**Introduction**

Every year the lives of more than 1.25 million people are cut short as a result of a road traffic crash. Between 20 and 50 million more people suffer non-fatal injuries, with many incurring a disability because of their injury. Road traffic injuries cause considerable economic losses to individuals, their families, and to nations.

The Road accidents in The United States in the year 2015 has been analyzed and the effect of various factors on accidents has been considered and some interesting insights are found which are explained below.

**Data Sources**

The Fatality Analysis Reporting System (FARS) provides information on all motor vehicle traffic crashes in the U.S.[**ftp://ftp.nhtsa.dot.gov/fars/2015/National/**](ftp://ftp.nhtsa.dot.gov/fars/2015/National/)**.**

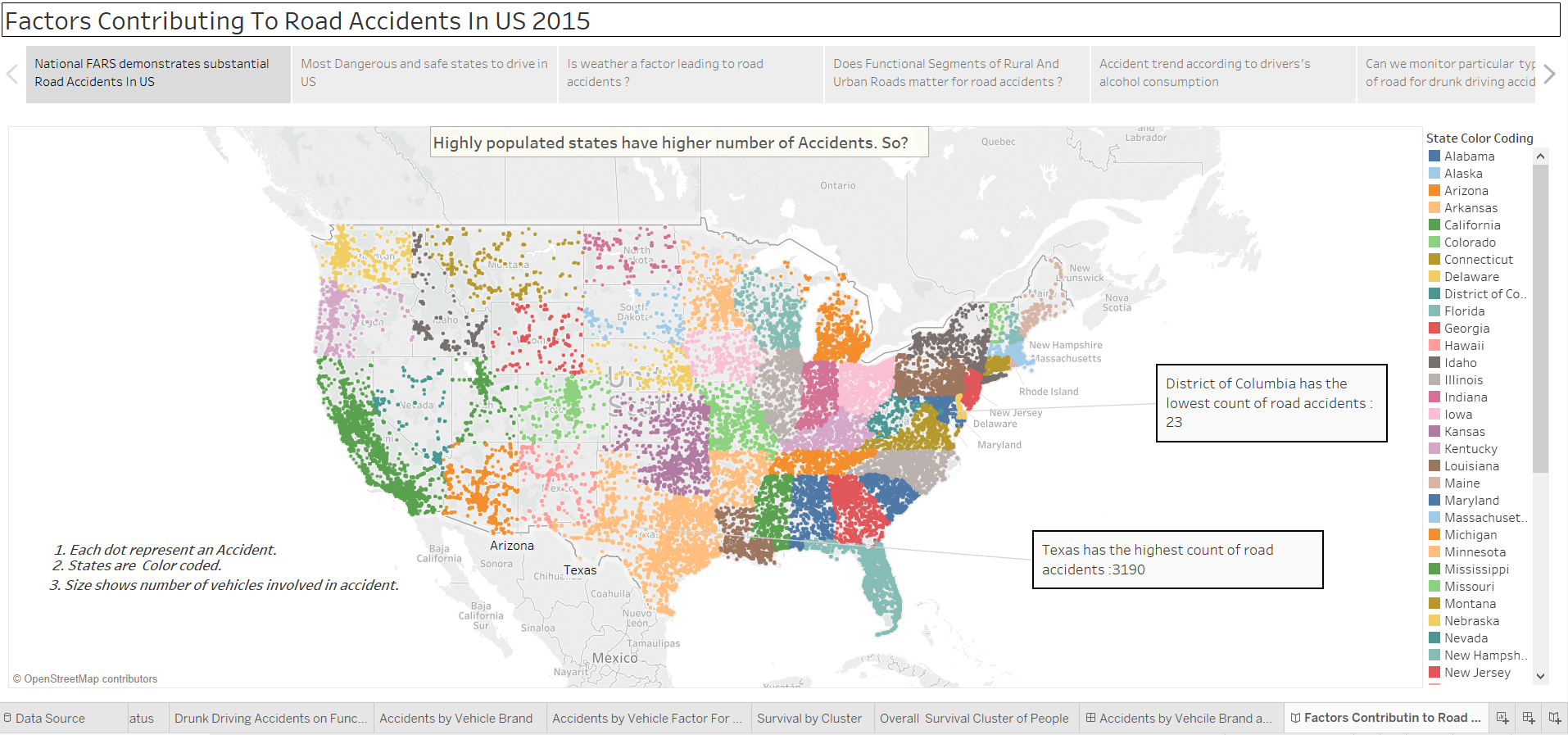
* **Primary data source is the file** Accidents.csv.
* **Secondary data sources are** Vehicles.csv, Factors.csv, Persons.csv ,Population.xlsx, Vehicle Make.xlsx

