Uplink

July • August 2002
Volume 14, Number 4

Published bimonthly by the National Institute for Computer-Assisted Reporting
www.nicar.org

Bits & Bytes

Training in Savannah

Investigative Reporters and Editors will offer computer-assisted reporting training during its regional conference in Savannah, Ga., Oct. 19-20. The hands-on CAR workshop will be 9 a.m. to noon Sunday, Oct. 20. There will be a limited number of registration slots open for the basic skills workshop. The investigative reporting sessions on Saturday will spotlight nut-and-bolts reporting skills and hot issues in the South.

For updated information on speakers and the conference schedule, go to ***www.ire.org/training/savannah02/***

Data updates:

The IRE and NICAR Database Library recently updated the following databases:

Federal spending

The Consolidated Federal Funds Report database, current as of fiscal year 2001, is a valuable tool for exploring federal expenditures in detail. Using the data, collected by the U.S. Census Bureau, journalists can look at federal assistance and overlay that information with demographics and housing about counties and cities. The database shows how

continued on page 2

RITALIN ROULETTE

State data pinpoints drug use by ZIP code

By Mark Skertic, *Chicago Sun-Times*

The story sounded so simple when my editor gave me the assignment: Do something on the amount of Ritalin and similar drugs used in the Chicago area to treat children for attention deficit/hyperactivity disorder (ADHD).

It turned out to be a months-long effort that involved the first time an Illinois newspaper using data from the state's prescription-monitoring program. Of course, when I got the assignment I didn't realize Illinois even collected the data.

Seventeen other states have similar programs – California, Hawaii, Idaho, Indiana, Kentucky, Massachusetts,

continued on page 18

SPOTLIGHT: CAMPAIGN FINANCE

Covering money, votes and power

By Aron Pilhofer, *IRE's Campaign Finance Information Center*

Campaign finance has been big news of late, whether it's the gushing generosity of Enron executives, the major parties' insatiable appetite for soft money or the rise of millionaire candidates funding $100 million campaigns.

And this fall comes the main event: On Nov. 6, the Bipartisan Campaign Reform Act – the first major overhaul of the federal campaign finance system in a quarter century – will kick in.

Although the new law only applies to federal candidates and committees, it

continued on page 17

SPOTLIGHT:

FOR MORE ON CAMPAIGN FINANCE SEE:

Bits & Bytes

continued from page 1

much money the federal government has spent. It covers federal expenditures or obligations for the following categories: grants, salaries and wages, procurement contracts, direct payments to individuals, other direct payments, direct loans, guaranteed or insured loans, and insurance.

Federal contracts

The federal contracts database provides details about more than 560,000 contracts, including the dollar amount, contracting agency and the type of product or service involved. The database contains information about each federal contract worth more than $25,000. For a few ideas about using the database see page 9.

Uniform Crime Reports

The FBI Uniform Crime Reports database is now current through 2000. Most of the data consists of the "index" crimes: murder, non-negligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny/theft, motor-vehicle theft and arson.

In addition, the database provides details about property crimes and includes the dollar values of the crimes. Additional data show age, sex and race information of people arrested.

The Supplemental Homicide Report table provides more detailed information on homicides, including demographic information about the victim and offender.

To order databases, contact NICAR at 573-884-7711, or go to ***www.ire.org/datalibrary/databases/*** for more information.

CENSUS UPDATE

As SF3 data flows, preparation still key

By Jeff Porter, *IRE and NICAR*

The U.S. Census Bureau's early missteps in the summer release of data show that even the agency itself can have trouble with the numbers.

So journalists, beware.

On June 25, the agency began releasing state-by-state Summary File 3 data – from the 2000 Census long form filled out by one of six households – starting with Vermont.

It became clear, though, that the data didn't correspond with a series of specifications or structure tables provided by the Census Bureau for users of Microsoft Visual FoxPro or Access database software. So the information didn't flow in correctly, creating angst for Census reporters. The bureau pulled back the data files, edited the specifications and structure tables then re-released the data on July 5.

Then it happened again. On the day the bureau's embargo was lifted, July 10, the agency reported:

"We have identified a *potential* problem with the Summary File 3 (SF-3) data files for the state of Vermont that the Census Bureau planned to release July 10, 2002. In accordance with our long-standing commitment to quality, we are postponing the release of all SF-3 files until we know whether any corrections or revisions are required. As we continue our review of these files, we will establish and issue a revised schedule for their release that allows us sufficient time to ensure the data's accuracy."

```
C:\sf1ascii>dir *.uf1

 Volume in drive C has no label
 Volume Serial Number is 1057-18CE
 Directory of C:\sf1ascii

DEGEO    UF1     8,161,404  09-13-01 12:04p degeo.uf1
DE00001  UF1     9,813,749  09-13-01 12:05p de00001.uf1
         2 file(s)     17,975,153 bytes
         0 dir(s)     134,721,536 bytes free

C:\sf1ascii>rename *.uf1 *.txt

C:\sf1ascii>dir *.txt

 Volume in drive C has no label
 Volume Serial Number is 1057-18CE
 Directory of C:\sf1ascii

DE00001  TXT     9,813,749  09-13-01 12:05p de00001.txt
DEGEO    TXT     8,161,404  09-13-01 12:04p degeo.txt
         2 file(s)     17,975,153 bytes
         0 dir(s)     134,721,536 bytes free

C:\sf1ascii>_
```

DOS commands to rename files

A data primer

The data files, finally released and a schedule that now might stretch later in the year than first intended, are massive. All told, the Summary File 3 release includes 813 detailed tables of Census 2000 social, economic and housing characteristics. A total of 484 tables address population; 329 address housing information. Then there's one more: a geography table, simply identifying for each record the level of geography, the name and codes of those geographic areas. For example, it might identify

one record as the state of Vermont with the code "50." So that makes 814 tables.

An impossible task? Not at all, in part because of the Access database specifications and the FoxPro and other structure tables that were problematic in the first release but were lifesavers when they were corrected.

First, the Census Bureau combined many of the tables down to a much more manageable collection of 76 data files, plus the geography table. So instead of 814 files to deal with, a journalist can take 77 files and find story after story, either national or local in scope.

Data, stories galore

Journalists can go back to the Census trough time after time to find or strengthen stories. The Census, after all, is simply counting people, their houses, their dollars. More than just Census stories live in the data. (For specific examples, see the related story in the *Uplink* May-June 2002 edition.)

But before one can pursue a story, one must pick up the data. Since most of the computer-assisted reporting world operates on Microsoft Access, let's focus on that.

To find a Summary File 3 Access 97 database with the corrected specifications, go to *www.census.gov/support/SF3ASCII.html* and download the Access file. If you're using a newer version, you can simply tell your computer to convert the Access database to the correct version.

Now, where's the data? Assuming you have a media password (or if your state's embargo is already lifted) go to *ftp://ftp2.census.gov/census_2000/datasets/Summary_File_3/* for the FTP version or *www2.census.gov/census_2000/datasets/Summary_File_3/* for an only slightly prettier Web interface to download the data.

With a caveat: Even on the late July data release, the bureau erred on the structure of the geography table, which is needed to tie specific Census numbers to a geographic area. Luckily, the geography table structure is the same item used in the earlier releases of Census data, so you can simply download the geography table specifications from the earlier release of Summary File 1 from *www.census.gov/support/SF1ASCII.html*.

Census files into Access

You'll need to unzip all those compressed files and, if you're an Access user, you might need to rename the text files. They all end with the suffix ".uf3." To change that suffix into ".txt" so that the software will recognize the files, visit the same Web page, where you can review the DOS procedure using the rename command.

