Figure 4.4. EXPLORATORY DATA ANALYSIS: Highway Deaths and Type of Car

EM8912: Kitchener-Waterloo Record, November 25, 1989, page F22

Type of car a factor in highway deaths, U.S. study reports

By Christine Russell
Special to the Record

The chance of dying in a traffic accident depends greatly on the type of car involved, with a more than eight-fold difference in death rates between the worst-ranked and best-ranked cars, a new study shows.

"This is the first time that fatality rates have been published" for individual types of cars, said Brian O'Neill, president of the Insurance Institute for Highway Safety, which evaluated 103 of the most popular cars. "I think this is the best measure to date of the safety of cars in serious crashes," he said.

About two-thirds of the differences in death rates can be accounted for by the size of the car and the age and sex of the driver, O'Neill said. Although the role of alcohol was not directly measured, it was accounted for in part because a large proportion of alcohol-related fatalities involve male or younger drivers.

A mid-sized specialty four-door, the Volvo 740-760, had the lowest death rate (0.6 per 10,000 registered cars), while a small sports car, the Chevrolet Corvette, had the highest (5.2). Of the cars with the 10 lowest death rates, only one – the Volkswagen Jetta four-door – was a small car.

By contrast, 12 of the 15 cars with the highest death rates were small, the study found, and the other three were mid-sized.

"This reinforces yet again the relative safety of large cars compared to small cars. .... But what we're also able to determine from this study is that even within the size classes, some cars are doing much worse or better than others," said institute president O'Neill, a statistician.

He noted that, for instance, among small four-door cars, the Chevrolet Chevette has a death rate nearly four times as high as the Volkswagen Jetta. And the safety records of a number of mid-sized cars "means that even though size is important, you don't have to be in a super-large car to be in a safe car, if the manufacturer builds the car appropriately."

The study also highlights the "very high death rates of performance and sports cars," said O'Neill. In addition, single-vehicle crashes are more likely to be fatal in small cars than in large cars, the study found.

The findings by the institute, a non-profit, scientific and educational group supported by the insurance safety groups and individual companies, are being published in a special Nov. 25 issue of its publication, *Status Report.*

The institute study computed death rates based on occupant deaths and car registrations in the three-year period from 1986-88, when more than 75,000 occupants of passenger cars died in crashes. This includes 1985-87 car models grouped by vehicle series of similar construction and wheelbase. Results were presented only for cars with at least 150,000 registrations during the period. Some models, such as the Chevrolet Chevette and Spectrum, have since been discontinued or modified.

The institute used data from the federal National Highway Traffic Safety Administration's Fatal Accident Reporting System and registration counts compiled by a commercial firm, R.L. Polk. The average death rate among the 103 models was two per 10,000 registrations.

In general, drivers under age 30 and male drivers are at greater risk of a fatal accident, regardless of the type of vehicle. They are also more likely to drive certain types of cars.

To compare how various types of cars were performing, the institute developed a model to calculate predicted death rates for each type of car, taking into account the expected proportion of young and male drivers as well as the car size, as reflected by its wheelbase. "The difference between a car's expected and actual death rate tells us how much better or worse that car is performing than we would expect, given its size and driver mix," O'Neill said.

Among the 103 vehicle series, 58 performed better than predicted during the study period and 39 performed worse. In six cars, predicted and actual rates were the same.

O'Neill noted that there was no single answer to account for the various differences found in the 103 cars, but in some cases, like the Volvo, improved design or structural differences might give a better safety record. In the case of sports cars, the high speed at the time of a crash might make fatalities more likely than expected, suggesting the need for higher-speed crash-protection features.

At the time of the study, O'Neill noted, the Toyota Cressida, which had one of the lowest death rates, was the only car included that required automatic restraints as a standard feature. For all new 1990 model cars, the government requires either air bags on the driver's side or automatic seat belt restraints for both front-seat passengers.

"This will improve the safety of virtually all cars, but some will still be better than others," O'Neill said. In the short term, he added, it is likely that the differences between certain cars will be accentuated because air bags are initially being offered in more of the larger cars.

The institute's study has not yet been reviewed by outside groups. Michael Finkelstein, associate administrator for research and development at NHTSA, noted that his agency has attempted to compare death rates for different cars but never published the results.

"We were never satisfied that we knew what the differences were the result of," he said. "My general caution is that fatalities are the result of a lot of factors, ranging from who the drivers are to differences in how the vehicles are used. Somewhere in the equation is how well the cars are engineered. We'll do a careful examination to see what we can learn from (the institute study)."

1 On the basis of what is said in the article EM8912 reprinted above, what do you infer was the *principal* question(s) to be answered by the Institute's investigation?

2 What *data sources* were used in the investigation?

- Were the data from these sources continuous or discrete? Explain briefly in each case.
- What control did the Institute have over the *accuracy* of these data?
  - How might this control, or lack of it, affect the answers from the investigation? Explain briefly

(*continued overleaf*)

- What arithmetic operation was used by the Institute to *combine* the data from the different sources?
  - What *name* is given to the resulting combination in the article EM8912?
  - Should this combination be modelled by a continuous or a discrete variable? Explain briefly.

3 The investigation made use of the *number of vehicle registrations* of certain car models over a 3-year period, whereas one of the investigations described in Figure 4.3 on page 4.11 used the *number of baseball games won* by major-league teams over an 18-year period. Discuss briefly which of the two data sets is likely to provide answers with *less* severe limitations.

4 Explain briefly what you understand by the statement in the *third* paragraph of the article EM8912 reprinted overleaf on page 4.13: *About two-thirds of the differences in death rates can be accounted for by the size of the car and the age and sex of the driver.*

5 Explain briefly what is meant by the *second* sentence of the third paragraph: *Although the role of alcohol was not directly measured, it was accounted for in part because a large proportion of alcohol-related fatalities involve male or younger drivers.*

6 What index of *car size* was used in the investigation? Identify the place(s) in the article EM8912 where this information is given.
- Suggest why this measure was used; what are its advantage(s) and disadvantage(s)?
- Briefly discuss one *other* measure of car size that might reasonably have been used in the investigation.

7 In the fourth paragraph of the middle column, the article EM8912 reprinted overleaf on page 4.13 mentions that *Results were presented only for cars with at least 150,000 registrations during the period.* Suggest reason(s) why this restriction was imposed.

8 How are the limitations of the investigation's answers made *more* severe by the matter mentioned at the end of the fourth paragraph of the middle column of the article EM8912: *Some models ..... have since been discontinued or modified.*
- What *general* matter does this illustrate about answers from data-based investigating?

9 In the third-last paragraph of the article EM8912 reprinted overleaf on page 4.13, Brian O'Neill of the Insurance Institute for Highway Safety mentions the improved safety likely to result from the installation of *air bags*, initially on larger cars. Suggest possible reason(s) why air bags (or automatic seat belt restraints for both front-seat passengers) might *not* increase the safety of car occupants in accidents.

10 In the last paragraph of the article EM8912 reprinted overleaf on page 4.13, Michael Finkelstein of the NHTSA is quoted as having said of another set of data on the same topic as the Institute's: *We were never satisfied that we knew what the differences were the result of.* On the basis of what is said in the article:
- outline an essential ambiguity involved in the analysis of *either* data set;
  - identify the best single-paragraph summary of this matter in the article;
- outline how Brian O'Neill might defend the assignment of reasons for the differences among cars in the Institute's *Report.*

The article EM8912 reprinted overleaf on page 4.13 is also used in Statistical Highlight #99.

1995-04-20