

Databases

Databases, Part I

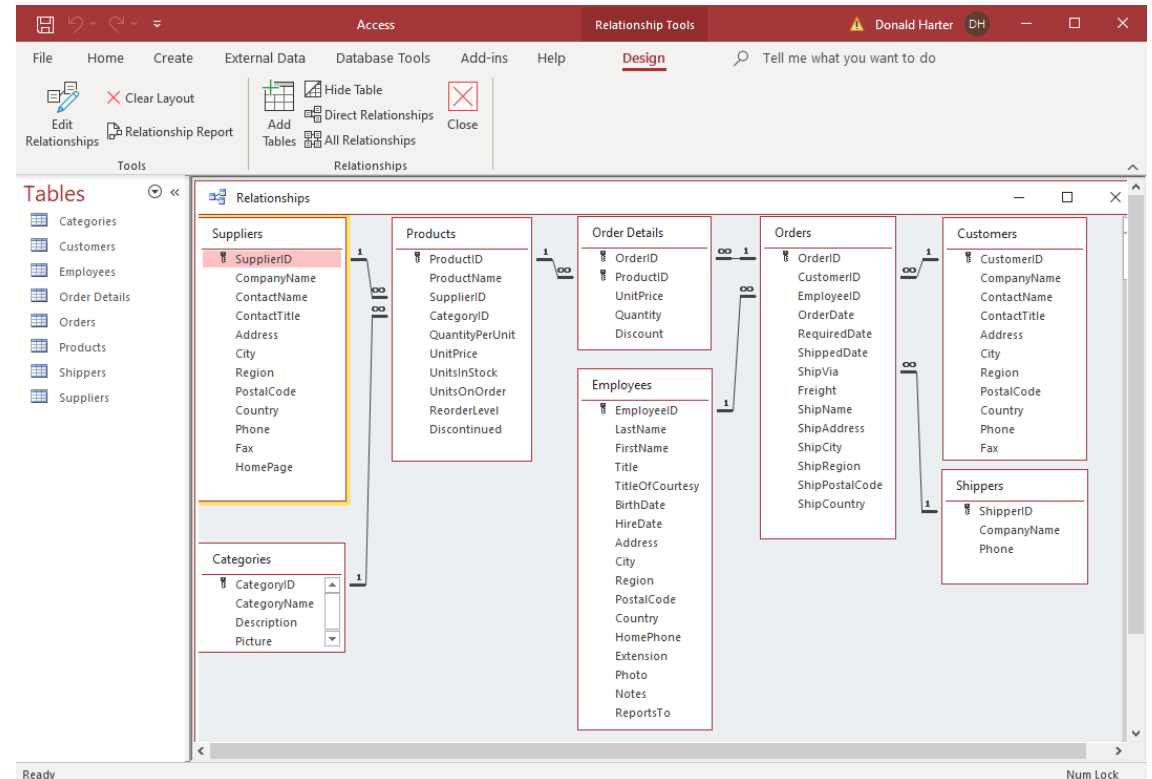
- A database is a collection of tables
 - Each table is somewhat like a spreadsheet in Excel.
 - A column in a database is called a field.
 - A row in a database is called a record.
 - Key fields uniquely identify each row (record) in a database.
- Relationships
 - Tables are related through common fields between two tables.
 - This relationship is called a join.
 - Joins can include inner join, left join, right join, and outer join.

Databases (cont.)

Description

- Each box in the diagram is a table.
- Entries in the table are fields.
- The key field is identified with a picture of a key.
 - Some tables require multiple keys to uniquely identify a record.
- Lines between tables are relationships.
 - The 1 to ∞ means a one-to-many relationship.

Sample database



Business Analytics: Databases

The End

MS Access (Importing Data)

MS Access: Importing Data

MS Access can import data from:

- Files: Excel, HTML, XML, text
- Other databases: Access, SQL server, Azure, dBase
- Online services: SharePoint
- Other sources: ODBC database

Business Analytics: MS Access (Importing Data)

The End

MS Access (Creating Relationships)

Relationships

- A relationship connects two tables based on a common field
- Types of relationships
 - Inner join: include only records where the joined fields from both tables are equal
 - Left join: include all records from the left table and only those records from the right table where the joined fields are equal
 - Right join: include all records from the right table and only those records from the left table where the joined fields are equal
 - Outer join: include all records

Business Analytics: MS Access (Creating Relationships)

The End

MS Access (Simple Queries)

Simple Queries

- Queries retrieve data from one or more tables.
- Grouping collapses records (rows) based on a common value of a field.
- When records are grouped, calculations can include sum, average, and count of the records.
- Records can be sorted in ascending or descending order.
- Filters can limit the records displayed to only those satisfying certain conditions.

Business Analytics: MS Access (Simple Queries)

The End

MS Access (Fixing Dirty Data)

Fixing Dirty Data

- Dirty data are data that are incorrect or inconsistent.
- Dirty data in fields that are used to relate two tables in a join are especially problematic because they jeopardize the relationship between tables.
- Identifying dirty data in joined fields can be done by comparing left joins and right joins.

Business Analytics: MS Access (Fixing Dirty Data)

The End

MS Access (Complex Queries)

Complex Queries

Complex queries can include:

- Multiple filters
- AND, OR, and NOT Boolean operators
- Calculations

Business Analytics: MS Access (Complex Queries)

The End

SQL Overview

SQL, Part I

What is SQL?

- SQL is a Structured Query Language that works across different databases.
- There are international standards for SQL, and SQL has been designated the standard language for relational databases.
- Although standards exist, there are some variations in implementation across database platforms.

SQL, Part II

SQL commands for queries include:

- SELECT: identify fields to be reported
- AS: creates an alias for reporting labels
- FROM: identify tables to be used
- WHERE: filter criteria
- GROUP BY: clustering criteria
- HAVING: additional criteria based on a calculation
- ORDER BY: sorting criteria
- IN: part of a list
- BETWEEN: range of numbers
- IS NULL: empty values
- LIKE: matches string of characters

SQL, Part III

Order of SQL commands

- SELECT
- FROM
- WHERE
- GROUP BY
- HAVING
- ORDER BY

Business Analytics: SQL Overview

The End