ABC INSURANCE COMPANY

As of 2016, there were over 268 million [registered vehicles](https://www.statista.com/statistics/183505/number-of-vehicles-in-the-united-states-since-1990/) on the roads in the United States. In 2015, there were 32,166 [fatalities](https://www.statista.com/statistics/191499/vehicle-crashes-by-severity-in-the-united-states-since-1991/), 1,715,000 injuries and 4,548,000 car crashes which involved property damage. So, while many of us feel secure in our vehicles, the statistics indicate the importance of automobile insurance and in most cases, auto insurance is required by law. Auto insurance is important because it not only covers any physical damage that may occur in an accident, but also any damage or injury that might be caused because of a vehicular accident or which may be done upon oneself or one’s vehicle by another vehicle or accident – a falling tree for example(Car insurance in the U.S. - Statistics & Facts)

The Insurance industry is immensely data intensive. Historically, their data has been largely fragmented and underutilized. Also, insurance industry goes through a lot of combining structured and unstructured data, which enables the industry to generate powerful insights. With an incredible amount of data flowing in from multiple new digital channels, the insurance industry is undergoing a paradigm shift in the way they function – right from product planning to pricing, introduction, marketing, customer self-service and claim processing.

The objective of this project is to build a “Risk Analytics model” to understand the renewal potential and claim propensity of Existing Customers under Personal Auto Insurance Lines.

**The question to answer is:**

Will the policy holder initiate a claim (Yes/No) for this policy in the next policy year?

Data dictionary:

|  |  |
| --- | --- |
| Sr No | Serial Number |
| ClaimStatus | Indicates whether the policy holder has made a claim or not. 1 indicates a claim and 0 indicates no claim |
| ClaimFrequency | Gives the number of claims claimed |
| Premium | Premium in $1000 |
| Billing\_Term | How often the premium is paid, i.e. once a year = 1, Three times in a year = 3, or 6 times in a year = 6 |
| Renewed | Indicates whether the policy has been renewed or not. 1= Renewed, 0 = not renewed |
| DOB1 | Date of Birth of the main policy holder/ main driver |
| DOB2 | Date of Birth of the second driver |
| DOB3 | Date of Birth of the third driver |
| DOB4 | Date of Birth of the fourth driver |
| DOB5 | Date of Birth of the fifth driver |
| Number\_of\_Driver | Count of the number of drivers in the policy |
| AgeUSdriving\_1 | How long the driver has been driving |
| AgeUSdriving\_2 | How long the driver has been driving |
| AgeUSdriving\_3 | How long the driver has been driving |
| AgeUSdriving\_4 | How long the driver has been driving |
| AgeUSdriving\_5 | How long the driver has been driving |
| Amendment | Number of changes made to the policy during the year (may it is: No. of changes made to the policy till date from the date of buying) |
| CoverageLiability | The three numbers represent (in the $ thousands) the liability limits for per-person bodily injury, bodily injury for all persons injured in any one accident, and property damage liability. Coverage liability represents the state's financial responsibility law i.e. the minimum requirement. the first figure is the amount to be paid for injuries per head, the second figure is the amount for all injuries and the third is the amount per vehicle damaged. For example, say you live in Ohio and hold the minimum amount of coverage, which is 25/50/25 (raised in December 2013). This means that the minimum liability limits in this state are $25,000 for injuries to one person, $50,000 for all injuries incurred and $25,000 for property damage for one vehicle in an accident. |
| CoverageMP | MP stands for Medical payments coverage pays the reasonable expenses an insured person incurs for medical and funeral services within three years of an accident. |
| CoveragePD\_1 | PD is an abbreviation used in the car insurance industry. PD stands for Property Damage. It is a type of liability coverage and is part of the foundation of all state required auto insurance laws. Property damage insurance covers any damages to someone's property. This could mean a mailbox, someone's car, or even personal property in the person's car that was damaged as a result of the collision. PD is stated as two parts here, where the first part is the maximum payment that will be made for a single property in a single accident and the second number is the maximum that will be covered for all properties damaged in the accident. For example, the policy might show that you have Property Damage coverage of $25,000 per property, with a maximum of $50,000 per accident. |
| CoveragePIP\_CDW | Personal Injury Protection (PIP)-This a package of first-party medical benefits that provides broad protection for medical costs, lost wages, loss of essential services normally provided by the injured person (i.e. childcare, housekeeping), and funeral costs. It is usually associated with a no-fault auto insurance system.   Personal injury protection coverage, in addition to medical and funeral expenses, pays 80 percent of an insured person’s loss of income. If the injured person is not employed, the coverage pays reasonable expenses for obtaining family or household services the insured person normally would have performed. |
| CoverageUMBI | Uninsured/Underinsured motorist bodily injury |
| CoverageUMPD | Uninsured/Underinsured motorist property damage coverage |
| DistanceToWork\_1 | Distance to work for the first driver |
| DistanceToWork\_2 | Distance to work for the second driver |
| DistanceToWork\_3 | Distance to work for the third driver |
| DistanceToWork\_4 | Distance to work for the fourth driver |
| DistanceToWork\_5 | Distance to work for the fifth driver |
| DriverAssigned\_1 | Count of drivers assigned to the first vehicle. 1 to max 5 |
| Engine\_1 | Engine specification size in litres for the first vehicle |
| ExcludedDriverName\_01 | First person declared as an excluded driver |
| ExcludedDriverName\_02 | Second person declared as an excluded driver |
| ExcludedDriverName\_03 | Third person declared as an excluded driver |
| ExcludedDriverName\_04 | Fourth person declared as an excluded driver |
| ExcludedDriverName\_05 | Fifth person declared as an excluded driver |
| ExcludedDriverName\_06 | Sixth person declared as an excluded driver |
| ExcludedDriverName\_07 | Seventh person declared as an excluded driver |
| ExcludedDriverName\_08 | Eighth person declared as an excluded driver |
| ExcludedDriverName\_09 | Ninth person declared as an excluded driver |
| ExcludedDriverName\_10 | Tenth person declared as an excluded driver |
| ExcludedDriverName\_11 | Eleventh person declared as an excluded driver |
| ExcludedDriverName\_12 | Twelfth person declared as an excluded driver |
| ExcludedDriverName\_13 | Thirteenth person declared as an excluded driver |
| ExcludedDriverName\_14 | Fourteenth person declared as an excluded driver |
| ExcludedDriverName\_15 | Fifteenth person declared as an excluded driver |
| ExcludedDriverName\_16 | Sixteenth person declared as an excluded driver |
| ExcludedDriverName\_17 | Seventeenth person declared as an excluded driver |
| ExcludedDriverName\_18 | Eighteenth person declared as an excluded driver |
| ExcludedDriverName\_19 | Nineteenth person declared as an excluded driver |
| ExcludedDriverName\_20 | Twentieth person declared as an excluded driver |
| GaragedZIP\_1 | Zip code of the place where the first vehicle is parked. |
| MaritalStatus\_1 | Marital Status of the first driver. M - Married or S - Single |
| MaritalStatus\_2 | Marital Status of the second driver. M - Married or S - Single |
