Uplink

May • June 2006
Volume 18, Number 3

Published bimonthly by the National Institute for Computer-Assisted Reporting

www.nicar.org

CENSUS

Making the switch to ACS

By Paul Overberg
USA Today

Once again, the Census Bureau's American Community Survey faced a budget challenge that could have killed it. It wasn't until a congressional conference committee hashed out details in mid-November that it was funded at sustainable levels for this year.

Nothing is certain in Washington, but this means it's increasingly likely that after years in pilot, ACS finally will take wing and replace the long form of the 2010 Census.

Each year, 3 million households will get the ACS survey. Each summer, data from the preceding year will be released for big places (over 65,000 people). By summer 2008, data from 2005-07 will be released as three-year averages for places of at least 20,000 people. Finally, in summer 2010, data from 2005-09 will be released as five-year averages for places as small as census tracts (3,000 people) and – in some cases – block groups (600 people). Every year after that, new annual data

continued on page16

SPOTLIGHT: TOXIC LIVING

Rolling hazards in rail cargo

By David Danelski, *The (Riverside, Calif.) Press-Enterprise*

A freight train derailment last year left several tank cars containing deadly chemicals twisted and dented just yards from homes in densely populated San Bernardino. The accident raised obvious questions about the transport of hazardous materials by railroad through our region.

Our investigation ultimately found ineffective track inspections, misidentified tank cars, and communications problems between railroad and police officials that allowed hundreds of residents to return home before it was safe to do so.

The *Press-Enterprise* covers Riverside and San Bernardino counties, an area

continued on page 10

SPOTLIGHT:

For more about toxic living see:

- Tracking mercury pollution in West Virginia, p. 6
- Data and other resources for reporting, p. 8
- Recent stories from IRE, p. 9
- The May-June 2006 *IRE Journal*

HEALTH CARE

ER gridlock exposed

By Augusta Brennan Jones, *WBAL-Baltimore*

If you have visited a hospital emergency room, you've probably experienced a wait.

But I'll bet it wasn't as long as the time I rushed my husband, John, to the hospital. It was a Monday night last July. John is a cardiac patient, and he was having neck and jaw pain.

That visit, which stretched to more than eight hours, inspired our investigative unit at WBAL to examine ER waiting times using freely available data from a Maryland state agency.

Shortly after I got John to Northwest Hospital outside Baltimore, he was called to the triage unit where they checked his vital signs and determined he could wait. So they sent him back to the waiting room with dozens of other sick-looking patients. After about an hour, he still had not seen a doctor; that's when a nurse announced there

continued on page 18

Bits & Bytes

Storm season

Just in time for storm season, the IRE and NICAR Database Library has updated its Storm Events database through Dec. 31, 2005. The data contains information about last season's hurricanes, including Hurricane Katrina.

The database includes more than 600,000 records of U.S. storm events ranging from hail and flooding to hurricanes and tornadoes. Each record contains details about fatalities and injuries, damage to crops and property, and remarks about the storm. Many of the storm records include latitude and longitude data, which can be used to map the locations using geographic information system (GIS) software.

In addition to information on recent storms, the database has archival information as far back as 1950.

To order the storm events data, contact the database library at 573-884-7711 or download an order form at *www.ire.org/datalibrary/orderform.*

Deadly roads

In February, the database library analyzed fatal vehicle accidents data for a WCBS-New York story.

Assistant administrator Megan Means analyzed the Fatality Analysis Reporting System (FARS) Data, available from the database library, to count fatalities on highways in New York and New Jersey. She found the majority of accidents occurred in daylight during fair weather con-

continued on page 4

INSIDE NICAR

CAR: It's a starting point

By David Herzog, *NICAR and Missouri School of Journalism*

Hundreds of journalists were on hand for the computer-assisted reporting day at the annual IRE conference in June in Fort Worth, Texas. The one-day focus on CAR is an annual rite for many journalists and, for others, an introduction to the use of databases, spreadsheets and Internet resources for reporting.

For many newcomers, it's easy to get overwhelmed by the possibilities of CAR and all of the jargon (think GIS, SQL, regression analysis, intranets). But those journalists who persist can develop new reporting tools that can help them produce better news stories. That's what it's all about, right?

In each issue of *Uplink*, you read about these news stories. And while you might notice differences in the software, techniques and obstacles overcome, you'll see one common thread: CAR provided the starting point for even more reporting

Speaking of *Uplink*, we're working on publishing the newsletter electronically. If you have any suggestions for use, please let me know. I'd love to hear them.

* *

The March-April *Uplink* "Mapping it Out" feature incorrectly identified the location of a village near Peoria, Ill., that was struck by a tornado. It is South Pekin.

About our contributors ...

Augusta Brennan Jones is executive producer-special projects at WBAL-Baltimore, where she manages the station's I-Team. She began her broadcast career as a producer at WRTV-Indianapolis.

David Danelski covers air pollution and does investigative projects for *The Press-Enterprise* in Riverside, Calif. He is an alumnus of the 2004 Advanced CAR Statistics Boot Camp.

Paul Overberg, a database editor at *USA Today*, helps coordinate its demographic coverage. He began using data from the American Community Survey nine years ago and has been invited to participate in several user workshops on development of its products.

