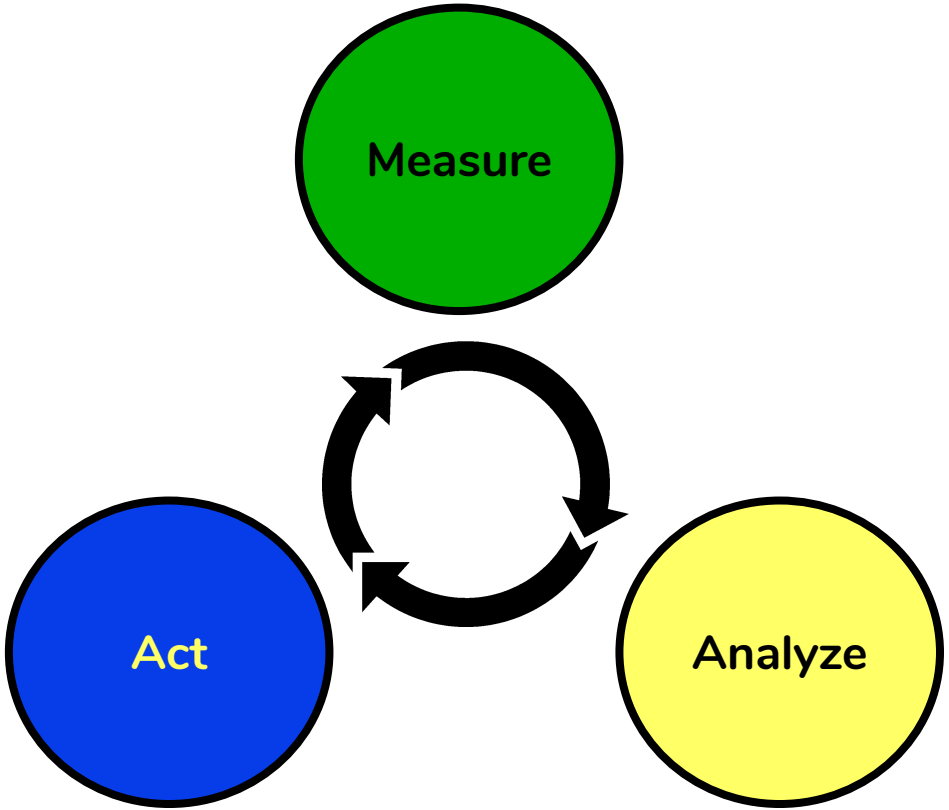
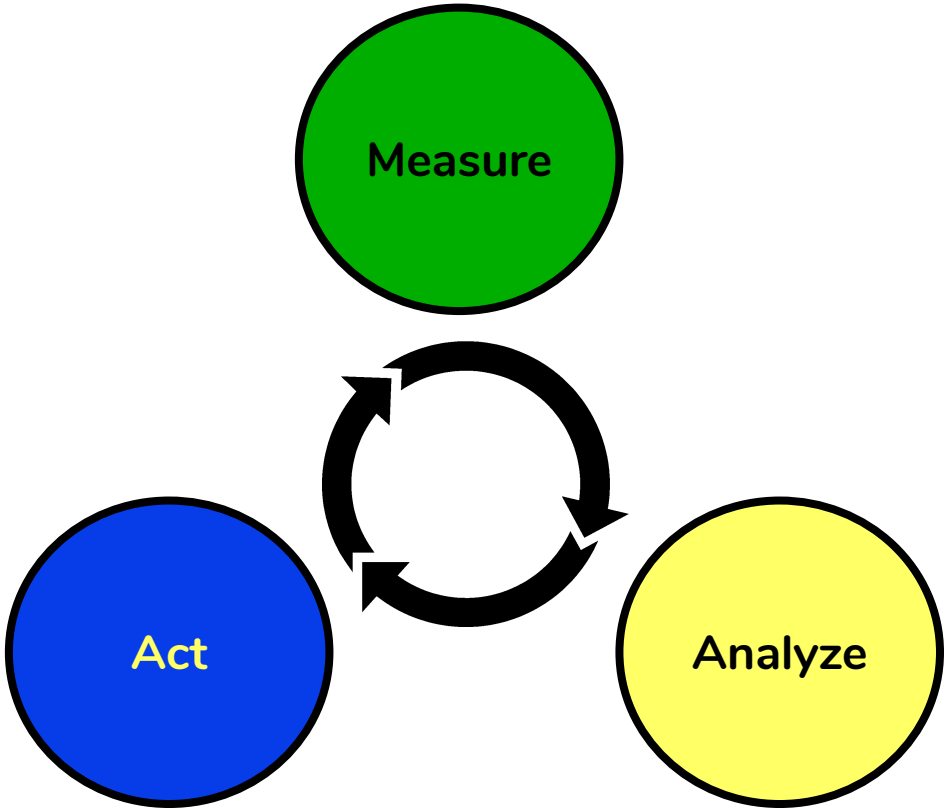
ORIE 3120

Lecture 2: SQL #1

Remember the operational improvement cycle?



