**Removing Nulls and Data type conversion**

* The dataset contains the vehicle crash information. The dataset was loaded in to tableau.
* Research was done to find out the meaning of each and every attribute in the dataset. The columns that did not make any sense were hidden to focus on the other attributes.
* Filter was applied at the data source level for the crash year column to filter out any nulls.
* Filter was also applied at the data source level the speed limit column to filter out any speed limit greater than 85.
* Filter was applied at the data source level for the BAC\_score column to filter out any data points that had a Bac\_score of more than 10. It is assumed that they are nulls as they are not relevant to the normal Bac\_score.
* Filter was applied at the data source level to filter out any vehicle that had a year of less than 100
* Several fields were in different format. For example the year was considered as text. Conversions were made so that each column pertains to the relevant data type.

**Calculate Fields**

* Few Calcuated fields were created from the attributes in the data set by using conditional statements as this gives a way to categorize the data points and provide better visualization.
* The field AGE\_GROUP wa screated using the formula

if[Driver\_Age] <= 2 then "Infants"

ELSEIF [Driver\_Age] >2 AND [Driver\_Age]<=12 then "Children"

ELSEIF [Driver\_Age]>12 AND [Driver\_Age]<=17 THEN "Teenagers"

ELSEIF [Driver\_Age] >17 and [Driver\_Age]<=25 then "Youth"

ELSEIF [Driver\_Age] >25 AND [Driver\_Age]<=40 THEN " Adults"

ELSEIF [Driver\_Age] >40 AND [Driver\_Age]<=60 THEN "Middle Age"

ELSE "Senior"

END

* The field BAC\_Level was created using the formula

if [BAC\_score] >= 0.08 then "illegal"

else "legal"

end

* The field HOV was created using the formula

IF [Occupants] <= 1 THEN "Single Occupant"

ELSE "Multiple Occupants"

END

* When going through the data the attribute speed limit is recognized as the speed limit of the road when the crash occurred and not the speed of the vehicle itself. So based on this a calculated field was created to categorize the road types using this formula

IF [Speedlimit] <= 20 then "Low speed"

ELSEIF [Speedlimit] >20 and [Speedlimit] <= 30 then "Urban(30mph)"

ELSEIF [Speedlimit] >30 and [Speedlimit] <= 35 then "Unpaved(35mph)"

ELSEIF [Speedlimit] >35 and [Speedlimit] <= 45 then "Rural paved(45mph)"

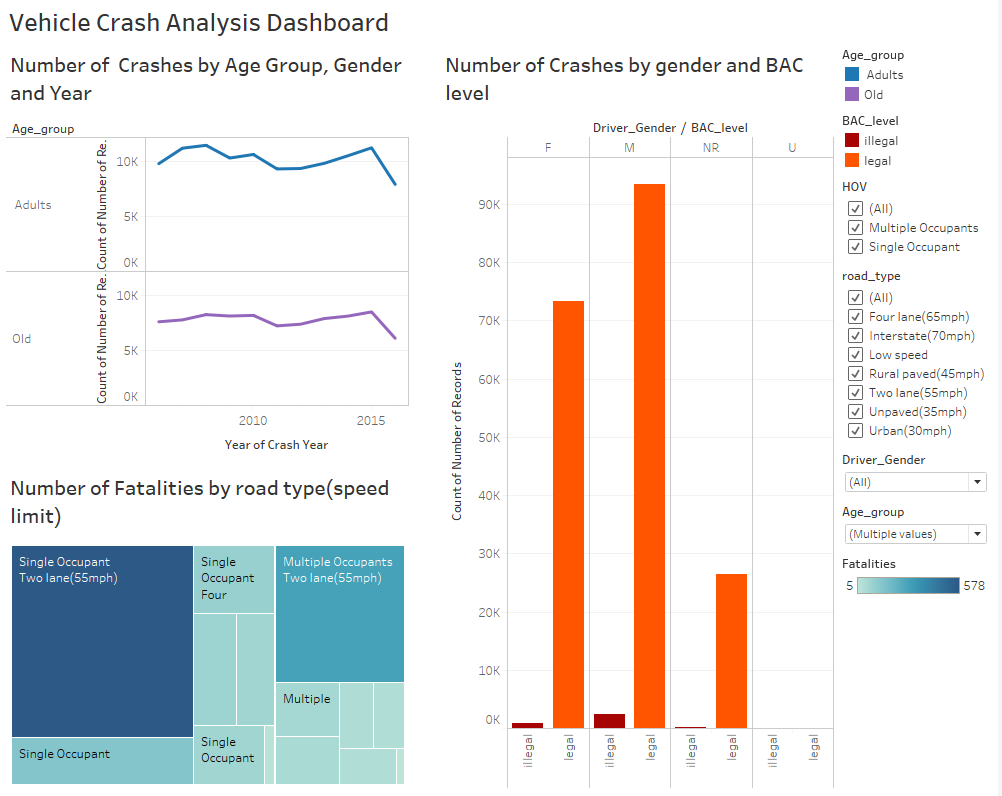
ELSEIF [Speedlimit] >45 and [Speedlimit] <= 55 then "Two lane(55mph)"

