IRE
N·I·C·A·R
Uplink
July/August 1999
A newsletter of the National Institute for Computer-Assisted Reporting

INVENTING A SCHOOL SYSTEM

Busting busing

By Ted Mellnik
The Charlotte Observer

What if kids went to their neighborhood school? That's the big question in Charlotte.

After the landmark 1971 Supreme Court decision upholding court-ordered busing here, the city was touted as the place where busing worked. But support has eroded over the past decade. Now a court challenge to race-based assignment policies threatens to dismantle Charlotte's system of busing for racial integration.

To explore the issues and challenges that would accompany a move to neighborhood schools, *The Charlotte Observer* designed a mock school system. In our system all kids went to the school (elementary, middle or high) closest to where they lived.

Our model system is simplistic and probably impractical, but it gave us a way to spotlight potential changes and trends. We believed that if we fully explained our methods and assumptions, our model could contribute to our community's debate over neighborhood schools.

The stories ran in January 1999 as the first five parts of "Deciding Desegregation," a continuing series. Reporter Debbie Cenziper and I described the forces of change acting on the schools. We compared schools in our system to existing schools, tracking ethnic diversity, concentrations of poverty, home-to-school distances and campus crowding.

Geography lesson

The key ingredient of the Deciding Desegregation series was geographic data on each of about 94,000 students.

We regularly use student data at *The Observer* because of the range of possibilities it offers for analysis. Most often we deal with results of tests such as the SAT and North Carolina's end-of-grade and end-of-course exams. The data has an entry for each student with key facts, including information we used for the Deciding Desegregation series such as a student's race, gender, grade in school and school attended.

But we still needed to find out where students lived.

Geographic data is a set of coordinates, like latitude and longitude. These sorts of data would allow us to find a spot on earth near each student's home and display it as a dot in a computer mapping program.

School officials have on hand geographic

Conintued on page seven

DADS PAY LITTLE

Welfare fails

By Jim Hopkins
The Courier-Journal

For the first time since welfare reform began in Kentucky, a growing body of evidence shows that former welfare recipients in the state's chronically poor Appalachian region are failing to achieve one of reform's most important goals: finding jobs that pay a living wage.

Welfare recipients in Kentucky's mountain counties also are more burdened by lax enforcement of child-support laws. Although welfare reform requires mothers to find work, many fathers who owe child support pay nothing, *The Courier-Journal* found in a six-month investigation.

Our inquiry came down to people like Sheila Stone, a former welfare recipient trying to support four children on a $5.50-an-hour bank job. Her ex-husband hid from the law because he owed more than $40,000 in child support. Stone and the authorities didn't know where her ex-

Continued on page two

Inside Uplink

This issue features several stories noted in the annual "Best CAR stories this year" panel at the National IRE Conference.

Ted Mellnik of *The Charlotte Observer* tells how its series evaluated school busing by inventing a school district.

Alan Miller of the *Los Angeles Times* explains how to find 10-year-olds giving thousands to politicos. See page 3.

Jack Dolan of *The Hartford Courant* found slow ambulances that led to deaths. See page 4.

Linda Goldston of the *San Jose Mercury News* writes about her first CAR story on trading exotic zoo animals. See page 8.

Bennett Cunningham of KTVT-TV (Dallas) found pharmacists giving away deadly drugs to the wrong people. See page 10.

NICAR's Training Director Tom McGinty explains new software that converts .PDF files into usable database files. See page 16.

PAGE 15
On the Internet

PAGE 18
Stats and Such

PAGE 20
Handouts of the Month

Uplink

July/August 1999
Volume 11, Number 6
A newsletter of the National Institute for Computer-Assisted Reporting

EDITOR
Brant Houston
brant@nicar.org

MANAGING EDITOR
Cindy Eberting
cindy@nicar.org

SENIOR CONTRIBUTING EDITORS
Len Bruzzese
len@ire.org
Jessica Larson
jessica@nicar.org

COPY EDITOR
Jessica Larson

ART DIRECTOR
Wendy Gray

STAFF
Chen Feng-Lan
Britton Clap
Jo Craven
Dawn Fallik
John Green
Seth Jayson
Tom McGinty
Mareesa Orth
Ted Peterson
Tom Philipp
Noemi Ramirez
Laura Ruggerio
Aaron Springer
Neil Warner

Uplink is published every month by the National Institute for Computer-Assisted Reporting, 138 Neff Hall Annex Columbia, MO 65211. (573) 882-0684. Subscriptions are $40 for IRE members, $60 for nonmembers. Postmaster: Please send address changes to NICAR. Send e-mail to jgreen@nicar.org NICAR is a joint effort of Investigative Reporters and Editors and the University of Missouri School of Journalism. NICAR services include hands-on newsroom training in computer-assisted reporting, special academic and advanced training in data analysis.

From page one:

Welfare

husband was. As part of our investigation, we tracked him down and eventually reported his return to jail.

Reaching our conclusions was anything but simple, however. We fought for electronic databases from the state welfare agency responsible for reform. We relied extensively on data from the Census Bureau, the U.S. Bureau of Economic Analysis and AutoTrak, a for-pay service that locates people. And we used paper records from the welfare agency and the state's employment services division to build databases of our own.

The face of welfare

From the beginning, we suspected the biggest challenge for Kentucky's welfare reform efforts would be the tens of thousands of women and children in the state's 49 Appalachian counties, home to some of the nation's poorest residents. Indeed, although Appalachia has 30 percent of the state's population, it has nearly half of all Kentucky's welfare cases.

One of our chief sources of data was a statewide telephone survey of 522 former welfare recipients conducted by the University of Louisville. We persuaded Kentucky welfare officials to give us the raw data from the survey, which the university conducted under a state contract. The survey asked a host of questions, from the reasons recipients left welfare to whether life was better without government assistance.

We divided the responses into three groups: Appalachian counties, other rural counties and urban counties. Using Paradox and Excel to analyze the results, we immediately saw significant differences in the three regions. For example, only 41 percent of the Appalachian women said they left welfare because they found a job, compared with 55 percent in the urban areas. Examining state welfare records received on paper and fed into a separate database, we found that child support was collected by only 45 percent of the women in Appalachia, versus 53 percent in the rest of the state.

The next step involved downloading state job-growth projections from the Internet, then sorting those into the same geographic groups. Our analysis showed that the Appalachian counties were expected to have half the job growth of urban counties over the next six years. This suggests there would be even more competition in the region for future jobs as additional women try to quit welfare. And where jobs are created in Appalachia, a disproportionate share will be in the low-wage ends of retailing and the service industry, our analysis showed.

The state's biggest welfare database turned out to be too unwieldy ... Learning to navigate our way through the data would have taken more time than we had.

We decided to tell the story through the experience of Stone, 30, and her ex-husband, Antoine Stone, 32, a formerly well-paid computer programmer evading an arrest warrant for failing to pay child support. Relying partly on AutoTrak, a service I first learned about through NICAR, we found Antoine Stone in Detroit, where he agreed to an interview.

We spent five days with Sheila Stone and her children to document the reality of trying to move off welfare, then contrasted that with the life Antoine Stone led on the run. Our hope was that readers would be pulled into the series by telling Sheila's story one day, then Antoine's the next.

Dealing with obstacles

The state's biggest welfare database turned out to be too unwieldy; it consisted of more than 80 individual, linked datasets, some of which included confidential information. Learning to navigate our way through the data would have taken more time than we had. In the end, we settled on using the university's survey results and on less detailed paper records that were more accessible.

Jim Hopkins can be reached by e-mail at hopkinsj@courier-journal.com.

Kiddie donors

By Alan C. Miller
Los Angeles Times

They are the most precocious of political donors: youngsters who are not old enough to drive a car, never mind cast a vote, yet they dispense $1,000 checks to candidates for public office.

The suspicion, of course, is that these school-age donors are the beneficiaries of a generous onetime boost in their allowance from parents looking to circumvent federal contribution limits. At $1,000 per election, family bundles are a tempting – and potentially illegal – way for wealthy givers to double or triple their largess to a favored candidate.

Federal law places no minimum age on donors but requires that the funds be "owned or controlled exclusively by contributors" and that they "give knowingly and voluntarily."

Consultant helps

Our curiosity about the extent of this practice was piqued by contributions from three students in a single family to a U.S. Senate candidate from California in 1998. So, we retained independent researcher Dwight Morris and his Campaign Study Group in Virginia, to provide a definitive set of numbers on federal contributions and a road map for further reporting.

The FEC regards student giving as such a potentially serious loophole that it has repeatedly urged Congress to ban donations by minors.

Starting with the Federal Election Commission records of contributions of $200 or more between Jan. 1, 1991 and Dec. 31, 1998 – a total of 3.9 million donations – Morris searched for donations to federal candidates or party committees by those identified as "students."

He then scrubbed the initial hits to locate additional money from those identified as students for some contributions but not others (donations often include no information listed for occupation).

Conversely, he eliminated other donors who were listed as students but who proved to be housewives taking a college course or professionals who had gone back to school.

The tally: 8,876 student donations totaling $7.5 million. And the trend lines showed the numbers were growing.

Using this database, Morris compiled a list of major student donors, including many of the 163 who gave a total of $5,000 or more. This was our starting point.

Since the FEC does not list ages, we had to learn more about these donors to determine whether they were law school students who might be contributing themselves or preadolescents whose largess was less likely to be their own.

Our strategy was to eliminate the older students through public records before approaching them and their parents.

Profiling givers

Using online newspaper searches and voting and motor vehicle records, we built files on the top 50 donors who gave $13,700 to $93,250 each. This included a 15-year-old who donated $20,000 to the Democratic Congressional Campaign Committee.

But we still needed more truly young givers that we believed were out there.

We then focused on students who gave smaller sums and were therefore more likely to include youngsters in affluent California counties.

Through the FEC records, we identified those whose donations in recent election cycles came at about the same time that their parents gave the legal maximum to the same candidate. We then eliminated those who were eligible to vote by checking their names with county registrars of voters.

Underage donating

The final step was to nail down a specific age for those not registered to vote or licensed to drive.

