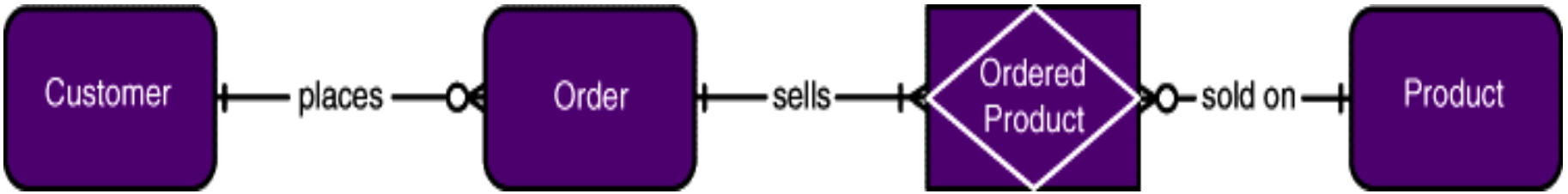
Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Typical DBMS Architecture



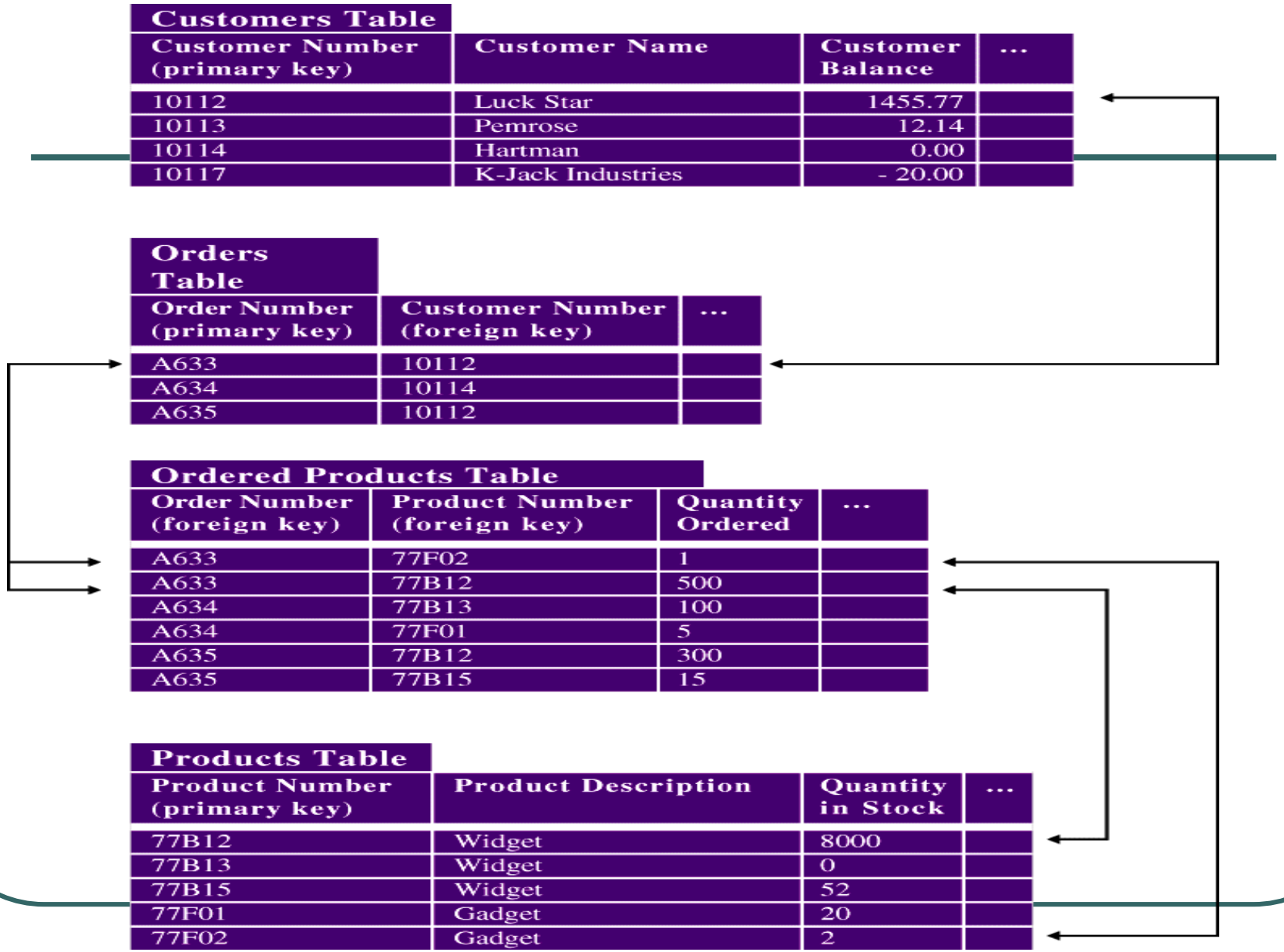
Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Logical Data Model



Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Physical Data Model (Relational Schema)



Customers Table

Customer Number (primary key)	Customer Name	Customer Balance	...
10112	Luck Star	1455.77	
10113	Pemrose	12.14	
10114	Hartman	0.00	
10117	K-Jack Industries	- 20.00	

Orders Table

Order Number (primary key)	Customer Number (foreign key)	...
A633	10112	
A634	10114	
A635	10112	

Ordered Products Table

Order Number (foreign key)	Product Number (foreign key)	Quantity Ordered	...
A633	77F02	1	
A633	77B12	500	
A634	77B13	100	
A634	77F01	5	
A635	77B12	300	
A635	77B15	15	

Products Table

Product Number (primary key)	Product Description	Quantity in Stock	...
77B12	Widget	8000	
77B13	Widget	0	
77B15	Widget	52	
77F01	Gadget	20	
77F02	Gadget	2	

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Goals of Database Design

- A database should provide for efficient storage, update, and retrieval of data.
- A database should be reliable—the stored data should have high integrity and promote user trust in that data.
- A database should be adaptable and scalable to new and unforeseen requirements and applications.

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Sample Physical Data Types

Logical Data Type to be stored in field)	**Physical Data Type Microsoft *Access***	**Physical Data Type Microsoft *SQL Server***	**Physicall Data Type Oracle**
Fixed length character data *(use for fields with relatively fixed length character data)*	TEXT	CHAR (size) *or* character (size)	CHAR (size)
Variable length character data *(use for fields that require character data but for which size varies greatly--such as ADDRESS)*	TEXT	VARCHAR (max size) *or* character varying (max size)	VARCHAR (max size)
Very long character data *(use for long descriptions and notes--usually no more than one such field per record)*	MEMO	TEXT	LONG VARCHAR *or* LONG VARCHAR2
Integer number	NUMBER	INT (size) *or* integer *or* smallinteger *or* tinuinteger	INTEGER (size) *or* NUMBER (size)
Decimal number	NUMER	DECIMAL (size, decimal places) *or* NUMERIC (size, decimal places)	DECIMAL (size, decimal places) *or* NUMERIC (size, decimal places) *or* NUMBER

