

A discussion of R's mtcars dataset variables

Soren Heitmann

Sunday, February 22, 2015

The following table is a discussion of variables in the R mtcars dataset. It was developed in response to the Coursera Regression Models class in the Data Science Specialization taught by Prof. Brian Caffo for the course project assignment: "You work for Motor Trend, a magazine about the automobile industry. Looking at a data set of a collection of cars, they are interested in exploring the relationship between a set of variables and miles per gallon (MPG) (outcome). They are particularly interested in the following two questions: 'Is an automatic or manual transmission better for MPG' and 'Quantify the MPG difference between automatic and manual transmissions'"

Understanding the data is key to interpreting possible models and variable interactions. This table was developed as a resource to assist thinking about how to develop a model through greater understanding of the data.

Please note, I am not a car expert by any means. Discussions below were compiled by doing online research while completing the course assignment. Please feel free to provide comments, corrections, considerations (comment box is available below in Rpubs navigation bar).

Table: Discussion of mtcars variables

#	Var	Description	Comments
0	name	Model of Vehicle	The car types are a mix that includes sedans (Datsun, Ford, Honda,...), luxury sedans (Mercedes, Cadillac,...), muscle cars (Javelin, Challenger, Camaro...) and high-end sports cars (Porsche, Lotus, Maserati, Ferrari...)
1	mpg	Miles/US Gallon	mpg is the determinant of fuel efficiency
2	cyl	Number of cylinders	Data includes vehicles with 4,6,8 cylinder engines.
3	disp	Displacement (cu.in.)	Displacement measures overall volume in the engine as a factor of cylinder circumference, depth and total number of cylinders. This metric gives a good proxy for the total amount of power the engine can generate.
4	hp	Gross horsepower	Gross horsepower measures the theoretical output of an engine's power output; notably, <i>gross</i> rating is of the engine in an isolated environment outside any specific vehicle. When installed in a car, exhaust systems, carburetor, alternator, power systems, etc all influence the power that actually gets to the drive train. Moreover, according to online sources, in the early 1970s, regulatory changes influenced how gross horsepower was measured. As this dataset is from the early-mid 1970s, it's unclear if hp metrics may be used as reliable comparators of engine power across models as it's uncertain how manufacturers are reporting.

#	Var	Description	Comments
5	drat	Rear axle ratio	The rear axle gear ratio indicates the number of turns of the drive shaft for every one rotation of the wheel axle. A vehicle with a high ratio would provide more torque and thus more towing capability, for example. Transmission configuration can often influence a manufacturer's gearing ratio.
6	wt	Weight (lb/1000)	The overall weight of the vehicle per 1000lbs (half US ton)
7	qsec	1/4 mile time	A performance measure, primarily of acceleration. Fastest time to travel 1/4 mile from standstill (in seconds).
8	vs	V/S	Binary variable signaling the engine cylinder configuration a V-shape (vs=0) or Straight Line (vs=1). V==0 and S==1. Configuration offers trade offs in power/torque, design usage in terms of space/size of engine and performance or center of gravity of vehicle. The geometry and placement of the engine, as influenced by its cylinder head, can have numerous knock-on influences on the vehicle beyond the technical engineering considerations of the cylinder angle.
9	am	Transmission Type	A binary variable signaling whether vehicle has automatic (am=0) or manual (am=1) transmission configuration.
10	gear	Number of forward gears	Number of gears in the transmission. Manual transmissions have either 4 or 5 forward gears; Automatic either 3 or 4
11	carb	Number of carburetors	The number of carburetor barrels. Engines with higher displacement typically have higher barrel configuration to accommodate the increased airflow rate of the larger engine; in other words, more capacity is available for an engine when it may need it versus constraining power output with limited barrels. A vehicle may have multiple physical carburetors, but it's less common; this metric is the sum of the number of carburetors and the number of barrels inside the carburetor.