Here, we found a wide disparity in the

Continued on page five

To learn more about campaign finance coverage, attend the Money in Politics seminar, Sept. 17-19 in Des Moines.
Hands-on NICAR classes included. See *www.ire.org/training/iowa.html* for more info and online registration.

You can also order tapes from past campaign finance panels.

- From the 1999 IRE Conference:
"Soft money goes Stateside: Get ready for Campaign 2000"
- From the 1999 NICAR Conference:
"Giving at Home"
"Federal Campaign finance: Before and after the elections."
- From the 1998 NICAR Conference:
"State and Local Campaign Finance"

To order a tape, call the IRE Resource Center at (573)882-2364 or visit *www.ire.org/store.*

Check out *www.nhtsa.dot.gov/people/injury/ems/StateEMS/nasemsd.list.htm* to find a listing of each state's EMS office and contact information.

Read or watch other stories about ambulance response times, available through the IRE Resource Center at *www.ire.org/resourcecenter* or call (573)882-3364. These stories include:

A 1995 story on WTVT-TV (Tampa, Fla.) that showed slow ambulances may have led to deaths in a Florida county. Ask for story #12822.

A 1995 investigation by the *Post-Bulletin* (Rochester, Minn.) of the Rochester Fire Department and its false alarms, response times and how its firefighters spend their time. Ask for story #12970.

A 1995 story on WKRN-TV documenting slow response times and an ambulance shortage. Ask for story #12834.

Cardiac crapshoot

By Jack Dolan
Hartford Courant

When Mike McIntire handed a cardboard box of diskettes across the divider and said, "I don't know if you can do anything with these," I knew I'd taken the right job.

It was my first day at the *Hartford Courant*, my first job since leaving NICAR. McIntire, my new colleague on the investigative team, was working on a story about deathly slow response times by American Medical Response, one of the nation's largest private ambulance companies. AMR made several promises to Hartford officials to wrest the 911 business away from smaller, local companies including responding to 90 percent of serious medical emergencies within six minutes.

Disks to database

McIntire had heard plenty of stories about AMR failing to meet that standard, but he knew he needed hard data to see if the anecdotes were part of a larger pattern. To that end, he got two years' worth of 911 call records from the Hartford police, who dispatch the ambulances. Each month of records came on a separate diskette as a tab-delimited text file. Those were the disks in the box Mike handed over the divider.

Before long, we'd appended all of the files into a single FoxPro table, calculated the response times, and sorted the 50,000 calls so that the one with the longest response time was on top. Then we picked through the data for what doctors say are indisputable emergencies: cardiac arrest, chest pains, and difficulty breathing.

We found that 34 percent of these most vulnerable patients waited more than six minutes for their ambulance – three times the number the AMR contract allowed. Hundreds of verified heart attack victims waited longer than ten minutes. Some waited longer than twenty minutes. Over two years, the number of people put at serious risk by long response times reached into the thousands.

Emergency medicine experts across the country told us that dangerously long response times are bound to happen once in a while. That's why ambulance companies, when negotiating contracts, make sure to give themselves a cushion 10 percent of the time; 34 percent shocked the experts. We were told that if we wanted to discover what happened to those who waited the longest, the best place to start checking was the morgue.

Joining death records

So we requested a database of all recorded deaths in Connecticut from the Department of Health. Among the fields in the death data were the name and address of the deceased and the date of death. The 911 records had no names of victims, but they did have the date of the call and the victim's address.

Once we had the hard numbers that showed just how long AMR's response times are, the rest of the CAR work just helped make an already strong story stronger.

We joined the databases on two criteria. The first was where the addresses matched. The second was when a death was recorded within one day after a wait of more than 10 minutes after a 911 call for someone suffering from one of those three most serious emergencies. A sad roll call of silent victims filled the screen. Through interviews with these victims' families, police, EMTs and AMR executives, we discovered that, like everyone else who waited too long for an ambulance, most of the dead fell into three distinct categories.

Mapping the calls

First, some seemed to have to wait too long for the ambulance because of where they lived. We grouped all of the addresses of calls that took longer than six minutes into census tracts, then mapped the tracts using ArcView. Each tract was a different shade depending on the percentage of calls that took too long – the darker the color, the longer the response time.

The darkest tracts were on the extreme northern, southern, and western edges of the

Continued on page five

From page four:

Response times

city. We expected a lot of long response times in the north, since it is the poorest part of Hartford. But the west is the wealthiest, and the darkest-colored southern tract is home to the mayor and many long-time Democratic city bosses. So much for the old adage about city services being best on the mayor's own street.

The explanation, it turned out, is that AMR retrenched on an agreement to keep a certain number of ambulances based on the edges of the city at all times. Instead, the company reduced the number of ambulances available, rotating those that remain between the city's hospitals, which are located in the center of Hartford.

A second apparent predictor of whether an ambulance would come quickly was the time of day. We expected to find that most of the longer calls happen in the early evening, when the 911 call volume is highest. Our logic was simple: if two people are waiting for one ambulance, somebody has to wait longer.

Instead, we found response times are longest at mid-morning when 911 call volume is very light. Several EMTs explained that the peak hours for the lucrative business of shuttling non-emergency, well-insured patients home from the hospital are between 10 a.m. and noon, just after doctors have made their rounds and discharged patients. Therefore, AMR is loath to send an ambulance off on a 911 call during those hours since the company has a much better chance of getting paid for the high priced taxi service.

A third contributor to long ambulance response times was the medical ignorance of some police officers thrust into the ambulance dispatcher role, often with very little medical training. We found 20 percent of confirmed cardiac arrest cases were not initially classified by police dispatchers as emergency calls, so the ambulance driver was not put on the highest alert.

Since our investigation, the *Courant* ran several stories on AMR's poor performance in Connecticut and the company's problems holding onto contracts in other parts of the country.

Story prompts suit

Finally, in early June, the state of Connecticut reached an anti-trust settlement with AMR. The settlement forced the company to give up 30 ambulance licenses to its competitors and give up 40 percent of the Hartford market.

In Hartford, ambulances are dispatched by the police and not by the ambulance company. That was key to our investigation, because it meant the dispatch database was a public record.

Once we had the hard numbers that showed just how long AMR's response times are, the rest of the CAR work just helped make an already strong story stronger.

Jack Dolan can be reached by e-mail at dolan@courant.com

See related tipsheets on mapping from the 1999 IRE Conference panel "Mapping as an Investigative Tool," Tipsheet #971 and Tipsheet #970.

These tipsheets are available online at *www.ire.org/resourcecenter*

From page three:

Donors

national public availability of birth records. Fortunately, we were able to access them in Los Angeles County.

Dr. Scott Stolnitz, a Marina del Rey dentist, initially told me he wanted nothing to do with our story. When I subsequently called back to say that his daughter was going to be prominently featured, he said, "You know nothing about my daughter."

I replied that I knew she was 10 years old when she gave $1,000 to Republican presidential candidate Lamar Alexander the same day that both her parents also gave the maximum.

"OK," he said, "Here's the story." (Which turned out to be that he decided to make the contribution on his daughter's behalf from a checking account that he funds.)

With our cases in hand, further reporting revealed that the FEC regards student giving as such a potentially serious loophole that it has repeatedly urged Congress to ban donations by minors, based on the presumption that contributors below age 16 are not making contributions on their own behalf.

Yet, since 1975, the oft-criticized watchdog has assessed just one fine arising from minors' donations.

Alan Miller can be reached by e-mail at alan.miller@latimes.com

Check out Tipsheet #795 from the 1998 NICAR panel "State and Local Campaign Finance." This tipsheet is available online at *www.ire.org/resourcecenter*

The National Association of Home Builders' Web site offers contact names, addresses and phone numbers for each state's local associations at *www.nahb.know-where.com/nahb/*

Check out *www.BuildingOnline.com/buildingindices.shtml* for a list of various architect, builder and contractor group Web sites and contact information.

See related tipsheet from the 1995 IRE Conference for the panel on community development block grants. This packet includes phone numbers of deputy inspectors general, a chart of available grants, a key to CDBG and Home Reports and a sample audit report from the office of the inspector general. Tipsheet #355 can be ordered from the Resource Center at (573) 882-3364.

WARNING HOMEOWNERS

Shoddy shacks

By Ames Alexander
The Charlotte Observer

Looking for a story that will bring about an avalanche of phone calls and e-mail from grateful readers? Have we got one for you.

In December, *The Charlotte Observer* published "Home Buyer Beware," a six-day series that detailed how Charlotte's housing boom had swamped builders, overwhelmed inspectors and left scores of homeowners with costly and dangerous problems.

Across the Carolinas, thousands of families have discovered serious construction flaws in their new homes, ranging from shifting foundations to walls in danger of collapse. The homes belonged to people like Alan and Marie Nelsen, who vacated their new, half-million-dollar house in an exclusive lakefront neighborhood after an engineer reported that it was uninhabitable. People like Michael and Chona Hoy, who say serious construction defects led to a 1996 blaze that destroyed their new house.

How could it possibly happen? Our research showed that builders are routinely hiring workers with little or no experience. North Carolina's laws favor builders over buyers, and overworked building inspectors are missing serious flaws.

Checking inspectors

During our eight-month project, reporter Rick Rothacker and I accompanied building inspectors on their daily rounds, watched houses get built, pored through public records and interviewed more than 400 homeowners, builders, inspectors and others. And with the help of database editor Ted Mellnik, we used a variety of computer files to add power, authority and detail to our reporting.

Local officials said they couldn't provide precise statistics on the workloads of inspectors, so we got the raw data – four million computer records on all building inspections conducted in Mecklenburg County since the 1970s. Among the conclusions the computer data helped us draw:

• In Charlotte, construction flaws were widespread. The data showed that public building inspectors were rejecting home builders' work in roughly 4 of every 10 inspections.

• Building inspectors were swamped. Inspectors in Mecklenburg were averaging 18 inspections per day. As a general rule, experts say, code officials should average about 10 inspections a day. On some days, code enforcers were doing more than 70 inspections – and spending an average of less than eight minutes on each of them.