**Insights**

1. Wyoming is the most dangerous state to drive in.
2. Cloudy, Rain, Snow and Fog are the prominent bad weather conditions across the top three dangerous states to drive.
3. Most of the accidents occurred in Principal Arterial followed by Collector & Local segment of roads.
4. Drink & Drive accidents are less on weekdays and spiked during weekends around 1 am - 2 am, whereas nondrinking accidents spiked during evening hours of 4 pm - 5 pm during weekdays.
5. At peak hours of Drink & Drive accidents during 1 to 2 AM, Rural Roads should be monitored more for collector segment & Urban Roads for another segment of roads.
6. People in the 2nd cluster have highest survival rate than the most and those in the 5th cluster has highest death rate.
7. People seating on the Front Seat are at high risk of not surviving the accident than the ones seating in other seats.
8. The ratio of number car accidents to the number of vehicles sold by a maker is very high for Suzuki and non-human factor for accidents is Headlights.

**Visualizations**

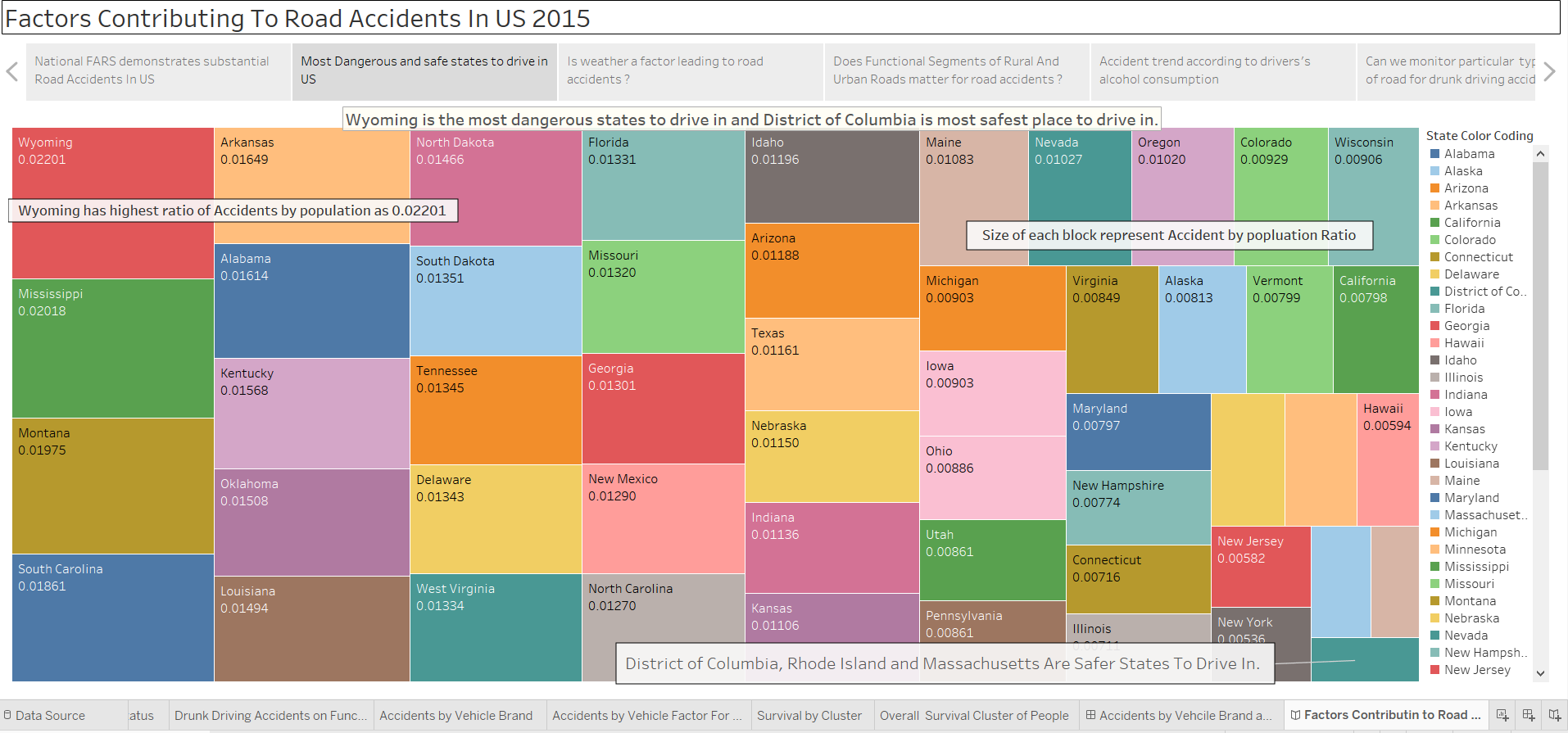
1. National FARS demonstrates substantial Road Accidents In US:



The states are color coded i.e. each state is represented by different color. The density of points represents the density of accidents taking place in the state.