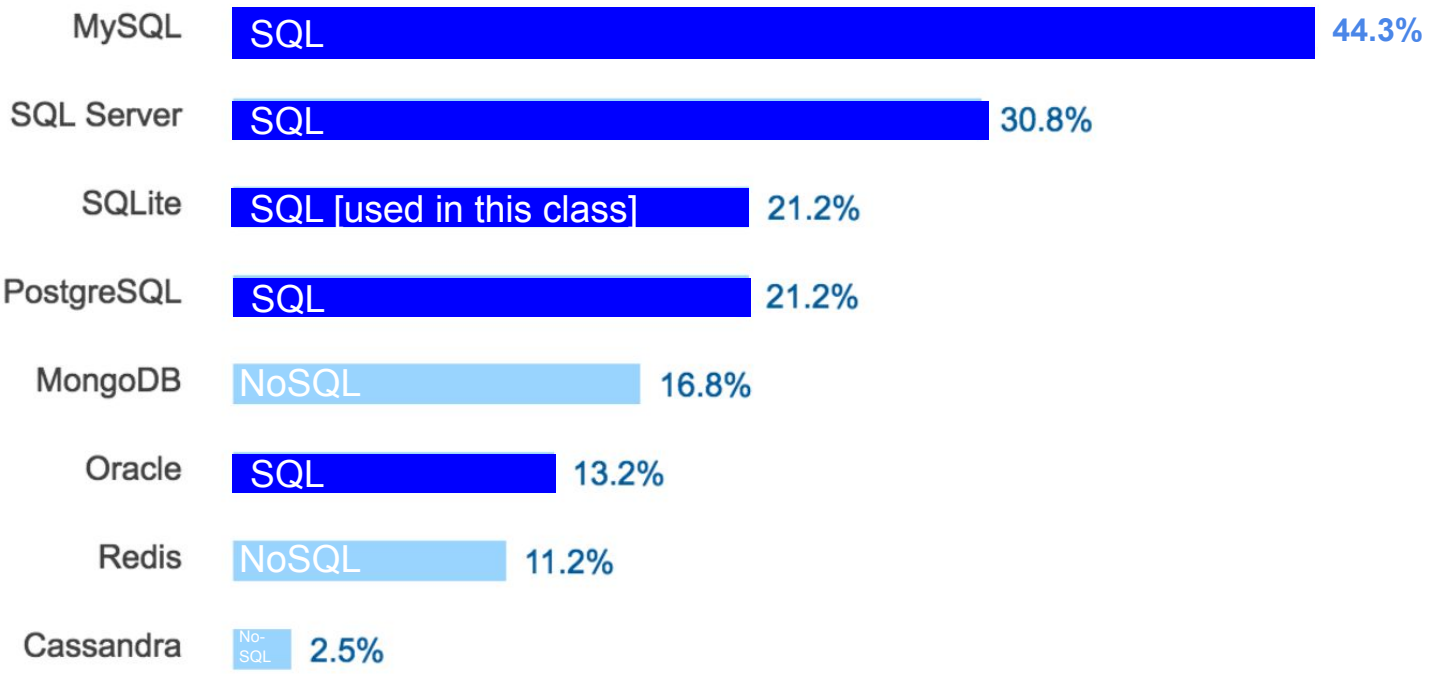
Measuring & analyzing involves data



Structured Query Language (SQL) is a language for manipulating data

- SQL is not a single software application made by a single company.
- Rather, it is a standard, which is used, packaged, and adapted by many software companies.
- Some SQL databases are free; some are not.
- Some SQL databases have non-standard features.

