10 – Introduction to Relational Databases (RDB)

LBSCI 700 | Spring 2019

Queens College, CUNY

10-database.pdf

Database Intro

Keys

Integrity Rules

Normalization Rules

DB Queries

DB Implementation: MS Access

Database Intro

Database Intro

- Data vs. Information

- Changing data into information
  - Organize data so that it can be viewed in a useful form
  - Requirements of this process

Data into Information: Identify Context

- Data
  - Athy, Annrei M 12****64
  - Cooper, Aisha M 23****12
  - Diggle,Robert 23****22
  - Iannuzzo,Jessica 23****12
- Context
  - Class Roster
  - Course LBSCI700, Section: 2, Spring 2019

DB – What and Why?

A **database** is **shared**, **integrated** computer structure that stored a collection of end-user data and **metadata**.

--- Database Systems: Design, Implementation, & Management: Rob & Coronel

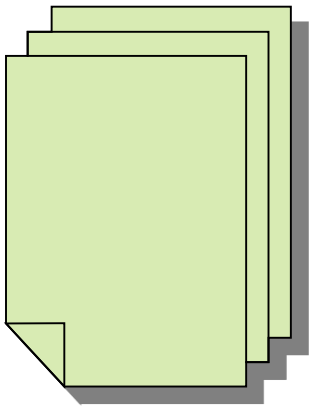
- Front End/Back End
- Spreadsheet vs DB
- Integrate with website – (interactive)
- Examples
  - Collect Data (AARP Membership)
  - Generate reports (Store Locator)
  - Searchable (ALA Accredited Programs)

DB Anatomy

Table

- Columns
- Rows

- Fields
- Records



Example: Table

Table name: STUDENT

STU_NUM	STU_LNAME	STU_FNAME	STU_INIT	STU_DOB	STU_HRS	STU_CLASS
321452	Bowser	William	C	Saturday, February 12, 1972	42	So
324257	Smithson	Anne	K	Tuesday, November 15, 1977	81	Jr
324258	Brewer	Juliette		Tuesday, August 23, 1966	36	So
324269	Oblonski	Walter	H	Sunday, September 16, 1973	66	Jr
324273	Smith	John	D	Friday, December 30, 1955	102	Sr
324274	Katinga	Raphael	P	Thursday, October 21, 1976	114	Sr
324291	Robertson	Gerald	T	Wednesday, April 08, 1970	120	Sr
324299	Smith	John	B	Wednesday, November 30, 1983	15	Fr

Database Systems: Design, Implementation, & Management: Rob & Coronel

- 8 rows & 7 columns
- Row = record
- Column = field
  - has specific characteristics (data type, format, value range)

DB Software

Relational DB (RDB)

Commercial

- Oracle
- Microsoft SQL Server

Open Source

- MySQL
- SQLite

Keys

Keys

Keys are used to ensure that each row in a table is uniquely identifiable

Primary – unique identifier (Each record must be unique)

Candidate keys - combination of fields used to identify a database record without any extraneous data. A table may have one or more candidate keys. One of these candidate keys is selected as the table primary key.

Foreign Keys – relating information in other table

1:M Relationship

One-to-many (1:M):

Most common.

The primary key table contains only one record that relates to none, one, or many records in the related table.

-Parent-child (sibling) example

1:1 Relationship

One to one (1:1):

Rare.

A special case of the 1:M

Both tables can have only one record on either side of the relationship. Each primary key value relates to only one (or no) record in the related table. Most one-to-one relationships are forced by business rules and don't flow naturally from the data. In the absence of such a rule, you can usually combine both tables into one table without breaking any normalization rules.

- Spouse example;
- PROFESSOR chairs DEPARTMENT

Relating Tables with Keys: Example

- Provides a logical "human-level" view of the data and associations among groups of data (i.e., tables)

Customer_ID	Customer_Account	Agent_ID
1224	4556	23
1225	4558	25

Agent_ID	Last_Name	First_Name	Phone
23	Sturm	David	334-5678
25	Long	Kyle	556-3421

