Group Number: 08

* Tejas Patel - 0734912
* Soorya Suresh - 0735168
* Sunny Singh – 0733910

**Automobiles Accidents**

This dataset is all about accidents that occurred in Canada during the years 1999 to 2017 and accidents has major impact in every phase of the world. we will try to find enriching factors associated to accident fatalities and for that we will use machine learning which will help us to predict the fatalities.

Our hypothesis says that accident is causing a lot of fatalities because of weather, as we are aware Canada is known for its weather and due to snow many drivers are not able to drive properly and its difficult to drive in such situation as it requires a lot of attentiveness. Some lags in it and they end up being getting hit or damaging infrastructure and due to this there are long traffic jams on the road which increases the pollution and waste a lot of time of other people. We will try to find the vital effect behind accident and how we can reduce this by implementing safety for such reasons and try to predict the fatal injuries because of the accidents.

We will be working on the dataset which is having in total of 1,48000 rows and 23 features and in this data we need to do data wrangling to make it useful for analysis as in this data we have columns that are not that useful for model to do analysis so we have to remove that and also there will be some unknown values in the data which has to be removed. And this data is in categorical format so we have to modify the data according to our need and create a useful dataset.

We will be building lot of models for doing the analysis on the data like we will use regression models like linear, decision tree, XGBoost, Random forest and we will also try to build the classification models on this data as we have data which is categorical and we will also work on that side of the data too. We will also use tableau software to build graph and do visualization of the data and find some valuable insights from the graphs.