• Whether a house passed an inspection depended in part on the inspector. In 1997, one inspector failed 62 percent of the buildings he inspected. One of his colleagues, though, failed just 17 percent.

Our research showed that builders are routinely hiring workers with little or no experience.

The Observer also conducted a survey of Charlotte-area residents who'd recently bought new homes and found that more than 30 percent wouldn't recommend their builder to a friend. To find all those homeowners, *The Observer* obtained databases from area counties with information about new house permits. Meanwhile, a North Carolina Department of Labor database allowed us to show exactly how brisk business had been for the region's home builders.

Helpful sites

We also used a variety of Web sites. Among the handiest was one maintained by the National Association of Home Builders: *www.nahb.com* . The group collects reams of information on new homes and the companies that build them – including survey data on how builders cope with the labor shortage and how much profit they're pocketing.

Some advice for reporters who hope to delve into construction issues:

• By all means request your county's building inspections data, but be prepared for some fancy computer work. For our project, Mellnik untangled the unwieldy hierarchical database provided by Mecklenburg County and put it in a form that allowed me to analyze it with Microsoft Access.

Conintued on page seven

From page one: # Schools

data for each student, drawn from their home addresses, in order to analyze attendance zones. In our data request for the school district, we argued successfully that with names and addresses stripped out, the geographic coordinates would not be invasive of student privacy: the coordinates could not be converted to addresses and used for purposes such as mass mailing or telephone solicitation. We also argued that releasing the data would serve an important public purpose because of high community interest in where kids go to school.

To this basic information we added four items for each student record: straight-line distance from home to school attended; ID of the nearest school; distance to the nearest school; and the census block group of the student's home. We used the student data, computer maps of school locations and the desktop computer-mapping software Atlas GIS for these calculations.

Time on the bus

One of the first patterns to emerge was the disparity in the way the burden of busing is distributed. Blacks bear more than whites.

That is, we found black students travel a longer distance to school. Also, more whites than blacks (54 percent compared to 32 percent) attend the closest school.

When we attempted to measure the impact of so-called forced busing by counting students who attend neither their nearest school nor a magnet school, we found 42 percent of blacks are subjected to forced busing compared to 23 percent of whites.

We built one part of the series around this existing difference, as well as the fact that in our model, distance to school dropped twice as much for black students as for whites when compared with their distances to current schools.

The death of desegregation?

Our other findings included:

• Under neighborhood schools, more than half (64) of the schools would be largely one race, compared to 19 now. Our measure was a school-level ethnic diversity index or the chance that two students picked at random from a school would be of different races.

• Neighborhood schools would create a group of highly impoverished campuses concentrated in the inner city. We measured poverty by assigning to each student the poverty rate of their home census block group.

• Under neighborhood schools, the number of students on overcrowded campuses would double and building classrooms to house them would cost hundreds of millions of dollars. Meanwhile, classrooms in some schools would stand empty. We defined overcrowding as 120 percent or more of building design capacity, not including mobile units on site.

In response to the series, we received more than 500 phone calls, e-mail messages, faxes and letters from readers, most of them appreciative. Our series has been the only attempt to explore the potential impact of neighborhood schools here. When questions arose about neighborhood schools during a series of community meetings conducted by school officials, our series was cited.

You can find the series online at: *www.charlotte.com/observer/special/deseg/index_jan.htm*

Ted Mellnik can be reached by e-mail at Tmellnik@charlotte.com .

Check out *www.schoolbusfleet.com/Resource/index.htm* to find surveys, government bills, statistics and other resources for reporting on busing stories.

See related tipsheet from the 1995 IRE Conference for the panel on desegregation of school systems. This tipsheet contains tips on understanding the dissimilarity index, the exposure index and counts and percentages. Tipsheet #311 can be ordered from the Resource Center at (573) 882-3364.

An audiotape of the panel is also available for $11 from Sound Images, Inc. at (303) 649-1811 or on email at soundimages@soundimages.net. Ask for Tape #IRE95-6.

From page six: # Home buyers

• Be ready to invest some time. It may make sense to attend a course on the building code, as I did. And there's no substitute for actually watching the nails get driven.

• You may also want to get your graphics artists involved from the start. William Pitzer's superb series of graphics showed how houses are built and what often goes wrong, adding depth and context to the project.

• If you tackle a story like this one, you'll probably find your time was well spent. After reading our series, Mecklenburg County officials agreed to hire more inspectors and take other steps to reduce the high inspection failure rate. State lawmakers introduced several bills to better protect home buyers. And more than 250 readers contacted the Observer to thank us and share their stories.

Ames Alexander can be reached by e-mail at aalexander@charlotte.com

First Ventures is a monthly column featuring a a reporter's first experience with CAR. For those interested in pursuing CAR skills, a NICAR bootcamp is the perfect plact to start.

Upcoming NICAR training:

Boot Camp
Columbia, Mo.
Aug. 8 - 13

Hands-on CAR classes
IRE Regional Conference
Los Angeles
Sept. 9-12

Intermediate Boot Camp
Columbia, Mo.
Sept. 17-19

Mapping Boot Camp
Columbia, Mo.
Oct. 21-24

Boot Camp
Columbia, Mo.
Jan. 3 - 8, 2000

Animals to go

By Linda Goldston
San Jose Mercury News

The zoos were making me mad.

I wanted animal disposal records from all of the accredited zoos so I could track their surplus tigers, monkeys and nyalas after they left the zoo grounds.

Most of the zoos weren't cooperating, and I wasn't about to settle for a dik-dik studbook and records from a couple dozen of the smaller zoos.

Even so, for a while I became known as "Ms. Scanner" in the office as I faithfully scanned every animal transaction record I'd been able to get my hands on into Excel – thinking all the while I'd plead, pry, pout and use my Southern accent to get the rest.

Records from my FOIA requests to the U.S. Fish & Wildlife Service and the U.S. Department of Agriculture already filled enough boxes to turn my desk, my car and my living room into forts. But I was still missing huge parts of the puzzle: Where did all of the animals go?

Seeking zoo records

One zoo director told me no reporter had asked for studbooks before, yet I was demanding them with a vengeance after learning all of them were public records, thanks to agreements signed with other zoological groups around the world. The studbooks recorded the genetic histories of animals in a species but they also contained such goodies as showing which animals left an accredited zoo and went to such dealers as Earl Tatum and Buddy Jordan.

I became very fond of a three-letter initial recorded after many of those animals were listed as going to dealers in the studbook – LTF, or "lost to follow-up." Lost to me, the public and the zoo itself. Lost in the regulatory black hole that governs zoo and exotic animals.

Enough already!

I had notebook after notebook of handwritten lists of where certain animals went after leaving the zoos, but the whole picture was far from complete.

So I decided, hell, I'll try to buy the same database the zoos use themselves to record their animal transaction records.

The first time I called the International Species Information System (ISIS) in Apple Valley, Minnesota, I simply said I wanted to buy a set of the CD-ROMs that hold all of the goodies.

The guy I spoke with said I didn't qualify – I needed to have a collection of animals. Ah, never fear, I said, I have a dog and a cat. The ISIS man laughed half-heartedly and said, um, no.

This ol' warrior will add one thing: A database is a remarkable tool – one that I came to admire and deeply appreciate. I might even do it again.

Fine.

About that time, a pair of wild ducks had taken up residence in the moat surrounding two sides of the Mercury News and hatched more than a dozen ducklings to make their way in the chlorinated waters of the moat.

So I called the ISIS man back: "Hey, now I have a dog, a cat, a mama duck, a papa duck and 14 babies," I said. "Surely that qualifies as an animal collection."

Um, no, that won't do, the ISIS man said, and there was no half-hearted laugh this time.

But as soon as I checked the list of ISIS members on the Internet and saw a few of the same dealers listed that I was trying to track, I had one of those moments investigative reporters know well. It's the moment when you say, OK, by god, I don't know how, but I WILL get the information.

Sources offer data

It took a little while longer, but I obtained two copies of the database from separate sources. It enabled us to verify the integrity of the data on each set.

Now all of you trained observers will have noticed that I slipped the first "us" into my narrative. I spent most of my career getting to the core of things by myself – well, me

Continued on page nine

From page eight: **Zoos**

and all those boxes that accumulate. This time, I needed something I'd heard about but figured I'd never need to use – the expertise to make sense of those CD-ROMs. And luckily, I didn't have to go far.

Jennifer LaFleur, database editor at the *Mercury News*, and a former NICAR training director, took one look at my bewildered expression and said, OK, I'll help you, dummy. She actually said, "Oh, OK," but you learn to joke early and often on these things.

Sgt. LaFleur drove me nuts for the first two weeks we were working with the data because I kept wanting to make printed copies of the 52 million animal transactions, which had to be made one copy, one transaction at a time. I wanted to hold them in my hands, add them to the forts of documents that now surrounded my desk, the inside of my car and my living room.

But Jennifer kept saying no.

No, no, no, you don't need to print them out, she said over and over, her voice rising each time.

Well, dammit, I did need to, I still want to now and probably will print out copies of many of the animal transactions – one at a time – before it's all over. But I agreed to be a good kid and set about trying to identify who the recipients of zoo animals were in the computerized records. There was no manual that said so and so is a dealer or that NBJ was dealer Buddy Jordan's NBJ Ranch in Texas.

Explaining the data

I also located a terrific source from the zoo world who could help us understand the ISIS data. That made it possible for Jennifer and me both to make sense of this complex and quite literally awesome database. The source agreed that since we had the data, it would help no one if we misinterpreted them.

And that's when Jennifer put this collection of millions of records into a database that we could search by individual zoo or pull out numbers for all of the accredited zoos.

The data analysis derived from it was the crowning touch to the culmination of all that work: A four-day series called "Zoo Animals To Go" by yours truly in the San Jose Mercury News, published February 7-10, 1999. For the record, this ol' warrior will add one thing: A database is a remarkable tool – one that I came to admire and deeply appreciate. I might even do it again.