Eric Sagara has reported for the *Tucson Citizen* for five years and has covered the crime, higher education, business and wildfire beats. He currently covers primary and secondary education. He is an alumnus of the 2005 Advanced CAR Statistics Boot Camp.

Duane Schrag is special projects reporter at the *Salina* (Kan.) *Journal*. He has worked for the Harris Enterprises newspaper group, based in Hutchinson, Kan., since 1984. Previously, he was editor and publisher of the Chanute (Kan.) Tribune.

Ken Ward Jr. is a reporter for *The Charleston* (W.Va.) *Gazette* and has received national reporting awards for stories about environmental topics. He is chairman of the Society of Environmental Journalists' First Amendment Task Force.

FIRST VENTURE

Stats help deflate big crowd count

By Eric Sagara, *Tucson Citizen*

If you have ever strained to count a crowd of people milling around a concert, parade or festival, you know how hard it can be to get an accurate head count.

The difficulty of crowd counting came up briefly at an IRE and NICAR Advanced CAR Statistics Boot Camp at Arizona State University, and I learned there is a way to go beyond official estimates and find a more accurate estimate for readers.

Within days of my completing the Boot Camp, editors wanted to find a way to accurately count the number of people that attended La Fiesta de los Vaqueros Rodeo Parade.

Touted as the world's longest, non-motorized parade at a length of 2.1 miles, it marks the beginning of the Tucson Rodeo, the first event of the season on the U.S. professional rodeo circuit.

Organizers say the parade and the events over the ensuing days are a major economic boon to the community with an estimated 200,000 people attending the parade alone every year.

As a native Tucsonan who had never been to the rodeo or the parade, I found that figure hard to believe. It meant that as many as one in every four people living in the county attended the opening day event.

Last year, our count put attendance at roughly a third of the parade committee's official estimates, which ranged from 175,000 to 180,000.

Stats camp instructor Steve Doig, a professor at Arizona State University, talked about several different methods of approaching the problem, including using of aerial photography and software designed specifically for the task.

We decided the parade did not warrant the expense of chartering an airplane or purchasing software. Instead, I opted for a notebook, an Excel spreadsheet and my mountain bike.

continued on page 4

SPOTLIGHT: TOXIC LIVING

Endangered data helps find deadly mercury

By Ken Ward Jr., *The Charleston (W.Va.) Gazette*

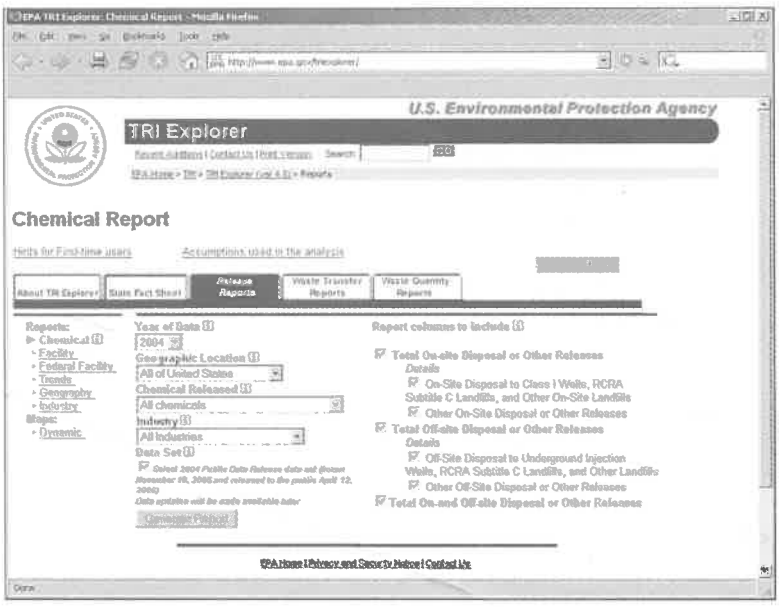
At the annual computer-assisted reporting conferences and on the NICAR listserv, I always hear about how bad government data is. When the conversation steers toward the U.S. Environmental Protection Agency's Toxic Release Inventory, the criticisms seem especially harsh. And they're probably mostly true.

There are lots of problems with the TRI. But right now, the biggest problem is that we might lose it. The EPA is proposing major cutbacks, including a plan to make the now-annual data collection an every-other-year deal for the chemical industry and other manufacturers.

I also know that some of the best stories I do are based in some way on TRI, or at least use some nugget from the TRI for that crucial "computer-assisted paragraph" that makes the story solid.

I wanted to get our own data and develop our own news.

TRI is still the best basic set of pollution numbers we have. With that in mind, here's how TRI helped me turn what could have been a boring daily story from a news release into a more meaningful Sunday and Monday package of stories.



The press release appeared in my e-mail inbox in January 2005. It was one of those things we environmental reporters get all the time. A national environmental group had analyzed some data and published a report.

This time, the group was called Oceana. The topic was mercury pollution – not from coal-fired power plants, from a little-known, but major source called chlor-alkali plants.

I've grown a bit tired of these reports, in large part because they make me feel guilty. These groups are doing our jobs, I tell myself. I should have found that data and done this report as a project or a Sunday story!

