**Car Insurance**

|  |  |  |  |
| --- | --- | --- | --- |
| Sr No | VARIABLE NAME | DEFINITION | THEORETICAL EFFECT |
| 1 | INDEX | Identification Variable (do not use) | None |
| 2 | TARGET FLAG | Was Car in a crash? 1 = YES 0 = NO | None |
| 3 | TARGET AMT | If car was in a crash, what was the cost | None |
| 4 | AGE | Age of Driver | Very young people tend to be risky. Maybe very old people also. |
| 5 | BLUEBOOK | Value of Vehicle | Unknown effect on probability of collision, but probably effect the pay-out if there is a crash |
| 6 | CAR AGE | Vehicle Age | Unknown effect on probability of collision, but probably effect the pay-out if there is a crash |
| 7 | CAR TYPE | Type of Car | Unknown effect on probability of collision, but probably effect the pay-out if there is a crash |
| 8 | CAR USE | Vehicle Use | Commercial vehicles are driven more, so might increase probability of collision |
| 9 | CLM FREQ | #Claims (Past 5 Years) | The more claims you filed in the past, the more you are likely to file in the future |
| 10 | EDUCATION | Max Education Level | Unknown effect, but in theory more educated people tend to drive more safely |
| 11 | HOME KIDS | #Children at Home | Unknown effect |
| 12 | HOME VAL | Home Value | In theory, homeowners tend to drive more responsibly |
| 13 | INCOME | Income | In theory, rich people tend to get into fewer crashes |
| 14 | JOB | Job Category | In theory, white collar jobs tend to be safer |
| 15 | KIDSDRIV | #Driving Children | When teenagers drive your car, you are more likely to get into crashes |
| 16 | MSTATUS | Marital Status | In theory, married people drive more safely |
| 17 | MVR\_PTS | Motor Vehicle Record Points | If you get lots of traffic tickets, you tend to get into more crashes |
| 18 | OLDCLAIM | Total Claims (Past 5 Years) | If your total pay-out over the past five years was high, this suggests future pay-outs will be high |
| 19 | PARENT1 | Single Parent | Unknown effect |
| 20 | RED CAR | A Red Car | Urban legend says that red cars (especially red sports cars) are riskier. Is that true? |
| 21 | REVOKED | License Revoked (Past 7 Years) | If your license was revoked in the past 7 years, you probably are a riskier driver. |
| 22 | SEX | Gender | Urban legend says that women have less crashes than men. Is that true? |
| 23 | TIF | Time in Force | People who have been customers for a long time are usually safer. |
| 24 | TRAVTIME | Distance to Work | Long drives to work usually suggest greater risk |
| 25 | URBANICITY | Home / Work Area | Unknown |
| 26 | YOJ | Years on Job | People who stay at a job for a long time are usually safer. |

* **Data Info**
* The training dataset contains 8161 observations of 26 variables (one index, two responses, and 23 predictor variables).
* Each record (row) represents a set of attributes of an insurance company individual customer that are related to their socio-demographic profile and the insured vehicle.
* The binary response variable TARGET\_FLAG has 1 if the customer’s car was in a crash, and 0 if not.
* The continuous response variable TARGET\_AMT defines the cost related to the car crash if it happened.
* Important Links:

Source:

<https://www.kaggle.com/datasets/xiaomengsun/car-insurance-claim-data?resource=download&sort=recent-comments>