Most businesses store their data using SQL



Data from a 2017 StackOverflow survey

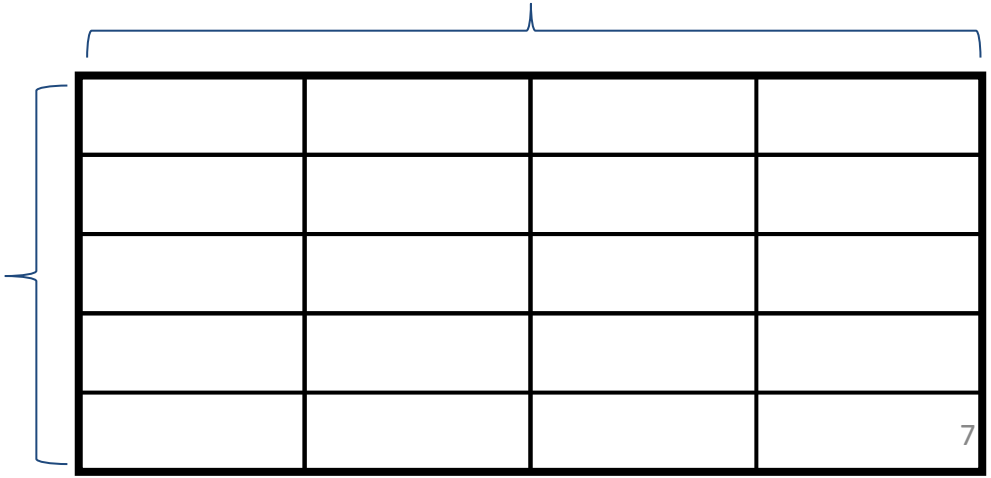
SQL is for “Relational” Databases

- A relational database consists of tables.
- Tables are logical units which are related to one another.
- This allows the data to be broken down into smaller, manageable units.
- By having common keys among tables, data from multiple tables can be joined to form one large set of data.

A table consists of fields and records

Fields: a fixed number of columns, each column having a prescribed data type (integer, single, double, text, date …) and length

Records: an unlimited number of rows, each row containing data in each column of the prescribed type



Data Types

- String types (alphanumeric characters)
  - Fixed length
  - Varying length
  - Large amount of text
- Numeric types (number values)
  - Decimals
  - Integers
- Date and time types
- Unstructured binary data (e.g., images, audio)

Null

- A null represents a missing value in a record in a specific field
- It is not zero
- It is not a space
- It is nothing
- A field with a null value has been left blank during record creation
- Sometimes this is fine, sometimes it is a problem

