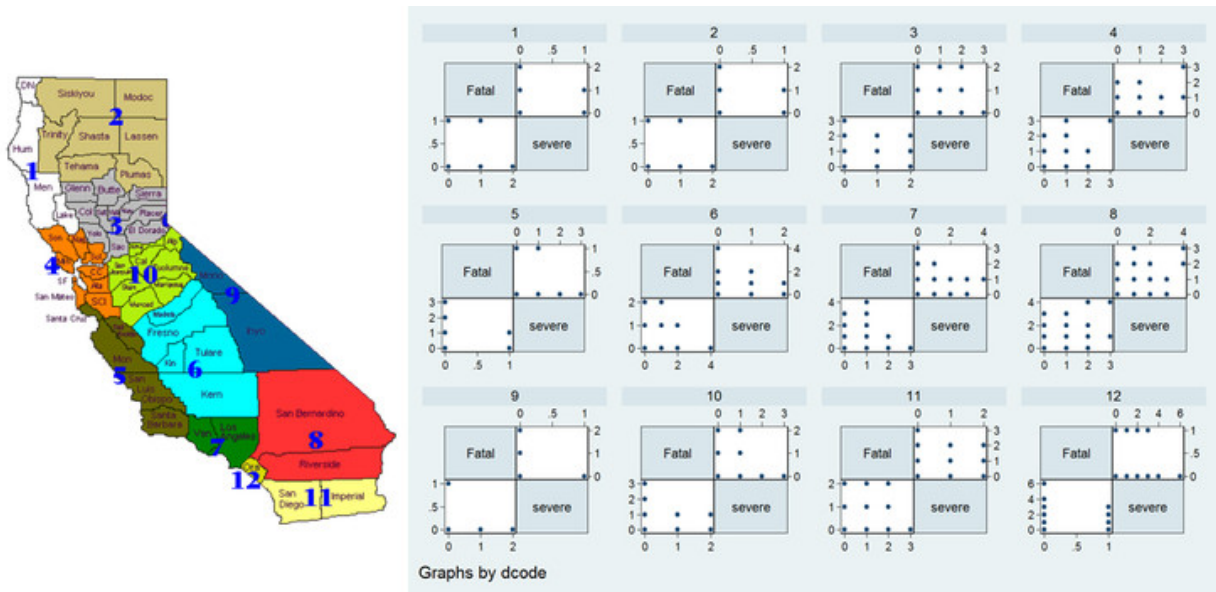# Engineers to study traffic data to determine causes of fatal accidents

April 19 2016, by Rebekka Coakley

This graphic details how fatal and severe accident patterns in California co-occur by district. California has 12 transportation districts, and the graphic presents a matrix of fatal and severe accidents on roadway segments for each district. In general, as the district number increases, the location moves southward within the state. The matrix shows geographical dissimilarities in the patterns of co-occurrence of fatals and severe injuries. Credit: Penn State

Traffic accidents are a leading cause of death in the U.S., especially for people under the age of 25. Computer scientist Kamesh Madduri and civil engineer Venky Shankar are developing a high performance

computing-based framework for the comparative evaluation of traffic safety on statewide transportation networks, which they hope will find patterns related to fatalities and severe injuries.

"Large accident datasets contain unique patterns of fatalities and severe injuries that cannot be unmasked with limited samples. This is because fatalities and severe injuries are relatively rare—less than 5 percent—of all accidents," said Shankar, a civil engineering professor at Penn State. "What interests me about this research is it is fundamental in terms of the computational methods being put to use to extract patterns relating to fatalities and severe injuries. It is a tool for discovery, and has the potential to provide unique insight into the contexts in which fatalities and severe injuries occur."

Madduri and Shankar will take traffic accident data collected from police reports from the state of California and create software, using algorithms that can read these big data sets.

"We really want to mine data for an entire state, and further do multi-state comparisons," said Madduri, an assistant professor of computer science and engineering in Penn State's School of Electrical Engineering and Computer Science. "As it is right now the dataset is too large. Statistical analysis of large accident datasets involves hundreds of variables, and the state-of-the-art in computational methods for estimating complex fatality models from large datasets is nonexistent. We'll develop software to do this."

The computational tools the two create will be able to abstract immense amounts of accident data, from which they hope to find common causes of fatal accidents.

"Big data is big inference in this case. This is what we are after—how to infer with high confidence the factors associated with fatal and severe

accidents. The second thing we want to do is to make this technique portable. We are going to work on accident data from the state of California, and then test to see if the tools we develop can retain the 'learning' and extract patterns in data from other states," said Shankar. "And we want to do this fast—we are talking orders of magnitude. Right now, it takes days to do this type of analysis for a dataset with just a few thousand observations on a reasonably powerful desktop. The bottom line is if we can create a tool that is fast, and provides quick-response analysis for national safety policy, we can move one step closer to reducing fatalities and severe injuries by one half."

This would be a massive step in national safety policy.

According to Madduri, their idea is preliminary and high-risk—no one has explored this area of data.

"We'll be working with large, multivariate datasets and there is a substantial gap between theory and practice in this area," said Madduri.

Provided by Pennsylvania State University

Citation: Engineers to study traffic data to determine causes of fatal accidents (2016, April 19) retrieved 25 April 2024 from https://phys.org/news/2016-04-traffic-fatal-accidents.html