But this one caught my eye. It said there were only nine of these chlor-alkali plants in the country, but that, as an industry sector, they rivaled the nation's 500 coal-fired power plants as a source of mercury pollution.

And, the report told me, one of these plants was in my home state of West Virginia.

PPG Industries operates it at Natrium, a dot on the map along the Ohio River in our Northern Panhandle.

Chlor-alkali plants make chlorine by pumping salty water through vats of pure mercury. Some mercury is directly discharged through vent stacks, but huge amounts of it are believed to simply evaporate out of the facility – and still more mercury is somehow "lost" into the environment. The process is more than 100 years old, but it is fast being replaced by newer and cleaner technology.

I knew right away I had to do a story. But I didn't want to just do a quick daily that rewrote the conclusions from the Oceana report. It seemed to me this was a rich topic that deserved more attention from me and my paper.

I talked to one of my editors, who had already seen a wire story on the Oceana report. He was thinking the

same thing I was. We decided right away this had the potential for a big Sunday take-out piece. We set a couple of goals: First of all, my editor wanted me to go to Natrium and visit the plant and find out what it was all about. Second, I wanted to get our own data and develop our own news about this plant's pollution, rather than just quoting from Oceana.

The Oceana report, published in January 2005, used 2002 numbers from EPA's Toxics Release Inventory. By the time I started my story, the 2003 data was out.

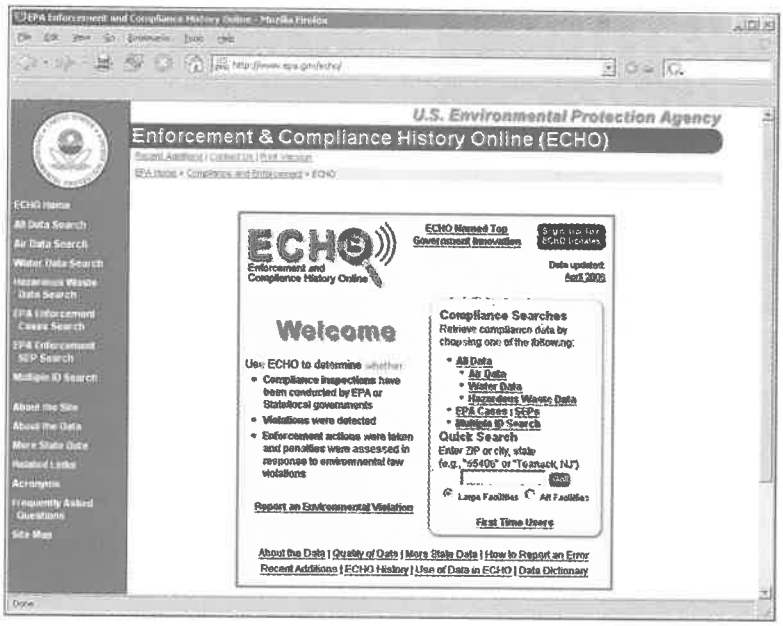
I went to EPA's TRI Explorer Web site (*www.epa.gov/triexplorer*) for data about U.S. facilities that reported discharging mercury into the air or water and about all coal-fired power plants and chlor-alkali facilities.

When I'm using TRI Explorer, I like to first view the results of my queries on the screen. Then, I download the results as comma-separated text files that I can easily open in Microsoft Excel spreadsheet. After I did that, I ranked the top emitters of mercury, both in West Virginia and nationally. I had to do some addition first, though. EPA reports air discharges in two categories, stack emissions and fugitive emissions. For every facility, I used Excel to add the two together for total air emissions.

(Journalists can also purchase the data for a processing fee from the IRE and NICAR Database Library. See *www.ire.org/datalibrary/databases/toxic* for more information.)

During the course of my reporting, I learned the state's water pollution permit for the PPG facility was up for renewal. So, I decided to find out about PPG's water discharges of mercury.

I did a lot of the standard reporting. I went over to the state Department of Environmental Protection and reviewed their permit files going back a couple of rounds of renewals, I interviewed the agency permit engineers and inspectors and – eventually – got the company to give me a tour and several lengthy interviews.



The guts of both stories came from ... simple bits of computer-assisted reporting.

I also used a computer resource that I don't think journalists use enough. I went to EPA's ECHO system, a site (*www.epa.gov/echo*) that provides access to the agency's Permit Compliance System. There, I downloaded the discharge information that PPG reported as part of its water pollution permits and put that into Excel.

We ended up with not just a nice Sunday take-out piece, but with a Sunday-Monday package that ran on 1A both days.

The Sunday story examined the PPG plant and its mercury pollution, as a local example of a little-noticed polluter. The Monday story looked at our state agency's record of poor enforcement at this facility.

The stories weren't published until mid-August, about eight months after the Oceana report came out. The delay was largely because of repeated problems getting a date to visit and tour the PPG plant. But, the additional time gave me a greater chance to learn more about mercury, and to play more with the emissions data.

I used a lot of tools to produce these stories – including spending a lot of time reading boring Federal Register notices and a few thick reports about mercury's health effects. And, of course, I visited the PPG plant and talked to the people who work there.