M:M/M:N Relationship

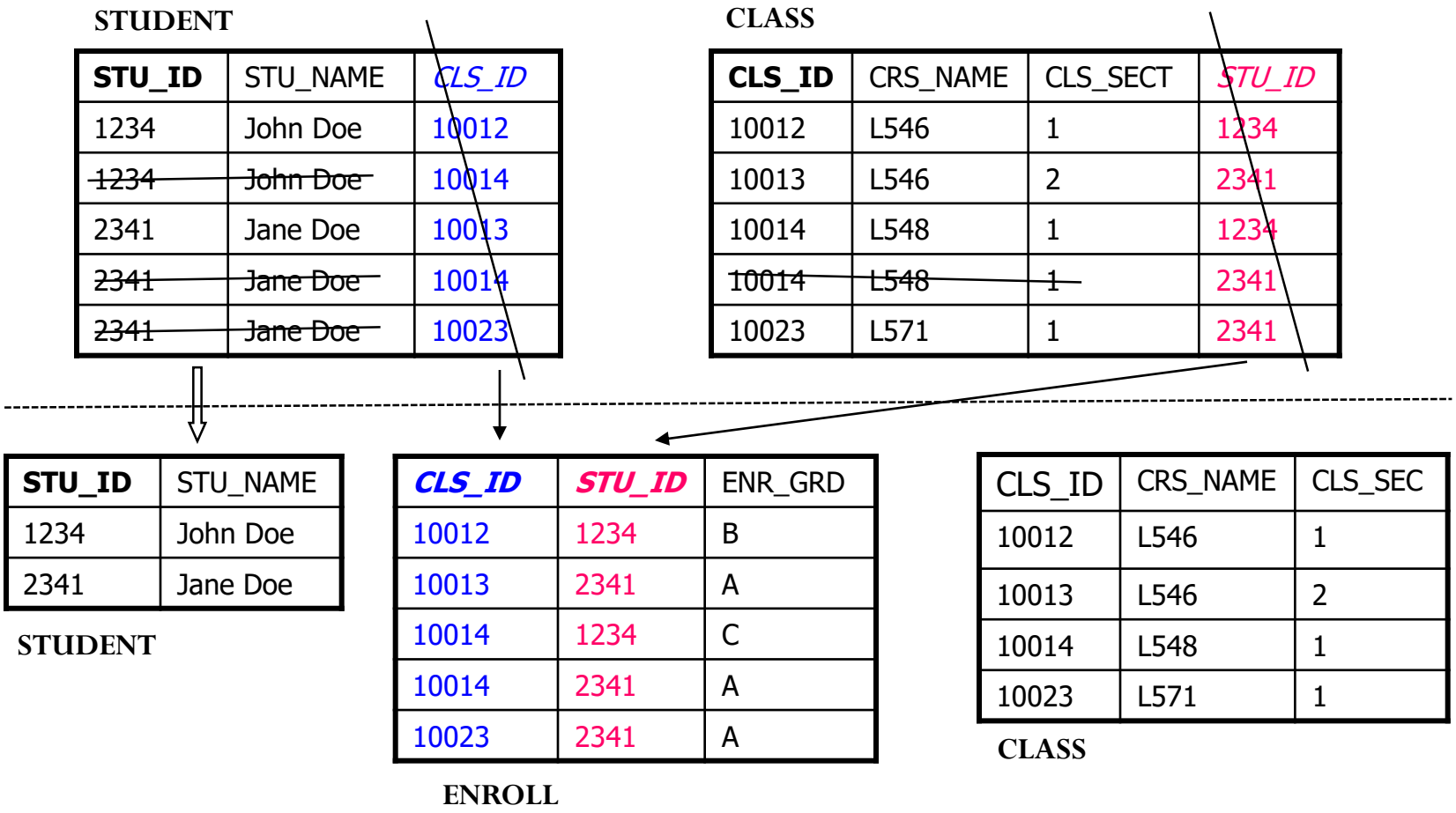
Many-to-many (M:M/M:N):

Each record in both tables can relate to any number of records (or no records) in the other table. Many-to-many relationships require a third table, known as a bridge table, because relational systems can't directly accommodate the relationship.

- Course and Students Example

M:N to 1:M Conversion

M:N to 1:M Conversion



1. Move the foreign key columns to create a bridge table & add attributes if needed.
2. Collapse the duplicate records in remaining tables.

Integrity Rules

Integrity Rules

- **Entity Integrity**
  - Each entity has unique key
- **Referential Integrity**
  - Foreign key value is null or matches primary key values in related table
- Most RDBMS enforce integrity rules automatically.