ELSEIF [Speedlimit] >55 and [Speedlimit] <= 65 then "Four lane(65mph)"

ELSE "Interstate(70mph)"

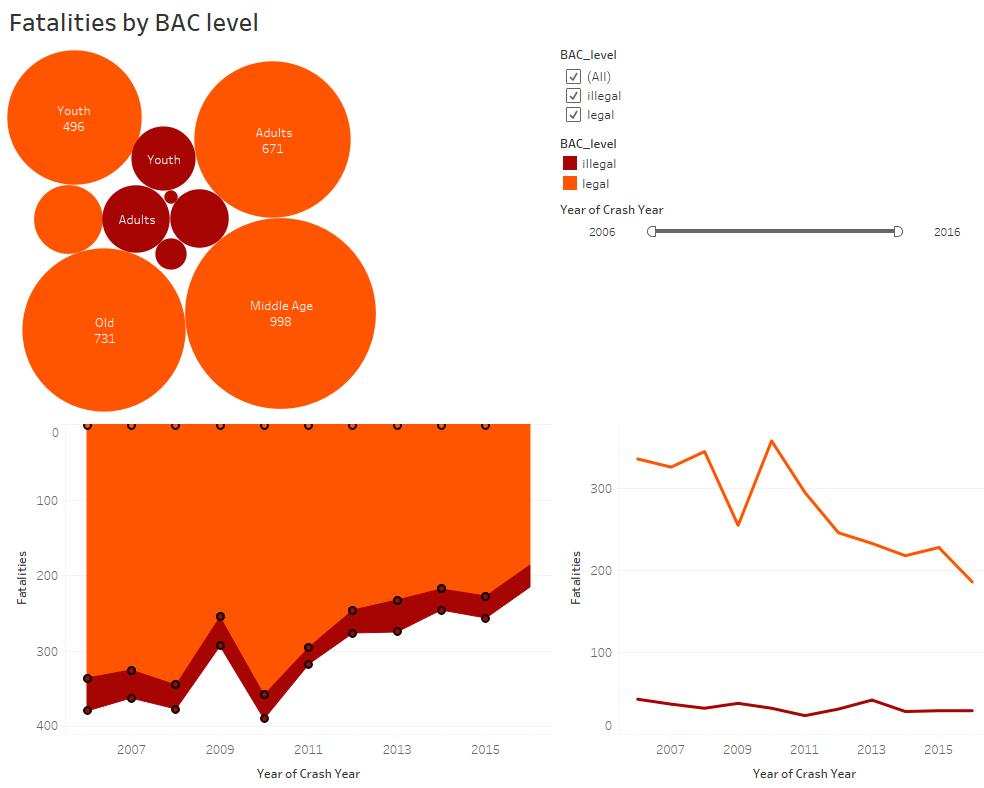
END

**Tableau Dashbards for the Vehicle Crash Analysis**



The above dashboard has three charts.

* A line chart the shows the Number of Crashes (which is count of number of records ) for all the age groups from 2006 to 2016. There were no month or date data in the dataset so it cannot be drilled down to quarter, months or any other time. Also, the trend shows a drop in the number of crashes in 2016. The main reason for this can be because the dataset contains only few months of data for 2016.
* A side by side bar chart that shows the number of crashes by driver gender and legal or illegal BAC score
* It has a heat map that shows the number of fatalities by road type and the type of road. The maximim number of fatalities was with Single Occupants on a two lane(55mph) roads and the minimum was Multiple Occupants on Unpaved(35mph) roads.
* It also has filters which allows the user to filter out the data points in the viz by just clicking on it.
* The filters were set to reflect on all the sheets.



The above dashboard has three charts which almost shows the same information.

* A bubble chart which shows the number of fatalities by age group.
* A line chart which shows the number of fatalities from 2006 to 2016.
* An area chart which shows the number of fatalities from 2006 to 2016. The vertical axis which shows the number of fatalities is reversed to give a different view.
* All the three charts shows the records with legal Bac score in orange color and illegal Bac score in dark red.
* It also has a filter to view the data for the desired year.