Linda Goldston can be reached by e-mail at lgoldston@sjmercury.com

See related tipsheets from the 1999 IRE Conference panel "Zoo Animals to Go," Tipsheet #1016; and Selected Contacts for "Dying to Fly," Tipsheet #1015. These tipsheets are available online at *www.ire.org/resourcecenter*

NERD NOTES

Zoo data can be a bear

By Jennifer LaFleur
San Jose Mecury News

One key problem with the zoo project was that there was no government/public record database that kept track of animal transactions. But not every database is a government database – keep in mind that associations, non-profit agencies and other organizations keep databases. And although you may not have a public access right to the data, you may be able to use your sources to get data. One caution: VERIFY, VERIFY, VERIFY. We did this with the zoo database by getting two different copies of the data from different sources and comparing them.

Once we had the data, the work began. We didn't receive a straightforward database with lots of documentation; we had to go into the "canned" software written by the ISIS people and find the databases we needed. And there were hundreds of data tables to sort through.

Again, by using sourcing, Linda was able to get some help understanding the database.

Also, the database simply included the names of entities that received animals. There was no field that designated someone as a dealer or an auction. All of that information came from documents and reporting. Linda had to go through the database and manually create a designation for dealers, auctions, hunting ranches, etc.

Reporting and sending samples of data runs to zoos helped us do a better analysis in the long run. Our first analysis looked at any animal transaction going to particular deal-

Continued on page eleven

See related tipsheet from the 1998 IRE Conference for the panel on investigating science and medicine. Tipsheet #828 is available online at *www.ire.org/resourcecenter*

An audiotape of the panel is also available for $11 from Sound Images, Inc. at (303) 649-1811 or on email at soundimages@soundimages.net. Ask for Tape #IRE98-28.

ERRING PHARMACISTS

Dealing drugs

By Bennett Cunningham
KTVT-TV (Dallas)

Pharmacists are among the most trusted professionals in the world. When we get sick, they dispense medicines to help us feel better. But when a pharmacist makes an error, it can be fatal.

Our three-month hidden camera investigation found a disturbing case. A doctor in Texas prescribed a painkiller to a healthy mother. The few pills she ingested almost took her life.

The drug was not a painkiller; it was a drug for diabetics. Today, the patient is an invalid, unable to speak, eat and think for herself. The pharmacist dispensed the wrong drug. You might think the pharmacist who made the dispensing error was banished from his profession and leading an ignominious existence. You couldn't be more wrong.

Board unaware

Our investigative report found that the Texas State Board of Pharmacy, which licenses and disciplines pharmacists, wasn't aware of many near-fatal mistakes. The reason is that when a patient sues, the lawsuit is against the pharmacy, rarely the pharmacist, which shields the pharmacist's name from the disciplinary board. Only if the patient files a complaint does the state get involved.

This story changed the way pharmacists in the state of Texas fill prescriptions

To make this story work, we needed a database. The idea was to track pharmacists found in lawsuits and determine if the state ever took disciplinary action. An open record request to the state yielded a plethora of great data. We asked for the full name, license number and date of birth of every pharmacist disciplined in the state of Texas since 1995. We requested the data on a floppy disk in either Access or Excel format. It wasn't a problem and a week later the data were cleaned and we were on our way. Sounds easy, but the bitter pill was just around the corner.

The basement of the Dallas County Courthouse is a musty old hovel. The ceiling is lined with fluorescent lights and the rows of yellow case files go on forever. The story was in here. I printed out a list of cases from the county computer. All of the lawsuits concerned large retail pharmacies. I whittled the list down by isolating the cases to large damage claims.

Poring over paper

For two weeks, a researcher and I read the case files. We read every deposition, every interrogatory and every exhibit. Our mission: to find the name of each pharmacist involved. By the end of our research, we had volumes of legal pads filled with information on pharmacists. Now we had to figure out what to do with all of it.

We already had the data from the state. Next, we needed to find out if one of the pharmacists mentioned in any of the lawsuits was simultaneously disciplined by the state. After attending so many NICAR seminars, the phrase "mapping out your data" rang in my head. We kept it simple. We typed all the names of the pharmacists we found in the lawsuits into the database, then asked one basic question – which ones don't match? If there was a match between the state's list and our list derived from the lawsuits, the pharmacist was most likely disciplined from events arising from the case (we'd double-check this later).

Once the database determined whether there was a match, we asked other questions. Every time we entered a pharmacist's name, we entered other pertinent information on that same record. So, for example, we knew how many male pharmacists were addicted to controlled substances. We knew how many were convicted of crimes. The possibilities seemed endless. Only our thoroughness and our creativity limited the queries.

Complex data

There were complications. The mass of data was so large that sometimes it made querying nearly impossible. Once, I had to print out the entire database on 500 sheets

Continued on page eleven

From page nine: # Zoo data

ers, auctions or hunting ranches. After talking with some sources, we realized that any animal can have multiple transactions, so we narrowed the database to look at just the last transaction for each animal, which is basically where the animal ended up.

One caution: VERIFY, VERIFY, VERIFY. We did this with the zoo database by getting two different copies of the data from different sources and comparing them.

The next set of analyses showed that some of the animals included things like "domestic sheep" and "domestic pigs," and even though we really enjoyed "Babe," we filtered out domestic animals. We wanted to concentrate the analysis on exotic animals.

We got another database from U.S. Fish and Wildlife that lists endangered and threatened species. We matched these data against the taxonomy of the animal transactions to find endangered animals.

CAUTION: Matching data from two different sources can be extremely tricky and will still require some grunt work to double-check the matches by hand.

Once we had come up with our analyses by zoos, by recipients of animals and by species, we did the most important step to any analysis. We each went back to the very beginning and recreated every analysis. You can't take a chance that you missed something the first time. You must be able to duplicate your work.

Jennifer LaFleur can be reached by e-mail at jlafleur@sjmercury.com

From page ten: # Drugs

of paper because I couldn't get a query to work. After taping and stapling all the corresponding sheets together, I laid them out on my office floor. Then I did it the old fashioned way – I counted. I just couldn't get the computer to do it right. The mistake I made was organizing my data improperly. Remember mapping? Well, I didn't do a very good job. My data were not consistent, resembling a Word document rather than a well-oiled database. I cluttered the database with so much extraneous information that certain queries rendered the program useless.

Planning databases

Reporters trying to do a story like this should have an easy time if they plan their databases wisely. It is easy to plug in the names and license numbers for people and then check them against each other. The hard part comes in the details. That's where the database devil lives. He thrives in confusion and bad planning. He waits for you to find a new nugget for your story and then bam! You are toast. Gather, map and pray – three words to live by. No story should ever scare you. Always be ambitious in picking stories. Despite the fact that a poorly planned database can leave you in the database jail, good planning will land you on GO, and the rewards are endless.

Reporters trying to do a story like this should have an easy time if they plan their databases wisely.

The aftereffects of this story changed the way pharmacists in the state of Texas fill prescriptions. A new education program is underway and all pharmacists must triple-check every prescription before it leaves the counter.

Bennett Cunningham can be reach by e-mail at Bennettc@ktvt.com

Order the tape of the 1999 IRE Conference panel "Hard-Hitting Animal Issues: zoos, exotics and pets – the laws and loopholes," which featured this story.

To order a tape, call the IRE Resource Center at (573)882-2364 or visit *www.ire.org/store*.

Bad drivers fund city

By Bill Torpy
The Atlanta Journal-Constitution

Pine Lake, a town of about 800 residents some 10 miles east of downtown Atlanta, was the epitome of the age-old Southern ticket trap. However, it used computer technology to track its prey. The city raised 78 percent of its $1.1 million budget last year from traffic tickets – or put another way, the city pulled in a dollar from traffic tickets for every dime paid by residents in property taxes.

Pine Lake is not the only city in Georgia where traffic fines provide key revenue, according to an *Atlanta Journal-Constitution* study of city and county budget records. Statewide, there were 18 police departments that have operated in the black consistently since 1989, and several other metro Atlanta jurisdictions ranked high as vigorous traffic-enforcement towns.

But Georgia's undisputed traffic-ticket champion was Pine Lake, where enforcement on just one 1,500-foot section of road last year generated $860,000 worth of fines and fees. And a strong majority of the tickets were issued to African-Americans traveling through the predominantly white town.

City residents don't complain: the aggressive ticketing equates to low taxes and plush services. Residents don't even have to bag their leaves; a truck rolls by to vacuum them from the curb.

Drivers protest

Although residents love the arrangement, many motorists who pass through the town take offense. A protest encampment has attached itself to the traffic court. The district attorney was asked by black residents to investigate alleged racial inequities in the ticketing. And legislators have threatened to go after the city's charter.

The investigative package launched by the AJC into the Pine Lake ticketing situation is a perfect example of combining database research with traditional reporting to increase a story's impact and push for changes.

A state legislator offered the tip after introducing me to three citizen-activists who had set up a semi-permanent protest against Pine Lake officials for what they termed as predatory policing. At first glance, the 1998 city budget numbers that the protesters provided (and the city verified) were startling and the story was a natural. I wrote about the protest to get it on the record and started on a longer news-feature.

After the first report, I heard that our computer-assisted reporting editor, David Milliron, was digging into numbers concerning ticket traps throughout Georgia. We quickly married our efforts. Milliron's computer-based package provided the numbers I needed for authority and a yardstick for comparison's sake.

A strong majority of the tickets were issued to African-Americans traveling through the predominantly white town.

Milliron had been compiling data collected by Georgia's Department of Community Affairs. (Georgia has been tarnished in the past by ticket traps and the state compiles data on fines, forfeitures and court fees to keep tabs on those towns and counties that may be overzealous.)

With the records, we went back 10 years to show year-by-year statistical aberrations. Pine Lake turned out to be the poster child for aggressive, ticket-happy policing in Georgia.

Dicing the numbers

There were several ways to slice and dice the numbers – fines based on a per capita basis, fines as compared to a percentage of the police department's overall budget, fines as compared to the entire government budget, and so on. But no matter how we sliced the numbers, Pine Lake always came out on top.

The analysis gave us a benchmark: It showed the state average in per-capita fines was $35. Twenty-five cities collected twice

Continued on page thirteen

You can read a copy of the AJC's ticket-scam story from the IRE Resource Center at *www.ire.org/resourcecenter* or by calling (573) 882-3364.
Ask for story #15458

Other stories on traffic ticket problems on-hand at IRE include:

- A 1997 piece in the *Philadelphia Daily News* proving police cutbacks lead to an increase in traffic deaths.
Story # 14589.