STU_ID	STU_LNAME	STU_FNAME	*DEPT_CODE*
12345	Doe	John	245
12346	Dew	John	243
22134	Dew	James	

DEPT_CODE	DEPT_NAME
243	Astronomy
244	Computer Science
245	Sociology

*The basic building blocks of all data models are entities, attributes, relationships, and constraints

DB Efficiency

Normalization

- **Normalization of database tables**

-Reduce repetitive entries (minimize data redundancies)

-Levels (1 NF, 2 NF, 3NF)

Normalization: First Normal Form

First normal form (1 NF) Rules:

- Each cell has a single value
- Each row/record needs to be unique

Steps

1. Eliminate repeating groups (columns) from the same table.
2. Identify primary key.
3. Identify all dependencies
   - Partial dependency
   - Transitive dependency

Normalization: 1NF example

Covert an ill-organized table to 1NF

Full Names	Physical Address	Movies rented	Salutation
Janet Jones	First Street Plot No 4	Pirates of the Caribbean, Clash of the Titans	Ms.
Robert Phil	3rd Street 34	Forgetting Sarah Marshal, Daddy's Little Girls	Mr.
Robert Phil	5th Avenue	Clash of the Titans	Mr.

FULL NAMES	PHYSICAL ADDRESS	MOVIES RENTED	SALUTATION
Janet Jones	First Street Plot No 4	Pirates of the Caribbean	Ms.
Janet Jones	First Street Plot No 4	Clash of the Titans	Ms.
Robert Phil	3rd Street 34	Forgetting Sarah Marshal	Mr.
Robert Phil	3rd Street 34	Daddy's Little Girls	Mr.
Robert Phil	5th Avenue	Clash of the Titans	Mr.

Normalization: Second Normal Form

Converting to 2NF is done only when the 1NF has a composite primary key

- Second Normal Form (2NF) Rules
  - It is in 1NF
  - There are no partial dependencies (single column primary key)

- Conversion to 2NF Steps
  1. Write each key component (w/ partial dependency) on separate line
  2. Write original (composite) key on last line
  3. Each component is new table
  4. Assign corresponding dependent attributes after each key

1NF (**FULL_NAMES**, **PHYSICAL_ADDRESS**, MOVIE_RENTED, SALUTATION)

↓

MEMBERS (**MEMBERSHIP_ID,** FULL_NAMES, PHYSICAL_ADDRESS, SALUTATION)
MOVIES (**MEMBERSHIP_ID**, MOVIE_RENTED)

Normalization: 2NF example

Composite Key

Robert Phil	3rd Street 34	Daddy's Little Girls	Mr.
Robert Phil	5th Avenue	Clash of the Titans	Mr.

Names are common. Hence you need name as well Address to uniquely identify a record.

TABLE1: MEMBERS

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3rd Street 34	Mr.
3	Robert Phil	5th Avenue	Mr.

TABLE2: MOVIES

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

Normalization: Third Normal Form

- Third Normal Form (3NF)
  - It is in 2NF
  - There are no transitive dependencies

- Conversion to 3NF Steps

1. Start with 2NF format
2. Break off the transitive dependencies and create separate tables

MEMBERS (**MEMBERSHIP_ID,** FULL_NAMES, PHYSICAL_ADDRESS, SALUTATION)

↓

MEMBERS (**MEMBERSHIP_ID**, FULL_NAMES, PHYSICAL_ADDRESS)

SALUTATION (**SALUTATION_ID**, SALUTATION)

Normalization: 3NF example

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3rd Street 34	Mr.
3	Robert Phil	5th Avenue	Mr.

Change in Name

May Change Salutation

TABLE1: MEMBERS

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION ID
1	Janet Jones	First Street Plot No 4	2
2	Robert Phil	3rd Street 34	1
3	Robert Phil	5th Avenue	1

TABLE2: MOVIES

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

TABLE3: SALUTATION

SALUTATION ID	SALUTATION
1	Mr.
2	Ms.
3	Mrs.
4	Dr.

Examples from https://www.guru99.com/database-normalization.html

DB Queries / Search for next time …

SQL - modify and retrieve data

Commands:

- **Select** - Fetches data.
- **Insert** - Inserts one or more rows of data.
- **Update** - Modifies existing row(s) of data
- **Delete** - Deletes rows of data

DB Implementation: MS Access

Outline

- Access Table
- Access Queries
- Access Reports

Recap

Database Intro

Keys

Integrity Rules

Normalization Rules

DB Queries

DB Implementation: MS Access

Last Things

About Usability Studies

- Continue …

ToDo

- Start homework
  - -- Look for email
  - -- Check Bb weekly folder
- Note any questions from reading and homework

* Next Monday, April 15, Asynchronous Online