

# Road Accident Data Analysis

## Using Python and Tableau

---

**Prepared by:** Vaishnavi M Jadhav

---

### PROJECT OVERVIEW

---

The project "Multi-Dimensional Data Analytics of Road Accidents in India" focuses on analyzing road accident data from across India using various data visualization and analytical techniques to uncover key insights and patterns. The goal is to better understand the causes, frequency, and distribution of road accidents with respect to factors like time, location, vehicle type, and driver behavior. The analysis leverages tools such as Python, Pandas, and Matplotlib, and presents findings through interactive visualizations and dashboards. This project aids policymakers, traffic authorities, and researchers in identifying high-risk areas and implementing data-driven strategies to enhance road safety in India.

### TOOLS & SKILLS USED

---

- **Python** – Core programming language used for data processing and analysis.
- **Pandas** – For data cleaning, manipulation, and handling large datasets efficiently.
- **NumPy** – For numerical operations and array processing.
- **Matplotlib & Seaborn** – For creating static visualizations and exploratory data analysis.
- **Jupyter Notebook** – For organizing code, visualizations, and insights in a readable format.

### DATA EXPLORATION STEPS

---

- **Data Collection:** Gathered road accident datasets from official sources.
- **Data Loading:** Loaded data using Pandas for inspection.
- **Data Cleaning:** Handled missing values, duplicates, and inconsistent formats.
- **Data Preprocessing:** Created new features and converted data types.
- **Univariate Analysis:** Explored individual variables using visualizations.
- **Bivariate & Multivariate Analysis:** Analyzed relationships between multiple factors.
- **Geographical Visualization:** Mapped accident data across Indian states.
- **Time Series Analysis:** Tracked accident trends over time.
- **Outlier Detection:** Identified unusual patterns and anomalies.
- **Insights & Interpretation:** Summarized key patterns and actionable findings.

## KEY INSIGHTS

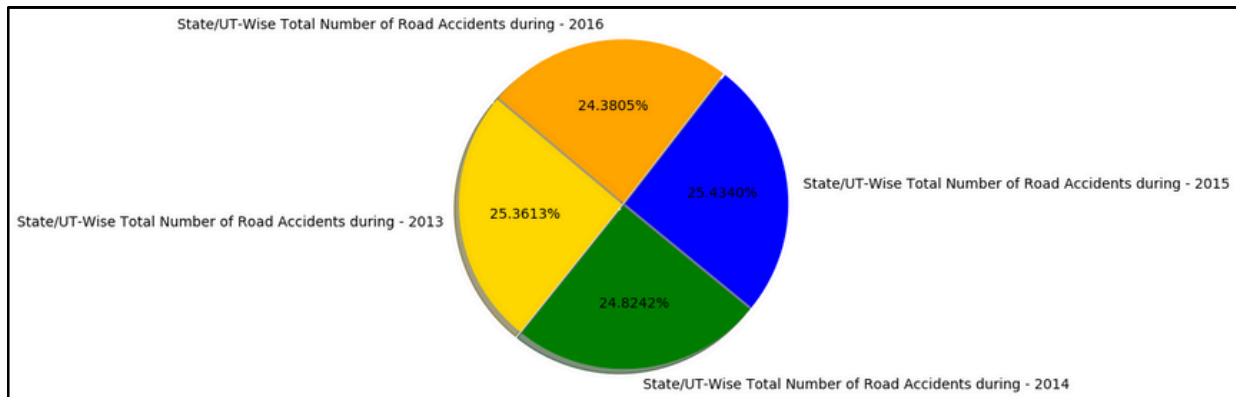
---

- **High-Fatality States:** States like **Tamil Nadu, Uttar Pradesh, and Maharashtra** consistently report the highest number of accidents and fatalities.
- **Time-Based Trends:** Most accidents occur during peak traffic hours, especially between **6 PM to 9 PM**, indicating increased risk during evening commutes.
- **Monthly Patterns:** A spike in accidents is observed during **festive months** and **monsoon seasons**, pointing to weather and travel activity as contributing factors.
- **Vehicle Involvement:** **Two-wheelers** and **trucks** are among the most frequently involved in road accidents.
- **Age Factor:** **Young adults (18–35 years)** are the most affected demographic, indicating the need for better driver education and awareness.
- **Urban vs Rural:** **Rural areas** often see higher fatality rates despite fewer accidents, suggesting poor emergency response and infrastructure.
- **Cause Analysis:** **Speeding, drunk driving, and violation** of traffic rules are major causes of accidents across most regions.
- **State-Wise Variations:** Some states have better **accident-to-fatality ratios**, implying more effective road safety measures or medical facilities.

