

Paper 112-26

Pulitzer Programming: How Investigative Reporters are Using SAS™

Stephen K. Doig, Cronkite School of Journalism, ASU

Richard O'Reilly, The Los Angeles Times

ABSTRACT

A new breed of newspaper reporters who practice "precision journalism" are using SAS™ software to do major investigative stories based on analyses of large data sets. These stories have examined votes in the disputed 2000 presidential election, damage from Hurricane Andrew, homicide patterns, census data, school test scores, municipal election fraud and many other issues.

INTRODUCTION

For most of the history of journalism, reporters have had to rely on anecdotal evidence and the pronouncements of experts when writing stories about serious community problems. However, the seed of a quiet revolution was planted in 1972 when reporter Philip Meyer wrote "Precision Journalism." Meyer's novel premise was that journalists should learn to use social science methods such as survey research and statistical analysis to add authority to their reporting. His vision started to become practical in the early 1980s when a handful of reporters around the country began to use the first personal computers.

The authors of this presentation were among those pioneers, and among the first to go beyond simple spreadsheets and database programs to using SAS software for tackling analyses of large data sets. In the past decade, a growing number of reporters have started using SAS for a wide variety of important stories. And the demands of working with 2000 Census data are encouraging even more of them to add SAS to their "computer-assisted reporting" toolbox.

Here are just a few examples of the major stories that have been done using SAS:

HURRICANE ANDREW

In August, 1992, this major hurricane slammed ashore just south of Miami, directly killing 15 people and causing \$25 billion in damage, the most expensive natural disaster in the history of the country. More than 25,000 homes were destroyed by the storm and another 100,000 damaged. In the aftermath, *The Miami Herald* began to probe why the damage was so extensive despite the fact that South Florida claimed the toughest building code in the nation. Using SAS, author Doig (then research editor of *The Herald*) processed more than 50,000 damage reports to discover the smoking gun: The newer the home, the more likely it was to be destroyed, thus indicating the effects of how the code had been weakened over the years. SAS also was used to quantify the impact of campaign contributions from the building industry and to sift through millions of building inspection records. The resulting 16-page report, titled "What Went Wrong", was a key element of the Pulitzer Prize for Public Service won by *The Herald* in 1993.

LOS ANGELES HOMICIDES

A team of investigative reporters at *The Los Angeles Times* undertook to examine every homicide committed in Los Angeles County from 1990 through 1995 to search for a wide range of patterns. Author O'Reilly used SAS to painstakingly assemble a database of court files, merged with police, district attorney, coroner and vital statistics data. Among the findings, murders of whites were more likely to be reported by *The Times* and more likely to be solved by police than those of minorities. This 20-

month enterprise, published in 1996 as "And Justice for Some", won medals from Investigative Reporters & Editors, and the Society of Professional Journalists, and was a runner-up for a Pulitzer Prize.

MIAMI VOTE FRAUD

In November, 1997, the city's mayoral election was won thanks to a lopsided margin of absentee ballots in the favor of the winning candidate. *The Miami Herald* began writing a series of stories exposing allegations of vote buying and irregularities in how absentee ballots were requested and filled out. Database editor Dan Keating used SAS to merge more than a dozen datasets with fuzzy-match techniques to help reporters pinpoint numerous instances of illegal voting. As a result of the Herald's work, a judge threw out the election results, and the Herald was awarded a Pulitzer prize.

SOUTH FLORIDA JUSTICE SYSTEM

In 1993, The Miami Herald did an exhaustive eight-month examination of the criminal justice system, concluding that the region had more crime and less punishment than any large metro area in the country. SAS was used to transform court databases built around separate criminal charges into one that showed what happened to individual defendants, allowing a first-ever look at how and why so many felons were escaping meaningful punishment. The resulting series, "Crime and No Punishment," won the Investigative Reporters & Editors medal in 1994.

2000 PRESIDENTIAL ELECTION

The contested results of the dead-heat Florida presidential vote have prompted many news organizations to undertake investigations of the election, and SAS has been used extensively. For instance, Geoff Dougherty of *The Miami Herald* examined the voter rolls to find several hundred felons who voted illegally. In a study done for *The Herald*, author Doig used the results from more than Florida 5,000 precincts to project that Al Gore would have won by about 23,000 votes if 180,000 ballots hadn't been invalidated. And Dan Keating, now working for *The Washington Post*, used SAS to show that Florida's minority voters were the most likely to have cast ballots that were invalidated.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Prof. Stephen K. Doig
Cronkite School of Journalism,
Arizona State University
Tempe, AZ 85287-1305
480-965-0798
steve.doig@asu.edu

Richard O'Reilly
Director of Computer Analysis
The Los Angeles Times
202 W. 1st St.
Los Angeles, CA 90012
213-237-7001
richard.oreilly@latimes.com