the *Montreal Gazette* in May 1989

Double Trouble in Moose Jaw School

(caption to a photograph in the Montreal Gazette in May 1989 showing six sets of twins)

Every morning, teachers at Prince Arthur school in Moose Jaw, Saskatchewan see double—and it's not because of what they were up to the night before. Six pairs of identical twins attend the school, which has an enrollment of 375. Identical births occur once in 270 births.

I use this example in class to illustrate how one can visualize the Poisson distribution using the very useful cell occupancy approach. I ask students to think of randomly assigning the approximately 10,000 twins in 2,700,000 births in Canada in a space of five to six years to 7,200 schools of size 375 each, and to imagine how many schools will receive 0, 1, 2, . . . Students quickly agree that if there are to be some O's, then there must be some 2's and 3's, and a few even bigger clusters, in order to have an average of 10,000/7,200 = 1.39 twin pairs per school. (The Poisson distribution should also apply in the United States; the numbers of schools and twin pairs would be 10 times bigger, but the mean per school would remain the same.) I use a microcomputer to simulate this distribution in real time, so that students see the twin pairs "piling up" one by one in the various schools. Because I cannot represent all 7,200 schools on the computer screen, I scale down the problem by a factor of 20, and use a grid of 7200/20 = 360 boxes or cells to represent schools.

At any stage of the assignment process, the numbers in each of the cells represent the number of times that each cell has been "visited." For each of the N = 10,000/20 = 500 "visits" (each one representing a twin pair), the target cell is chosen randomly (with replacement) from the numbers 1 to 360, and that cell's occupancy is updated (incidentally, when the run is completed, we are left with a table of random numbers with a Poisson distribution). Finally, in order to stimulate discussion of the enormous difference between using *named* and *unnamed* towns, I ask students if the headline would be any less remarkable if it read "Double trouble in 'Anytown Canada' school" and what the implications of a "write the headline first, fill in the name after the fact" policy would be. This example is particularly helpful in teaching students how to think about random disease clustering in epidemiology (imagine "twin pairs" changed to "childhood cancers").

This general-purpose cell occupancy program can also be used, with selectable numbers of cells and numbers of visits, to simulate duplicate birthdays, and to show such phenomena as the "multiple events" discussed in section 7.1.3 of Diaconis and Mosteller. It is programmed in a still-rather-crude Excel macro available from the author.

{Births Case 3 from "Jumping to coincidences article American Statistician. Aug 1992. Vol. 46. 197-202