An Introduction to Computer-assisted reporting

What is computer-assisted reporting and why should I care?

Since the mid-1990s most major newspapers have been using computer-assisted reporting in their journalism. For more than a decade, at least one Pulitzer Prize each year has been based on data analysis. In other words, CAR is not a passing fad. It's here to stay.

Computer-assisted reporting can help most reporters. Many great stories buried in data that cannot be done otherwise.

Don't worry. If the thought of data overwhelms you – forget that it's data; think about it as a huge library of public records.

You probably have gotten a copy of a report after a serious automobile accident on a dangerous segment of freeway. What if, in addition to the reporting you've always done, you could look at every other accident that has happened in that same spot and put the story in perspective?

"But the highway department can give me those numbers," you say. They probably can but:

- There may be a delay for them to run your request.
- They may not be available when you need them

Whether you're covering an accident or another event or topic, data can help you tell a better story.

Databases give you more to the story. Rather than writing based on a few analogies, you'll know about the whole population or all of the accidents or all of the crime incidents.

Your best analogies are in the data. The examples that illustrate the point of your story are part of what you'll find as you sift through your data. Data analysis doesn't just mean finding overall numbers – it gives you a huge list of possible subjects for stories.

Contrasts are in the data. Because you have an entire "population" you have the highs and lows so you can compare the various ends of the spectrum.

You know what an "outlier" really looks like. When newspapers analyze school test scores to see how schools perform – the interesting stories are in the "outliers" – the schools that don't fit the pattern. You can only find these statistical trends by analyzing the data.

For more information:

- ✓ CAR in Canada: ww.carincanada.ca
- ✓ IRE Resources Center and ExtraExtra: www.ire.org
- ✓ The Centar za istraživačko novinarstvo: www.cin.ba
- ✓ FarmSubsidy.org

You have more powerful figures that weren't just fed to you by a government agency.

You can make connections you might not otherwise be able to make. One of the most powerful CAR tools is joining databases to find the common records. Some newspapers have joined voter data to death records to find dead voters. Others have matched teachers to criminal conviction data.

You have authority. By doing your own data analysis, you're producing the research, not just rehashing someone else's work.

You have at hand tools to do your work better and provide insight. Databases and spreadsheets can be great tools for staying organized or keeping track of your paper records. Other tools include:

- GIS: Geographic Information Systems allow you to map data to find geographic trends.
- Statistical tools: Some stories call for more complex analyses.
- Online databases: When you analyze data, you end up with a lot of information that won't necessarily fit into the newspaper. Putting data online can make that information available for readers.

How we analyzed truck safety

For a 2006 series on trucking safety, we obtained several different databases for our analysis, including federal and state accident data, federal and state truck inspection data, registration databases and databases tracking enforcement. As with many government databases, we had to do extensive cleaning and integrity checking on all these databases.

Having both federal and state data allowed us to make comparisons to verify the data, but we did other integrity checks as well. For example, by looking at changes over time, we found a significant drop in accidents from Houston. Further research showed that the state did not have all the accident records from the city of Houston, so we had to supplement that data with local records. We found duplications that the state did not know existed. Each pass we made at the data introduced new questions. We did dozens of different analyses for this project. Among them:

- ✓ Accidents: Our analysis was done using database software to crunch accidents by company, road, characteristics of causation and other factors.
- ✓ We used statistical software to look at the causation data to see if there were any relationships between accident characteristics.
- ✓ We examined inspections to see what sorts of inspections were being done and where they were being conducted.
- ✓ We matched complaint and compliance review data to see if companies with accidents were being investigated.

We used several different programs for this project, including Microsoft Access, Excel, ArcView, SPSS, SQL server and FoxPro for analysis. We used Ultraedit and Monarch for data cleaning. We also used mapping software to plot where fatal accidents happened.

-- Jennifer LaFleur

Texas gets the short shrift on national arts grants

Exploring arts grant data lead to a Page One Sunday story about how Texas doesn't get its fair share of federal arts grants from National Endowment for the Arts, or NEA.

A quick glance at annual grants lists going back 15 years showed that other big states such as New York fare much better.

As an exercise in database massage, we integrated U.S. Census data with the NEA grant lists to come up with per capita funding for each state. Texas ranked below Guam and American Samoa in per capita funding.

My only tip is to make sure you use the latest available data. Sometimes, a phone call to a governmental entity will reveal newer data than it has up on its web site.

--Scott Parks
The Dallas Morning News

CAR Resources: Books about computer-assisted reporting

The New Precision Journalism. By Philip Meyer. Indiana University Press, Bloomington, 1991

The Reporters Handbook: An Investigator's Guide. By Brant Houston, Len Bruzzese and Steve Weinberg. St. Martin's Press, New York, Fourth Edition, 1996

News and Numbers. By Victor Cohn. Iowa State University Press, Ames, 1989.

Numbers in the Newsroom: Using Math and Statistics in News. By Sarah Cohen. Investigative Reporters and Editors.

Computer-Assisted Reporting: A Practical Guide. By Brant Houston. St. Martin's Press, New York, Third Edition, 2003.

See links for more books….