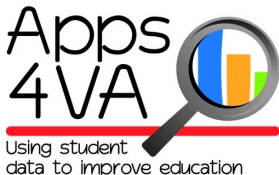# Apps4VA at JMU
Student Projects Featuring VLDS Data

Dr. Chris Mayfield

Department of Computer Science
James Madison University

VLDS Insights — June 30, 2015

One minute version

- 250 students from JMU Computer Science have participated in Apps4VA (Spring 2013, Spring 2014, Spring 2015)

- 66 prototype apps have been developed over the years
  http://apps4va.cs.jmu.edu/

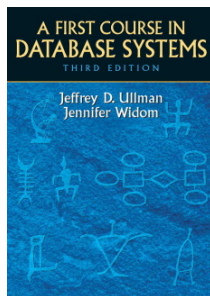- Always looking for new ideas, new partnerships, new data!

Special thanks to Carole Ottenheimer (CIT),
Bethann Canada (VDOE), Brooke Bell (CIT)

2012

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 $\rightarrow$ working DB application

1. Create E/R models from application descriptions

2. Convert database models into relational schemas

3. Import real data and enforce integrity constraints

4. Identify redundancies in designs and remove them

5. Write sophisticated database queries using SQL

6. Evaluate query performance and create indexes

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/ (competition video)
- ▶ About : http://www.apps4va.org/about.html (VLDS 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 I tell my students

Introduction to VLDS data sets

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:

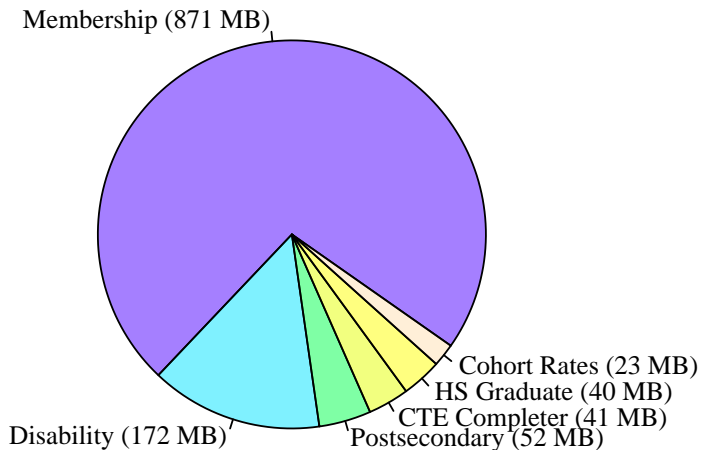- http://www.doe.virginia.gov/statistics_reports/research_data/data_elements.shtml

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?

As of Jan 2014: 5.1 GB of CSV files

(3.9 GB test data + 1.2 GB other stats)

Membership (871 MB)

Cohort Rates (23 MB)
HS Graduate (40 MB)
CTE Completer (41 MB)
Postsecondary (52 MB)

Disability (172 MB)

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

Available data

	fallmem	dec1cnt	testdata	ctecomp	hsgrad	dropout	cohort	psenroll	psachv
2006-2007	X	X	X		X				
2007-2008	X	X	X	X	X		X	X	X
2008-2009	X	X	X	X	X		X	X	
2009-2010	X	X	X	X	X		X	X	
2010-2011	X	X	X	X	X		X	X	
2011-2012	X	X	X	X	X	X	X		
2012-2013	X	X	X	X	new	new	X		
2013-2014	new	new	new	new	new		new		

Other data sets:

- ▶ Composite index
- ▶ Div/sch directory
- ▶ Health indicators
- ▶ Teacher salaries
- ▶ SAT/ACT scores

See Williams 2013 Insights
"A Walk to Remember:
Guided Tour of Virginia's Data"

2013

VLDS Query Generator!

http://apps4va.cs.jmu.edu/RealityCheck

2014
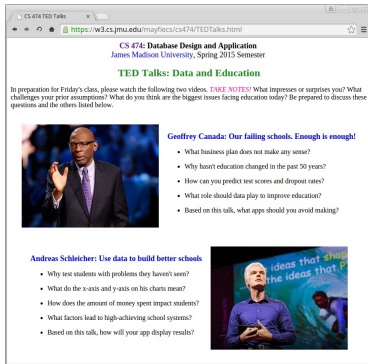
2015

The secret ingredient

Project Vision

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

Example research questions

- ► "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?"

- ► "What is the return on investment of school budgets in terms of student test scores?"

- ► "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