| MaritalStatus\_3 | Marital Status of the third driver. M - Married or S - Single |
| MaritalStatus\_4 | Marital Status of the fourth driver. M - Married or S - Single |
| MaritalStatus\_5 | Marital Status of the fifth driver. M - Married or S - Single |
| Occupation\_1 | Occupation of the first driver |
| Occupation\_2 | Occupation of the second driver |
| Occupation\_3 | Occupation of the third driver |
| Occupation\_4 | Occupation of the fourth driver |
| Occupation\_5 | Occupation of the fifth driver |
| Relation\_1 | Relationship of the first driver with the main policy holder. Only Self |
| Relation\_2 | Relationship of the second driver with the main policy holder |
| Relation\_3 | Relationship of the third driver with the main policy holder |
| Relation\_4 | Relationship of the fourth driver with the main policy holder |
| Relation\_5 | Relationship of the fifth driver with the main policy holder |
| Rental\_1 | first vehicle (If rental is allowed) |
| Sex\_1 | Gender of the first driver M - Male, F - Female |
| Sex\_2 | Gender of the second driver M - Male, F - Female |
| Sex\_3 | Gender of the third driver M - Male, F - Female |
| Sex\_4 | Gender of the fourth driver M - Male, F - Female |
| Sex\_5 | Gender of the fifth driver M - Male, F - Female |
| Surcharge1Unit\_1 | First surcharge for the first vehicle. Y - Yes, N- No |
| Surcharge2Unit\_1 | Second surcharge for the first vehicle Y - Yes, N- No |
| Surcharge3Unit\_1 | Third surcharge for the first vehicle Y - Yes, N- No |
| Towing\_1 | first vehicle Towing and labor cost coverage is an optional coverage that you can add to your car insurance that typically protects you against some of the costs and hassles associated with common roadside breakdowns like dead batteries, flat tires or even an embarrassing lockout. (Some insurers may automatically fold this coverage into their policies, so be sure to ask.) |
| Units | Number of vehicles covered in the policy |
| VehicleInspected\_1 | first vehicle inspected. 1 - Vehicle was inspected, 0 - Vehicle was not inspected |
| ViolPoints1Driver\_1 | First time the first driver is scoring a violation point. |
| ViolPoints1Driver\_2 | First time the second driver is scoring a violation point. |
| ViolPoints1Driver\_3 | First time the third driver is scoring a violation point. |
| ViolPoints1Driver\_4 | First time the fourth driver is scoring a violation point. |
| ViolPoints1Driver\_5 | First time the fifth driver is scoring a violation point. |
| ViolPoints2Driver\_1 | Second time the first driver is scoring a violation point. |
| ViolPoints2Driver\_2 | Second time the second driver is scoring a violation point. |
| ViolPoints2Driver\_3 | Second time the third driver is scoring a violation point. |
| ViolPoints2Driver\_4 | Second time the fourth driver is scoring a violation point. |
| ViolPoints2Driver\_5 | Second time the fifth driver is scoring a violation point. |
| ViolPoints3Driver\_1 | Third time the first driver is scoring a violation point. |
| ViolPoints3Driver\_2 | Third time the second driver is scoring a violation point. |
| ViolPoints3Driver\_3 | Third time the third driver is scoring a violation point. |
| ViolPoints3Driver\_4 | Third time the fourth driver is scoring a violation point. |
| ViolPoints3Driver\_5 | Third time the fifth driver is scoring a violation point. |
| ViolPoints4Driver\_1 | Fourth time the first driver is scoring a violation point. |
| ViolPoints4Driver\_2 | Fourth time the second driver is scoring a violation point. |
| ViolPoints4Driver\_3 | Fourth time the third driver is scoring a violation point. |
| ViolPoints4Driver\_4 | Fourth time the fourth driver is scoring a violation point. |
| ViolPoints4Driver\_5 | Fourth time the fifth driver is scoring a violation point. |
| ViolPoints5Driver\_1 | Fifth time the first driver is scoring a violation point. |
| ViolPoints5Driver\_2 | Fifth time the second driver is scoring a violation point. |
| ViolPoints5Driver\_3 | Fifth time the third driver is scoring a violation point. |
| ViolPoints5Driver\_4 | Fifth time the fourth driver is scoring a violation point. |
| ViolPoints5Driver\_5 | Fifth time the fifth driver is scoring a violation point. |
| ViolPoints6Driver\_1 | Sixth time the first driver is scoring a violation point. |
| ViolPoints6Driver\_2 | Sixth time the second driver is scoring a violation point. |
| ViolPoints6Driver\_3 | Sixth time the third driver is scoring a violation point. |
| ViolPoints6Driver\_4 | Sixth time the fourth driver is scoring a violation point. |
| ViolPoints6Driver\_5 | Sixth time the fifth driver is scoring a violation point. |
| ViolPoints7Driver\_1 | Seventh time the first driver is scoring a violation point. |
| ViolPoints7Driver\_2 | Seventh time the second driver is scoring a violation point. |
| ViolPoints7Driver\_3 | Seventh time the third driver is scoring a violation point. |
| ViolPoints7Driver\_4 | Seventh time the fourth driver is scoring a violation point. |
| ViolPoints7Driver\_5 | Seventh time the fifth driver is scoring a violation point. |
| ViolPoints8Driver\_1 | Eighth time the first driver is scoring a violation point. |
| ViolPoints8Driver\_2 | Eighth time the second driver is scoring a violation point. |
| ViolPoints8Driver\_3 | Eighth time the third driver is scoring a violation point. |
| ViolPoints8Driver\_4 | Eighth time the fourth driver is scoring a violation point. |
| ViolPoints8Driver\_5 | Eighth time the fifth driver is scoring a violation point. |
| Year\_1 | Year of manufacture of the first vehicle |
| Make\_1 | Make of the first vehicle |
| Model\_1 | Model of the first vehicle |
| Zip | Zip code |
| Total\_Distance\_To\_Work | Total Distance to work of all the drivers combined |
| NoLossSigned | Whether statement of No loss has been signed or not. 1 - yes and 0 - No |
| Type | Different types of auto insurance viz, A, AP, DP, FC, P, REN, RET, VD, XFR |
| CancellationType | Type of cancellation viz, NP, INS |
|  |  |