## About Dataset

### Context

This dataset consist of data From 1985 Ward's Automotive Yearbook. Here are the sources

Sources:

1) 1985 Model Import Car and Truck Specifications, 1985 Ward's Automotive Yearbook.  
2) Personal Auto Manuals, Insurance Services Office, 160 Water Street, New York, NY 10038  
3) Insurance Collision Report, Insurance Institute for Highway Safety, Watergate 600, Washington, DC 20037

### Content

This data set consists of three types of entities: (a) the specification of an auto in terms of various characteristics, (b) its assigned insurance risk rating, (c) its normalized losses in use as compared to other cars. 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.

Note: Several of the attributes in the database could be used as a "class" attribute.

### Inspiration

Please bring it on whatever inferences you can get it.

Questions I had:

Question: Does the data suggest that cars that are more expensive often show a lower depreciation over time?

Answer: It would seem that the depreciation for higher priced cars was lower than those cars that were not high end

Question: What do you feel was missed during the analysis?

Answer: This data set is very small compared to other datasets which makes it difficult to see if there were significant changes in the data compared to larger data sets.

Question: Were there any variables you felt could have helped in the analysis?

Answer: More safety statistics could have made it much more interesting to use than just setting the safety levels by a numbering scheme.

Question: Were there any assumptions made you felt were incorrect?

Answer: I came into this with no expectations, so I made no assumptions.

Question: What challenges did you face, what did you not fully understand?

Answer: I faced an issue with CDF setup and also hypothesis estimations.