

Introducing Undergraduate Database Students to K-12 Education Research

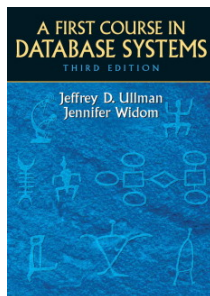
Chris Mayfield (JMU), Carole Ottenheimer (CIT),
Bethann Canada (VDOE), Brooke Bell (CIT)

SIGCSE 2014



JMU database course

- ▶ Database modeling
- ▶ Relational algebra
- ▶ SQL programming
- ▶ Normalization theory
- ▶ Transactions, ACID
- ▶ Indexes, views, etc.
- ▶ 3-tier applications



Group project objectives

English language description → working DB application

1. Create **E/R models** from application descriptions
2. Convert E/R models into **relational designs**
3. Identify **redundancies** in designs and remove them
4. Import data into a DBMS and **enforce integrity** constraints
5. Write sophisticated **database queries** using SQL
6. Evaluate **trade-offs** of different ways of phrasing a query
7. Implement a **web-based** interface to the database

What I used to do

Find a publicly available data set that:

1. is large enough to require indexing
2. exposes students to research topics

Each team creates their own front-end

For example:



(Credit: T. M. Murali, Virginia Tech)

Then along came Apps4VA



- ▶ Home : <http://www.apps4va.org/> (watch video)
- ▶ About : <http://www.apps4va.org/about.html> (watch video)
- ▶ Ideas : <http://www.apps4va.org/idea-bank.html>
- ▶ Data : <http://www.apps4va.org/data.html>
- ▶ Blog : <http://www.apps4va.org/blog.html> (see Oct 2012)

What is the data about?

Divisions – number and name

- ▶ 099 Jefferson County

Schools – number and name

- ▶ 0010 Flat Hat High

Students – *names withheld*

- ▶ Grade code
- ▶ Race code
- ▶ Gender
- ▶ Disability?
- ▶ Limited English?
- ▶ Disadvantaged?



For each **school year** (e.g., 2008–2009)
and each **level code** (STATE, DIV, SCH)

Detailed information

Dataset descriptions:

- ▶ http://www.doe.virginia.gov/statistics_reports/research_data/index.shtml

Attribute dictionary:

- ▶ http://www.doe.virginia.gov/statistics_reports/research_data/data_elements.shtml

Aggregation

- ▶ *The datasets are compiled using all the possible combinations of all the demographics about students so each row within the dataset contains a rate or count in addition to the demographics used to arrive at the rate or count.*

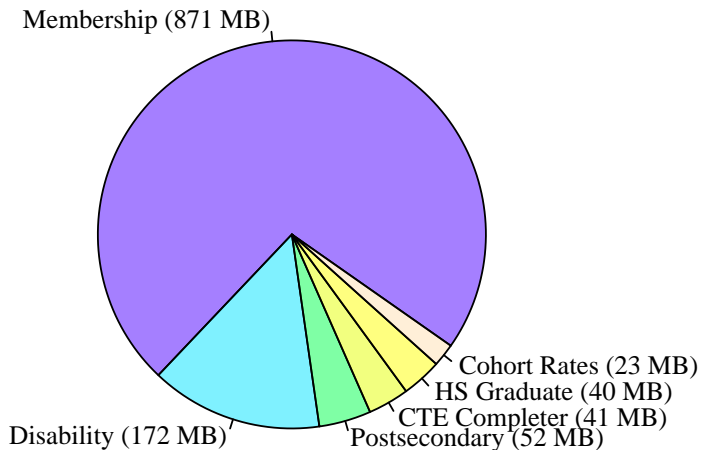
Suppression

- ▶ *Within each dataset, rows were withheld if deemed that the number of students in the group could lead to the identification of a single student. In most cases, student groups of 9 or less are suppressed.*

How much data is there?

5.1 GB of CSV files

(3.9 GB test data + 1.2 GB other stats)



The VLDS datasets

1. `fall_membership` 10,794,438 rows
Number of students enrolled in VA public schools each Sep 30th
2. `dec_child_count` 2,673,579 rows
Number of students with disabilities receiving special education
3. `sol_test_data` 37,897,923 rows
Avg scores and pass/fail rates for English, History, Math, Science
4. `cte_completer` 567,477 rows
Number of students who completed approved CTE course sequences
5. `hs_graduate` 515,614 rows
Number of high school graduates and completers of similar diplomas
6. `annual_dropout` 27,555 rows
Summer and term dropouts in grades 7-12 (do not return by Oct 1st)
7. `ontime_cohort` 302,516 rows
On-time graduation rates for students entering 9th grade together
8. `postsec_enroll` 770,541 rows
HS grads who enrolled in public higher ed in VA and earned credit
9. `postsec_achieve` 29,653 rows
HS grads in postsec institutions nationwide and estimated credit

The secret ingredient

Rather than simply expose students to research topics, help them to conduct their own educational research.

Example proposals:

- ▶ “What is the return on investment of school budgets in terms of student test scores?”
- ▶ “To what extent does the math and science gender gap exist in Virginia public schools?”
- ▶ “What is the relationship between medical health indicators of a community and student performance?”
- ▶ “How are graduation rates related to school funding and overall wealth in the community?”
- ▶ “Are K-12 students being prepared for real world jobs? What jobs are in demand in their area?”

Success stories

- ▶ Partnership with VDOE
- ▶ Students relate to the data
- ▶ Education research component
- ▶ Running example in course
- ▶ Portfolio for job interviews
- ▶ Increased rigor and dedication

Before and after Apps4VA

- ▶ Same course/instructor
- ▶ 60 students each year

Year	D	W	F	Rate
2012	4	6	4	23%
2013	7	1	1	15%