# Investigating the loss of Insurance Renewal

**Description:**

Investigating loss of Insurance Renewal is taken as part of data analysis; here the objective is to investigate the losses by insurers based on various strategies like policy cancellation, theft of vehicles and accidental cases. The losses are validated for the non renewal Motor vehicle policies.

**Project Description:**

To investigate in this process, we need to gather datasets for theft of vehicles, policy cancellation and subsequent claim cases. The analysis can help the insurer decide what parts of the information to focus on, to show you where they are losing customers. The parts of Information here mean that why the loss is caused and why not the insurance is renewed. Need to cover all the common phases of data analysis like data collection, cleaning, implementation and visualization.

**To define the collection of datasets for the mentioned strategies include:**

1. **Theft cases**: Dataset should have the base details like state, make and model of the vehicle & year of manufacture and count of thefts cases.
2. **Policy Cancellation cases**: Dataset should contain common details like type of coverage, premium paid, vehicle age, NCD availed, age of the insured etc.
3. **Subsequent claim cases**: Dataset should contain the details like age of Insured, sex, number of children, annual premium amount, total number of claims availed.

**Implement the most common data analysis techniques like:**

**Data Collection**: Raw data to be collected through various websites for the defined strategies like Theft cases, Policy Cancellation and Subsequent claim cases.

**Data Processing & Cleaning**: To check for invalid data and correct the same and make as an excel sheet or create as data table to proceed further steps.

**Model and Algorithm Implementation**: Supervised learning technique using K means (with multiple clustering) and decision tree for single condition. And also implement linear regression representation and random forest algorithm for all columns in datasets.

**Data Visualization and reporting the analysis results**: Here we need to abstract the data sources, collect the data, and then filter those data and show the data to end users graphically as charts, graphs and reports like Pie charts, scatter plots and bar charts.