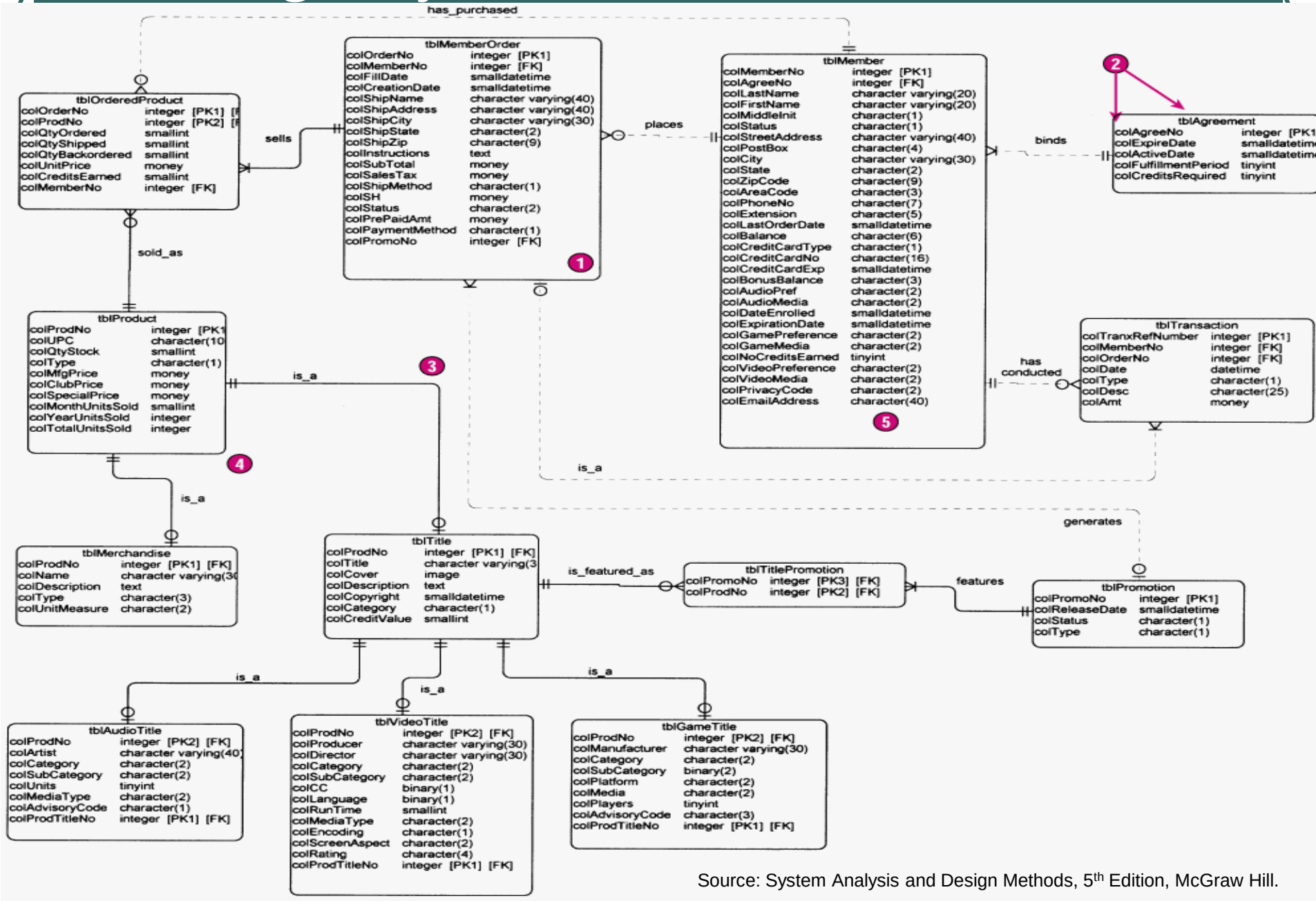
Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Sample Physical Data Types (concluded)

Logical Data Type to be stored in field)	**Physical Data Type Microsoft *Access***	**Physical Data Type Microsoft *SQL Server***	**Physicall Data Type Oracle**
Financial Number	CURRENCY	MONEY	*see decimal number*
Date (with time)	DATE/TIME	DATETIME *or* SMALLDATETIME *Depending on precision needed*	DATE
Current time *(use to store the data and time from the computer's system clock)*	*not supported*	TIMESTAMP	*not supported*
Yes or No; *or* True or False	YES/NO	BIT	*use CHAR(1) and set a yes or no domain*
Image	OLE OBJECT	IMAGE	LONGRAW
Hyperlink	HYPERLINK	VARBINARY	RAW
Can designer define new data types?	*NO*	*YES*	*YES*

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

SoundStage Physical Database Schema



Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

A Method for Database Design

- Review the logical data model.
- Create a table for each entity.
- Create fields for each attribute.
- Create an index for each primary and secondary key.
- Create an index for each subsetting criterion.
- Designate foreign keys for relationships.
- Define data types, sizes, null settings, domains, and defaults for each attribute.
- Create or combine tables to implement supertype/ subtype structures.
- Evaluate and specify referential integrity constraints.

Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

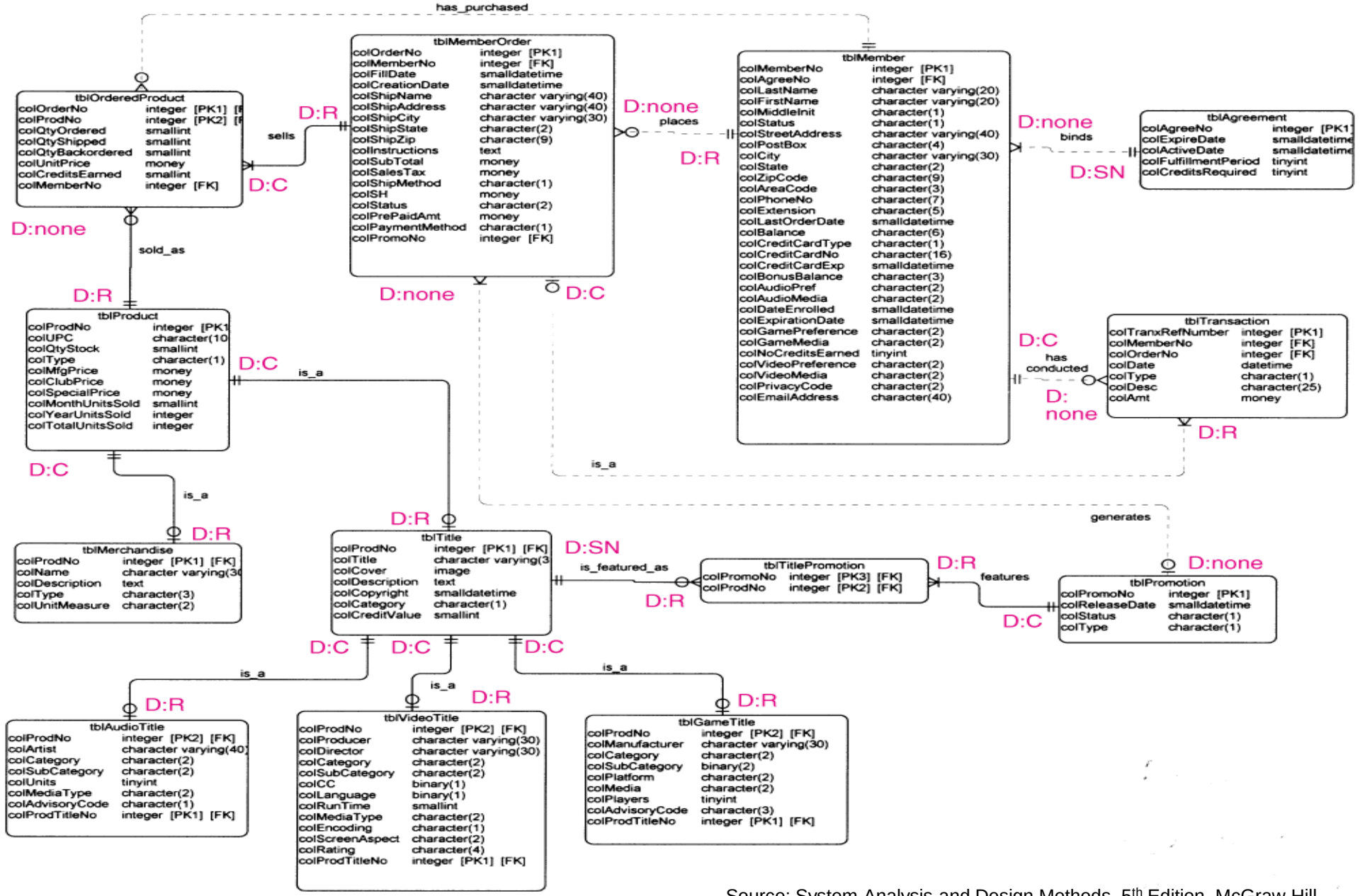
Database Integrity

- Key integrity
- Domain integrity
- Referential integrity

A **referential integrity** error exists when a foreign key value in one table has no matching primary key value in the related table.

- No restriction
- Delete: cascade
- Delete: restrict
- Delete: set null

SoundStage Referential Integrity Constraints



Source: System Analysis and Design Methods, 5th Edition, McGraw Hill.

Database Distribution and Replication

Data distribution analysis establishes which business locations need access to which logical data entities and attributes.

- The analysis drives distribution decisions:
  - Centralization
  - Horizontal distribution (also called partitioning)
  - Vertical distribution (also called partitioning)
  - Replication

Summary

- Conventional files and modern, relational databases.
- Fields, records, files, and databases.
- Modern data architecture, e.g. files, operational databases, data warehouses, personal databases, and work group databases.
- Roles of systems analysts, data administrators, and database administrators
- Architecture of a database management system
- Relational database: entities, attributes, and relationships from a logical data model.
- Transform a logical data model into a physical, relational database schema.
- Generate SQL code to create the database structure in a schema.