**Observation:** Highly populated states like Texas, California have a higher number of accidents. Highly Populated are most dangerous states to drive in.

2. Most Dangerous and safe states to drive in the US:

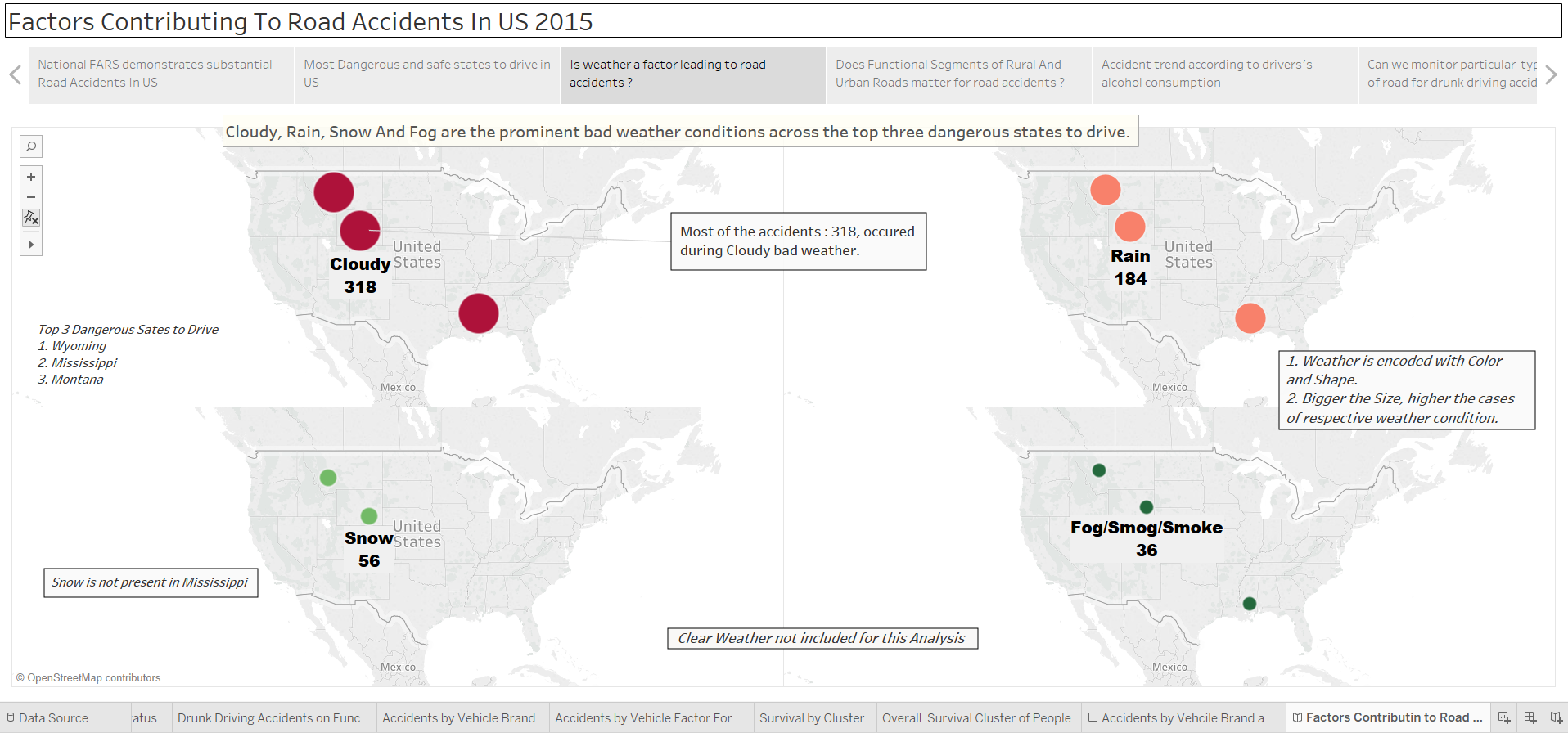


The above tree map has been shown using a calculated field (ratio) to estimate the most dangerous states to drive in. It is clearly visible on the map that as the ratio is decreasing (dangerous to safest), the size of the states is also decreasing.

**Observation:** Wyoming has the highest ratio of the number of accidents to the population.

Wyoming is the most dangerous state to drive in.