Once that's accomplished, you can use the Access database you downloaded earlier with these simple steps (found at *www.census.gov/support/SF1ASCII.html#Microsoft Access procedures,* complete with screenshots).

- Click "File," "Get External Data," then "Import" from menu
- Select text file and click on the "Import" button
- Click on the "Advanced" button
- Click on the "Specs" button
- Select matching import specification and click on the "Open" button
- Click on the "OK" button
- Select option to store data in a new table or in an existing table (then select matching table), then click on the "Next" or "Finish" button. If you choose an existing table, you're finished. That's highly recommended.
- If not, highlight updated table and then click on the Design button
- Right click on field LOGRECNO and then select "Primary Key"

Finding the geography

The field LOGRECNO is the key way to link to the geography table. That table is important not only because it identifies by name the state, county or tract, but it helps you decide which places to examine and how to add them up. For example, there's a field called GEOCOMP. That stands for "Geographic Component," a number code, which includes not only entire areas such as counties, but smaller areas within a county – the urban area, the rural area, and so on. So if you want to count all the people in a specific county, then you'll want to filter the records to only include GEOCOMP number "00." That's the code to include everything, not specific geographic component types.

To find detailed information on each of those geographic table fields, go to *www.census.gov/support/GeoFile.html* for general information then dive into details at *http://www2.census.gov/census_2000/datasets/Summary_File_3* to download the file "0SF3_geo_header.doc." While you're there, also download and read, if you haven't already, the file "0README_SF3.doc," "0SF3_File_Structure.doc" and "0SF3_table_matrix.doc."

Jeff Porter can be reached by e-mail at jeff@nicar.org

readme.txt

Web addresses

In addition to the Web pages listed in the accompanying article, these can help a journalist to understand and use the Census data:

www.2000census.org
Investigative Reporters and Editors offers training opportunities, tipsheets and helpful links.

http://cronkite.pp.asu.edu/census/
Created by Steve Doig, Knight Chair in Journalism for Arizona State University, who specializes in computer-assisted reporting. The Web site helps journalists and others make good use of Census data.

www.censusscope.org
An easy-to-use tool for investigating U.S. demographic trends, from the Social Science Data Analysis Network at the University of Michigan.

MAPPING IT OUT

This Uplink feature shows the latest uses of mapping in news reporting.

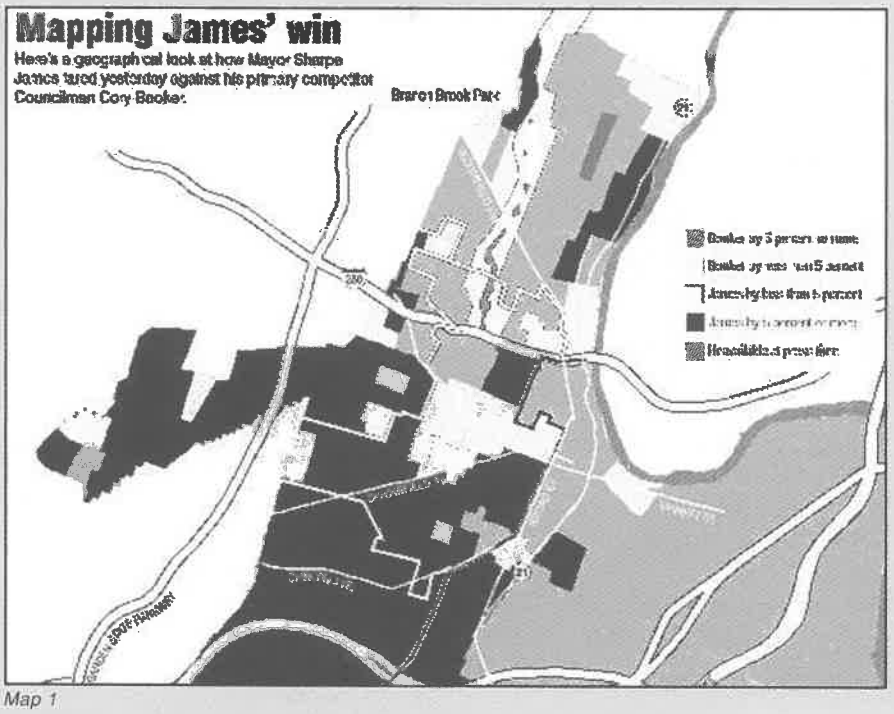
Plotting political moves

It's the political season – a great time to hunt for stories that show your readers and viewers the pockets of political power in your area. A number of journalists have turned to geographic information system (GIS) software to visualize campaign contributors and election results and find stories that otherwise might have been hidden inside tables of data.

Results and turnout

In early May, *The Star-Ledger* in Newark, N.J., mapped the results for the primary election race between incumbent Mayor Sharpe James and challenger Cory Booker, a city councilman. Using ArcView, CAR specialist Robert Gebeloff created a thematic map that showed the areas taken by each and by what margin. (See Map 1). For the analysis, Gebeloff mapped a polygon file showing the city voting districts and shaded the layer with four colors that displayed the victory margin percentage for each candidate. Graphic artist Angela Porter took the information generated by Gebeloff and created a map for publication. The voting district files came from the U.S. Census Bureau and had to be edited to reflect redistricting changes.

The day after the election, Gebeloff dug deeper and used ArcView to show how voter turnout contributed to the mayor's primary triumph. He recycled the voting district map and this time shaded it with just two colors: one for the districts won by James and another for the districts taken by Booker. Gebeloff next overlaid another copy of the voting district polygon file and displayed it as a density map that created a dot for every 20 voters casting ballots in the election. The map (See Map 2) showed that James fared well and drew a large turnout in the city's most densely populated wards, while Booker failed to generate as high a turnout in the areas he won.

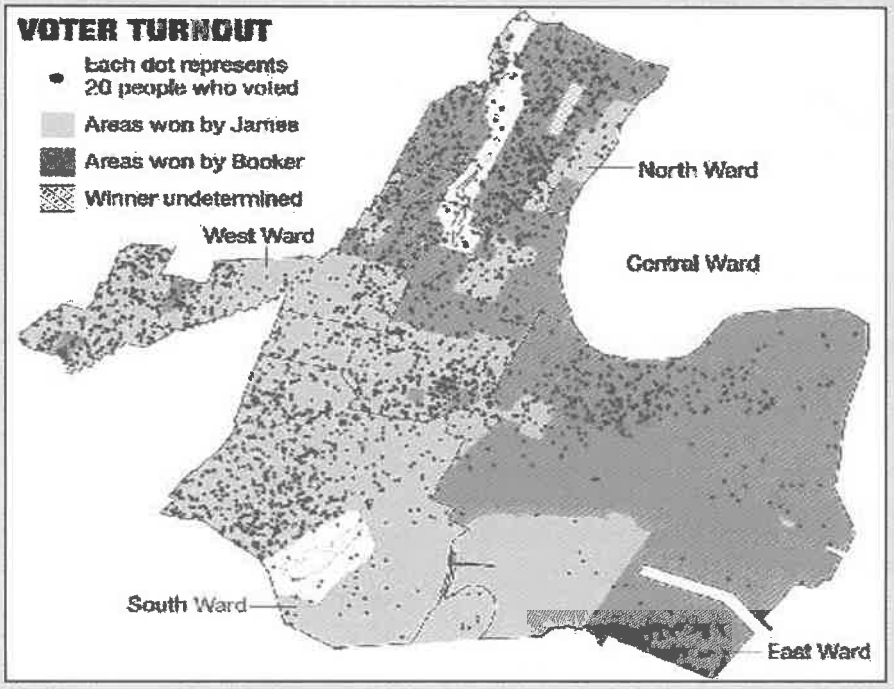


Map 1

Race and recall

Reporters at the *Flint* (Mich.) *Journal* used analytic mapping for a February story about the petition drive to secure a recall election of the city's black mayor. A team of three reporters spent weeks entering into an Excel spreadsheet key pieces of information, including the names and addresses of the 13,000 people who signed recall petitions. Mary Ann Chick Whiteside, the new media manager at the paper, plotted those addresses in MapInfo, using the MapMarker geocoding tool to create a point for each address. She manually created voter precinct maps in MapInfo, using a paper map as a guide. Then, with both layers open in MapInfo, she assigned the precinct number to each signer of the petition.