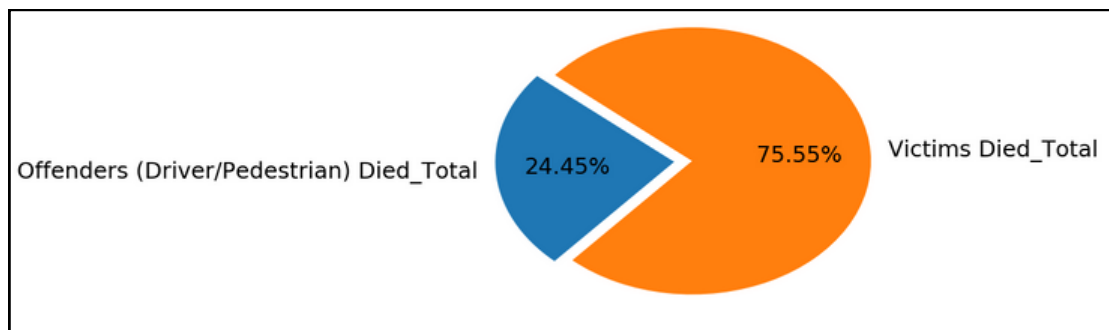
## SNAPSHOTS OF OUTCOMES

---

- Total Number of Road Accidents from 2014 to 2016 - Pie Chart

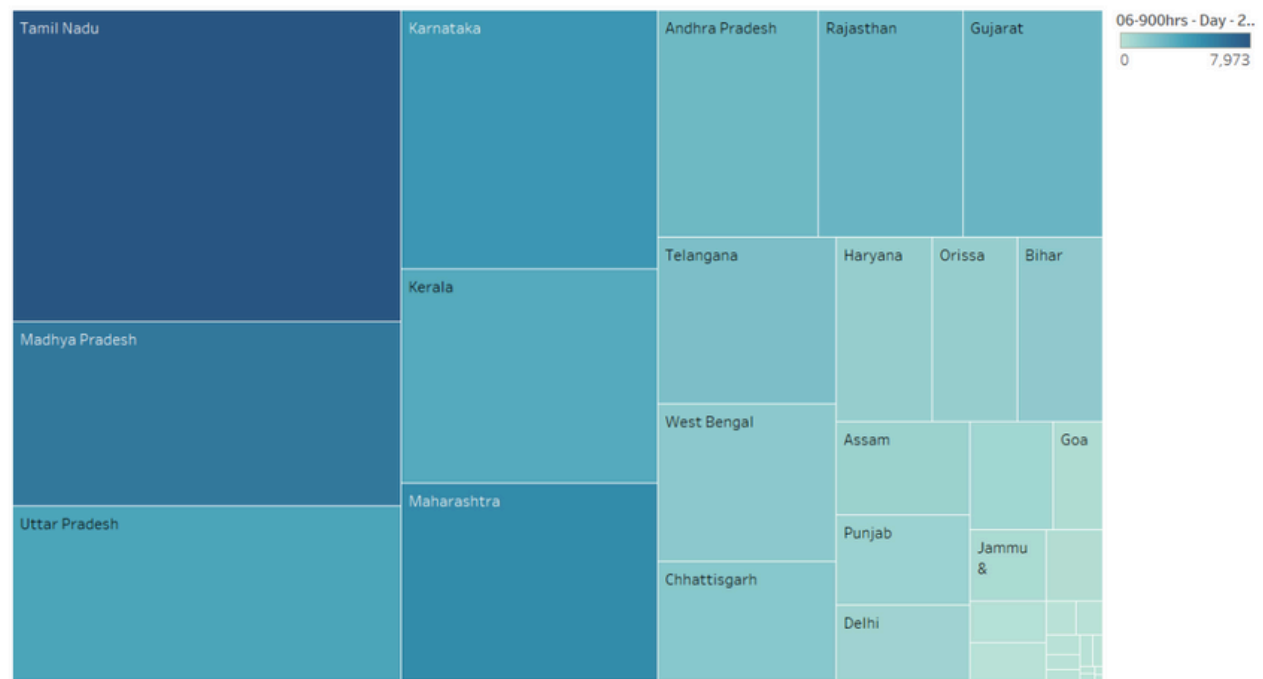


- Percentage of total offenders and victims who died in accidents.



- Accidents according to the TIME OF OCCURENCE.

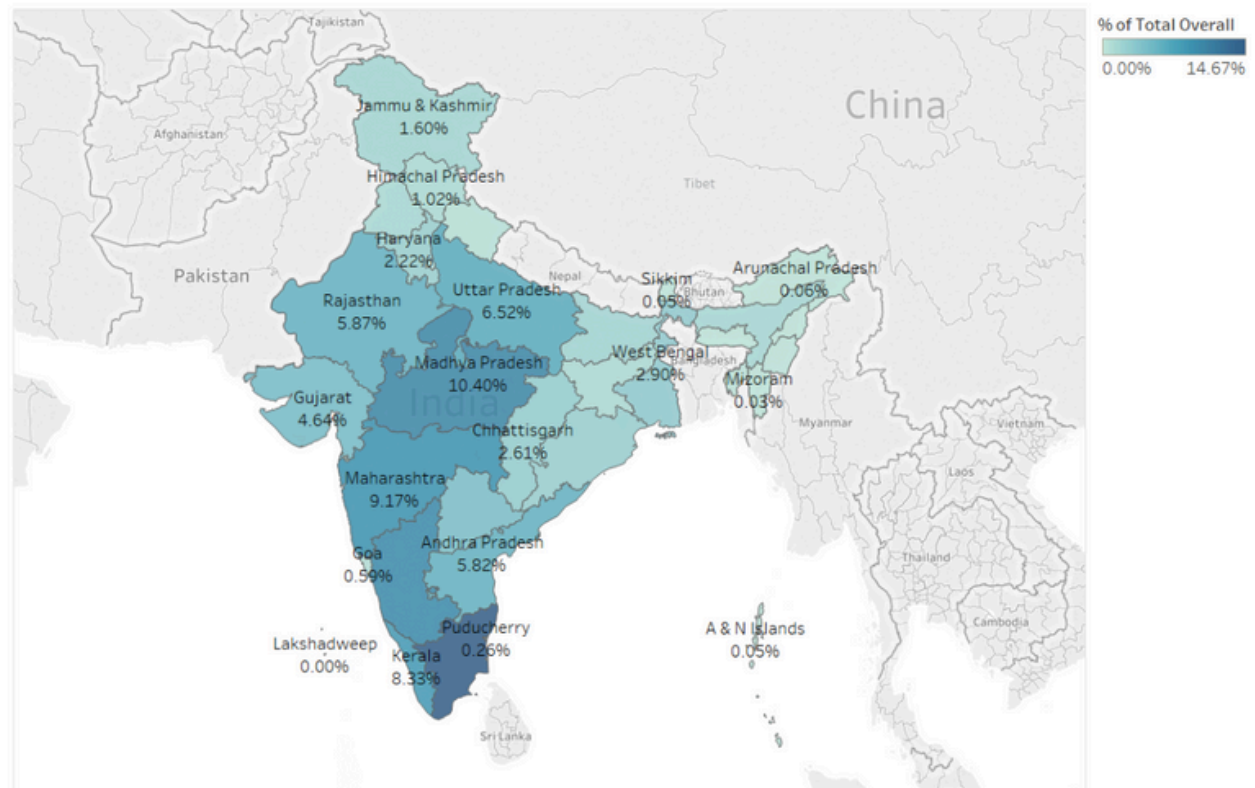
Treemap of people killed between 06-900hrs - (Day) - 2014 and 2016.



States/Uts. Color shows sum of 06-900hrs - Day - 2014. Size shows sum of 06-900hrs - (Day) - 2016. The marks are labeled by States/Uts. The view is filtered on States/Uts, which excludes Total.