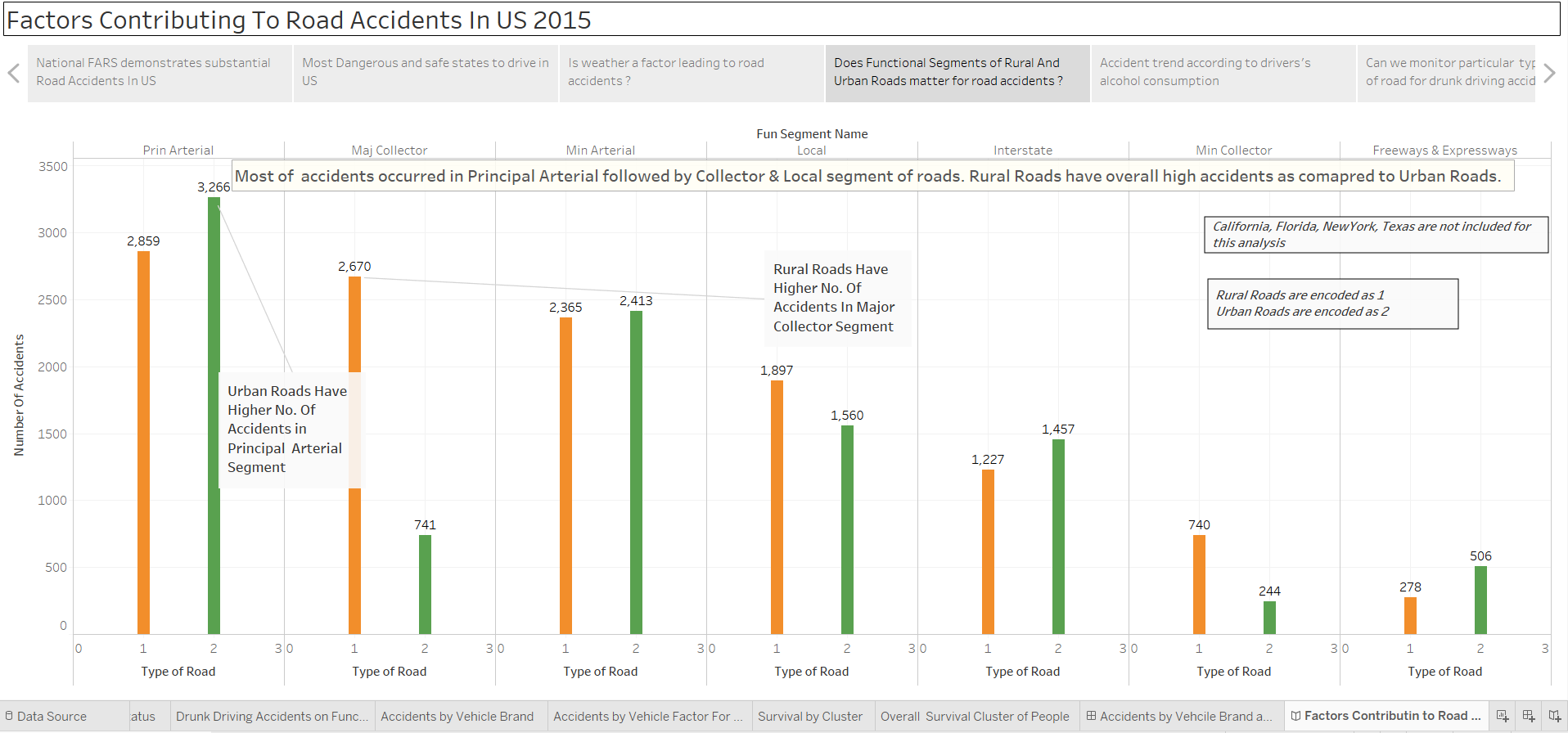
3. Is weather a factor leading to road accidents?



This includes multimaps to show the top 4 weather conditions affecting accidents in the most dangerous states found in the previous visualization. Color is representing different types of weather. The size of the circle represents the number of accidents taking place due to a weather.

**Insight:** Cloudy, Rain, Snow and Fog are the prominent bad weather conditions across the top three dangerous states to drive.

