PHASE 2 RESULTS

- ***Definition of Safety & Time Periods:***

To determine whether the roads are safer than before, we need to split our time frame into 2 periods, which are:

  - *2002-2014* --> Past
  - *2015-2023* --> Now

Also, we need to define what "safer" is. Fort his reason, we will be using 3 parameters:

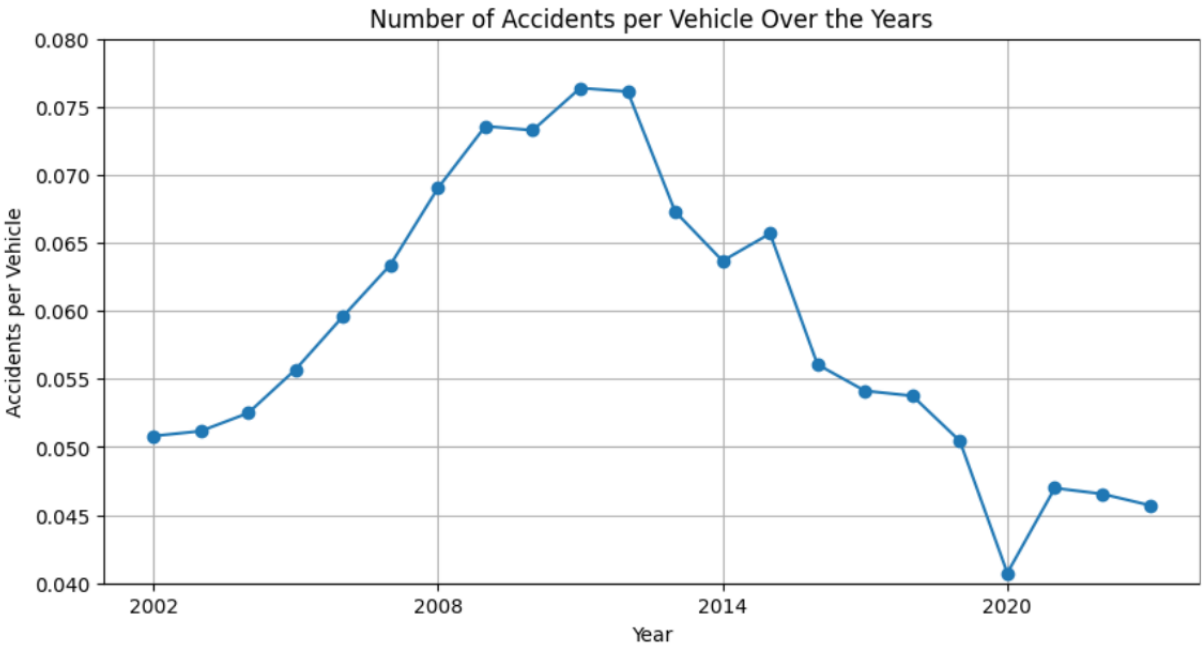
  - ***Accidents per vehicle***: Number of accidents that occured that year per the amount of vehicles that wer on the roads that year.