But the guts of both stories came from a couple of pretty simple bits of computer-assisted reporting that I did with an Internet connection and Microsoft Excel.

Contact Ken Ward Jr. by e-mail at kward@wvgazette.com.

Cargo

continued from page 1

in Southern California that serves as a railroad conduit for growing volumes of goods that have arrived from Asia at the ports of Los Angeles and Long Beach.

Much of our railroad coverage has focused on increasing rail traffic, toxic diesel exhaust from locomotives and long waits at rail crossings. The threat of potentially deadly chlorine gas escaping from a ruptured tank car became a new target for our reporting after the April 4, 2005 derailment.

The San Bernardino wreck, we reported, was a symptom of a national hazard that claimed more lives in 2005 than in the previous 20 years combined. San Bernardino had escaped a disaster. *(www.pe.com/digitalextra/metro/trains/inland.html)*

We started with a database analysis.

The Office of Hazardous Materials Safety, part of the U.S. Department of Transportation, collects data about most hazardous materials spills and other reportable hazmat incidents from all major modes of transportation, including railroads, trucks, boats and aircraft.

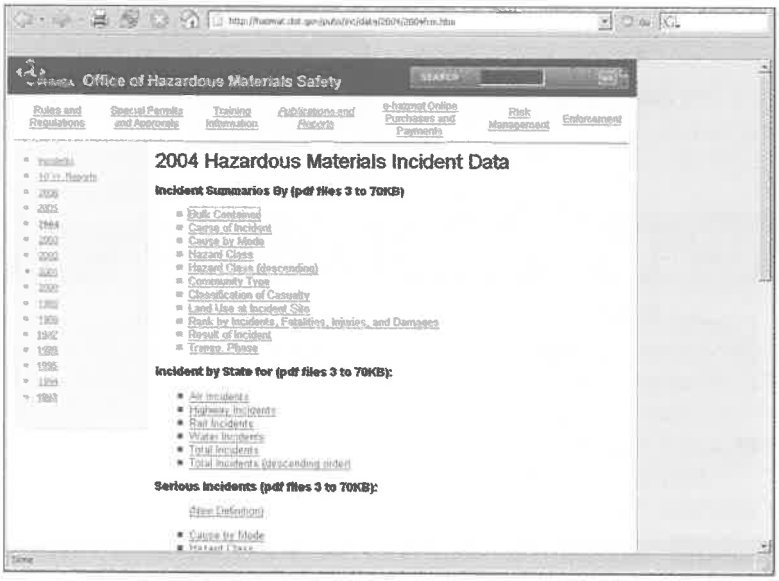
There is a lot information here. The user's guide, *http://hazmat.dot.gov/pubs/inc/userguide.pdf,* explaining the fields, codes and acronyms, is 52 pages. The files can be used to investigate a mode of transportation, a shipping company, or the number and types of problems in a particular city, county or state.

The DOT data is divided into three files for each year: One with basic incident information – locations, mode, commodity spilled, damage estimates, number of evacuees, injuries, death, etc. Another file provides details about the containers. A third file has a text field for incident narrative. Data back to 1993 is posted at *http://hazmat.dot.gov/pubs/inc/data/2004/2004frm.htm.* Note: The formats and fields were changed in 2005, making comparisons with previous years more difficult.

(The IRE and NICAR database library offers a more extensive version of the database to journalists. The data covers 1970-June 2005. See *www.ire.org/datalibrary/databases/databases.php* for more information.)

My first realization was that any meaningful analyses required a database manager. My experience was limited to exercises at workshops, and I wasn't confident enough to run queries.

So I bought "Access 2000 For Windows for Dummies" by John Kaufeld and spent the better part of two days going through the book, trying out the various Microsoft Access functions on the hazmat files.



The next step was to organize the data. With help from a DOT computer technician, I combined 13 files with annual data to create a large database file with all reported hazmat incidents from 1993 through 2004. This file had 188,325 records. Setting the criteria for just railroads cut it to 12,693 records.

My queries then yielded some fruit. Most importantly, we found out that San Bernardino County had more reported train hazardous materials incidents than any another county in the nation during the 11-year period we examined.

The DOT data told us that most of the hazardous material involved in the incidents was just passing through; our area was neither the origin nor the destination. In other words, our region faced the risk of catastrophe without reaping much economic benefit.

To measure the potential human impact of a catastrophic tank car rupture, we collaborated with Ray Carnes, a computer modeling expert with the Redlands, Calif.-based ESRI, which makes geographic information system (GIS) software. He calculated that 1.5 million people in our circulation area live close enough to railroad tracks to be at risk from a major chemical spill from a train.

He used ESRI's ArcGIS 9.1 software to generate the maps and analysis for the article. Railroad data came from the U.S. National Transportation Atlas. County boundary data came from Tele Atlas, a commercial vendor of mapping data, and 2005 residential block population data was extracted from updated U.S. Census information.

Carnes first generated a map document for visualization and analysis of the relationships among the

data. Within this document, he performed an analysis on the railroad data to produce polygons representing all areas within distances of a half-mile, one mile, three miles and five miles from any railroad line.