Finally she analyzed the petitioner data and the results were telling: more than 70 percent of the 10,000 valid signatures came from predominantly white neighborhoods. Less than 10 percent of the valid signatures came from predominantly black neighborhoods. (See the story on the Web at *www.mlive.com/news/fljournal/index.ssf?/fljournal/special/stories/recall_2.html*). After the reporters confronted recall leaders with the results of the analysis, the leaders admitted that they had targeted white neighborhoods for their drive.



Map 2

Mapping contributions

The Center for Responsive Politics used GIS to take a look at the source of 1999 presidential candidate contributions. Using Maptitude, the center created for each state a trio of three-dimensional maps showing the population of every ZIP Code and the amount of contributions to Democratic candidates and Republican candidates. The maps clearly showed that the contribution patterns often failed to mirror population patterns. You can see the maps on the Web at *www.opensecrets.org/states/presmap.asp*.

Election Machine

Any journalist who's interested in mapping election results should check out the Election Machine package developed by Neil Reisner of the *Miami Daily Business Review*. IRE and NICAR have made it available for download at *www.nicar.org/electmach/index.html*.

The Election Machine provides the tools you need to jumpstart your election coverage using GIS. You'll find Access, Excel and ArcView files and a PowerPoint slideshow that explains how to use the tools.

Reisner, the developer of the package, is a former national trainer for NICAR.

Would you be willing to share a mapping example with fellow journalists? Send an electronic copy of the map along with details to David Herzog at ***dherzog@nicar.org***

SPOTLIGHT: CAMPAIGN FINANCE

Data can help profile a candidate's street machine

By David Gulliver, *The Virginian-Pilot*

It goes like this: Millionaire decides to be politician, millionaire throws personal fortune into race, millionaire runs TV ads day and night.

Sometimes the millionaire still loses. Just ask Michael Huffington or Ross Perot – or Mark Warner. The cell phone mogul put $6 million of his money into a Senate run from Virginia in 1996. He lost.

Last year, Warner tried again – this time for Virginia's governor.

So by November, from Seinfeld at 7 to the news at 11 you'd see ads for Warner or his Republican opponent, Mark Earley.

And when Warner won in a state that hasn't gone Democratic since Doug Wilder in 1989, it was easy to figure he did it with just the ad blitz.

But we had a way to test it.

Data resource

In Virginia, newspapers have a great resource, the Virginia Public Access Project (VPAP).

It began when Virginia newspapers teamed up to build a campaign contributions database for the 1997 elections. David Poole, a former reporter, runs the group.

Now, VPAP takes in candidates' electronic data, codes it, makes it available on the Web site *www.vpap.org* and (for subscribers) via downloads.

The database grows every year, in size and in detail. For example, the name and address of every person or business getting campaign money is listed. Date and reason for payment also are listed.

Poole pointed out those features in midsummer, as I was importing a first round of data. We checked it then and saw Warner paying a lot of people in two cities near our region.

I planned to look at our cities when the post-election reports arrived, in mid-December.

Cleaning and querying

A look at what kind of organizations the candidates put into the field means getting a count of "street-level" employees. On paper, that's pretty simple – but we do this stuff on computers.

(For the record, I used SQL commands in Visual FoxPro 6.0 for most work, and exported tables to Excel to build charts.)

A field called "service," or the reason for payment, helped zero in on money for staffers. The campaigns did not respond to questions, so I had to develop criteria.

Earley, the Republican, coded most payments to field people as "Literature Drop/Flush" – as in "flushing the vote," like a hunter flushes game from the brush.

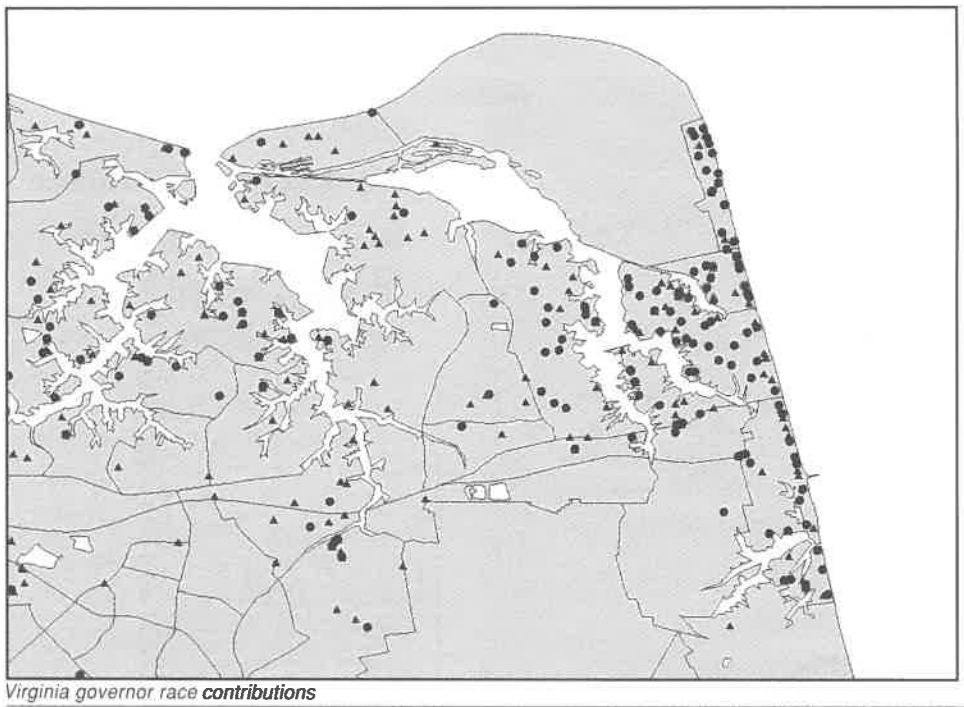
But about 100 entries were coded "Salary" or "Payroll." Some of the same people were coded under flush. Others were people who were first paid close to Election Day.

Those I counted as field employees. I excluded a handful of records – payments to a handful of top aides and to companies that managed his payroll. Warner's payroll was more complicated. He had two separate committees working to elect him, and had many employees multicoded — "Election Day Staff" and "Salary," for example.

I built his employee roster using the same criteria as I did for Earley. From there, it was simple to run queries to look at how each campaign built its operations.

Running a count of employees, overall and by city, let me compare the campaigns' staffing, statewide and by region.

Having the date of payment and date of report was important. I ran queries



Virginia governor race ***contributions***

that gave me how many employees received their first payment by each report date.

That allowed comparing how many people worked for each candidate at specific mileposts during the campaign.

Finally, addresses with several employees with different last names might help find some possible mischief – for example, dummy addresses or businesses funneling money.

Results showed surprises

The data let us debunk some myths about millionaires like Warner.