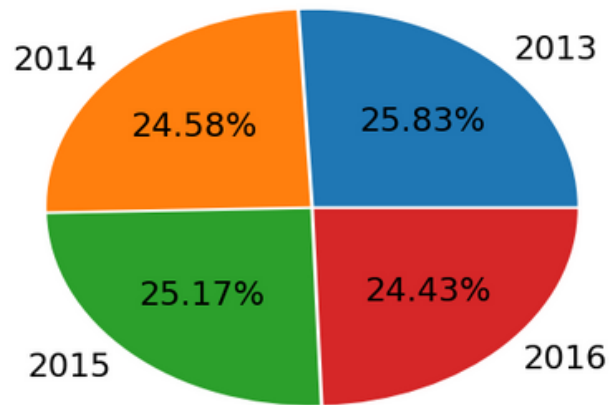
- Count of accidents plotted over a MAP using Tableau.

Overall Accidents

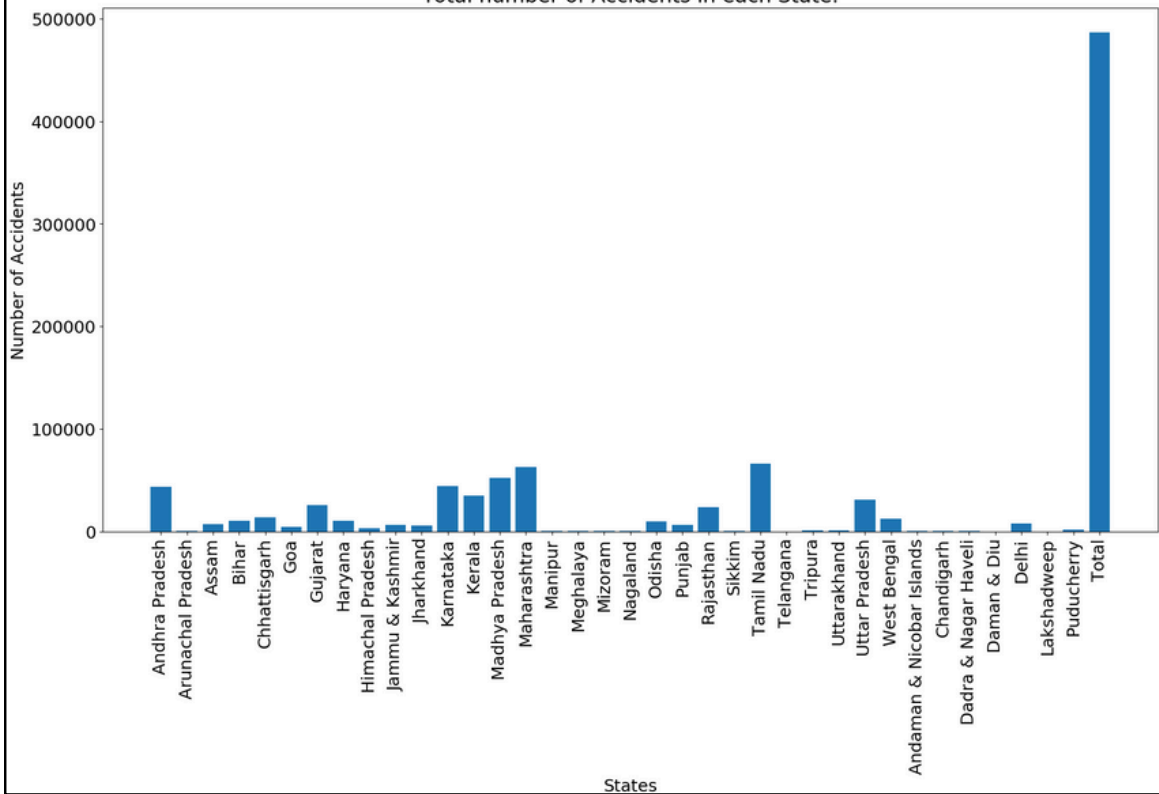


Map based on Longitude (generated) and Latitude (generated). Color shows % of Total Overall. The marks are labeled by State/ UT and % of Total Overall. Details are shown for State/ UT and State/ UT. The view is filtered on State/ UT, which excludes Total.

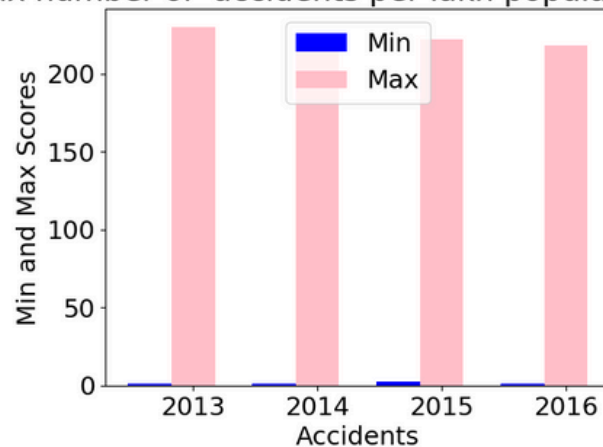
Mean Accidents per 1L population for each year.



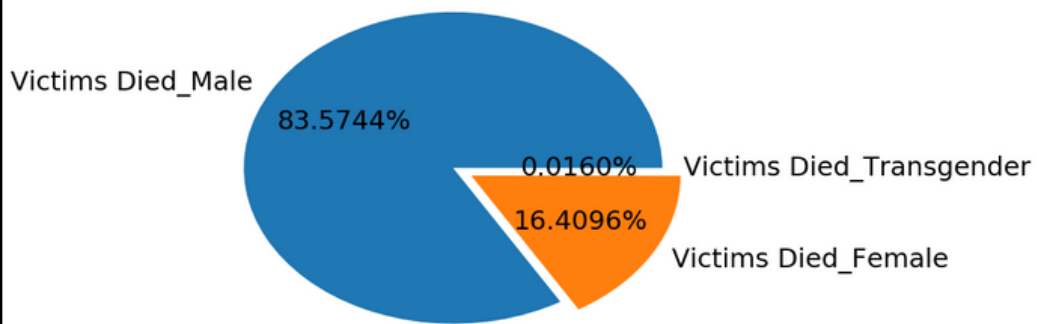
Total number of Accidents in each State.



