American New Cars and Trucks (2004)

Background

Specifications are given for 428 new vehicles in America for the 2004 year.

Data

The variables recorded include price, measurements relating to the size of the vehicle, and fuel efficiency. The dataset contains some missing data (shown by an asterisk *). Updates for this may be found in the database on the site http://www.autofiles.org/ and may allow filling in of some of the missing values. (note: On the http://www.autofiles.org/ database fuel efficiency for Open Road is labeled 'Road' and City is 'Mixt')

source: Kiplinger's Personal Finance, December 2003, vol. 57, no. 12, pp. 104-123

Variables

Vehicle Name = maker and model

Car Type = (1=regular, 2=Sports Car, 3=SUV, 4=Wagon, 5=Minivan, 6=Pickup)

Car Train Type = (1=Front Wheel Drive, 2=Rear Wheel Drive, 3=All Wheel Drive)

Retail Price = what the manufacturer thinks the vehicle is worth, including adequate profit for the automaker and the dealer in U.S. Dollars

Dealer Cost = what the dealership pays the manufacturer in U.S. Dollars

Engine Size = size in litres

Number of Cylinders = 4 6 or 8 (=-1 if rotary engine)

Horsepower = power of car measured in hp

City = fuel efficiency in litres per 100km in cities and on motorways

Open Road = fuel efficiency in litres per 100km on country/open roads

Weight = weight of car in kg

Wheel Base = length of wheel base in cm

Length = Length of car in cm

Width = width of car in cm

* indicates missing data

Questions

Explore the relationships between the variables (for example price vs fuel efficiency, weight vs fuel efficiency)

Can we determine the price of a car by a particular vehicle attribute?

Which gives the best indication of price?

Which gives the best indication of fuel efficiency?

How do 4 6 and 8 cylinder cars and their relationship to fuel efficiency compare?

Can you tell which cars are hybrid/electric by looking for any unusual data points?