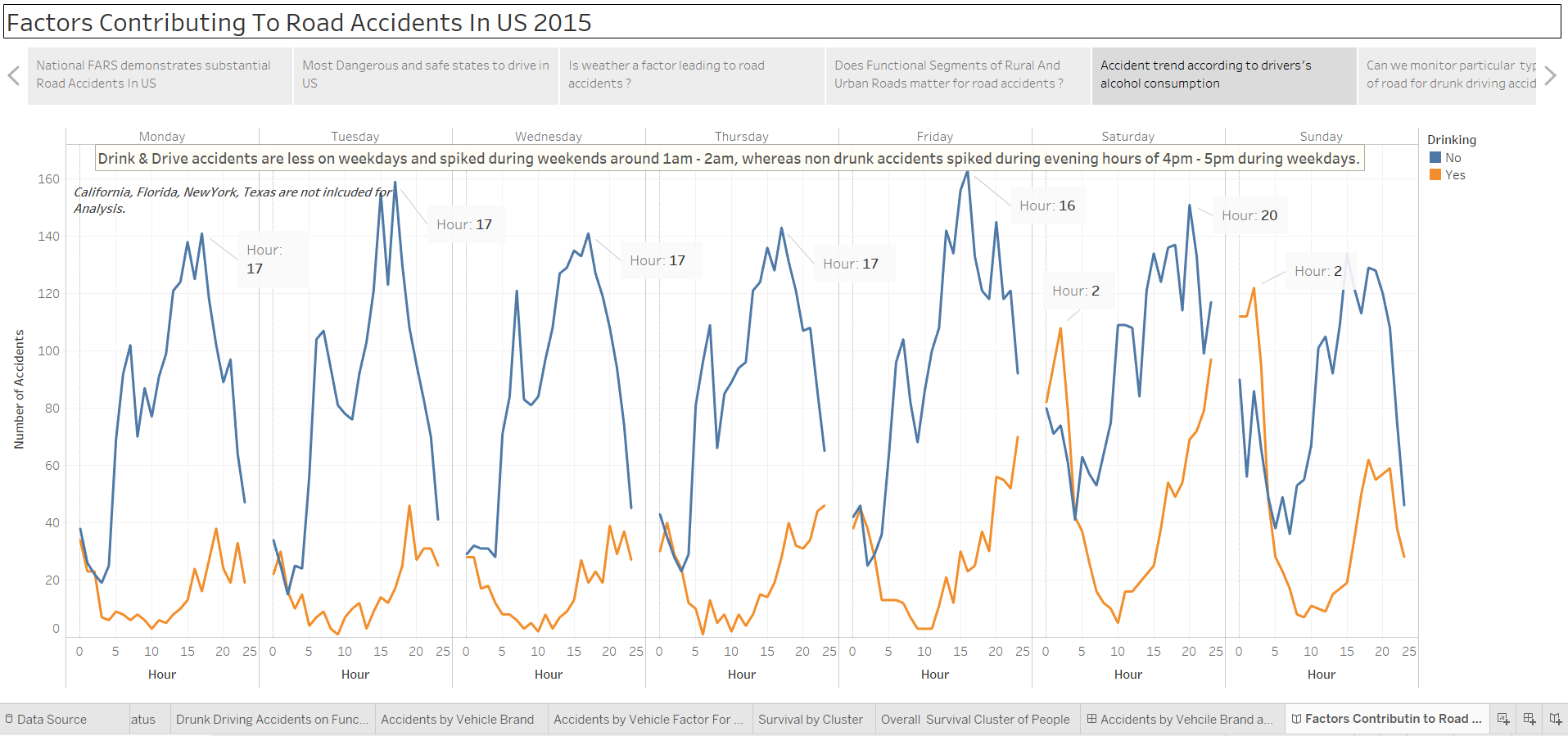
4. Does Functional Segments of Rural and Urban Roads matter for road accidents?



In this visualization, the number of accidents as per functional segments of the roads has been grouped using bar graphs. The green bar represents accidents which took place in urban roads and orange bar represents accidents which took place in rural roads. Sum of accidents is mentioned on top of the bars.