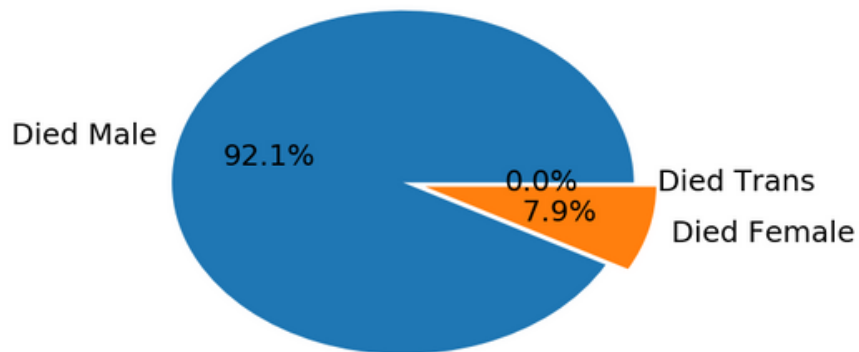
Min and Max number of accidents per lakh population in resp years



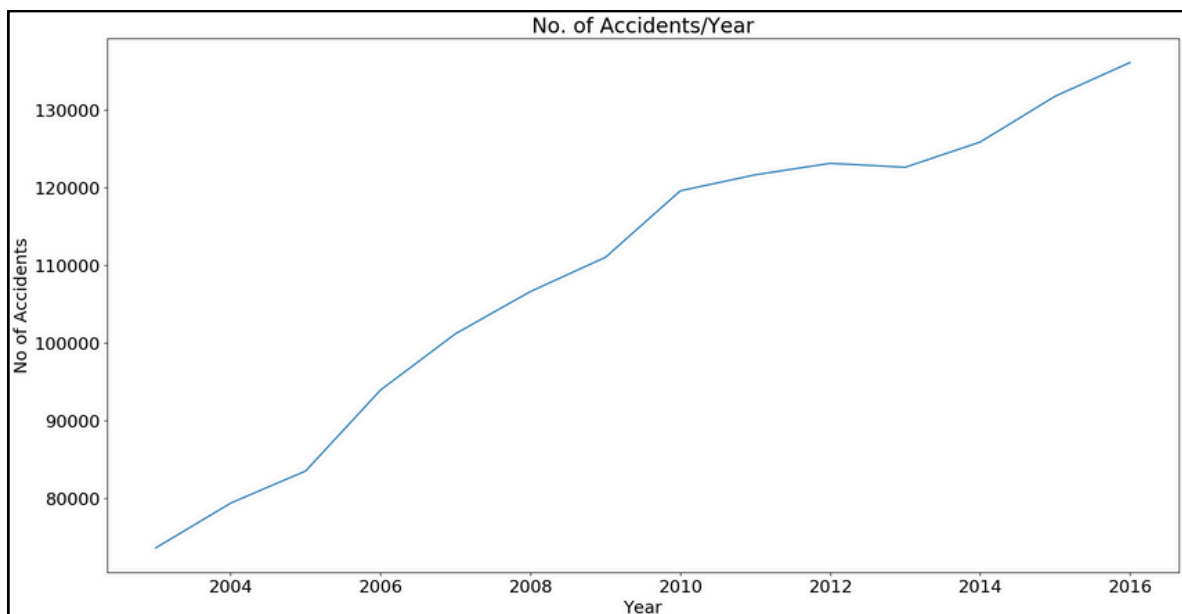
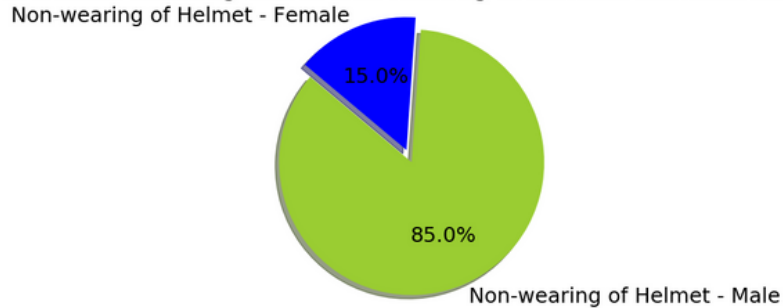
Percentage of Victims according to Gender who died.