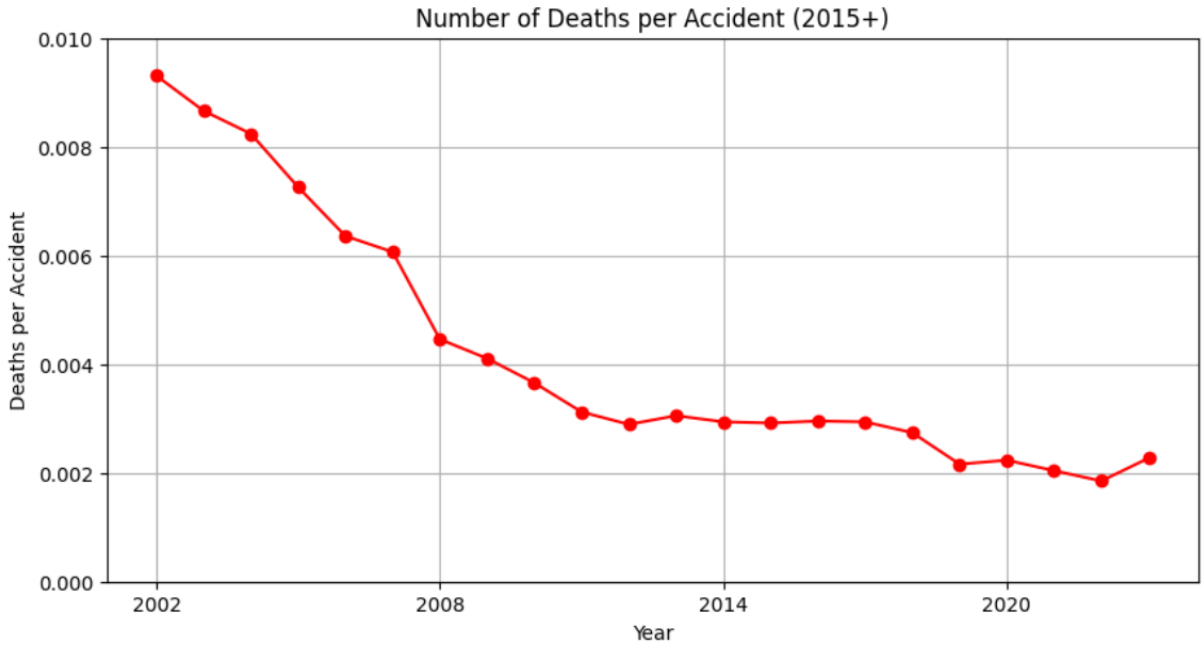
  - ***Deaths per accident***: This gives us the death rate of the accidents. (The number of deaths caused by the accident that occur in hospitals in the next 30 days after the accident have not been recorded between 2002-2014. Fort hat reason, to have a more consistent dataset, I have used only the number or fatalities recorded at the crash site)

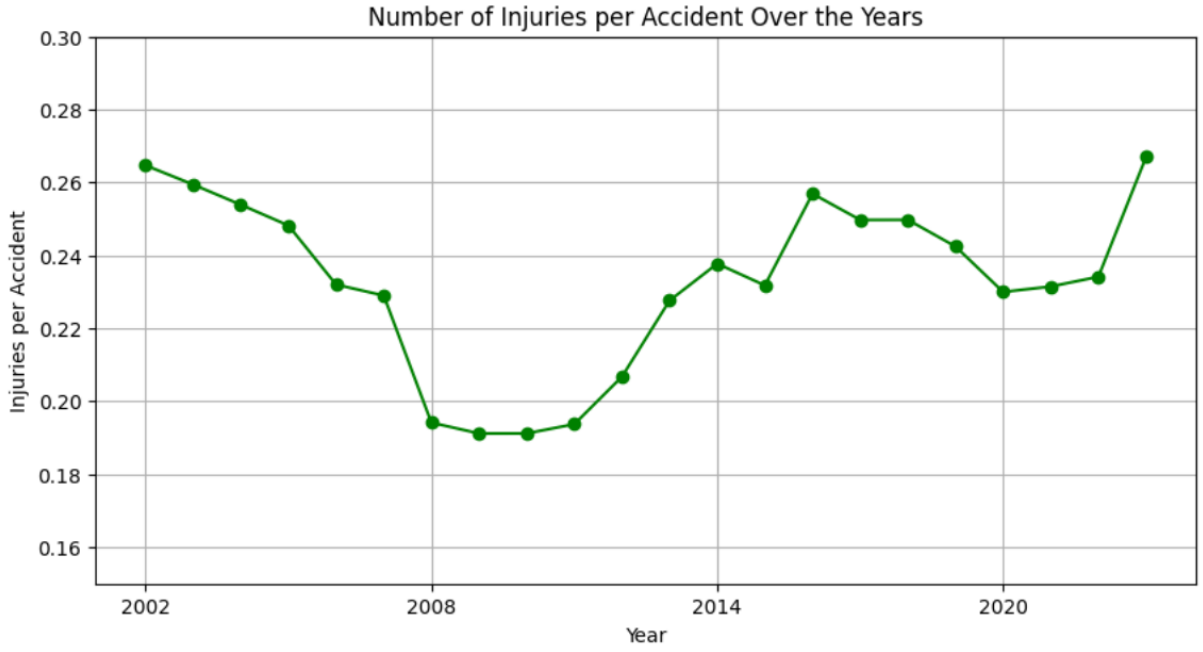
  - ***Injuries per accident***: Similar to the death rate of the accident, this gives us the injury rate of every accident