Here's an example table with a key

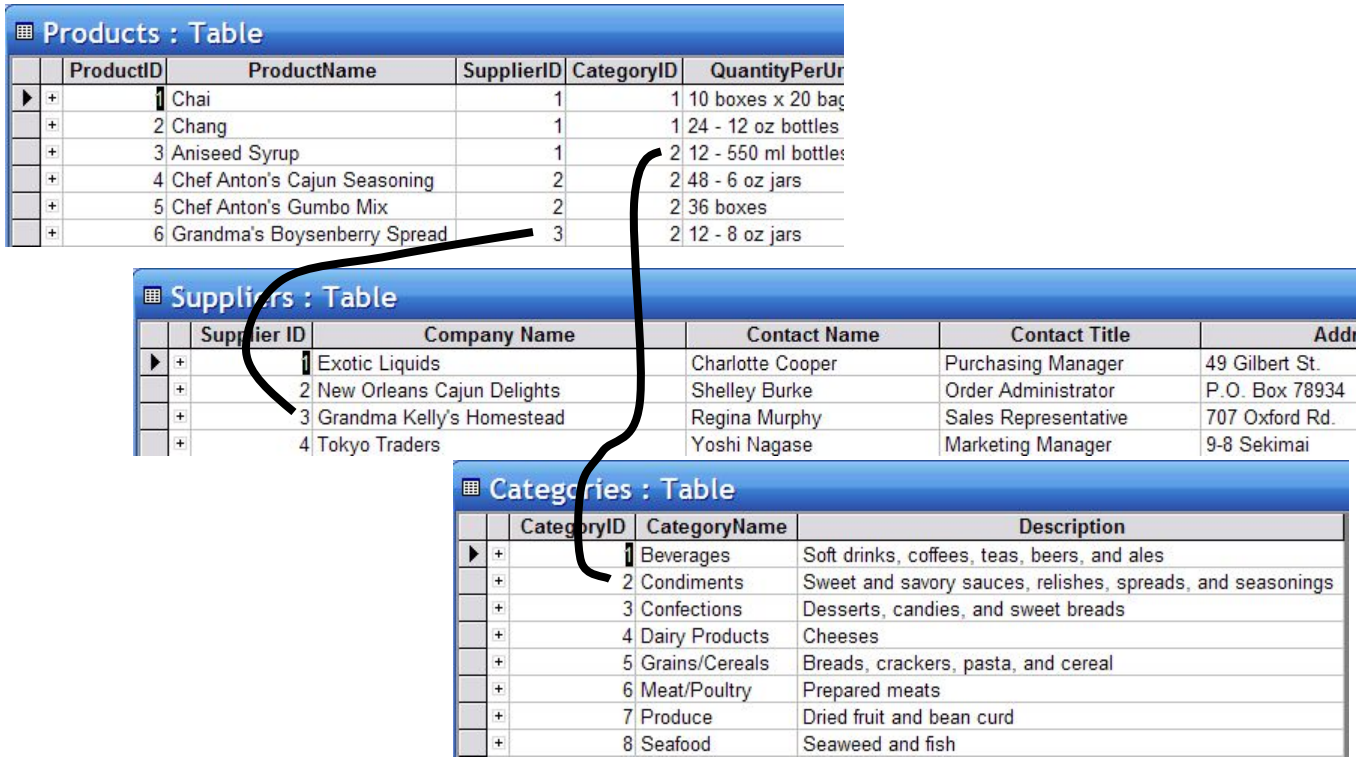
Products : Table

ProductID	ProductName	SupplierID	CategoryID	QuantityPerUnit	UnitPrice	UnitsInStock	UnitsOnOrder	ReorderLeve
1	Chai	1	1	10 boxes x 20 bags	$18.00	39	0	1
2	Chang	1	1	24 - 12 oz bottles	$19.00	17	40	2
3	Aniseed Syrup	1	2	12 - 550 ml bottles	$10.00	13	70	2
4	Chef Anton's Cajun Seasoning	2	2	48 - 6 oz jars	$22.00	53	0	
5	Chef Anton's Gumbo Mix	2	2	36 boxes	$21.35	0	0	
6	Grandma's Boysenberry Spread	3	2	12 - 8 oz jars	$25.00	120	0	2
7	Uncle Bob's Organic Dried Pears	3	7	12 - 1 lb pkgs.	$30.00	15	0	1
8	Northwoods Cranberry Sauce	3	2	12 - 12 oz jars	$40.00	6	0	
9	Mishi Kobe Niku	4	6	18 - 500 g pkgs.	$97.00	29	0	
10	Ikura	4	8	12 - 200 ml jars	$31.00	31	0	
11	Queso Cabrales	5	4	1 kg pkg.	$21.00	22	30	3
12	Queso Manchego La Pastora	5	4	10 - 500 g pkgs.	$38.00	86	0	
13	Konbu	6	8	2 kg box	$6.00	24	0	

Record: 1 of 77

Key: field, or ordered set of fields, whose data uniquely identify a record

Use keys to make relationships



Products : Table

ProductID	ProductName	SupplierID	CategoryID	QuantityPerUr
1	Chai	1	1	10 boxes x 20 bag
2	Chang	1	1	24 - 12 oz bottles
3	Aniseed Syrup	1	2	12 - 550 ml bottles
4	Chef Anton's Cajun Seasoning	2	2	48 - 6 oz jars
5	Chef Anton's Gumbo Mix	2	2	36 boxes
6	Grandma's Boysenberry Spread	3	2	12 - 8 oz jars