Warner didn't just win with ads. He built a massive street-level machine and outmanned his opponent in every way – city by city, and at every stage in the campaign.

By election day, he had almost 2,600 paid staff across the state – five times as many as the Republican. Warner spent $1.4 million on staff, more than double Earley's total.

Two months before the election, Warner had six times as many paid staff getting out his message. And in our region, Hampton Roads – also Earley's home turf – Warner still had twice as many paid staff.

That was to be expected. In October, a map showed more than 1,000 contributions from five cities to the two candidates. Warner received heavy support from traditional Republican turf like the Virginia Beach oceanfront – and even in Earley's home city, Chesapeake (bottom center of map; dots represent those who contributed to Warner, and triangles show Earley contributors).

Earley's aides said they had more unpaid volunteers than Warner – something impossible to test with the data – but admitted that Warner's paid army overwhelmed them.

One weakness was that previous data was not as detailed, so there was no way to compare "street-level" efforts in past campaigns. But professors, party bosses and ward leaders told us that there'd never been an effort like Warner's, and that Democrats seemed to have stolen the Republican playbook.

The last round of queries – looking for multiple-person addresses – paid off in an unexpected way.

The suspicious addresses turned out to be families with grandparents, parents, stepbrothers and sisters all under one roof.

And so it included a human angle – and some insider stuff too. The homeowner was a part-time teacher, who got a call from the teacher's union looking for volunteers. And so she, her parents and kids got paid to put brochures under windshield wipers after church for a few Sunday mornings.

Lessons from the data

This data was unusually clean. The story took two-plus days for data work, and two-plus days of reporting, writing and editing. If your own data needs considerable cleaning, leave more time for cleanup and verification.

If you don't have such detailed data, use the campaign's overall salary spending and office space expenses to measure its get-out-the-vote effort.

David Gulliver can be reached by e-mail at dgulliver@pilotonline.com

readme.txt

Resources

Tipsheet #1499 from the IRE Resource Center describes the Virginia Public Access Project's campaign finance data work. Members can obtain tipsheets at ***www.ire.org/resourcecenter/***. Members and nonmembers can also obtain copies by calling 573-882-3364.

data AND THE NEWS

Business and data

The WorldCom debacle is still unfolding after the company admitted that it had improperly accounted for almost $4 billion in expenses. Journalists can dig up information about that company and thousands of others in a database of federal contracts. A substantial part of WorldCom's income is from federal dollars.

IRE and NICAR offers the Federal Procurement Data System of federal contracts. IRE and NICAR can provide data for fiscal years 1992-2001.

Disasters in data

In the coverage of disastrous wildfires in Western states, data can provide some perspective. From the Data Library's Storm Events database: On May 4, 2000, for example, a fire in New Mexico began that eventually caused $1.5 billion worth of damage.

Despite its name, the Storm Events database includes much more than tornadoes, hurricanes and snowstorms. Obtained from the National Climatic Data Center, it's the official U.S. government database of many natural disasters around the country. The database covers 1950-2001 and includes such items as the type of disasters, states and counties, the dates when the disasters occurred, dollar amounts of damages, fatalities and injuries, and in many cases, a memo field with additional information.

SPOTLIGHT: CAMPAIGN FINANCE

Looking at the spending side of the ledger

By Hank Shaw, *The* (Fredricksburg, Va.) *Free Lance-Star*

Most campaign finance stories focus on who's giving money to candidates; after all, controversial political action committees such as the National Rifle Association or Planned Parenthood draw headlines, as do celebrity contributors and big state or federal contractors.

Know candidates by the friends they keep, the saying goes. Well, an even better way to "know" a candidate is by looking at the spending side of the ledger.

In Virginia, there are few competitive congressional races this year—or any other, for that matter. Carefully drawn districts see to that.

Nevertheless, a look at every area representative's reports as a matter of course can pay off. Consider the case of 7th District Rep. Eric Cantor, a first-termer sitting in one of the most Republican districts in America.

Cantor had no opponent for the first 15 months of the 2002 election cycle, until former Georgia Rep. Ben "Cooter" Jones said he'd challenge Cantor in early March. Jones played Cooter on "The Dukes of Hazzard" during the 1970s. He later served two terms in Congress.

Surprise in the data

But Jones' challenge would not have been reflected on Cantor's April 15 FEC report, which only covered until March 1—just two weeks before Jones announced.

So I expected a sleepy report. Not so. Cantor had already spent an astonishing $674,000, between his federal campaign account and a state leadership PAC. That rivaled NRCC Chairman and 11th District Rep. Tom Davis' war chest and blew away everyone else in the commonwealth or even the region.

Why would someone spend so much so early in an election year no one expected to be much of a challenge? I had an idea, and some number-crunching proved it right.

First thing was to download all of Cantor's Federal Election Commission reports from *www.fec.gov*.

Following the directions, I dumped Cantor's reports into an Excel spreadsheet, added them in order and stripped the contributor information off before formatting any data, leaving only the spending side.

Looking at the expenses

It proved to be a fairly large—but manageable—data set, with 465 entries. By way of comparison, Cantor's neighboring first-termer, 1st District Rep. Jo Anne Davis, lived far more frugally. Her list contained only 128 entries.

Just eyeballing the data, it was clear that Cantor's chief political consultant, Ray Allen Jr., had both feet in the trough. His name, and those of his various employees and companies, made the list numerous times.

A simple pivot table separated Allen items from the rest. Dropping the vendor name in the row section and the amount in the data section, the table showed how much money each of Allen's entities received, nine in all.

It turned out he'd received more than $300,000 in consulting fees, direct mail expenses and payroll from the Cantor campaign. That's more money than all but one other campaign in Virginia had even raised, let alone spent on one consultant.

Some of the examples for the article: $47,000 worth of catering expenses, $25,525 for two polls—one of which was conducted in March 2001 — and $5,162 on plane tickets for fund-raising trips.

Cantor also spent $282 for a night at the Minneapolis Hyatt and $380 for two nights at Boston's Logan Hilton.

Data can cement stories

None of this required any serious knowledge of Excel; database manager Access wasn't needed for any serious queries. Expect this to change, however, once Cantor begins fundraising for his race against Jones.

Cantor was always known as a silver-spoon candidate, and his consultant Allen is famous statewide as a spender of other people's money. This story cemented those reputations with hard data, where before all the evidence was merely anecdotal.

Cantor and Allen said they expect to raise at least $1.5 million to defeat Jones this fall. Bet on every penny getting spent, too.

Hank Shaw, who just left *The Free-Lance Star*, can be contacted by e-mail at scrbblr@hotmail.com

SPOTLIGHT: CAMPAIGN FINANCE

Boot Camp boosts campaign reporting

By Jim Miller, *The Modesto Bee*

Until I attended a NICAR Boot Camp in summer 2000, I knew almost nothing about how computers could help on stories. While still learning, I've become comfortable using CAR covering campaign finance.

A pretty straightforward story a few days before the March 5 primary focused on the location of candidate donors. Was the source of money within the district they were running for, or from donors elsewhere?

With some campaign contribution data already in Access because of earlier stories, I updated the information from the Federal Election Commission and California Secretary of State Web sites and added a field called "in-district." For each record, I entered "yes" or "no" based on the donor's address. It went pretty fast, despite the number of records.

Tracking money sources

California, like other states, redrew its political boundaries last year. Some contributions came from cities – such as Modesto – that had been divided like postwar Berlin. ArcView could have plotted donor addresses, but I decided to classify split-city contributions as in-district.

Access queries counted the numbers of in-district "yes" and "no." The results showed the breakdown of in-district and out-of-district contributions. Then I started calling candidates and experts.