- A 1996 series in *The Columbus Dispatch* revealing how traffic tickets became a billion-dollar business. Some cities collected more in traffic fines than taxes.
Story #13685.

- A 1995 *San Jose Mercury News* story showing how drivers get out of tickets by going to traffic school.
Story # 12812.

- A 1991 story in the (Belleville, Ill.) *News-Democrat* finds blacks are five times more likely to receive a traffic ticket.
Story #7959.

From page twelve: # Bad Drivers

that amount. But Pine Lake collected a whopping $1,055 per capita in 1998.

Originally Police Chief Mac McGlamery played down the issue of ticketing, saying his force was aggressive, but fair. When presented those figures, he kept saying "Are you sure? No. 1? In the state?"

Negotiating for data

To go deeper on the story, Milliron also negotiated for computerized ticket data from a number of police agencies. In Pine Lake's case, none of the police or city officials had the technical know-how to retrieve the data from their new computer system. The police chief said the agency had nothing to hide, and that Milliron could have the data if he could figure out how to get them out of the computer system.

So Milliron worked with the city's software vendor in South Carolina, who sent him a software upgrade necessary to write a report on the system, then export the data to diskette. Milliron performed the necessary software upgrades and collected all of the data in one afternoon at the police department. (He even showed the chief's secretary and court clerk how they could do the same for future requests.)

Now came the final data analysis.

State records made it nearly impossible to break down who was being ticketed and for what offense. That's where the raw ticket data from the city of Pine Lake proved helpful. That data included the usual ticketing demographics such as age, sex, race, offense, date, time, fine amounts, issuing officer and even final disposition.

No tickets issued

It also showed that almost no tickets were issued for moving violations. That was too messy in court, the chief said. In fact, the department hasn't issued a speeding ticket since McGlamery took office in 1997. Instead, the 10-officer department targeted yes/no offenses like out-of-date license plates, no insurance, and no driver's license. And the fines meted out in the small municipal traffic court were much larger than fines for similar offenses in the neighboring cities and counties.

To aggravate matters, Pine Lake is predominately white, but 87 percent of those ticketed were black. That, however, may not be as sinister as it sounds, as the town is surrounded by almost an entirely African American population and a statistical count of cars found a 4-to-1 ratio of black drivers to white.

Also, the court disposition data – included in the department's ticket database – showed that blacks and whites fared similarly in court when facing similar charges.

Public responds

The articles drew an avalanche of e-mail and telephone calls. The city responded by saying it would try to be kinder and gentler – but still firm – in how it deals with driving infractions.

The investigative package launched... into the Pine Lake ticketing situation is a perfect example of combining database research with traditional reporting to increase a story's impact.

But there were more than just promises. The city also cut its police force in half and has replaced its white municipal court judge with a black woman. Also, the city has overhauled its fine structure and is vowing to sell the new expanded police department building.

The changes did not totally stop the controversy swirling around the tiny city. Protesters still seek the ouster of the city's police chief and demand that the city stop writing tickets on the busy street running on the city's southern edge. The city has refused to do that.

But as time has rolled by, the protests are less frequent.

Bill Torpy can be reached by e-mail at btorpy@ajc.com

You can hear *The Atlanta Journal-Constitution's* CAR Editor David Milliron on the 1999 NICAR Panel "Cool tools to get past the drudge work and get back to your story."
See *www.ire.org/training/audio.html* to order a tape of the panel or call SoundImages, Inc., (303) 649-1811.

Check out this Tipsheet that detail other aspects of covering municipal judges and traffic courts:

Tipsheet # 1007 from the 1999 IRE Conference panel "Judging the Judges: Federal, state, and local" describes a 14-month probe of municipal judges. The investigation revealed wasteful spending, frequent absences and travel to conferences that had little to do with the largely traffic and misdemeanor court.

See related tipsheet from the 1999 IRE Conference panel "Getting More Out of the Court Beat Than the Dailies Your Editor Demands," Tipsheet #1024; as well as Tipsheet #736 from the 1997 IRE Conference for the panel on court records. These tipsheets are available online at *www.ire.org/resourcecenter*

An audiotape of the 1997 panel is also available for $6 from Gaylor MultiMedia, Inc. at (888) 310-3100 or online at *www.gaylor-multimedia.com*. Ask for Tape #70.

Justice is Blind

By Tony Mauro
USA TODAY

The Supreme Court, which still hands out quill pens to lawyers who argue before it, does not lend itself easily to computer-assisted reporting. But for two projects, USA TODAY used information technology to help the reporting on two major enterprise projects.

Legal prejudice?

The first was a March 1998 look at the law clerks who assist the justices in screening cases and drafting opinions. In addition to assessing their influence, we undertook the first-ever look at the demographic breakdown of the clerks hired by the current justices, dating back to 1972 in the case of William Rehnquist. While the names of the clerks were publicly available, nothing else about them was, and the court told us "You're on your own" in terms of tracking down this information.

We created a database with all the clerks' names and began, one by one, to figure out their background. Identifying gender was easy in most cases, and a small number of names (Asian and Hispanic) also made ethnic background obvious. But most names were not obvious, and we wanted to be doubly sure that our categorization was correct. Through online searches of legal directories and other sources, we were able to locate most of the clerks. One of the most useful free sites was the Martindale-Hubbell directory of lawyers around the country, at *http://lawyers.martindale.com/marhub/form* .

We called many of the clerks and asked them to give us their own ethnic or racial information and to do "peer reporting" on other clerks from their year at the court. We encountered some resistance but were eventually able to feel certain about our information.

To double-check, we sent our information to each justice individually. Only two justices responded. For the other seven, we did not interpret their silence as confirmation of our information, but after further checking we went ahead, and have not heard since that any of our information was wrong.

The survey indicated that the justices have hired extremely low numbers of minorities as law clerks. The stories triggered objections from civil rights groups and a debate over the reasons for and significance of the court's hiring practices. The following is a link to the original stories and follow-up stories on this issue.

http://www.usatoday.com/news/court/clerk/clerk000.htm

Docket declinations

The second project was a December 1998 look at the Supreme Court's docket, with special focus on the cases that the court decides not to decide-the cases the court passes up. Lawyers have complained with increasing frequency that the court, in reducing its caseload, has left many areas of the law in a state of confusion.

The court told us 'You're on your own' in terms of tracking down this information.

To document this, we solicited the help of U.S. Law Week, a BNA publication that lists, in capsule form, every case on the court's docket of paid cases. With their permission, we used their database to break down the cases the court declines to consider. This customized database indicated that the court declines substantially higher percentages of cases in certain categories than others.

Ultimately, we made only limited use of these data for two reasons.

First, the wide variance in numbers of cases in different categories made the "rejection rate" seem exaggerated in some instances.

Secondly, other reporting indicated that there are so many factors that go into the court's decision to decline a given case that relying on the numbers alone would lead to drawing unsound conclusions. The following are links to the main stories in this package.

www.usatoday.com/news/court/nscot933.htm

www.usatoday.com/news/court/nscot932.htm

Tony Mauro can be reached by e-mail at tonymauro@aol.com

Backgrounding by Web

The following are excerpts from a handout presented during the "Backgrounding the individual by Web and computer" panel at the 1999 IRE Conference.

By Wes Williams
KPNX - TV (Phoenix)

Start with as much info as possible:

Who are you searching for and why? Even if you're starting with just a name, it helps to know why you're looking for this person and what the story is about. This will give you other ideas on where to look for information.

Is your subject a lawyer? Check an online legal directory. Does his bio say where he or she went to school? Check the school's alumni Web page for any stories on him or her. Every nugget of info you have will help, but look especially for approximate age, geography and occupations.

The computer is often the fastest way to background an individual. The following is a list of Internet sources of information. These are in the general order I go through, but it varies each time according to the amount of info I already have, the type of info I'm trying to find and how long I have to get it.

Search engines

I've found good information before just by punching someone's name into a search engine. I often start with a quick keyword search through a few of my favorite search engine Web sites.

Pros: They're quick, free and can lead you to a lot of personal information if you're lucky.

Cons: You can end up with too many possibilities and get bogged down in trying to wade through all the John Does to find your John Doe.

Good places to start include:
AltaVista (*www. Altavista.com*)
InferenceFind (*www.inference.com*)
HotBot (*www.hotbot.com*)
Dejanews (*www. Dejanews.com*)

Public records

The best and most accurate information is in public records. Some are available on the Internet, some on dialup bulletin-board systems and some you need the individual databases from the appropriate agency.

If I already have a good ID on someone (name, DOB, address), I go to these places first to look for info relevant to the story. Here's a list of the ones I hit most often.

- County court: Check for Web page or dialup BBS to get access to court docket information.
- County Recorder: Find property records and more in these like liens, tradenames, etc.
- County Assessor: See who's paying the taxes on a property or what it's worth.
- Secretary of State/USS: These help when trying to nail down business associated with someone's name or finding some lines
- State Corporation Commission: Maybe yours is on the Web. Ours is a dialup BBS that charges a small fee.
- Professional registrations: Is your subject registered with a state agency (doctor, police officer, child care, center owner, contractor) Check for Web sites that have the info and get an electronic copy of the data so you can have a back up and do database-matching projects.
- Federal Bankruptcy District Court via PACER: This is slowly coming back to the Web, but for now, it's a dialup BBS in most places. See *www.uscourts.gov/PubAccess.html* or call 1-800-676-6856 for info. Feds charge 60 cents per minute for PACER access. You can get bankruptcy docket information and it's great on deadline.

