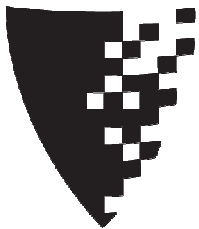


# SQL: Part II

Introduction to Databases

CompSci 316 Fall 2014



**DUKE**  
COMPUTER SCIENCE

# Announcements (Thu., Sep. 18)

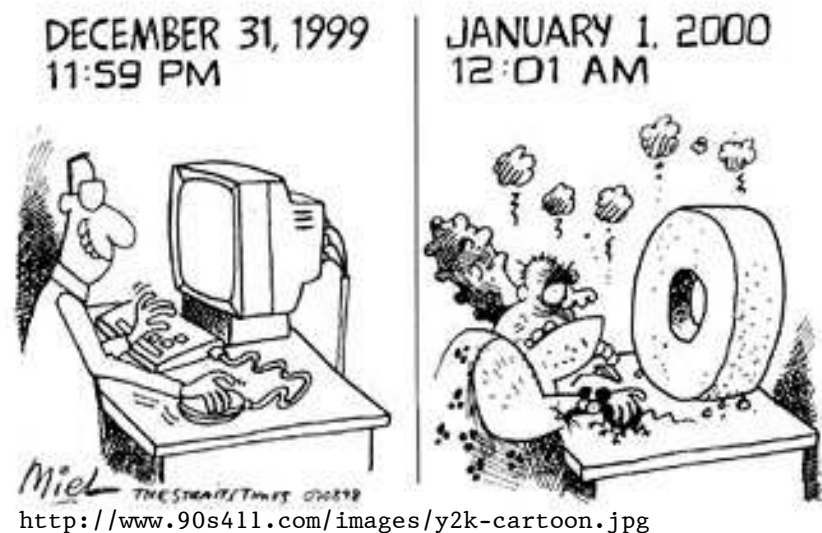
- **Homework #1 sample solution** to be posted on Sakai by tomorrow
- We are working on resolving the **websubmit** issue
  - Use Chrome and IE for now
- **Homework #2** due in two weeks

# Incomplete information

- Example: *User* (*uid*, *name*, *age*, *pop*)
- Value **unknown**
  - We do not know Nelson's age
- Value **not applicable**
  - Suppose *pop* is based on interactions with others on our social networking site
  - Nelson is new to our site; what is his *pop*?

# Solution 1

- Dedicate a value from each domain (type)
  - *pop* cannot be  $-1$ , so use  $-1$  as a special value to indicate a missing or invalid *pop*
  - Leads to incorrect answers if not careful
    - `SELECT AVG(pop) FROM User;`
  - Complicates applications
    - `SELECT AVG(pop) FROM User WHERE pop <> -1;`
- Perhaps the value is not as special as you think!
  - Ever heard of the Y2K bug? “00” was used as a missing or invalid year value



# Solution 2

- A valid-bit for every column
  - *User (uid, name, name\_is\_valid, age, age\_is\_valid, pop, pop\_is\_valid)*
  - Complicates schema and queries
    - `SELECT AVG(pop) FROM User WHERE pop_is_valid;`

[Click here to download full PDF material](#)