A state Senate candidate who received more than 90 percent of his money from outside the district explained it this way: He got into the race late and so needed to focus on big-dollar donors in Los Angeles and Sacramento. And, he said, statewide organizations such as unions and business associations that gave to him have local members. His outspent primary opponent, who raised most of his money in the district, predictably said the findings proved that he was more in touch with local issues.

Experts provided important perspective for the story. One noted that out-of-district money lets a candidate take positions different from those of local moneyed interests, i.e. farmworker rights in an agricultural area.

As it was, the candidate with heavy out-of-district support won the primary.

Sharing political foes

After the primary, there's another campaign finance story: Looking at how many donors had given to two opposing candidates for Congress during the candidates' time in the state Legislature.

Dennis Cardoza, a Democrat assemblyman, and Dick Monteith, a Republican state senator, became part of the biggest political story of the year when they ran for the Modesto-area congressional seat now held by Gary Condit. Both won their respective primaries.

Until then, Monteith and Cardoza were considered political allies, with similar positions despite belonging to different parties. There also was an apparent understanding of "don't help my opponents and I won't help yours" between the two.

But as they say in politics, if you want a friend, get a dog. Primary victories in hand, Monteith and Cardoza began to bad-mouth the other's legislative records.

Campaign contributions – from a compiled database of their campaign contributions from 1997 through 2001 – gave a good perspective on their histories in the Legislature.

continued on page 12

boot camp

continued from page 11

Labor-intensive data

Fortunately for my wrists, my employer paid a firm about $250 to enter approximately 1,500 records from 1997 and 1998. I entered 1999 data and downloaded 2000 and 2001 reports off the Internet into Excel. Queries from the database pulled information on matching donors.

There were considerable headaches, however. Over the years, campaign committee treasurers had listed contributors a variety of ways, such as Gallo Winery, EJ Gallo, E&J Gallo Winery, etc. One query found a donor match for Gallo Winery, for example, and more searches for "*gallo*" found other company contributions. Such wildcard searches helped on this story.

Also, after seeing a message on the NICAR-L listserv, a careful check of the data-entry firm's work found some errors.

Gradually, though, a list of dual donors emerged. There were dozens of them, totaling hundreds of thousands of dollars in contributions. During interviews, many of the donors said they liked both lawmakers and agreed with their voting records.

Cardoza and Monteith, of course, disagreed that a shared past of campaign donors implied similar political views. And campaign finance experts noted that businesses and interest groups often give to candidates of all stripes to get access.

Still, CAR gave some perspective on both candidates' newfound criticism of the other. More CAR campaign finance stories are planned as the November general election approaches.

Jim Miller can be contacted by e-mail at jmiller@modbee.com

SPOTLIGHT: CAMPAIGN FINANCE

Working with FEC electronic filing data

By Aron Pilhofer, *IRE's Campaign Finance Information Center*

From import to analysis, working with raw federal campaign finance data used to be a frustrating chore for even the most skilled CAR reporter.

But not anymore.

The Federal Election Commission's new electronic filing Web site makes viewing, importing and analyzing data fast enough for deadline reporting, and easy enough for anyone with basic CAR skills and a spreadsheet.

The FEC has been experimenting with electronic filing for years. But the 2002 election cycle is the first in which it is mandatory for political parties, PACs as well as candidates for president and U.S. House. For journalists, this is a boon.

For deadline, electronic filing data is simply better than the raw data used in previous cycles. It is more complete because it includes both contributions and expenditures – something journalists never had easy access to before.

It also includes full street addresses, rather than just city, state and ZIP Code. That makes the task of positively identifying contributors, their relatives and spouses much easier.

Viewing data

The FEC electronic filing site is located here: *http://herndon2.sdrdc.com/dcdev/*. From the main search page, you can find filings any number of ways: by committee, name, state, party, report type, the filing date or committee ID.

Once you've located the committee you are interested in, you'll see all the filings for the current cycle listed under the committee name in order of the date they were filed with the FEC. Here's a typical example:

FEC-18381 Form F3N - period 01/01/2001-06/30/2001, filed 07/31/2001 - MID-YEAR

The first code (FEC-18381) you can ignore. It's an internal code used by the FEC as a unique identifier. The rest tells you what form is being filed, whether it is a new or amended filing (amended filings will have an "A", new filings will have an "N"), what reporting dates it covers, when it was filed and the name of the reporting period.

It's important to note that just because a filing appears at the top of the list does not mean it is the most recent filing. Frequently, committees will amend or correct older reports, and those will appear on the list in order of the date they were filed with the FEC, not the filing periods they cover. The older reports won't be removed but will have a small "amended by" notation added along with a link to the updated report.

The bottom line: Make sure you're looking at the most up-to-date version of the report.

Once you've got the filing you want, you can view it by clicking on the "view" link just to the left of the label. This will produce an electronic version of the old paper records organized in exactly the same way with the summary pages first followed by the detailed schedules.

Download and import

To get a file onto your local computer for spreadsheet work, simply click on the "Download" link next to the one you want. The file will be saved with a ".fec" extension, but it is nothing more than a simple comma-delimited text file.

You'll also want to save yourself some typing and download a list of column headers from the FEC site. They are found in a self-extracting file, which you can download by clicking on the link at the top of the page titled "Format documentation for downloads." The headers file is an Excel document titled "e-filing headers.xls" and includes all the column heads for each of the FEC forms. Once you have your filing imported, you can simply copy and paste in the appropriate column head.

To begin importing, launch Excel, click File-> Open, navigate to the folder in which you saved your filings and open the filing. Unless you renamed it when you downloaded, it will be named something like "21232.fec".

Click on your file, and click the "Open" button and the text import wizard will open. In order for the data to be imported correctly make sure of the following:

- Use the "start import at row" selector to begin at row 3. The first two lines are not important for us. They are a header row that indicates what software the committee used, and the second line is the aggregate data. (Not all versions of Excel have the "start import" feature. If yours doesn't, you'll have to delete the first two rows by hand later on.)

- If you are importing expenditures (which have a different layout than contributions, but are in the same file) you'll want to open the filing in a text editor before importing into Excel and determine what line the disbursement data begins on, and use the "start import at row" selector to exclude contributions.

- By default, all the columns are formatted as the "General" datatype, which means Excel will guess what sort of format you want based on what type of data it thinks the column contains. This will cause problems for ZIP codes, which often have leading zeros. To avoid losing data, make sure the ZIP field is formatted as text.

- The date fields are unformatted and are listed as year, month then day (e.g., 20011231). To convert them to dates on import, change the column to the date

continued on page 14

FEC

continued from page 13

format and select the YMD option in the pulldown menu. Now, your dates will be formatted correctly.

Final touches

There are a few more steps you might want to take before you start crunching and munching the data.

■ Scroll down the spreadsheet until you see the "SA" codes in Column A change to "SB" (SA means schedule A, or, contributions, and SB means schedule B, or, disbursements). You don't want to have both in the same sheet, and if you scroll over you'll see some of the data was corrupted during the import because disbursements have a different layout than contributions even though they are in the same file. I find it easiest to simply note the row on which the disbursements begin, delete them out of the sheet and then re-import starting at the row on which the expenditure data begins.

■ Wouldn't it be nice to know what all these columns mean? Open the Excel headers file you downloaded earlier and find the row that matches the code in the first column of your data (e.g., SA, SB, etc). Find the appropriate column headers, then copy and paste into your spreadsheet.