**Insight:** Most of the accidents occurred in Principal Arterial followed by Collector & Local segment of roads.

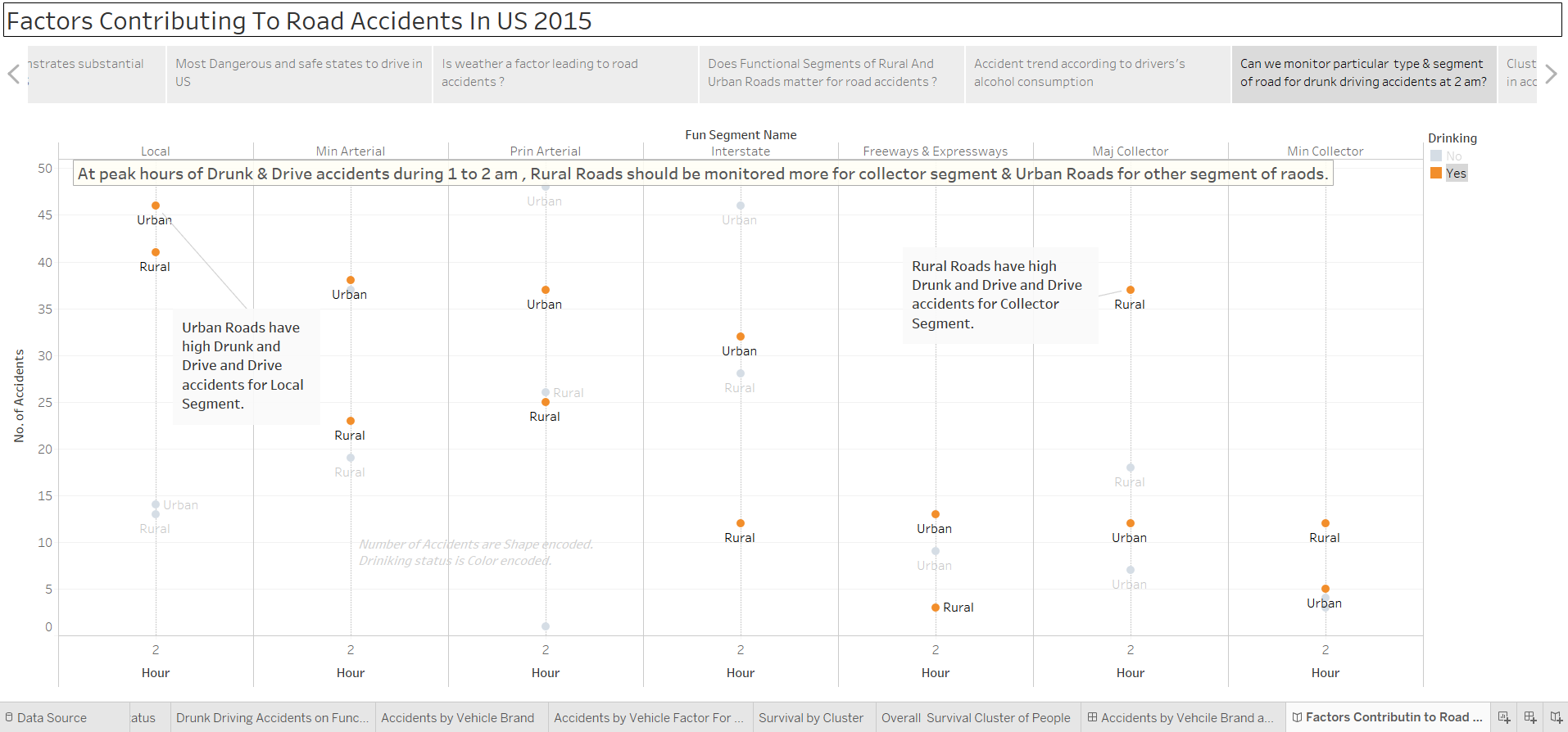
5. Accident trend according to drivers’ alcohol consumption



Trend lines are used to depict accidents taking place due to consumption and non-consumption of alcoholic drinks. The accidents are grouped by time and day of the week. The trend line in blue shows accidents taking place and orange trend lines show accidents taking place due to drinking and driving. Peak time (24-hr format) of accidents taking every week are annotated.

**Insight:** Drink & Drive accidents are less on weekdays and spiked during weekends around 1 am - 2 am, whereas nondrinking accidents spiked during evening hours of 4 pm - 5 pm during weekdays.

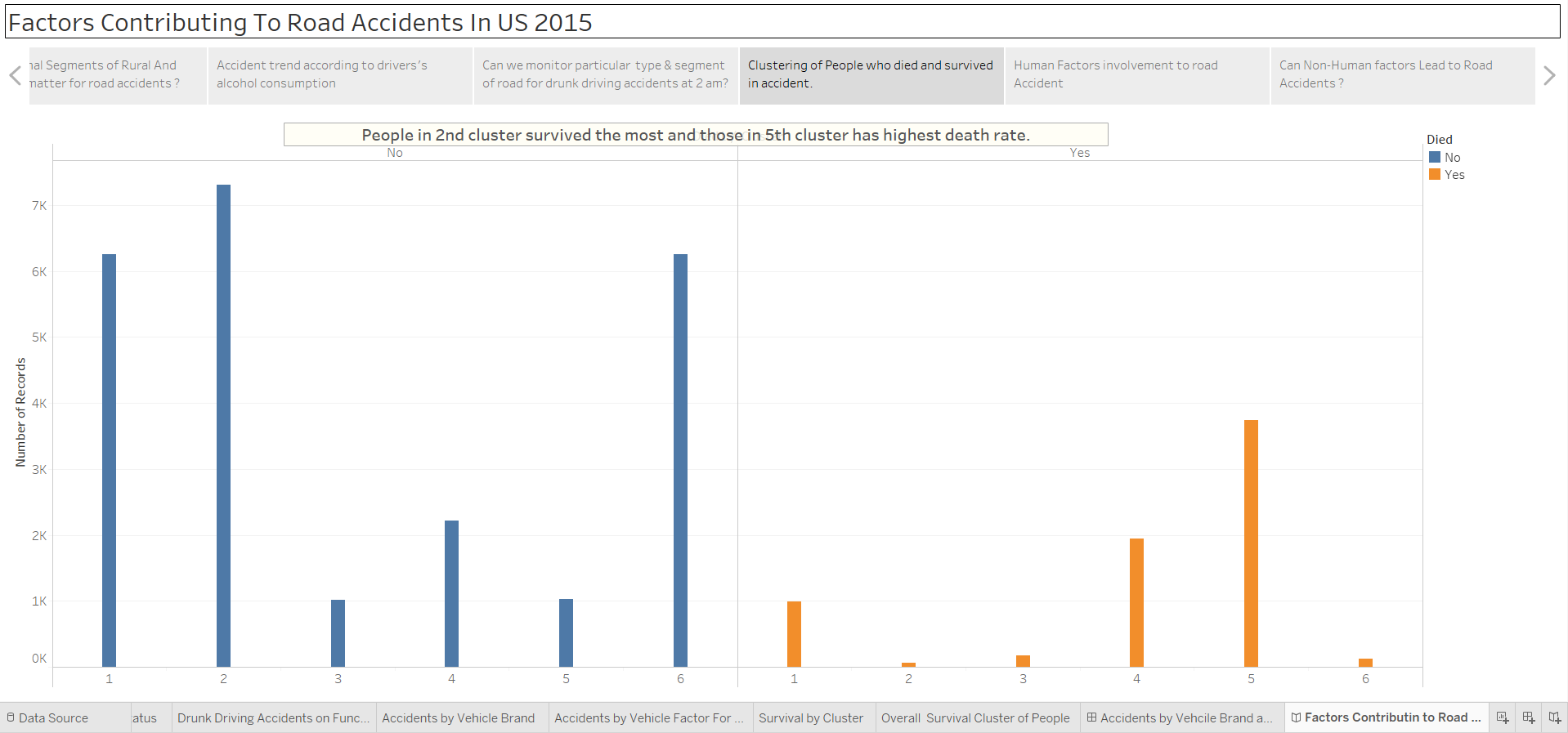
6. Can we monitor type & segment of road for drunk driving accidents at 2 am?