- ***Exploratory Data Analysis Results:***

When we calculate these three parameters and plot them over their values over the years, we get these results:







These three graphs show that there are some consistent decreasing trends for out *"Deaths per Accident"* parameter, whereas the other two graps show more varying trends over different time periods.

We also calculated some statistics to further understand the differences between the periods (These are direct results of the Python code):

ACCIDENTS PER VEHICLE:

Minimum rate of accidents per vehicle: 0.04074606861411521

Maximum rate of accidents per vehicle: 0.07638061228396507

Mean before 2015: 0.06404078110031187

Mean after 2015: 0.051122843773686644

Difference between means: -0.012917937326625231

Std dev before 2015: 0.009459407313590941

Std dev after 2015: 0.007312708410178602

DEATHS PER ACCIDENT:

Minimum rate of deaths per accident: 0.0018508350250657567

Maximum rate of deaths per accident: 0.009306989678859968

Mean before 2015: 0.005393015838183288

Mean after 2015: 0.002456105279608589

Difference between means: -0.0029369105585746987

Std dev before 2015: 0.0023702147622526423

Std dev after 2015: 0.00043061177593134827

INJURIES PER ACCIDENT:

Minimum rate of injuries per accident: 0.19118124528882247

Maximum rate of injuries per accident: 0.2669853044129375

Mean before 2015: 0.22535856957180853

Mean after 2015: 0.2437017416772691

Difference between means: 0.018343172105460576

Std dev before 2015: 0.02726406833327248

Std dev after 2015: 0.013051369215305177

- ***Hypothesis Testing:***

We determine our general hypotheses as such:

  - ***Null Hypothesis***: There is no difference between then and now in terms of safety
  - ***Alternate Hypothesis***: The roads are more safer now than before

Again, the time peroids will be 2002-2014 for before and 2015-2023 for now. To make a conclusion **we will make 3 t-tests**, **on our 3 safety parameters**: accidents per vehicle, deaths per accident and injuries per accident**. If at least 2 of our test accept the alternate hypothesis, then we will accept our general alternate hypothesis**(i.e. the roads are safer now). **Our t-tests will be one-tailed, meaning we are only checking if there is a decrease in the parameters we are researching.**

Here are the results of our hypothesis tests:

HYPOTHESIS TESTING:

Accidents per vehicle: t = 3.6071732282394318 p = 0.0008979869700369997

Deaths per accident: t = 4.36476894534469 p = 0.0003745251846628744

Injuries per accident: t = -2.1026450757684834 p = 0.9751884488536819

[Accidents per vehicle] --> Significant difference (p = 0.0009) → Reject H_0

[Deaths per accident] --> Significant difference (p = 0.0004) → Reject H_0

[Injuries per accident] --> No significant difference (p = 0.9752) → Fail to reject H_0

These results tell us that there are less accidents occuring per vehicle and there are less deaths occuring per accident in comparison with the past, whereas we cannot say that there is a decrease in injury rates of accidents between our two time periods.

In conclusion, since we rejected 2 of our 3 null hypotheses, we say that (by our definition of safety) the roads are safer right now than before 2015.