■ You may have already noticed that the name field is a single column, with individuals' first and last names delimited by the caret (^) character. There is a quick fix for this: Click on the Contributor Name column to select the entire thing, and select Insert-> Columns four times. You should have four blank columns to the right of your Contributor Name column. Now comes a nifty trick: Assuming your Contributor Name column is still selected, select Data-> Text to Columns and a wizard similar to the text import wizard will appear. In step 2, click the "other" delimiter option, and in the window type a single caret (shift-6). Not all campaigns will use four fields for names, but some do. It's better to add too many blank fields at the beginning because if you are not careful with the text to columns feature, it will over write data in your spreadsheet. Don't forget to type in the appropriate column headers, e.g. first name, last name, prefix, suffix.

Ready for stories

The electronic filing data gives you the ability to analyze in ways that would have been impractical using paper records: Calculate the number of out-of-state contributors. Look for evidence of bundling, kiddie contributors, spouses and other related individuals. There are any number of ways to look at the data.

At this point you could also import your spreadsheet into a database manager like Microsoft Access or FoxPro to compare it with previous filings or those of another candidate.

With electronic filing reporters can do more than write the standard deadline horse race story – who's raised the most, and has the most on hand. Now, you can bring some context and depth to your reporting, and do it on deadline.

Aron Pilhofer can be reached by e-mail at aron@ire.org

readme.txt

Resources

For more information about campaign finance, including data, training, tipsheets to help journalists and more, go to ***www.campaignfinance.org***. For a more detailed tutorial about electronic filings, especially for beginners, go to the IRE Resource Center Web site at ***www.ire.org/resourcecenter*** or call 573-882-3364

Beware: Data caveats and tips

■ The U.S. Senate exempted itself from the electronic filing requirement, so you will not find any Senate filings at the FEC electronic filing site. Those data are available only in the raw files available from the FEC's FTP site.

■ The FEC posts filings almost instantly on the electronic site, even before they have been reviewed by staff. Reports are frequently amended, so be aware that the data you download may be changed later on. Before you run with a story, it is a good idea to make sure the reports are the most up-to-date.

■ The "official" data is still maintained on the agency's FTP site (and from the IRE and NICAR Data Library, already processed into DBF format). If you are not working on a story that requires the absolutely most recent filings, those are the files you should use.

■ Just because the data is filed electronically does not mean it is clean. Expect to find mistakes. But committees are using software that standardizes names and addresses somewhat, creating fewer data-entry errors.

■ Before you dive in to the number crunching, print out a copy of the summary pages from the filing just to make sure your aggregate numbers are correct. If things don't add up, it's probably because the campaign has memo entries in the data with additional information about certain types of contributions. Filter those out using the Memo Code or Memo Text columns.

■ It's also best to use the codes in column A to break out the different types of contributions, and ignore the "Entity Type" column, which frequently is incorrect. The code tells you everything you need to know anyway – the schedule and line number on which the contribution will be aggregated.

Tech tip...

SPOTLIGHT: CAMPAIGN FINANCE

Setting up an automated FTP download

By Griff Palmer, *San Jose Mercury News*

When looking for an FTP client, power users often cast aside the lowly command line client bundled with Microsoft's Windows 32-bit operating systems in favor of jazzier GUI clients.

Beneath the Win32 client's drab exterior, though, is a program with scripting capabilities to rival programs from the Unix side of the tracks. Database software FoxPro, with its own scripting abilities and its ability to call other programs, makes a perfect complement for this deceptively plain-clothed client.

Any command that can be invoked from the FTP command line can be written into a text file. Whole strings of commands can be stored in a script and executed en masse.

Running a FoxPro script in concert with an FTP script, and throwing in some PKZip as needed, it's possible to automate complex data retrieval and importation tasks.

Key to this process is the fact that FTP can capture not only data files but also the contents of directory listings. A text capture of a directory listing, is a nice, structured file that can be parsed into a FoxPro file, edited and output to a text file that can, in turn be invoked as an FTP script.

Think about it for a second: You can capture a directory listing from a remote FTP site and use the file names from that directory listing to form the backbone of an FTP script. It can save a lot of typing.

Preparing for the script

Putting together such a script can be time-consuming. The effort may not be worth it for one-time jobs. But for FTP jobs you run regularly, these scripts can be real time-savers.

Let's go through a FoxPro/FTP/PKZip trifecta that goes to the FEC site, captures data files and unzips them. (From there it should be obvious how to add commands to the script to append the unzipped files into pre-structured tables.)

There are a couple of prerequisites: (1) a tweak to your FoxPro settings; (2) a shell file into which to append the directory listing once you've captured it.

To adjust our FoxPro settings, we've got to stop thinking scripts for a minute and delve into the Windows GUI.

Using your favorite method, browse to the FoxPro directory (FPW26) in a standard FoxPro 2.6 installation. Look for a file icon labeled "FOXRUN." Right-click on the icon and choose Properties. Click on the Program tab in the Properties dialog. Near the bottom of the dialog is a checkbox labeled "Close on exit." Make sure the checkbox is checked. This setting assures that control is returned to FoxPro after an external program has done its stuff. (The same process can work with later versions of FoxPro as well.)

Making directory listing

The shell file should contain one field wide enough to accommodate the widest path/filename listing in a directory display. For what we're doing here, 50 bytes will be wide enough. Create a table called dirlist.dbf with a single field called filenames.

We're going to break this into two scripts. The first script will go to the site and capture the directory listing. It consists of a text file called ftpscrpt.txt, containing these commands:

```
cd /FEC/electronic
ls /FEC/electronic text.txt
close
bye
```

The first line changes to the appropriate directory on the FTP site. The second line calls a directory listing of /FEC/electronic and redirects the output to a file called text.txt. Then the script ends the FTP session.

But how do we start the FTP session and invoke the script? The following command would do it from a DOS prompt:

```
ftp -s:ftpscrpt.txt -A ftp.fec.gov
```

FTP launches the client. The -s: switch, followed by a filename, causes the client to treat the contents of that file as a script. The -A switch

continued on page 16

Tech tip...

continued from page 15

forces the client to log you on as an anonymous user. (If you find yourself using a server that doesn't allow anonymous log-ins, just drop the -A switch and add a user name and password to your ftp script file.)

The command string and switches are followed by the address of the ftp site.

Last steps for program

This same string of commands and switches can be called from inside a FoxPro script. Create a program file. (I named mine, fittingly enough, ftp.prg.)

By simply adding the word "run" to the line above, the FoxPro script will launch the FTP session and carry out the commands in the ftp script file:

```
run ftp -s:ftpscrpt.txt -A ftp.fec.gov
```

Once the ftp script is finished, control will be handed back to FoxPro, and you'll have a text file called text.txt, containing the directory listing from the FEC site.