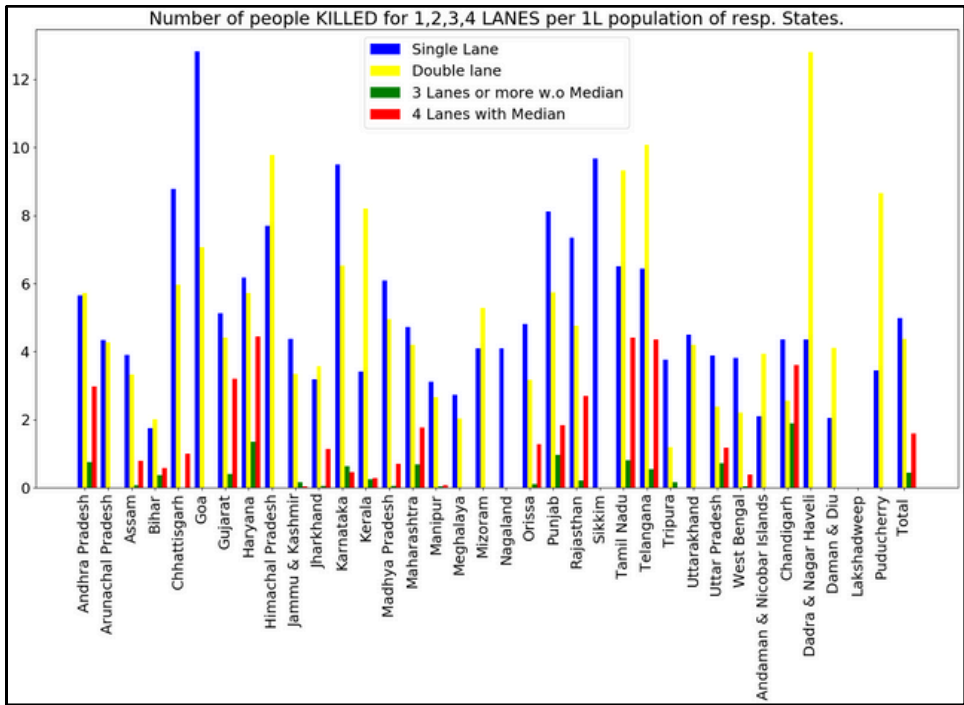
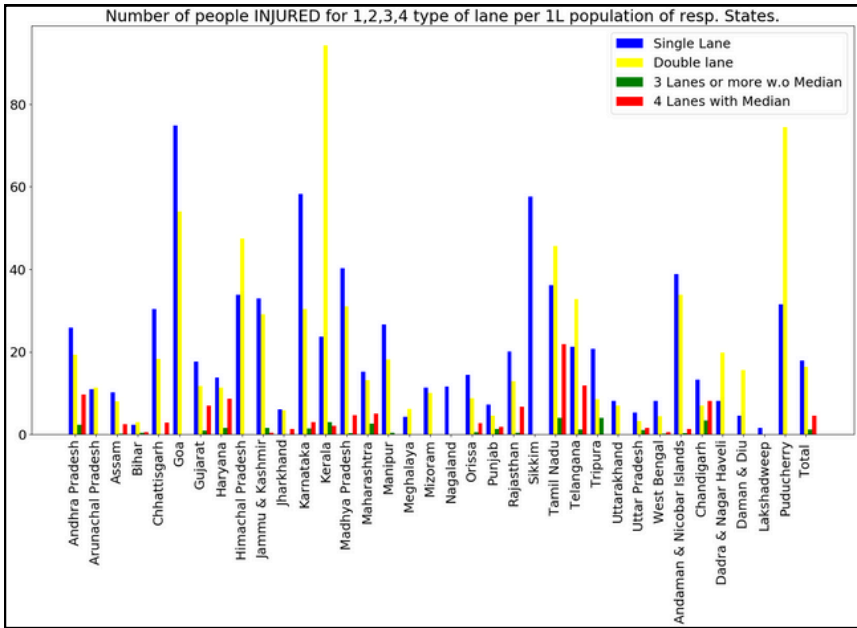
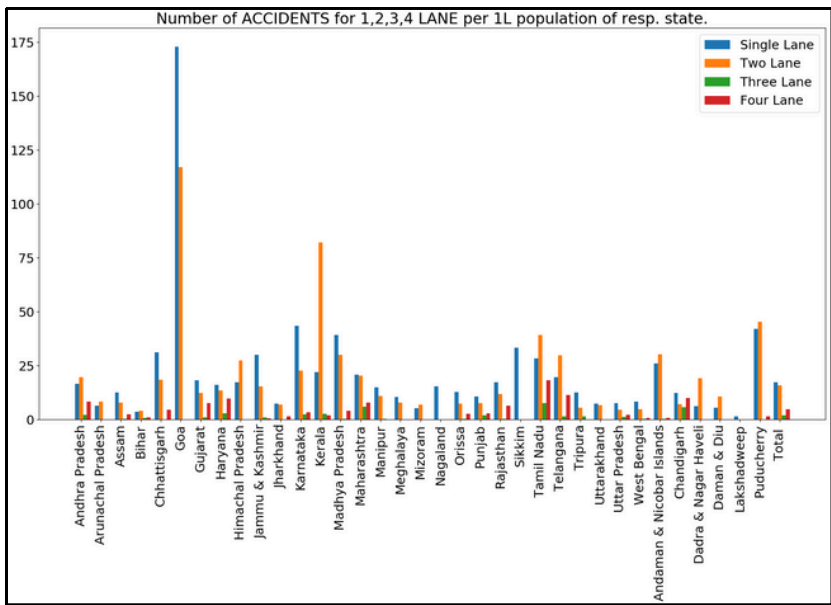


