

## **Data Description**

This data set consists of three types of entities:

- The specification of an auto in terms of various characteristics.
- Its assigned insurance risk rating: symboling (-3, -2, -1, 0, 1, 2, 3).
- Its normalized losses in use as compared to other cars: normalized-losses from 65 to 256.

The second rating corresponds to the degree to which the auto is more risky than its price indicates. Cars are initially assigned a risk factor symbol associated with its price. Then, if it is more risky (or less), this symbol is adjusted by moving it up or down the scale. Actuarians call this process "symboling". A value of +3 indicates that the auto is risky, -3 that it is probably pretty safe.

The third factor is the relative average loss payment per insured vehicle year. This value is normalized for all autos within a particular size classification (two-door small, station wagons, sports/speciality, etc...), and represents the average loss per car per year.

## **Class Attributes**

- make: alfa-romero, audi, bmw, chevrolet, dodge, honda, isuzu, jaguar, mazda, mercedes-benz, mercury, mitsubishi, nissan, peugot, plymouth, porsche, renault, saab, subaru, toyota, volkswagen, volvo
- fuel-type: diesel, gas.
- aspiration: std, turbo.
- num-of-doors: four, two.
- body-style: hardtop, wagon, sedan, hatchback, convertible.
- drive-wheels: 4wd, fwd, rwd.
- engine-location: front, rear.
- engine-type: dohc, dohcvt, l, ohc, ohcf, ohcv, rotor.
- num-of-cylinders: eight, five, four, six, three, twelve, two.
- fuel-system: 1bbl, 2bbl, 4bbl, idi, mfi, mpfi, spdi, spfi.