Databases

store some stuff

a database is:

a collection of information that is organized so that it can easily be accessed, managed, updated, and deleted

Database Actions

- typically you manage a database through a **Database Management System (DBMS)**
- DBMS's allow users and software to interact with databases
- Example DBMS's:
    - MySQL, PostgreSQL, SQLite, Sybase, Oracle, etc

SQL

- stands for **Structured Query Language**
- designed for managing data in a **Relational Database Management System (RDBMS)**
- includes within it a *definition language*, *data manipulation language*, and *data control language*
- basically means that SQL can do data inserts, queries, updates & deletes, schema creation and modification, and data access control

PostgreSQL (or Postgres)

- is an **Object-Relational Database Management System (ORDBMS)** following SQL standards
- this means that it focuses on **object-oriented models**, rather than simply relational design
- this is useful for creating databases to work with object-oriented programs
- allows the storage of functions, custom objects, language-specific Data Types, etc

Database Design

the process of producing a detailed **data model** of a database

- data models contain all the logical information needed to define a database and its relationships
- the most common way to present this design is through **entity-relationship models** or **ER models**

DB Design Process

1 Determine the Purpose of the DB

2 Find and Organize the needed information

3 Divide the data into database tables

4 Divide data in tables into columns

5 Specify Primary Keys

6 Set up relationships between tables
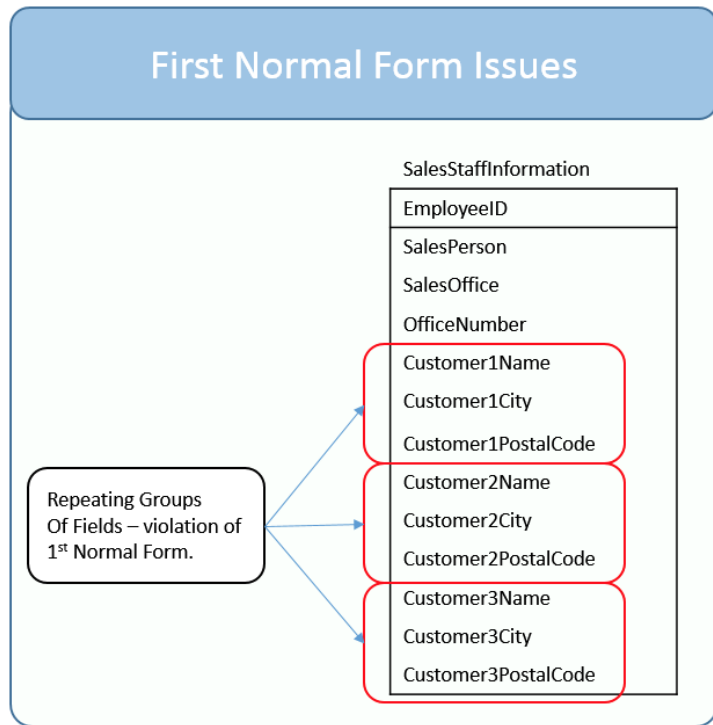
7 Refine your design

8 Normalize

1st Normal Form

a relation is in **first normal form** if the domain of each attribute is atomic, and each attribute contains a single value from the domain

The 5 following conditions must be met for 1st Normal Form:

1. No ordering in the rows
2. No ordering in the columns
3. No duplicate rows
4. Every cell contains one value from the domain (within limit)
5. All columns are regular (no hidden info in rows)

SalesStaff						
EmployeeID	**SalesPerson**	**SalesOffice**	**OfficeNumber**	**Customer1**	**Customer2**	**Customer3**
1003	Mary Smith	Chicago	312-555-1212	Ford	GM	
1004	John Hunt	New York	212-555-1212	Dell	HP	Apple
1005	Martin Hap	Chicago	312-555-1212	Boeing		

First Normal Form Issues

SalesStaffInformation

EmployeeID
SalesPerson
SalesOffice
OfficeNumber
Customer1Name
Customer1City
Customer1PostalCode
Customer2Name
Customer2City
Customer2PostalCode
Customer3Name
Customer3City
Customer3PostalCode

Repeating Groups Of Fields – violation of 1st Normal Form.

2nd Normal Form

a relation is in 2nd normal form if it:

1. Is in 1st Normal Form
2. All non-key attributes are fully functional dependent on the primary key

All non-key attributes in a relation cannot be dependent on a subset of the primary key – this is mostly just relative for composite keys.

TABLE_PURCHASE_DETAIL

Customer ID	Store ID	Purchase Location
1	1	Los Angeles
1	3	San Francisco
2	1	Los Angeles
3	2	New York
4	3	San Francisco

3rd Normal Form

a relation is in 3rd normal form if it:

1. is in 2nd normal form
2. There is no transitive functional dependency

Attributes in a table cannot be dependent on the state of other attributes in the table.

TABLE_BOOK_DETAIL

Book ID	Genre ID	Genre Type	Price
1	1	Gardening	25.99
2	2	Sports	14.99
3	1	Gardening	10.00
4	3	Travel	12.99
5	2	Sports	17.99

Why Normalize?

- minimize duplication within your database
- reduce potential for insertion, deletion, and update errors
- reduce potential for stale data
- avoid overly complicated queries
- minimize redesign for adding or removing data

FYI - There's more to Normalization

There are 8 Normal forms in total. In general, you need to normalize to at least 3rd Normal Form.

- 1NF - First Normal Form

- 2NF - Second Normal Form

- 3NF - Third Normal Form

- BCNF - Boyce–Codd Normal Form

- 4NF - Fourth Normal Form

- 5NF - Fifth Normal Form

- 6NF - Sixth Normal Form

- DKNF - Domain/Key Normal Form

ER Models

diagram that describes relationships between entities

nouns: entities

verbs: relationships

adjective: attribute

adverb: attribute for a relationship

Database Practice

Design a database to store the following information about PDX Code Guild's Students & their assignments:

- student name
- student email
- student phone number
- assignment title
- assignment code location
- latest commit for the assignment
- status of assignment
- assignment grade
- code reviewer(s)
- status of code review(s)