A filter has been applied for drink and drive accidents according to road segments and rural/urban roads at the peak time of accidents (2 am). Orange dot represents Drink & drive accident cases & blue dot represents accidents not involving alcohol.

**Insight:** At peak hours of Drink & Drive accidents during 1 to 2 AM, Rural Roads should be monitored more for collector segment & Urban Roads for another segment of roads

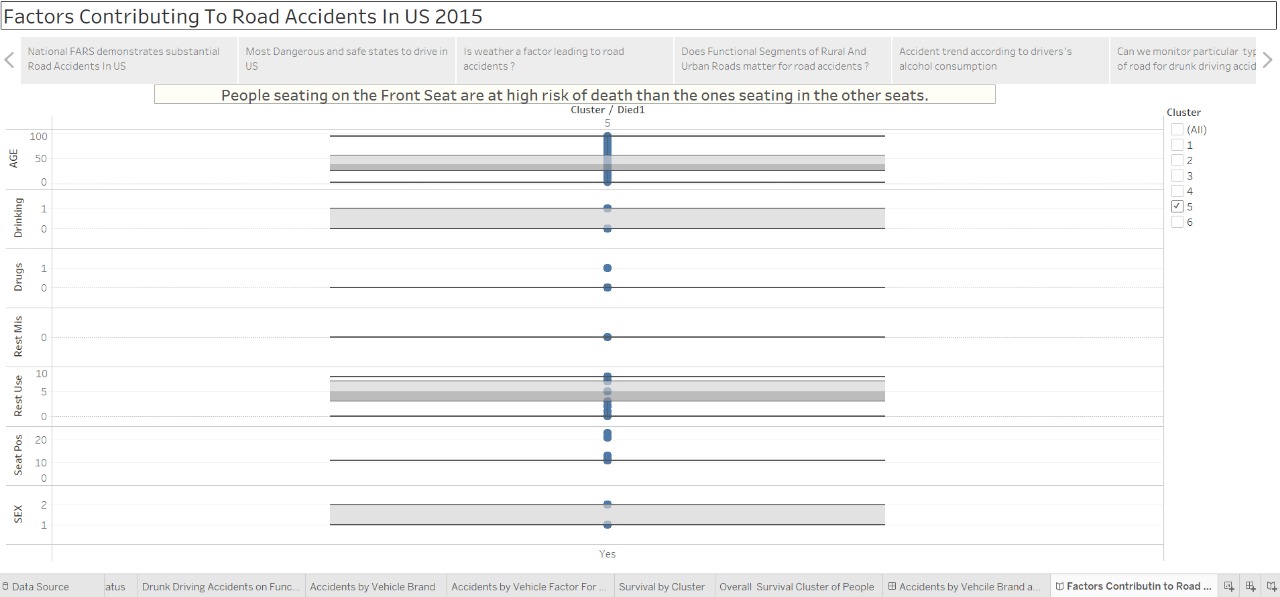
7. Clustering of People who died and survived the accident



Bar graphs are used in this visualization. Blue coded data shows the survival cluster and orange-coded data shows death clusters. Here the accidents are clustered in 6 clusters using bar graphs. The count of 6 clusters has been obtained from clustering analysis using R. Above is the plot of the total within sum square vs a number of clusters. The plot conveys that curve stagnates at a number of clusters equal to 6 and hence it has been used as an optimum number of clusters.