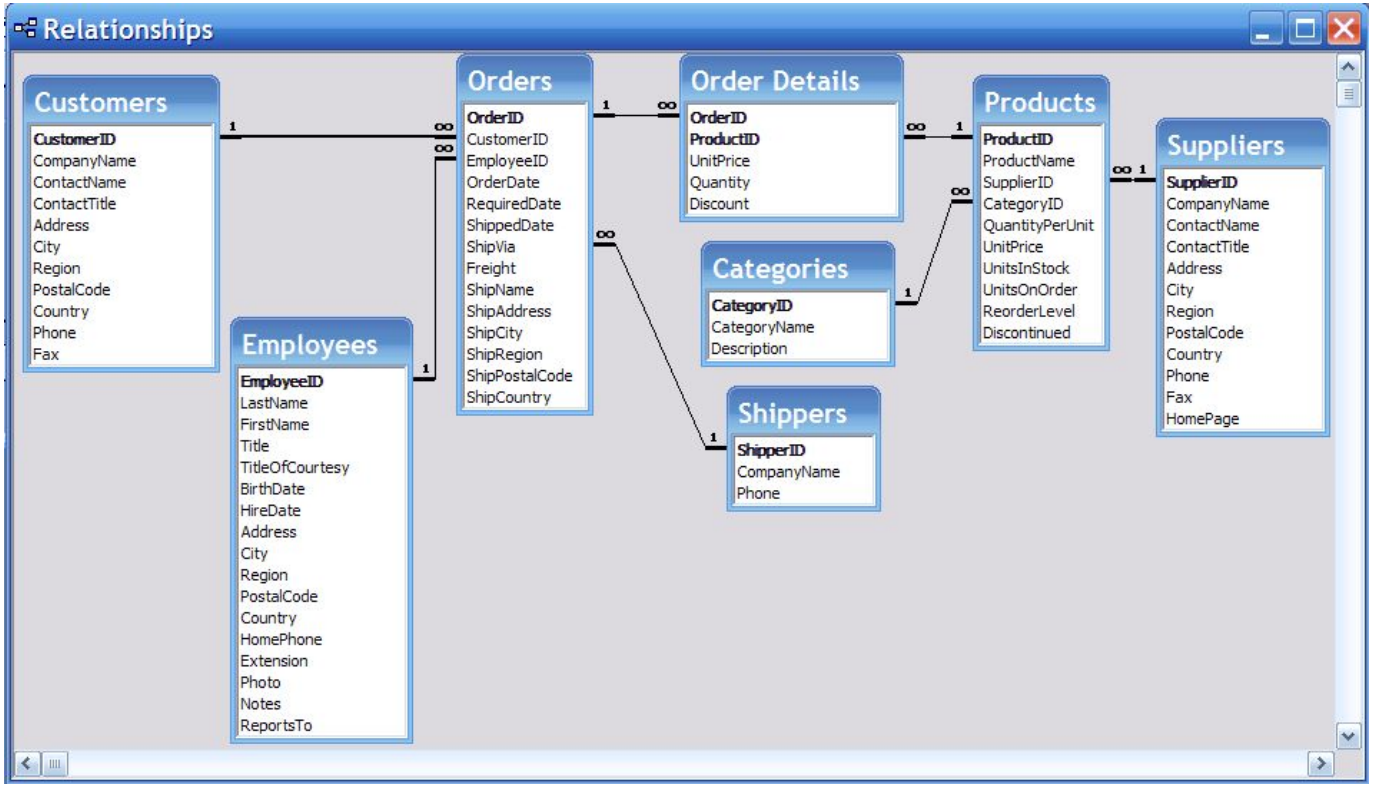
Suppliers : Table

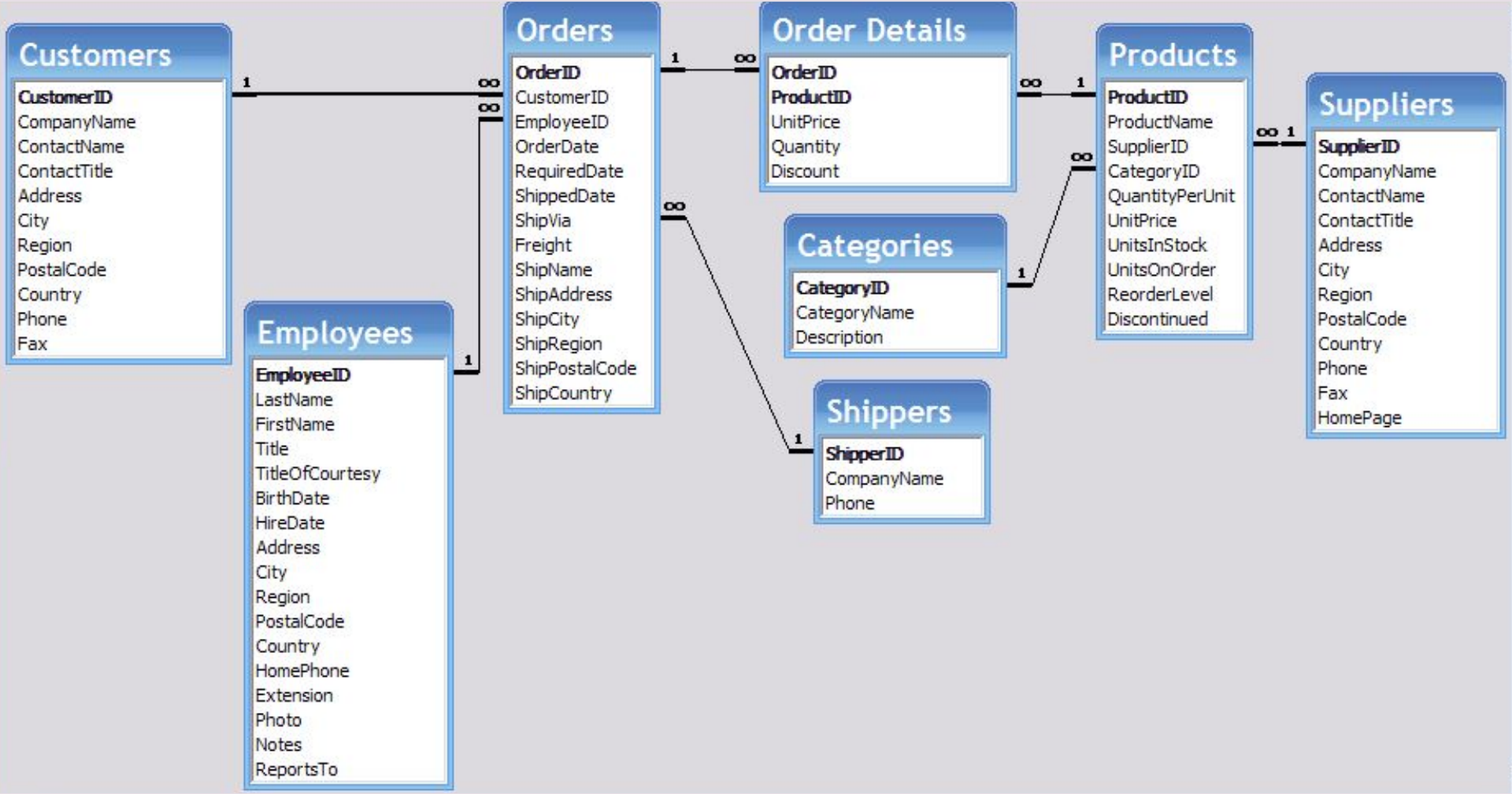
Supplier ID	Company Name	Contact Name	Contact Title	Address
1	Exotic Liquids	Charlotte Cooper	Purchasing Manager	49 Gilbert St.
2	New Orleans Cajun Delights	Shelley Burke	Order Administrator	P.O. Box 78934
3	Grandma Kelly's Homestead	Regina Murphy	Sales Representative	707 Oxford Rd.
4	Tokyo Traders	Yoshi Nagase	Marketing Manager	9-8 Sekimai

Categories : Table

CategoryID	CategoryName	Description
1	Beverages	Soft drinks, coffees, teas, beers, and ales
2	Condiments	Sweet and savory sauces, relishes, spreads, and seasonings
3	Confections	Desserts, candies, and sweet breads
4	Dairy Products	Cheeses
5	Grains/Cereals	Breads, crackers, pasta, and cereal
6	Meat/Poultry	Prepared meats
7	Produce	Dried fruit and bean curd
8	Seafood	Seaweed and fish

A database schema is a collection of tables related by keys




Customers
CustomerID
CompanyName
ContactName
ContactTitle
Address
City
Region
PostalCode
Country
Phone
Fax
Orders
OrderID
CustomerID
EmployeeID
OrderDate
RequiredDate
ShippedDate
ShipVia
Freight
ShipName
ShipAddress
ShipCity
ShipRegion
ShipPostalCode
ShipCountry
Order Details
OrderID
ProductID
UnitPrice
Quantity
Discount
Products
ProductID
ProductName
SupplierID
CategoryID
QuantityPerUnit
UnitPrice
UnitsInStock
UnitsOnOrder
ReorderLevel
Discontinued
Suppliers
SupplierID
CompanyName
ContactName
ContactTitle
Address
City
Region
PostalCode
Country
Phone
Fax
HomePage
Categories
CategoryID
CategoryName
Description
Employees
EmployeeID
LastName
FirstName
Title
TitleOfCourtesy
BirthDate
HireDate
Address
City
Region
PostalCode
Country
HomePhone
Extension
Photo
Notes
ReportsTo
Shippers
ShipperID
CompanyName
Phone

Avoid data duplication when creating a database schema

- Don’t put customer phone number into Orders table
  - Same phone number then occurs in hundreds or thousands of records
  - Wasted storage
  - Subject to error: person recording new order may make typing mistake and enter phone number incorrectly
- Instead set up Customers table (one record per customer) and include phone number field