Census blocks within these polygons were then selected using the Select by Location function. Each block was shaded in a color to represent its distance from railroad tracks. Population numbers were then summarized for each increment in distance from the railroad lines.

ESRI's MapStudio application was used to download additional data for the map – including a digital terrain model.

With the most spills in the nation and more than a million people at risk, it made sense for us to allocate substantial resources to the project.

"Several sources, including a lead emergency planner, said the same thing: It's not a question of whether a catastrophic hazmat train wreck will happen; it's a question of when. We wanted to write this story before then, when it could nudge public officials to be more prepared and to let our readers know of the danger passing by their homes, schools and workplaces every day," projects editor Cathy Armstrong said.

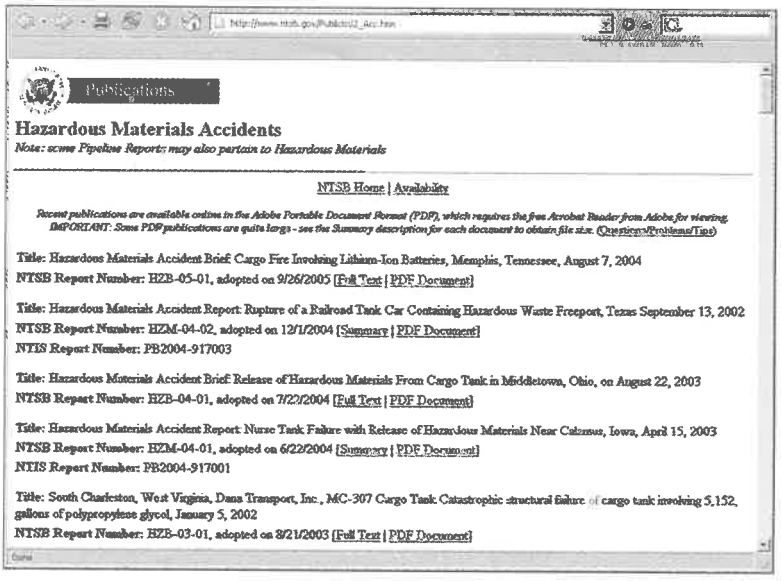
Four reporters were assigned to the project. We filed multiple Freedom of Information Act and state public-records law requests. We traveled to fatal accident sites in North Dakota and South Carolina. We read thousands of pages of reports by the National Transportation Safety Board (*www.ntsb.gov/Publictn/Z_Acc.htm*) and other agencies.

We also gathered 9-1-1 audio, private footage and law-enforcement videos for our online presentation.

Our computer analyses were among several key sources of information for our project. Perhaps the most compelling information came from the accident reports pried loose under FOIA and state public-records law.

But the project was built upon those initial analyses – the sheer volume of hazardous materials incidents here, and the huge number of people who live in the danger zone. Those were the foundations of our 10-page special section and nationally recognized online presentation.

Contact David Danelski by e-mail at ddanelski@pe.com.



Publications

Hazardous Materials Accidents

Note: some Pipeline Reports may also pertain to Hazardous Materials

NTSB Home | Availability

Recent publications are available online in the Adobe Portable Document Format (PDF), which requires the free Acrobat Reader from Adobe for viewing. IMPORTANT: Some PDF publications are quite large - see the Summary description for each document to obtain file size. (Questions/Problems/Tips)

Title: Hazardous Materials Accident Brief: Cargo Fire Involving Lithium-Ion Batteries, Memphis, Tennessee, August 7, 2004
NTSB Report Number: HZB-05-01, adopted on 9/26/2005 [Full Text | PDF Document]

Title: Hazardous Materials Accident Report: Rupture of a Railroad Tank Car Containing Hazardous Waste Freeport, Texas September 13, 2002
NTSB Report Number: HZM-04-02, adopted on 12/1/2004 [Summary | PDF Document]
NTIS Report Number: PB2004-917003

Title: Hazardous Materials Accident Brief: Release of Hazardous Materials From Cargo Tank in Middletown, Ohio, on August 22, 2003
NTSB Report Number: HZB-04-01, adopted on 7/22/2004 [Full Text | PDF Document]

Title: Hazardous Materials Accident Report: Nurse Tank Failure with Release of Hazardous Materials Near Calumus, Iowa, April 15, 2003
NTSB Report Number: HZM-04-01, adopted on 6/22/2004 [Summary | PDF Document]
NTIS Report Number: PB2004-917001

Title: South Charleston, West Virginia, Dana Transport, Inc., MC-307 Cargo Tank Catastrophic structural failure of cargo tank involving 5,152 gallons of polypropylene glycol, January 5, 2002
NTSB Report Number: HZB-03-01, adopted on 8/21/2003 [Full Text | PDF Document]

Tech tip...

Excel & VB help crack school finances

By Duane Schrag, *The Salina (Kan.) Journal*

All across the United States lawsuits have been filed in recent years accusing state governments of failing to provide a suitable education for their citizens. Many – perhaps most – of the cases turn on the question of adequate financing: Are schools being given enough money to educate students most at risk?

Kansas is one of those states. More than two years ago, a county district judge ruled that too little money was being spent to educate low-income, bilingual and special education students. During the litigation the State Board of Education – the principal defendant – commissioned a study to find out whether the Kansas school finance formula was appropriate and, if not, how much it would cost to implement one that was.