Now you can use your FoxPro script to work a little mojo on that text file. Everything below – including explanatory notes that begin with the "&&" character string — is in our Fox program file, ftp.prg:

```
run ftp -s:ftpscrpt.txt -A ftp.fec.gov
&& The ftp script we've already
invoked

use dirlist
&& Opens the shell file we created
earlier

append from text.txt type sdf
&& Appends the directory list into the
.dbf file

replace all filenames with "get
"+filenames
&& Adds an ftp "get" command before
each file listing.

go top
&& Goes to the top

insert blank before
&& Inserts a blank row

replace filenames with "binary"
&& Adds an FTP "binary" command,
critical for downloading binary files
such as .zips or .jpegs

replace filenames with "cd /FEC/
electronic"
&& Adds an FTP change-directory
command

go bottom
&& Goes to the bottom of the file

append blank

replace filenames with "quit"
&& Adds an FTP "quit" command

&& Suppose we didn't want all the
files in the directory listings, but
only those produced after certain
dates. This is easy with the FEC data
because the file naming conventions
are date-based. You can use the file
names to eliminate from the script
those files you're not interested in
getting. Suppose we don't want any
files dated before March 25, 2000:

delete for filenames='get' and
substr(filenames,21,8) < "20020325"
pack
copy to ftpscrp1.txt fields filenames
type delimited with blank
&& Now we copy our work out of the
.dbf file to a .txt file "delimited
with blank" is necessary to eliminate
leading spaces from the text file.

&& Et, voila! We've now created a
second FTP script that will go back to
the FTP site and retrieve the files
we've specified.

close all
run ftp -s:ftpscrp1.txt -A ftp.fec.gov
&& We run the newly created script
file, ftpscrp1.txt

run \pkzip\pkunzip *.zip
&& Then we unzip the files we've just
created.
```

Avoiding other problems

There is one issue here. The old PKZip v. 2.04g doesn't support long file names. When you run into long file names that can create problems. WinZip offers a command line interface, which theoretically should work as well from a FoxPro script and should overcome the long file name problem. Doubtless there are other command line-based compression programs out there that can be worked equally well into this scheme.

If you regularly run a FoxPro/FTP script combination, you might forget to erase text files and to clean out your shell file. A pre-existing file can make your script stop and prompt you for an overwrite — no big deal unless you're counting on that script running itself while you take a lunch break. Appending file names to a shell file that's already got data in it has the potential for creating nasty confusion.

You can simply add record-delete and file-delete commands to your FoxPro script to clean up detritus that may get in the way later.

In this example script, we used the "ls" command to capture the directory from the server. This gave us a short-form directory. However, there may be times when you want a long-form directory. For instance, if you want dates to key on for deciding which files to download and the site's file naming conventions aren't date-based, the date-time stamps in the long-form directory can be valuable.

Substituting "dir" for "ls" will give you a long directory.

Happy FTPing!

Griff Palmer can be reached by e-mail at GPalmer@sjmercury.com

readme.txt

Resources

To download a copy of the program ftp.prg, go to *www.nicar.org/techtips.html*. The Web page also features downloads and other supplementary materials for various articles that have appeared in past *Uplinks*.

power

continued from page 1

will affect all levels — county, state and local. For example, many observers predict that state and local committees will be the new conduits for the hundreds of millions in soft money the national parties will be prohibited from raising under the new law.

But first things first. Election 2002 is an especially crucial midterm election, with a closely divided electorate and a Congress reflecting that divide. Both parties believe they can win a majority in the House and Senate, and the national party committees are raising record amounts of hard and soft money.

In short, there will not be a dearth of storylines heading into the fall campaign season. And with that in mind, this issue of *Uplink* is intended to help add depth to your reporting on campaign finance.

■ David Gulliver of *The Virginian-Pilot* covers the trend of independently wealthy candidates.

■ David Herzog of IRE/NICAR shows how some journalists have used GIS software in their reporting to show visually, for example, the pockets of power.

■ Hank Shaw, formerly of *The* (Fredericksburg, Va.) *Free Lance-Star* writes about ways campaign finance data can be used to understand more about the interests backing a particular candidate, including the little used expenditure data.

■ Griff Palmer of the *San Jose Mercury News* demonstrates how to use FoxPro's scripting capabilities to automate some of the more laborious tasks, such as downloading and importing campaign finance data.

■ And I discuss the wealth of data to be found at the Federal Election Commission's new electronic filing Web site, and how easy it is to download, import and analyze it.

This is just a fraction of what's available through IRE, the Campaign Finance Information Center and other organizations.

For data, tipsheets and other information, visit the CFIC Web site at *www.campaignfinance.org*. Our Resource Center includes tipsheets written by some of the best in the business.

Tracker, our quarterly newsletter, is chock full of practical information and advice from reporters who know the issue inside and out. And of course, the NICAR Database Library has up-to-date data for every election cycle since 1991-92 in an easy-to-import format.

The FEC (*www.fec.gov*) has posted a number of easy-to-understand documents and reports on its Web site that will help reporters understand the complexities of the system. Links to your state elections offices can be found on the CFIC site (*www.campaignfinance.org/states/index.html*).

For data on the Web, Center for Responsive Politics (*www. opensecrets.org*) has one of the best searchable databases of federal data. The National Institute on Money in State Politics (*www.followthemoney.org*) offers the same sorts of lookups, but for state candidates and committees.

To learn more about how the law will change when the reform law takes effect, visit the Campaign Finance Insititute (*www.cfinst.org*), which has done an outstanding synopsis and analysis of the new law. Common Cause (*www.commoncause.org*) and Public Citizen (*www.publiccitizen.org*) also have written extensively on the subject.

Lastly, there is one trend that reporters at all levels should be aware of: the rise of political nonprofits, sometimes called Section 527 organizations. The groups, along with state and local party committees, are expected to be the loophole allowing soft money back into the federal system. They also are becoming more common in local and state elections. For an extensive history, download the Common Cause report "Under the Radar," available on its Web site. Public Citizen also has done three studies examining the problems of disclosure and policing of 527s, which ranges from poor to nonexistent. Those reports also are available on Public Citizen Web site.

The campaign finance issue will be big news for some time to come, and covering elections and the issues surrounding them is only going to get more complex as the parties and candidates must get more creative in finding ways to fund their elections. The resources above are a good start.

Aron Pilhofer can be reached by e-mail at aron@ire.org.

readme.txt

Training

IRE and NICAR Census seminars are being planned for journalists across the United States. For the latest information, go to ***www.ire.org/training/otr.html.***

Tipsheets

The tipsheets from the annual IRE conference in San Francisco are now available on the IRE Web site in the Resource Center section. You can search them at ***www.ire.org/resourcecenter/***. To find the tipsheets from this particular conference, search for San Francisco IRE Annual 2002.

Tipsheets are available on topics ranging from backgrounding public and private businesses to hospital and medical data sources to ten tips for writing tightly. If you have any questions about tipsheets, please contact the Resource Center at 573-882-3364 or ***rescntr@ire.org.***

drug use

continued from page 1

Michigan, Nevada, New Mexico, New York, Oklahoma, Pennsylvania, Rhode Island, Texas, Utah, Washington and West Virginia.

Each state's program is slightly different from the others, but the overall intent is the same: to collect information about the prescription of controlled substances. Ritalin, Adderall and other ADHD drugs are Schedule II narcotics, so details about their use is collected under these programs. For details on the programs, see the DEA's Web site at *www.deadiversion.usdoj.gov/pubs/program/rx_monitor/faqs.htm.*

In Illinois, every time a pharmacist fills a prescription for Ritalin, the state gets a record of that transaction electronically. The record includes the name and address of the person who received the drug, his or her birth date, and amount of the drug prescribed, which doctor wrote the script and where it was filled.

The discovery that Illinois collected data on where people who get prescriptions filled live was a potential gold mine.

Data pinpoints drug use

The Drug Enforcement Administration encourages states to do prescription monitoring because it's helpful in tracking doctor shopping – instances where the same patient visits several physicians in hopes of collecting multiple prescriptions. It's also a good way to see if some doctors are prescribing unusually large amount of certain drugs.

For reporters it's an excellent way of seeing exactly where some drugs are being used.

Many stories have been done using DEA point-of-sale data, which shows how much Ritalin and other Schedule II narcotics are dispensed by pharmacies. The best use of this data was in 2001 when *The* (Cleveland) *Plain Dealer* used it to measure Ritalin usage for every county nationally.