Public record databases

These sites below have public record databases already online:

- THOMAS, home of bills, laws and the Congressional Record (*www.thomas.loc.gov*)
- EDGAR Securities & Exchange Commission filings (*www.sec.gov/edaux/searches.htm*)
- Internet Non-Profit Center based on IRS records (*www.nonprofits.org/loc/index.html*)

Continued on page twenty-three

See related tipsheets from the 1999 IRE Conference panel "Backgrounding the Individual on Deadline," Tipsheet #1030; The Byte Chase, Tipsheet #967 These tipsheets are available online at *www.ire.org/resourcecenter*

Also see related tipsheets from the 1997 IRE Conference panel "Who is John Doe? How to Background Someone in a Hurry," Tipsheet #707. These tipsheets are available online at *www.ire.org/resourcecenter*

To download a copy of Aerial go to *http://www.ambia.com*

Trial is free for 30 days and then costs $149.

Aerial can also be ordered via fax at (703) 934-7154. Download the product order form off the Web site.

Other Aerial uses:

- Extract text from documents
- Copy tables to the clipboard
- Launch documents by title
- Navigate by page number

TECH TIP

Importing .PDF tables

By Tom McGinty
NICAR Training Director

There they are, staring you right in the face: the preliminary figures for the 1998 FBI Uniform Crime Report. It's all there on your screen – city-by-city tallies of murders, aggravated assaults, robberies, you name it – and all of the figures appeared with the click of a hyperlink on the FBI Web site. You're salivating. Images of spreadsheets are dancing in your head. Your editor is going to love this!

But wait a minute. Something looks funny. What's up with the strange menu bar that suddenly appeared at the top of your Web browser? Why does the whole Web page look so odd, as if it's a picture of a document? Oh no! Could it be? Aaarghhh! It's true: The FBI has published the statistics using Adobe Acrobat, a format that Web publishers love but computer-assisted reporters loath.

The bad news is, it's a pain to get data tables out of Acrobat documents and into a crunchable format. The good news is, it's not impossible. And that's true because of Aerial, a piece of software developed by Ambia Corporation that turns Acrobat documents into text. This tech tip will go over the nuts and bolts of using Aerial. But first, here's a little background.

The way it was

In the good old days of the Web – say, four or five years ago – most tabular data on the Internet were published in plain text that could be quickly imported into spreadsheet programs or database managers and manipulated at will.

Then along came those pesky HTML tables, which had pretty raised borders that gave them a 3-D look but were kludged up with HTML coding that made them nearly impossible to import. Microsoft's programmers solved that problem with Excel '97, which easily imports HTML tables.

But Acrobat also was on the scene and only growing in popularity with government agencies that put data-rich reports on the Web. It's an understandable, if annoying, phenomenon. With Acrobat software, bureaucrats can transform huge, typed reports and graphics into Acrobat Web documents with the click of a button. But unlike text or HTML tables, Acrobat tables cannot be imported into other programs. That's where Aerial comes in. And as the following directions will show, it's not too tough to install or use.

The bad news is, it's a pain to get data tables out of Acrobat documents and into a crunchable format. The good news is, it's not impossible. And that's true because of Aerial...

First, you have to be sure that you have Adobe Acrobat Reader 3.x installed on your computer. If you have a different version (either an earlier version or 4.0) the current version of Aerial will not work. Check your version of Acrobat and then:

New software

Go to *www.ambia.com*, and click on the "Products" link, then click on the "Adobe Acrobat Tools" link. Next, click on the "Aerial" link. If you found that you didn't have Acrobat Reader 3.x, you can click on the yellow and red "Get Acrobat Reader" link on the bottom left of this page and download a free version of Reader 3.x. Install that software and then go back to the Ambia Web. Once you get to the Aerial page, you can either download the "demonstration" copy, which is free and good for just 30 days, or you can buy the software for $149 (unfortunately, Ambia recently increased the price from $49).

It's probably a good idea to try the demonstration version to determine if you will really want to use the software. Once you get the file downloaded, go to the

Continued on page seventeen

From page sixteen: **Tech tip**

folder where you stored it and double click on the file you downloaded to begin the installation.

Once the installation software starts up, one of the first steps is to set the path where it will be installed. Aerial fills in a default path that probably won't work, so you should hit the "browse" button and find the folder where Acrobat Reader 3.x is located.

Open that folder, then open the "Reader" folder and then open the "plug-ins" folder. Your path probably will look something like this: c:\acrobat3\reader \plug_ins. Then just click on the Okay button and complete the installation.

Grab the table

Now that the program is installed, the fun part starts. Use the FBI Uniform Crime Report as your first test. Go to *http://www.fbi.gov/ucr.htm* and locate the link that reads "1998 Preliminary Crime Statistics (.pdf)." Instead of clicking on the link and opening the file in your Web browser, put your pointer over the link, click on your right mouse button and a drop-down menu will appear. Choose the option to "Save link as..." (To get the same drop-down menu on a Macintosh computer, hold the control button down and then click on the link.) This will bring up a "Save" dialogue box. Choose the folder where you want to save the file and click on "Save." Instead of opening the .pdf file in your browser, you have saved it to your hard drive.

Now start up Acrobat Reader (you can find it by going through the "Start" button and "Programs"). Under the "File" menu, choose "Open" and then open the file you just saved. Now go to the page that has the table you want to move into a spreadsheet or database manager. You need to highlight the table first, so on the Acrobat Reader menu bar, click on the icon with the letters "abc" inside it. That will give you the text cursor.Starting at the top left corner of the table, hold your mouse button down and drag to the bottom of the page. The text of the table should become highlighted as you go. (Even if a table spans several pages, you can only highlight one page at a time, so you'll have to repeat these steps.)

Converting tables

Once you have the entire page highlighted, go to the "Plug-Ins" pull-down menu, choose the "Aerial" option and then choose "Copy table to Clipboard." A new window titled "Select Clipboard Format" will pop up. As you probably already know, the best choice is the default choice: "Tab Delimited ASCII." Make sure that selection is chosen and then click "OK."

Using Aerial, you have turned the mirage-like data of an Acrobat document into real numbers that will slake your thirst for crunching.

You have just converted the once-unusable Acrobat table into a tab-delimited ASCII table that can be used any way you see fit. You can either paste it into a word processor, save it as text and later import it into Excel, or paste it directly into Excel. Give both methods a shot and see which one works best for you. Now go back to the UCR Acrobat document and get the second page of the table, etc.

Be cautious

Using Aerial, you have turned the mirage-like data of an Acrobat document into real numbers that will slake your thirst for crunching.

A note of caution: Although Aerial has worked flawlessly for me every time I have used it, I have heard other reporters say that they encountered dropped characters in the translation. Of course, it's always a good idea to check the integrity of your data once it's imported.

Tom McGinty can be reached by e-mail at tmcginty@nicar.org

CONTACT THE IRE RESOURCE CENTER AT (573) 882-3364 OR VISIT *WWW.IRE.ORG/RESOURCECENTER* TO SEARCH FOR TIPSHEETS

ALSO SEARCH THE LIBRARY OF MORE THAN 15,000 INVESTIGATIVE STORIES ONLINE AT *WWW.IRE.ORG/RESOURCECENTER/STORYDB.HTML*

THE RESOURCE CENTER'S WEB PAGES INCLUDE LINKS TO RESOURCES FOR SPECIFIC BEATS AS WELL AS INFORMATION ON BUYING AUDIO TAPES FROM CONFERENCE PANELS.

KEEP VISITING *WWW.IRE.ORG* FOR THE TIPSHEET OF THE WEEK AND THE STORY OF THE WEEK.

Reporters looking to find out more about statistics and stories can start at IRE's Resource Center.

Sarah Cohen of *The Washington Post* (and former NICAR training director) has put together several tipsheets on how reporters can use statistics safely.

The most recent tipsheet came from the 1999 IRE Conference panel "How to Assess Risk: What's really dangerous." The tipsheet adivses reporters on how to figure out what researchers are really saying with their studies, and was designed to help decipher the world of medical and environmental reporting.

IRE Tipsheet # 961 is available online at *www.ire.org/resourcecenter*

Raging to regress

This is the first of two parts on how to do a regression analysis. Next issue's column will deal with regression outliers, the T-test and making transformations.

By Neill A. Borowski
The Philadelphia Inquirer

How many times have you interviewed an "expert" who quickly described a process without going into detail? If you're any kind of interviewer, you made this expert stop and painstakingly explain how the process worked. The process might have been a controversial bond issue, a new communications technology or the treatment for a disease.

Now let's consider the rage to regress. Only at a NICAR session will you hear journalists in the bar late at night discussing how they "ran this independent variable with that dependent variable and came up with an r-square of ...."

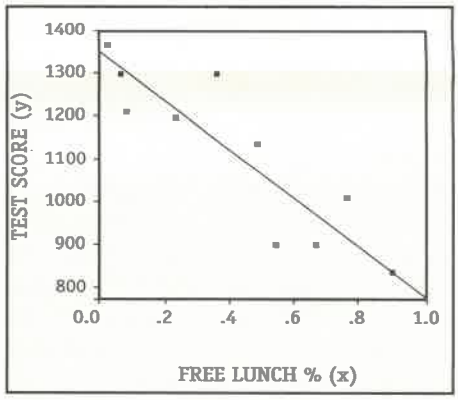
Yet, when we use this popular and insightful tool, we never press for an explanation of how the process works. We plug our variables into the black box (SPSS, SAS, Excel, etc.) and out pops the regression results. Many don't seem the least bit interested in how the black box did what it did.

Call me a math nerd, but I like to understand the statistical underpinnings. Obviously you can't hand-crank a regression analysis of thousands of records. But it still makes sense to understand the process that's going on inside the black box.

Our goal in simple linear regression analysis is to come up with a model: $y = ax + b +$ random error. The "y" is your dependent variable (such as test scores). The "x" is your independent variable (such as poverty rates). The "a" is the slope of the regression line and the "b" is the constant (the intercept of the vertical axis). This is the equation of the regression line.

The regression line is fitted from, in essence, many averages. The model tries to plot the line that best fits the data that you've provided. The first step is a scatterplot, simply taking each data point and drawing a point where its x value and y value intersect. The regression line drawn through the points is the best fit, based on the average lengths of the data points from the regression line.

These two charts (below) represent the percentage of students eligible for free lunch in a school (x) and that school's average test score (y). The other columns simply are x-squared, y-squared and x times y.