The study by consultants Augenblick & Myers proposed a new formula for school spending. The study's findings were presented as evidence in the case, along with the state's estimate of the financial impact of the formula. In ruling for the plaintiffs, the judge repeatedly said the cost of fixing school finance in Kansas would be at least $1 billion. However, the state had submitted a figure of $853 million. And the study estimated it at $236 million.

My story set out to answer the question: How much would it really cost to implement the findings? And how would it affect the 75 school districts in our readership area? I used Microsoft Excel, Access and Visual Basic programming to answer the questions.

Earlier, when I had asked the gurus at the state education department about that, they were dismissive. "Augenblick & Myers got it wrong," they said. I have worked with these people for years, and I took their word for it. If I had listened more closely to my instincts, we would have had this story earlier.

The Kansas Supreme Court later upheld the lower court's ruling, so the funding was in the news again. I requested figures on how the ruling would affect the 302 school districts in Kansas. The Kansas Department of Education sent me the information, with the stipulation that I not attribute it. "Even the school districts haven't seen this," I was told. That was a red flag.

Initially, the thought of calculating the cost on my own was too intimidating. But after looking closely at the formulas, I realized it was not out of the question. It required algebra, not rocket science. Calculating the cost involved:

- formulas to compute weighting factors for district size, low-income, bilingual, and special education. These factors assign more "weight" to low-income students, for instance, which increases state aid for them.

- enrollment figures for each school district, and the subgroups within each district.

The enrollment information was relatively easy to get. The state does a fairly good job of posting it and many other key tables online. Most importantly, the data is usually available as Microsoft Excel spreadsheets, so rekeying is not necessary. If possible, get this data in electronic, not paper, format. Transcription errors are too easy to make.

In my calculations, I had to make sure I was using full-time equivalents when those were required, headcounts when headcounts were required. I had to re-run the data after learning the special education figures posted online were incomplete – they excluded children in gifted programs.

The formulas themselves are straightforward – a series of If/Then statements, with assorted multiplication and division operations. Excel has a built-in If() function that I found too limited for my purposes. Instead, I wrote custom functions in Visual Basic.

Example 1 examines the adjustment for school size. This is how the consultant presented the formula for a base budget per pupil of $4,550.

Example 2 is the equivalent in Visual Basic. I named the function size_weight(), which takes as arguments the base budget per pupil (b) and the district's enrollment (e).

You may notice the base budget per pupil is a variable, not a constant – this made it possible to compute the cost for any base figure. For our story, we used $5,000, which was the study's original figure adjusted for inflation.

With the formulas in hand, and the raw numbers in spreadsheet format, it was a breeze to compute the weightings, and thereby create the exact budget figure for each school district. At this point, I knew how much

the proposal would cost, but that was only half the problem.

The other half of the problem was comparing the proposal to the status quo. Again, finding the raw data was not difficult – every school district's budget is posted online, in sufficient detail to extract the information needed for the story.

Extracting the correct figures required some work. In order to compare apples and apples, I needed to deduct expenses for transportation, capital outlay, food service and adult education. The budget is a massive spreadsheet – it consists of 14 separate sheets, each with 302 rows (one for each school district) and as many as 20 columns of figures. Spending for the services in question were scattered across those sheets.

First, I sought help from the local school district. I sat down with the business manager and made certain I knew which columns had to be deducted to account for the necessary change. There were dozens.

Once again, I wrote a Visual Basic function to systematically extract the necessary information. While it is possible to do this by hand, the number of operations (about 10,000) makes it highly likely a mistake will be made.

Now I could make the proper comparison. I had the true cost of applying the proposal, district by district. And I had the actual comparable cost under the existing formula. I used Microsoft's Excel and Access to manage the data, although relying solely on Excel would be possible.

A&M's model was never adopted, so it is impossible to say what it would have ended up costing. But in January 2006, an exhaustive study by the Kansas Legislative Division of Post Audit on the "true" cost of providing an education said an additional $400 million is required; this comes on top of an extra $143 million added last year. That put the additional spending right where we said A&M would have put it.

Perhaps the most important lesson this story reinforced for me is that journalists must do a tricky balancing act – know when to ask for information, know when to do the work yourself.

Sources for this story gave wildly different interpretations of what the A&M study said. Some who should have known it backwards and forwards were dead wrong. Only by reading it, re-reading it, discussing it, and re-reading it further was I able to offer the story's conclusions with confidence.

Contact Duane Schrag by e-mail at dschrag@saljournal.com

(Example 1)

```
Fewer than 430 students = {[(430 - Enroll)/10 X .01] X 4,550} + $5,852
For 430 - 1,300 students = {[(1,300 - Enroll)/80 X .01] X 4,550} + $5,358
For 1,300 - 11,200 students = {[(11,200 - Enroll)/600 X .01] X 4,550} + $4,550
More than 11,200 students = $4,550
```

(Example 2)

```
Function size_weight(b As Long, e As Long) As Long

If e < 430 Then
    size_weight = (((430 - e) / 10 * 0.01) * b) + (b * 1.2862)
Else
    If e < 1300 Then
        size_weight = (((1300 - e) / 80 * 0.01) * b) + (b * 1.178)
    Else
        If e < 11200 Then
            size_weight = (((11200 - e) / 600 * 0.01) * b) + b
        Else
            size_weight = b
        End If
    End If
End If

End Function
```

ER

continued from page 1

would be a six-hour delay.