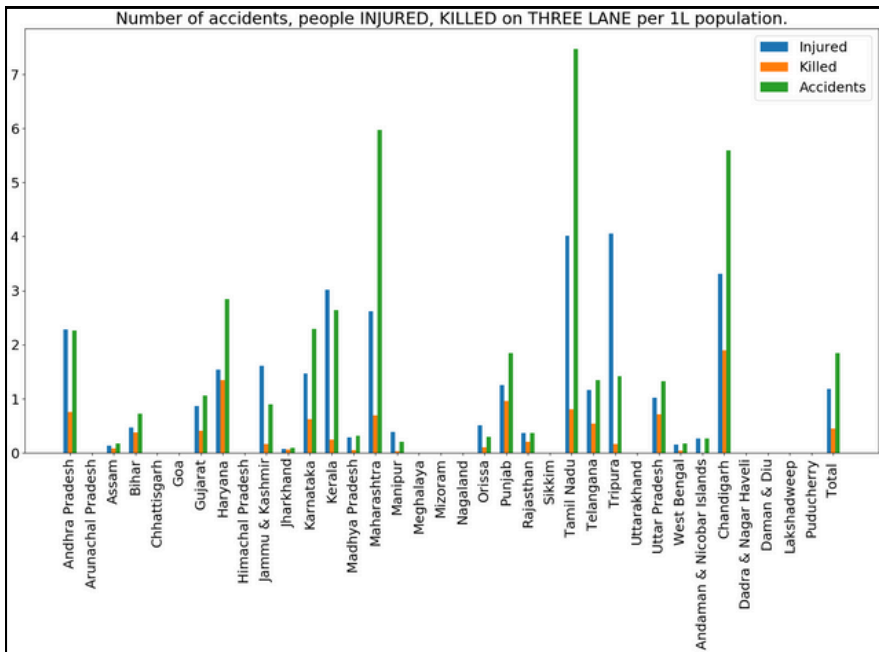
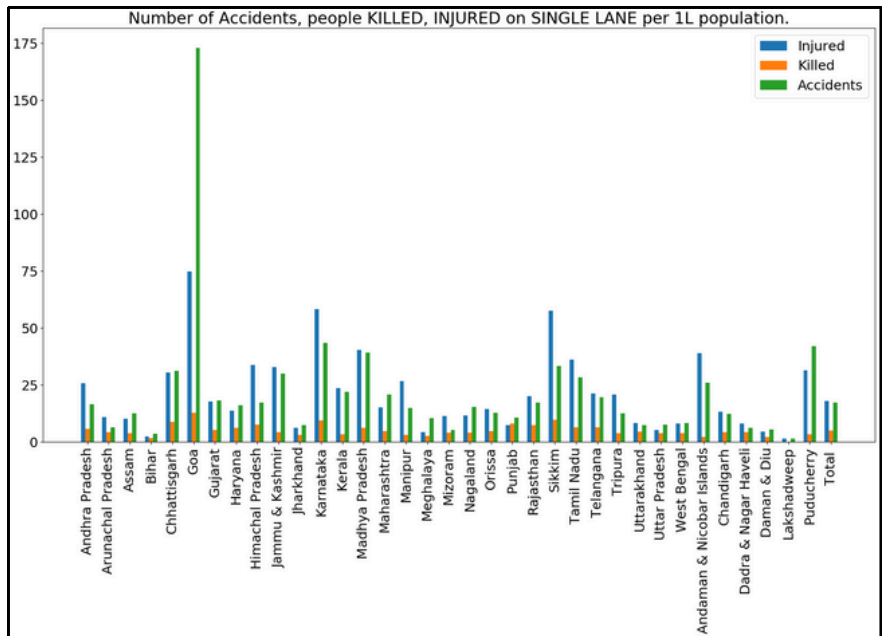
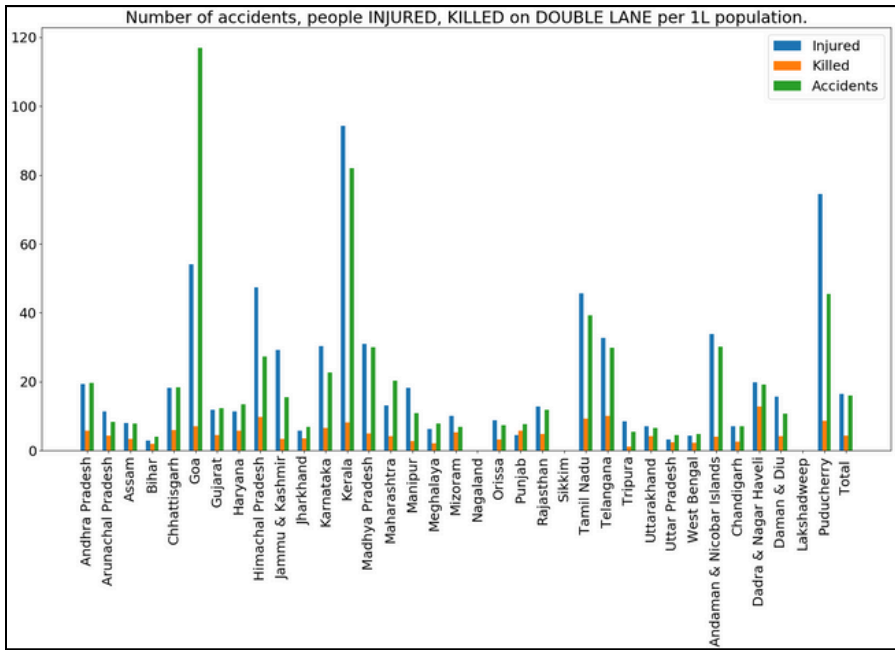
Percentage of offenders who died according to gender.



Percentage of Deaths occurring due to non-wearing of helmets between male and female.