**Insight:** People in 2nd cluster survived the most and those in the 5th cluster has highest death rate

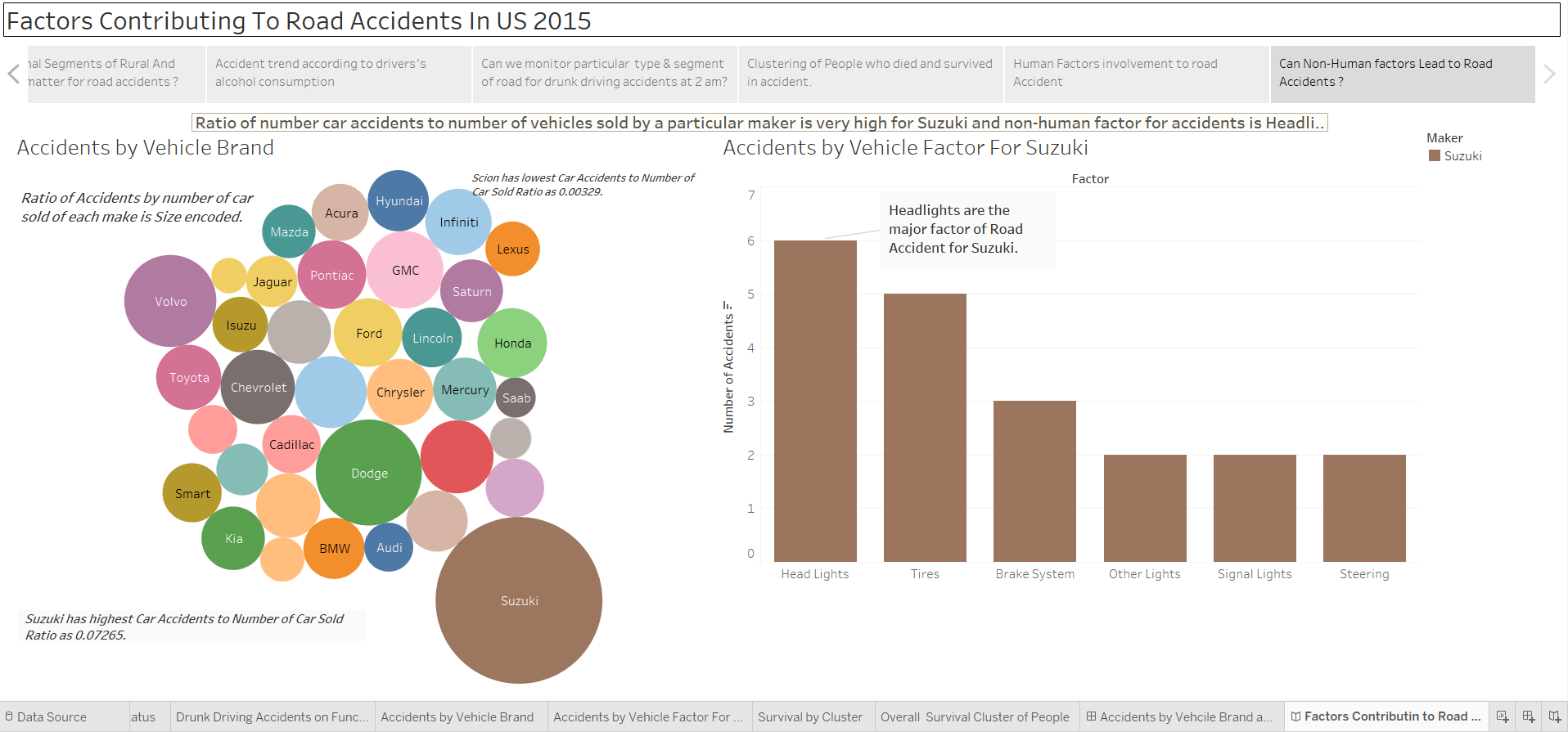
8.Human Factors involvement in Road Accident



Box plots are used for this visualization. Yes at x-axis represents whether the person survived or not. Y-axis has various attributes of the victim i.e. gender, seating position, restraint used, drugs use, drinking and driving and age.

**Insight:** People seating on the Front Seat are at high risk of death than the ones seating in the other seats

9. Can Non-Human Factors Lead to Road Accidents?



This dashboard has two visualizations. The bubble-chart depicts accidents by vehicle brand. The size of the bubble represents the calculated field(ratio) of a number of accidents by brand. Color represents different car manufacturer.

The second visualization has a bar graph showing factors affecting road accidents. This visualization is for Suzuki. A filter has been applied for Suzuki which is color-coded with brown color.

**Insight:** Ratio of number car accidents to a number of vehicles sold by a particular maker is very high for Suzuki and non-human factor for accidents is Headlights.