I said, "I'll take him to Sinai Hospital."

She said, "Sinai is on an eleven-hour delay."

I couldn't believe it. There didn't seem to be any option but to wait, and it gave me lots of time to wonder what's going on inside Baltimore ERs.

I was a bit bleary-eyed when I got to work the next day. John and I had spent more than eight hours at Northwest even though he had been "fast tracked." I began brainstorming with my unit about how we could investigate central Maryland ERs.

We realized there would be plenty of obstacles. Privacy laws prevented us from taking hidden cameras inside a hospital waiting room, and we didn't have the manpower to monitor two dozen hospitals on a daily basis. We tried calling all of the ERs on various nights. Some hospitals said it was against their policy to give out wait times, but others did. We heard five hours, six hours, seven hours and more. One frustrated nurse told me, "There are 37 patients in the waiting room, and some have been waiting for hours. It's a very long wait!" We knew we were on to something, but how could we accurately quantify it?

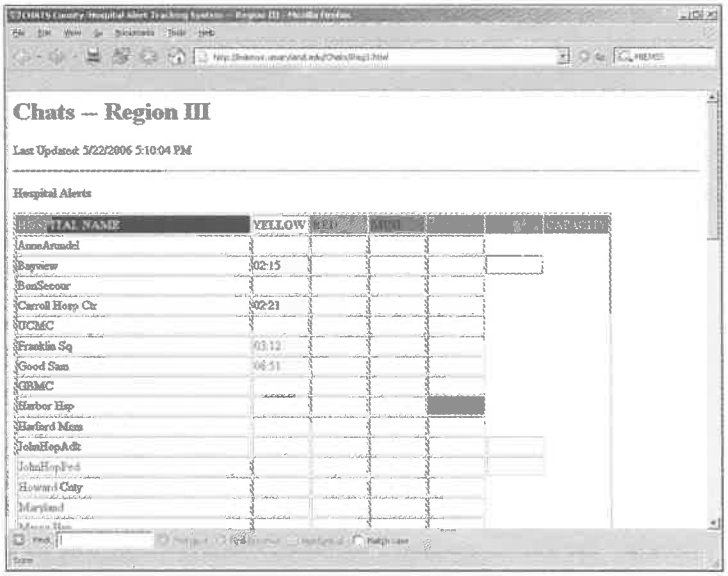
The answer came when we discovered a Web site maintained by the Maryland Institute for Emergency Medical Services Systems (MIEMSS). This state agency keeps track of when hospital ERs go on alert, so that medic units can be diverted. By monitoring the site, which reports in real time, we quickly became aware that many hospitals in our area were often on Yellow Alert. According to MIEMSS, that means the ER is experiencing an overwhelming load and patients with urgent medical needs should not be brought there. The Web site would prove to be a gold mine of information.

We found the state kept an archive of the Web site data. We asked the agency to provide us with 32 days of information. More specifically, we asked for the time when a hospital went on alert, what type of alert, when the alert was canceled, and the duration. The state was very cooperative. Within a matter of days, we had all the data. It arrived via e-mail at no cost in Microsoft Excel spreadsheet format, which was great for me. I'm no computer whiz, but have worked with Excel. There were 645 rows and eight columns of data.

At first, the duration of an alert was difficult to interpret. The state maintained that data in hours and tenths of hours (rather than hours and minutes), but agreed to run a conversion for us. Once we cleared that up, our first discovery was overwhelming. By simply sorting the data by hospital and duration of alert, we were able to calculate that the ER at Johns Hopkins Hospital was on Yellow Alert 48 percent of the time. Northwest's ER, where I took my husband, was on Yellow Alert 41 percent of the time, and seven other hospital ERs in our region were on Yellow Alert 22 percent to 28 percent of the time. We also found most Yellow Alerts were issued on Mondays (the same day of the week that I had been there) by sorting by type of alert and day of the week.

We also discovered two hospitals on Red Alert 24 percent and 30 percent of the time. Red Alert means a hospital cannot receive any patients, even those who are critically ill, because there are no cardiac-monitored beds available inside the hospital's inpatient units. But it doesn't necessarily mean that patients are turned away, as we first believed. Rather, they would likely be stabilized at that hospital, and then transferred to another facility. We did not become aware of this distinction until after our report aired. We issued a clarification.

What the data could not provide was



a sense of how long patients were actually waiting. For this, we sent WBAL producer Joyce Karp and photographers Chuck Cochran and Greg Marsh, inside separate ER waiting rooms, during what we thought would be busy times. Cochran and Marsh did not carry cameras but helped gather information. Not wanting to intrude, all three simply approached patients, introduced themselves, and asked how long they had been waiting. Then they asked if they could contact them at a more appropriate time. Several of these people later spoke on camera, and told stories of long waits, some as long as 10 hours.