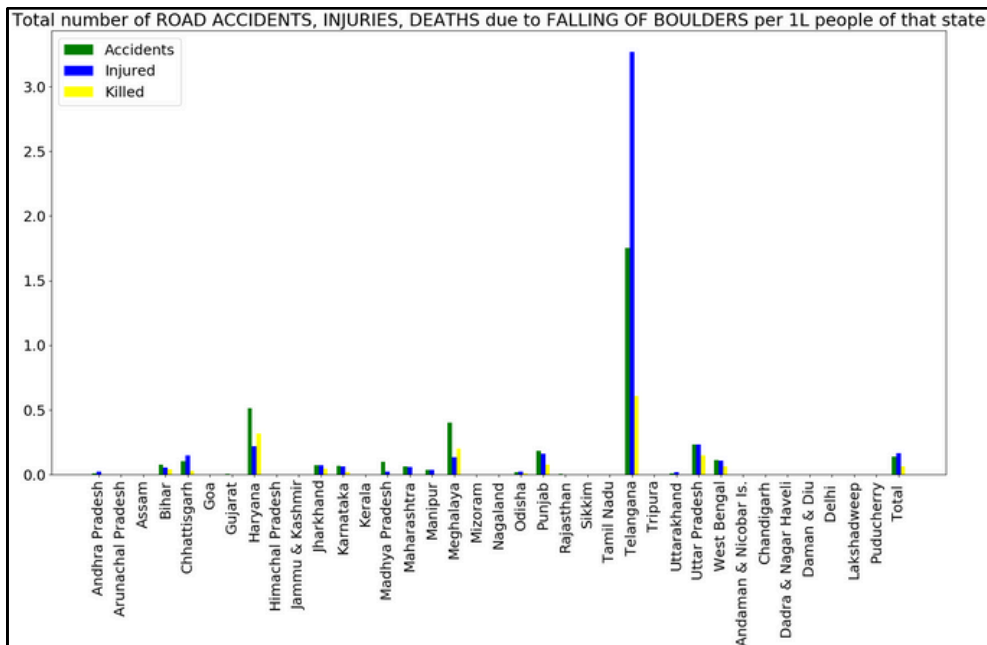
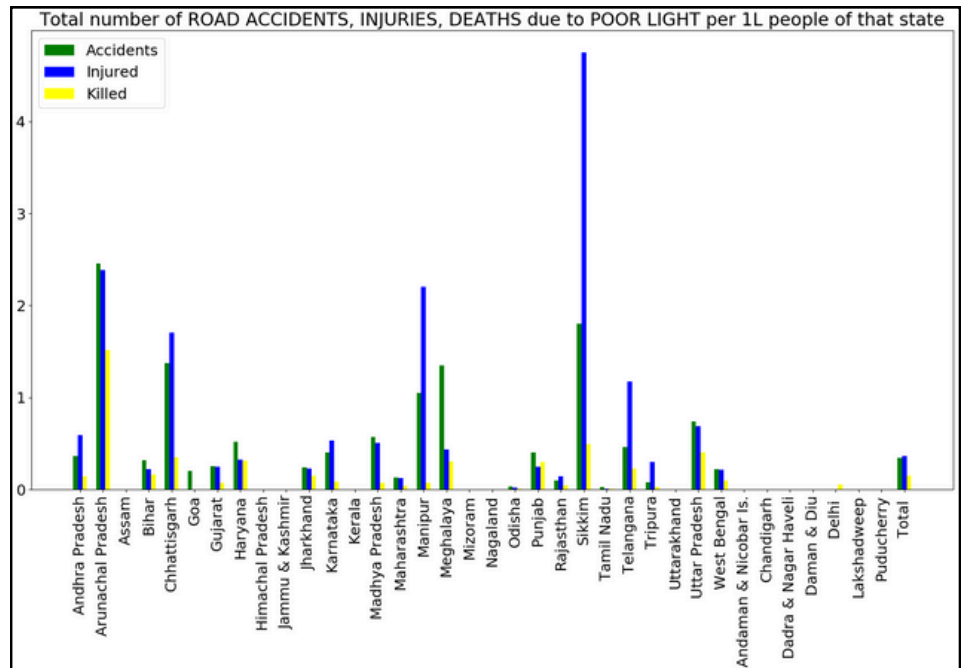
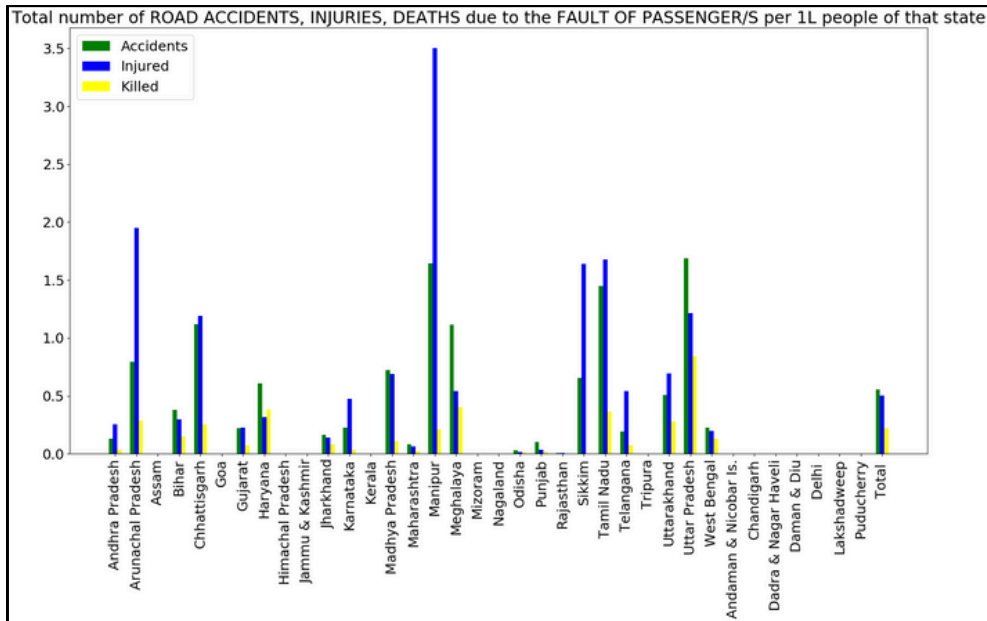




