

DATA DICTIONARY: HondaComplaints.xlsx

These data are from the consumer complaints submitted to the NHTSA. The target is *crash*, which is a binary outcome indicating whether the vehicle was involved in an accident.

| ATTRIBUTE | Type | DESCRIPTION |
|-------------|----------|--|
| Make | Binary | HONDA or ACURA |
| Model | Nominal | TL, ODYSSEY, CR-V, CL, CIVIC, or ACCORD |
| NhtsaID | Nominal | Case ID, ignore |
| Year | Nominal | 2001, 2002 or 2003 |
| State | Nominal | 57 entities, some errors, ignore |
| abs | Binary | Y or N (yes or no, antilock brake system) |
| cruise | Binary | Y or N (yes or no, cruise control) |
| crash | Binary | Y or N (yes or no, accident yes or no) |
| mph | Interval | 0 to 100 |
| mileage | Interval | 0 to 200,000 |
| Description | Text | Consumer Report |

Notes:

Ignore attribute "State" and "NhtsaID"

There are no missing values

The attribute "mph" has outliers, a few entries above 100 mph.

The attribute "mileage" has outliers, a few entries above 200,000