In addition, investigative reporter David Collins set up a ride-along with a paramedic, who told him that ER gridlock often prevents him from quickly getting patients into hospital beds. He said it's not unusual to find patients stacked up on gurneys in hallways, because no beds are available.

The danger, experts say, is the longer the wait, the worse a problem can become, and it makes patients more susceptible to catching colds, the flu, or a contagious disease. In fact, the state health department told Collins that there were two recent cases in which emergency room waits could have resulted in the deaths of two patients.

All of this, together with the data analysis, exposed how serious the gridlock was, and helped us build a solid investigation. And although the hospitals, health officials and other experts disagree upon the cause or solutions to long ER waits, none questioned our data. In fact, the president of a health care advocacy group told Collins that our investigation exposed a crisis in health care in central Maryland, and unless it is fixed and hospitals are held accountable, long waits in the ER will no longer be considered a crisis, but rather, a disaster.

Contact Augusta Brennan Jones by e-mail at abrennan@hearst.com.

IRE and NICAR Services

Investigative Reporters and Editors, Inc. is a grassroots nonprofit organization dedicated to improving the quality of investigative reporting within the field of journalism. IRE was formed in 1975 with the intent of creating a networking tool and a forum in which journalists from across the country could raise questions and exchange ideas. IRE provides educational services to reporters, editors and others interested in investigative reporting and works to maintain high professional standards.

Programs and Services

IRE Resource Center: A rich reserve of print and broadcast stories, tipsheets and guides to help you start and complete the best work of your career. This unique library is the starting point of any piece you're working on. You can search through abstracts of more than 20,000 investigative reporting stories through our Web site.
Contact: Beth Kopine, beth@ire.org, 573-882-3364

Database Library: Administered by IRE and the National Institute for Computer-Assisted Reporting. The library has copies of many government databases, and makes them available to news organizations at or below actual cost. Analysis services are available on these databases, as is help in deciphering records you obtain yourself.
Contact: Jeff Porter, jeff@ire.org, 573-882-1982

Campaign Finance Information Center: Administered by IRE and the National Institute for Computer-Assisted Reporting. It's dedicated to helping journalists uncover the campaign money trail. State campaign finance data is collected from across the nation, cleaned and made available to journalists. A search engine allows reporters to track political cash flow across several states in federal and state races.
Contact: Brant Houston, brant@ire.org, 573-882-2042

On-the-Road Training: As a top promoter of journalism education, IRE offers loads of training opportunities throughout the year. Possibilities range from national conferences and regional workshops to weeklong boot camps and on-site newsroom training. Costs are on a sliding scale and fellowships are available to many of the events.
Contact: David Donald, ddonald@ire.org, 573-882-2042

Publications

The IRE Journal: Published six times a year. Contains journalist profiles, how-to stories, reviews, investigative ideas and backgrounding tips. *The Journal* also provides members with the latest news on upcoming events and training opportunities from IRE and NICAR.
Contact: Brant Houston, brant@ire.org, 573-882-2042

Uplink: Bimonthly newsletter by IRE and NICAR on computer-assisted reporting. Often, *Uplink* stories are written after reporters have had particular success using data to investigate stories. The columns include valuable information on advanced database techniques as well as success stories written by newly trained CAR reporters.
Contact: David Herzog, dherzog@ire.org, 573-882-2127

Reporter.org: A collection of Web-based resources for journalists, journalism educators and others. Discounted Web hosting and services such as mailing list management and site development are provided to other nonprofit journalism organizations.
Contact: Matthew Dickinson, matt@ire.org, 573-884-7321

For information on:

Advertising: Megan Means, meganm@ire.org, 573-884-2360

Membership and subscriptions: John Green, jgreen@ire.org, 573-882-2772

Conferences and Boot Camps: Ev Ruch-Graham, ev@ire.org, 573-882-8969

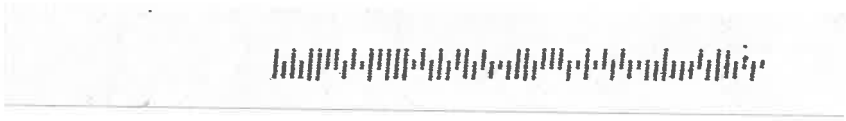
Listservs: Matthew Dickinson, matt@ire.org, 573-884-7321

Mailing Address: IRE, 138 Neff Annex, Missouri School of Journalism, Columbia, MO 65211

Investigative Reporters and Editors, Inc.
138 Neff Annex
Missouri School of Journalism
Columbia, MO 65211

NON-PROFIT ORG.
U.S. POSTAGE
PAID
Jefferson City, MO.
Permit NO. 89

Ira Chinoy
Philip Merrill College of Journalism
1117 Journalism BLDG
University of Maryland
College Park MD 20742-0001



Uplink Info

A newsletter of the National Institute for Computer-Assisted Reporting

NICAR is a joint program of Investigative Reporters and Editors Inc. and the Missouri School of Journalism.

NICAR services include supplying journalists with government databases, training programs, tipsheets and data analysis.

Editorial
573-884-7711

Uplink Address:
IRE-NICAR, 138 Neff Annex
Missouri School of Journalism
Columbia, MO 65211

Postmaster: Please send address changes to IRE-NICAR.

Uplink