FREE LUNCH % (x)	x^2	TEST SCORE (y)	y^2	xy
0.55	0.3025	900	810000	495
0.67	0.4489	900	810000	603
0.03	0.0009	1365	1863225	40.95
0.48	0.2304	1152	1327104	552.96
0.12	0.0144	1218	1483524	146.16
0.89	0.7921	856	732736	761.84
0.76	0.5776	1010	1020100	767.6
0.08	0.0064	1300	1690000	104
0.26	0.0676	1193	1423249	310.18
0.35	0.1225	1297	1682209	453.95
SUMS				
4.19	2.5633	11191	12842147	4235.64

Continued on page nineteen

From page eighteen:

Stats

Now the regression math – or "method of least squares" – begins.

Sxx = (The sum of the x squares) – (The sum of the x's, squared/n)

Or, in math notation:

$$Sxx = \sum x^2 - \frac{(\sum x)^2}{n}$$

Substituting from the table above:

$$Sxx = 2.5633 - \frac{4.19^2}{10} = 0.80769$$

Now the same for the y's (test scores):

Syy = (The sum of the y squares) – (The sum of the y's, squared/n)

Or, in math notation:

$$Syy = \sum y^2 - \frac{(\sum y)^2}{n}$$

Substituting from the table above:

$$Syy = 12{,}842{,}147 - \frac{11{,}191^2}{10} = 318{,}298.9$$

Finally, we need a third "average" – the average of the x's times the y's.

Sxy = (The sum of the x's times the y's) – (The sum of the x's times the sums of the y's)/n

Or, in math notation:

$$Sxy = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

Substituting from the table above:

$$Sxy = 4{,}235.64 - \frac{(4.19)(11{,}191)}{10} = -453.389$$

With these "averages" behind us, we now can find the components of the regression line.

$$a = \frac{Sxy}{Sxx} = \frac{-453.389}{.80769} = -561.34$$

Since a is the slope of the regression line, and because a is negative, the line is negatively sloped – meaning that it goes from the upper left to the lower right.

Intuitively, this makes sense. As x grows, there are more (a higher percentage of) poor students in a district. So, as we go from left to right on the horizontal (x) axis, test scores drop.

b = The average of the y's minus (a (from above) times the average of the x's)

$$b = \bar{y} - a\bar{x}$$
$$b = 1119.1 - (-561.34)(0.419) = 1354.3$$

So, the equation of the line, y = ax + b, is:

$$y = -561.34x + 1354.3$$

Plug in the x's from the table above and the y's that are yielded in this equation will be the predicted scores.

Now, we need to calculate the correlation coefficient (r) and, more importantly, the coefficient of determination (r^2).

$$r = \frac{Sxy}{\sqrt{(Sxy)(Sxy)}}$$

We already have all of the needed components, so we plug them in:

$$r = \frac{-453.389}{\sqrt{(0.80769)(318{,}298.9)}} = -0.894$$

$$r^2 = 0.8$$

That's a coefficient that suggests a strong relationship between income and test scores. It means that 80 percent of the variability in test scores is explained by the variability in income, according to the regression " model." The larger the coefficient (up to 1), the stronger the relationship between the independent and dependent variables.

Your regression model still may be off base. It could be totally ridiculous. The formula does as it's told. As you can see from the math, there is no mystical transmutation of the numbers. It's all a matter of averaging, squaring, substracting, etc. – all calculations designed to simply minimize the distance of each actual data point from the estimated line. Keep this in mind in your rage to regress. Don't stop. Just keep this in mind.

Neill A. Borowski can be reached by e-mail at nborowski@phillynews.com

Each year, IRE offers an advanced bootcamp on statistical analysis at the University of North Carolina in Chapel Hill.

The seminar is aimed at strengthening skills of reporters who want to move beyond CAR. Under the watchful eye of UNC Professor Phil Meyer, a CAR guru who teaches "Precision Journalism," reporters learn basic statistical programs.

Contact (573) 882-0684 for more information or watch *www.ire.org/training* for dates.

Order the tape of this panel "Investigating Sprawl: Techniques and Revelations" from the IRE Resource Center at (573) 882-3364 or visit *www.ire.org/store*.

The IRE Resource Center has several series about Sprawl. To order a copy of a story call (573) 882-3364 or e-mail your request to *rescntr@nicar.org*.

The stories include:

- The Kansas City Star's series on urban sprawl that was a finalist for the Pulitzer Prize. Ask for story #13004.
- A 1999 Time Magazine about nationwide efforts to fight against sprawl. Story #15398.
- A 1999 Detroit News piece on the strain sprawl causes on city services. Story # 15461.
- A 1998 Los Angeles Times story that showed politicians, who vowed to control sprawl, actually encouraged it. Story # 15040

Sprawling around

The following is a handout from the "Investigating Sprawl: Techniques and Revelations" panel at the IRE National Conference in June.

By Rose Ciotta
The Philadelphia Inquirer
and Greg Reeves
The Kansas City Star

Sprawl is not new unless you haven't done the story in your town. Americans have been moving to the suburbs since before World War II, but the impact of this exodus has often been hidden. Here's a list of data sets to help you dig out how sprawl has affected your coverage area:

1) **U.S. Census** – There's valuable info here even as we wait for the first Census 2000 data on Dec. 31 of next year. The Census Library in Washington, D.C. can send you population and other demographic data for your area back to pre-Civil War days. Charting county-to-county population moves by census tract since World War II gives the basic picture of sprawl for most areas.

www.census.gov/population/www/documentation/twps0027.html

www.census.gov/population/www/estimates/countypop.html

2) **U.S. Census Building Permits** – Place-level data on the number of housing units from single family to doubles, triples and buildings with more than five units built every month or every year. Data are collected in a mail survey of localities that issue the building permits. Dates to 1980 easily; some earlier years are available on paper. Can be compared to units in structure in census. Annual and monthly data dating to 1980 are available on diskette from the bureau's Construction Division at (301) 457-1321.

www.census.gov/ftp/pub/const/www/c40index.html

3) **Local Property Roll including Sales Data** – Tracking number of housing units and property values year by year in your area can provide real-life contrast between urban and suburban trends. Also, property taxes in some areas become grossly inequitable between newer and older homes because of the lack of updated reassessments. This is an important public policy question and can make for a high-impact story. Property assessment databases in most areas are public record, and should be part of your basic CAR data set. Data on the relationship of assessment to sales will help you pinpoint areas usually in the inner ring that are losing their value. The result is often that owners of lower valued homes pay a higher proportion of taxes than expensive homes. Look for public ownership of buildings. As population shifts, more residential property is abandoned in the inner city and inner ring. Use the roll to examine who is selling and who is buying the land where development is occurring.

Sprawl is not new unless you haven't done the story in your town ... but the impact of this exodus has often been hidden.

4) **HMDA (Home Mortgage Disclosure Act) Data** – Who applies for home loans by census tract, income and race as well as who gets approved and denied. Issued annually by the FFIEC (Federal Financial Institutions Examination Council) – one loan application, one record. This data will allow you how to illustrate the impact of sprawl on minorities. The data shows where banks have loaned money for home purchases and the demographics of the buyer including race and income. The purpose of the disclosure law is to help the feds determine whether banks are serving local housing needs and to identify discriminatory lending patterns. Last year, 7,925 banks, savings associations and credit unions reported 16.4 million loan applications for 1997.

www.ffiec.gov

5) **FHA and VA Mortgages** – Say hello to your local recorder of deeds office. Pick a subdivision that sprang up during the big move to the suburbs. Look up the mortgage on each parcel – federally subsidized Federal Homes Administration and Veterans Administration mortgages are easy to spot. How much of a role did the federal government

Continued on page twenty-one

From page twenty: **Sprawl**

play in your area in contributing to the move to suburbia?

6) **IRS Migration Data** – Track yearly movement of people from counties. These data are especially useful if your area has seen the movement of people from a city county to a suburban county. Provides demographic information on the persons moving, including age range and education.

Issued yearly, with 1996-1997 the latest available. Includes the number of tax returns (which can be used to approximate the number of households), number of personal exemptions (which can be used to approximate the population), and, for the first time, adjusted gross income (AGI) and median AGI. (202) 874-0410.

www.irs.ustreas.gov

7) **Highway Volume Data** – Most states and counties keep track of traffic volume. Looking at travel over time can reveal changing residential patterns. Also examine the type of road, what it was designed to carry and the volume it does carry.

8) **FBI Uniform Crime Reports** – Annually issued data reveal the frequency of crimes by cities and towns. If you analyze these data over time, they will likely reveal increasing crime in the inner city. Suburban residents often cite increasing crime to justify their move to the suburbs.

www.fbi.gov

9) **Campaign Finance** – Contributions by builders to local officials making decisions about new developments, to county-level officials who often drag their feet on countywide planning and to state lawmakers who often refuse to impose planning requirements on localities. Also look for contributions from environmentalists lobbying for smart growth initiatives.

10) **Business Property Tax Breaks** – Look for property data to show special exemptions for new businesses that locate in new residential areas. Calculate how much in local property taxes they are saving.

11) **County and ZIP Business Patterns** – County-level data that reflect the number of businesses by category, number of employees and payroll value from 1988 through 1996 for county data. ZIP code level data have been published annually since 1964, with 1994-95 the latest year available on the Census Bureau Web site. These data are especially useful to show how businesses move from the inner ring to the outer ring and the growth of businesses in new residential areas.

www.census.gov/epcd/cbp/view/cbpview.html

http://tier2.census.gov/zbp/zbpinfo.htm

12) **American Housing Survey** – A wealth of information about the nation's housing – apartments, single-family homes, mobile homes, vacant housing units, household characteristics, income, housing and neighborhood quality, housing costs, equipment and fuels, size of housing unit, and recent movers. National data are collected every other year, and data for each of 46 selected metro areas are collected about every four years, with an average of 12 areas a year. The national sample covers an average of 55,000 homes. Each metropolitan area sample covers at least 4,800 homes. The Census Bureau conducts the survey for HUD. You can't get tract-level data from AHS, but you can divide your metro area into several parts, and report hundreds of data items, including estimated cockroaches-per-household.

www.census.gov/hhes/www/ahs.html

3) **Agricultural data** – Track the (shrinking) agricultural acreage in your area over time. The U.S. Department of Agriculture puts out the Census of Agriculture with data on farms from which at least $1,000 of agricultural products were produced and sold in a year. Data is available on a county and state level. The census is taken every five years – during the years ending in "2" and "7."

www.nass.usda.gov/census/

14) **Other Data** – These may include school enrollment and cost per pupil, construction and commercial building permits, contributions by builders, voter registration data or censuses of governments and government finance by the Census Bureau. Local and regional planning departments and agencies often have much data on urban and suburban growth. Some newspapers have also conducted polls on public attitudes toward development and quality of life issues.