But that data has severe limitations when you want to measure usage by ZIP code, as we did in Chicago. In many areas, people don't get their prescription filled in the same ZIP code where they live. In Chicago, for example, point-of-sale data shows an unusually high Ritalin-usage rate in the downtown Loop area. But that's only because thousands of people come to work, get their children's prescriptions filled at pharmacies in the parents' office buildings, and then they take it home to the suburbs in the evening.

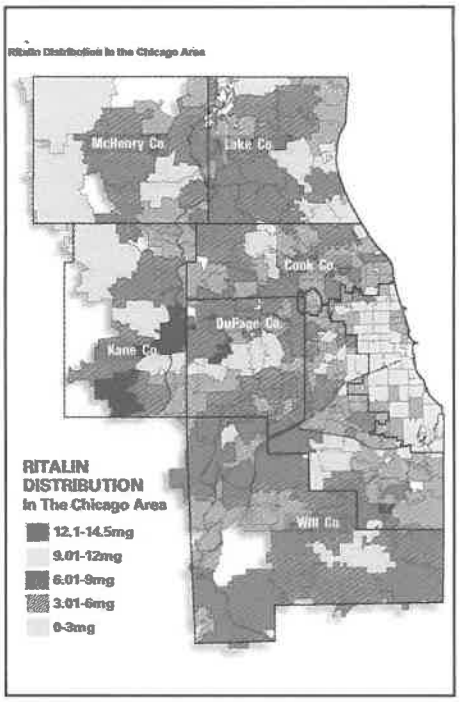
The discovery that Illinois collected data on where people who get prescriptions filled live was a potential gold mine. All we had to do was get our hands on the data.

Negotiations and threats

Illinois' Department of Human Services initially told the *Chicago Sun-Times* it would cost $6,000 to provide information on where people were being prescribed Ritalin. When I called back to ask why the charge was so outrageously high, I was told that someone was going to have to write a program to extract that information from the prescription-monitoring database. (It was at this point that we realized Illinois even had a prescription-monitoring program).

Fine, we told them. Just send us the entire database and we'll extract what we need. That, the state told us, was impossible.

We insisted they were mistaken, and after several more telephone calls, a few letters and promises of legal action, the state agreed with us – to a point. Based on our negotiations with them, the department agreed to give us portions of the data we were after.



Names of patients and doctors weren't provided, nor were pharmacies. The state also refused to provide data for all counties in Illinois, arguing that in some counties one person had received a particular drug – and we could theoretically track down that patient, thus violating his right to privacy.

In the end, we agreed to request data only from the ZIP codes in the six counties in northeast Illinois – the region where the vast majority of the state lives and where the *Sun-Times* sells most of its newspapers. A few weeks later, a CD packed with data arrived in my mailbox.

False steps and analysis

There were a few problems. The data didn't include dates of birth, only patient ages (another concession we agreed to because of the state's fears that patients could be identified). It looked like there were a lot of babies receiving unusually large doses of some drugs.

A call to the state explained the prob-

lem. Anytime a record lacked a DOB, the programmers inserted a zero. That made analyzing the amount of drugs given infants impossible.

At about this point the project was sidelined by Sept. 11 coverage. Ultimately, we asked the state to re-run the data, fix the age field (those missing DOBs were now labeled with 999) and update it with more numbers collected since our last request. Soon after, we had a CD with 641,735 records. Each line was 98 characters long.

The data imported into Paradox (version 9.0) without a problem. The record layout the state provided made parsing the lines of data simple. But that didn't make it easy to read.

The data had no drug names, only National Drug Code numbers. The number tells you exactly what the drug is and how strong a dose was prescribed. For example, 9042769 is a 10-milligram dose of methylphenidate hydrochloride (the generic form of Ritalin) in tablet form.

The state's data had 457 different NDC numbers. That doesn't mean there were 457 different drugs. In many cases, the same drug has multiple NDC numbers. For example, the NDC number 7811749 also indicates the patient received 10-milligram methylphenidate hydrochloride tablets.

To make sense of it all, we downloaded the NDC dictionary from the DEA (PC and Mac versions are at *www.deadiversion.usdoj.gov/arcos/ndc/ndc_dic.htm.*) By joining this data to the state's, we had a database that showed at a glance what drugs were prescribed and at what dosage.

Finally, we needed to know how many people lived in each ZIP code and how many fell into different age groups. We used Census ZIP Code Tabulation Area population. It's available from American Factfinder on the Census Bureau Web site *www.census.gov.*

With this new, enhanced database – it now included the state's data, drug names and population figures – we did some relatively simple calculations. Looking at Ritalin, for example, we added up all the milligrams of Ritalin prescribed to people under age 18 in each ZIP code, then divided by the number of people under age 18 living there.

This ratio helped us learn where the largest proportions of certain populations were receiving the drug.

Mapping the concentration

Using ArcView 3.2, we opened ZIP code shape files for Illinois, and linked our state data to that on the "ZIP Code" fields. Then we mapped the ratio.

Even someone with only a passing familiarity with the Chicago area could see the patterns: ZIP codes in the north and far west suburbs – traditionally some of the area's wealthiest – were the same places were large concentrations of the under age 18 population received these drugs.

Most of the city, as well as the south suburbs, had relatively low usage rates.

We ran a lot of those ratios – for specific drugs and also for all forms of methylphenidate. We also used it to examine usage of ADHD drugs such as Adderall and Dexedrine. Generally, the results were consistent – wealthier areas have larger concentrations of the drugs being used.

At this point a lot more traditional reporting began, with calls to doctors and parent groups, as well as visits to schools and meetings of support groups for parents with ADHD children.

And while many of them had anecdotal observations about where usage was high or low, we had concrete numbers that helped focus the story on ADHD in a way that hadn't been done locally before.

Mark Skertic can be reached by e-mail at mskertic@suntimes.com

readme.txt

Web address

These stories are on the *Sun-Times'* Web site at ***www.suntimes.com/special_sections/ritalin/***

Resources

The IRE Resource Center provides stories and tipsheets to help journalists. Search the story and tipsheet databases at ***www.ire.org/resourcecenter*** for order numbers. Call 573-882-3364 to order these and other investigative stories and tipsheets.

Story

The (Cleveland) *Plain-Dealer* compiled 1996 Census projections and data from the Drug Enforcement Administration showing the breakdown of Ritalin sold in grams in the nation's 3,141 counties. The results show that the use of and prescribing of Ritalin varies from state to state but also by region. The analysis found high Ritalin use in the Northeast and upper Midwest, and even disparities in use within Ohio. Attitudes and theories on Ritalin use vary with some believing children in cold, isolated areas might be more prone to use the drug, some think military families are more inclined to use it, others believe children and families with attention disorders "migrate" to resort areas where the drug is more readily available.

Tipsheets

Beth Marchak of *The Plain Dealer* offers guidelines for using Drug Enforcement Administration and other federal agency data to report on Ritalin consumption in this 2002 tipsheet. (Tipsheet #1572)

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

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

Uplink

Uplink Info

A newsletter of the National Institute for Computer-Assisted Reporting

Editor
Brant Houston
brant@ire.org

Managing Editor
Jeff Porter
jeff@nicar.org

Asst. Managing Editor
David Herzog
dherzog@nicar.org

Art Director
Lisa Triefenbach
lisa@ire.org

Editorial Phone
573-884-7711

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

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

Director of Publications
Len Bruzzese
len@ire.org

Advertising Coordinator
Pia Christensen
pia@ire.org

Subscription Administrator
John Green
jgreen@ire.org

Subscriptions
IRE members $40, nonmembers $60

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

Postmaster: Please send address changes to IRE-NICAR.