Rose Ciotta and Greg Reeves can be reached by e-mail at rciotta@phillynews.com or greeves@kcstar.com

See related tipsheet "Suburban Development: Documenting the Impact of Urban Sprawl," from the 1999 NICAR Conference panel "A Brief List of Development-Related Public Records," Tipsheet #860; as well as Tipsheet #686 from the 1997 IRE Conference for the panel on development.

An audiotape of the 1997 panel is also available for $6 from Gaylor MultiMedia, Inc. at (888) 310-3100 or online at *www.gaylor-multimedia.com*. Ask for Tape #27.

See these other handouts on editing CAR stories. You can read some online at *www.ire.org/resourcecenter.* Others you can order at (573) 882-3364 or by e-mail rescntr@nicar.org.

From the 1999 IRE Conference:

"Writing the Story" provides ideas for getting people in and numbers out while building a CAR narrative. (Tipsheet #968)

From the 1999 NICAR Conference:

"How to Ask Questions When You Don't Know the Data" provides a list of the top ten questions editors should ask during a CAR project. (Tipsheet #907)

"The Development of CAR" includes another list of questions for editors working with CAR stories. (Tipsheet #950)

HANDOUTS OF THE MONTH

Editing CAR work

The following are excerpts from a handout presented during the panel: "Latest Tactics for Writing and Editing CAR Stories" at the 1999 IRE Conference.

By Jeffrey Kummer
St. Paul Pioneer Press

Here's a checklist of dos and don'ts that other editors – whether they are computer-savvy or not – may find helpful at each stage of a CAR story:

Concept

• Brainstorm, discuss and critique stories thoroughly before you begin. CAR is time-consuming and expensive. Understand what you, your reporters and your newspaper or station is getting into.

• Listen to all ideas. Beware of any proposal that begins with "Let's take a look at…" Such stories are doomed unless they're better defined.

• Don't "data dive." Just because there are million of records in those databases, it's easy to assume there has to be a story there some place.

• Questions to ask:

What do you think a computer analysis would show?

What are you expectations based on?

Who have you talked to?

What other evidence do you have?

What will CAR add to this story that we couldn't get in other ways?

Reporting

• Keep talking – perhaps daily, at least weekly, depending on the story's time frame. An occasional memo or draft that includes a working lede and nut graf can help keep the editor and reporter on the same track. It can also serve as an early warning system for a reporter gathering lots of information but losing sight of the story.

• Get reporters off the computer and out of the newsroom. Don't wait until the entire analysis is done to bounce findings off of sources, locate experts or talk with the real people behind the data you are examining. Those forays into the field can reinforce findings or poke holes in faulty conclusions.

• Allow enough time. CAR is not a shortcut. It is time consuming, especially for beginners. Trying to do CAR on deadline can convince novices their first CAR project should be their last.

• Team when possible, especially on stories where timing is an issue. Dividing the work between two or more reporters can get the story into print or on the air sooner.

• Don't forget to involve copy editors, artists and photographers early. You can eliminate the "deadly data" syndrome from many CAR stories by using charts, summary boxes, etc.

Writing

• Leave enough time for the writing and editing. When it's time to write, work with the reporter to take a final inventory of the information gathered. Evaluate the strengths and weaknesses of the material – what you will use and what you don't.

• Be on the lookout for CAR fatigue. Those nuggets of information that once seemed so revealing start losing their luster as the project goes on. Help reporters remember what will be of interest to someone getting this information for the first time.

• Tell the story with people and events, not with data. If the reporter can't find examples of the trends revealed by the analysis, question the findings.

• Discuss presentation, but resist formulas. At this point, it is no longer a CAR story. It's simply a story to be told in as compelling a way as possible.

• Don't burden readers or viewers with process. In most cases, they won't care how many thousands of electronic records we filtered and sorted. Our motto might be: We do CAR so you don't have to.

Editing

• If you have CAR skills, help be a backstop to the reporter's work. Most reporters welcome another set of eyes and hands on the data.

An experienced CAR reporter I recently hired told me: "It drove me nuts that editors at my previous paper had zero CAR knowledge. As a result, every time I had a

Continued on page twenty-three

From page twenty-two: **Editing**

CAR story in the paper, I was hanging out there."

• See the story through a reader's eyes. Is it interesting? Does it hold together? Are its points clear and its focus sharp?

• Do you understand the data the reporter has used and are you comfortable with the conclusion the reporter has reached?

• Question the reliability of the data the same way you would the veracity of any other key source that was to be the basis for a story. Why did you choose these records? Who has checked each calculation for errors? What data integrity checks were done? How current is the data? Have the data been used elsewhere in similar stories? Were there any problems or pitfalls that arose in those cases?

• If you don't have CAR skills, ask the reporter to show you how he or she came up with key findings. If the reporter can't recreate those numbers with you sitting next to him or her at the computer, be concerned.

• Attack the numbers. Get as many of them out of the story as possible. Can some of them be delivered better in a graphic? Can percentages be recast in easier-to-read terms, such as "four out of five doctors?"

After publishing

• Do the post-mortem. Sit down with the reporter and walk through the process. What would you each do differently next time? Where were all the flaws in the data? Who should we have talked to earlier? What did we wish we'd have gotten that we were unable to get this time around?

• Is there a plan for following the story? Can you come back in a year or two and see what has changed? Is the methodology documented well enough that another reporter could redo the analysis later and be able to compare the findings?

• Are there other uses for the data? Once the CAR story is done, work with the reporter or your CAR coordinator to make the data available to the rest of the newsroom. It's great to do a project on repeat drunk drivers. But now that you've got driving records in digital form, those could be an ongoing resource for other reporters and editors.

Jeffrey Kummer can be reached by e-mail at jkummer@pioneerpress.com

In May, NICAR held its first-ever Editors' Boot Camp, an intensive three-day workshop tailored to the needs of newsroom managers. These boot camps teach editors what they need to know to make CAR successful. NICAR is planning another editors' seminar. Watch *www.ire.org/training* for details and dates.

From page fifteen: **Backgrounding**

• Social Security Death Index (*www.ancestry.com/ssdi/advanced.htm*)

• Aircraft Owners and Pilots. Watch for outdated info. (*www.landings.com*)

• Docfinder with links to state medical boards on the web. (*www.docboard.org*)

• MilitaryCity.com. Subscription-based search for military members and veterans. Free for Gannett companies.

• Property Records Online, University of Virginia page. Lists where to find property record Web Sites (*www.people.virginia.edu/~dev-pros/Realestate.html*)

Private people finders

These are the ultimate tools for quick and thorough research. The catch: They're very expensive and can waste your time when you come up with a hundred John Doe listings and you don't even have a middle initial. These are great for locating hard-to-find people, relatives, crisscrossing addresses and finding non-local leads. If I'm in a hurry and it's worth the cost, this may be my first stop instead of my last. These services are usually not very good for finding people under age 22.

• Autotrack charges $2 a minute and an average in-depth search takes about 8 minutes. (*www.dbtonlin.com*)

• CBD Infotek charges $17 for an in-depth Discovery Plus! Report (*www.cdb.com*)

• Knowx Web-based service, search cost ranges from free to $1.50 per individual database, detail records range from 95 cents to $6.95 per record. (*www.knowx.com*)

Wes Williams can be reached by e-mail at williamsaz@worldnet.att.net

Bits, Bytes and Barks

NICAR data update

Updates of the following data sets are now available through NICAR's data library. Prices and years available can be found at *www.nicar.org* or by calling 573-884-7711.

• The Uniform Crime report consists of six separate databases from the FBI gathers crime information from law enforcement agencies around the country. All of the databases provide the region, state, county, city, MSA and reporting agency identifier. Those who want to see a slice of the dataset can go to *http://www.nicar.org/data/fbi* and download a section.

• The National Bridge Inventory System (NBIS) 1998 includes maintenance records collected by the Federal Highway Administration for assessing the nation's bridges. It is also used by the Department of Defense and the Federal Emergency Management Agency. State inspectors assess bridges and their reports are submitted annually to the FHA. This is a great database for CAR beginners. More information can be found at *http://www.nicar.org/data/fhwa/index.html#nbis*

• Federal Contracts Data 1998 offers a database of Individual Contract Action Reports (ICARS) created by the Government Services Agency. Each record contains information about the agency granting the contract, the contractor, the place of performance and the contract amount.

This year, NICAR is including a FIPS code table and a table of U.S. Representatives so that reporters can match contracts to local representatives. The FIPS table isn't always filled in, so matches will not appear 100 percent of the time, but it's a good place to start. Upcoming in August: Hazardous Materials, National Boat Registration information and IRS 990 databases.

Tipsheets and stories

In conjunction with NICAR data updates, the IRE Resource Center has put together a list of stories and tipsheets. Not only do the stories help reporters see what kind of information is found in the data sets, but the tipsheets point out potential hazards and give good contacts to help put the information in perspective.

We have tipsheets and stories to go with the FBI, Bridges and IRS 990 datasets. We also have stories to match with the Federal Contracts Data.

You can search the resource center database at *www.ire.org/resourcecenter* to see what's offered and to order copies of the tipsheets.

IRE Regional Conference

It's not too late for reporters and editors interested in the IRE Regional Conference coming up in Los Angeles on September 9-12, 1999.

For updates on that conference and other IRE or NICAR gatherings, see *www.ire.org/resources/conferences.2*

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

NON-PROFIT ORG.
U.S. POSTAGE
PAID
PERMIT NO.286
COLUMBIA, MO. 65211