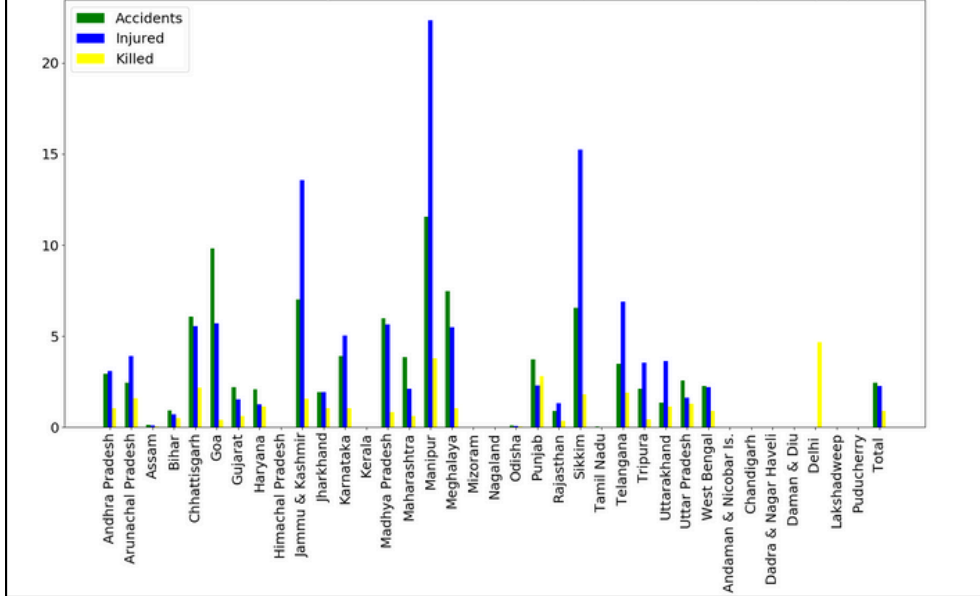




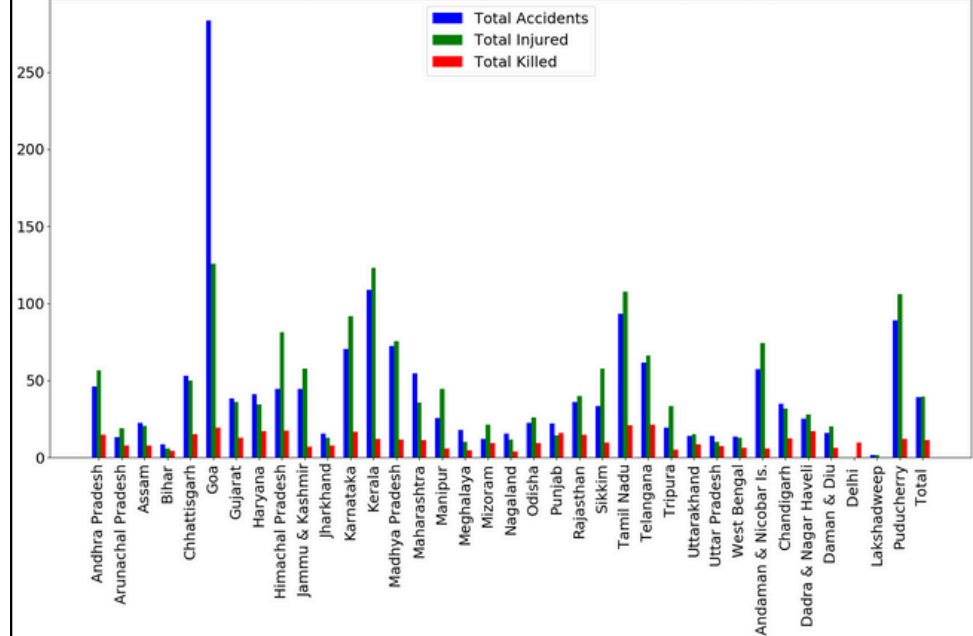




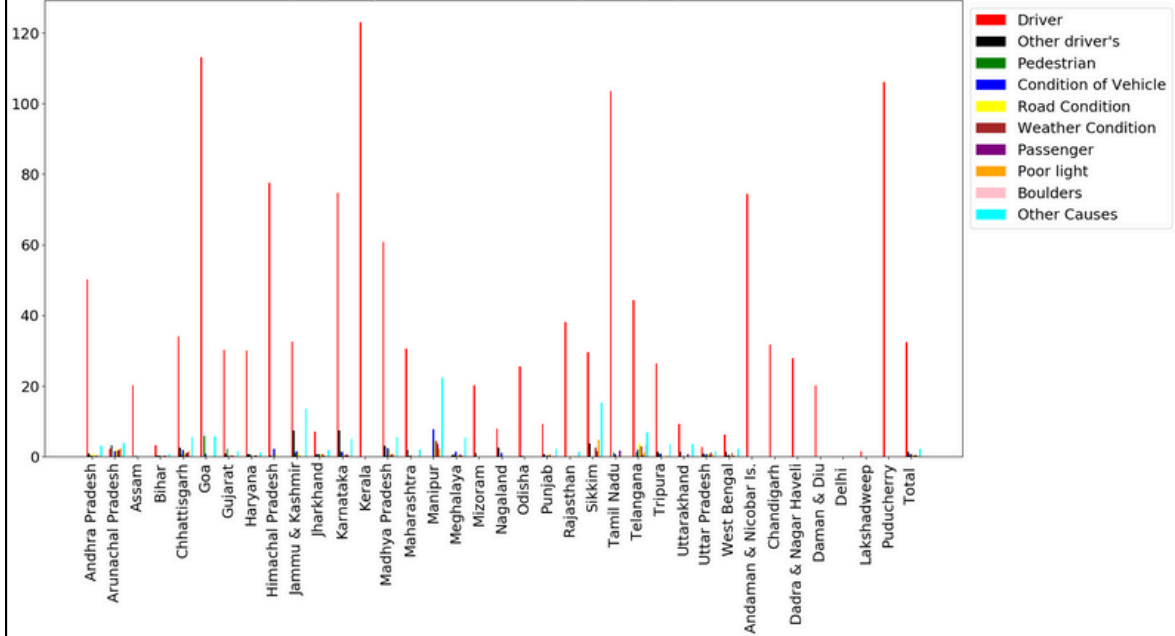
Total number of ROAD ACCIDENTS, INJURIES, DEATHS due to OTHER/UNKNOWN CAUSES per 1L people of that state

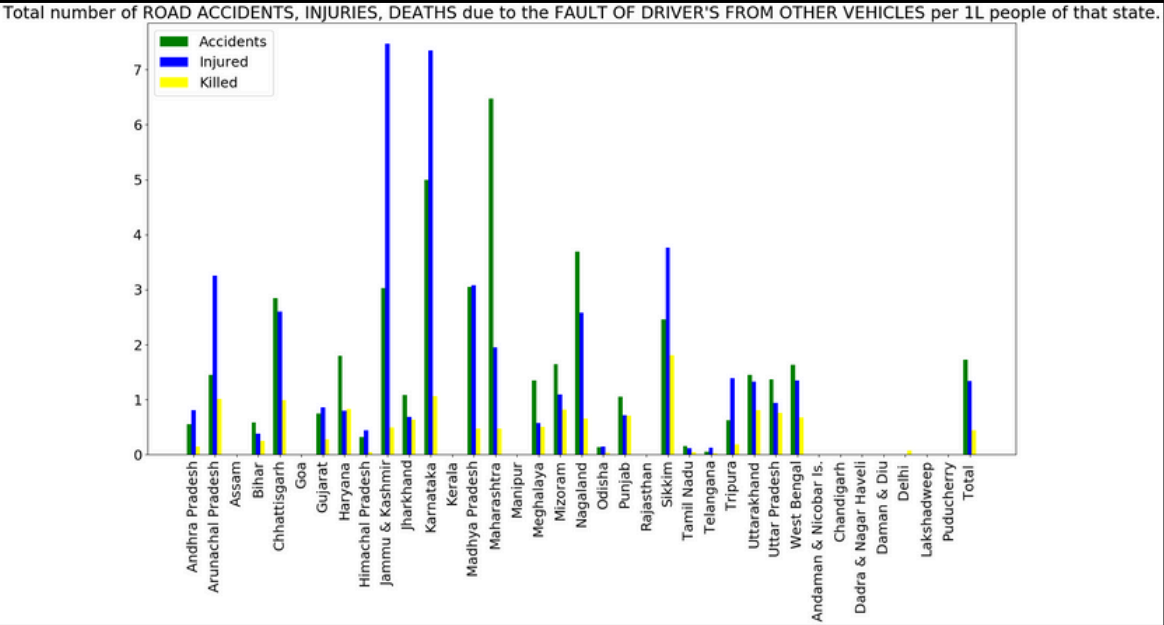
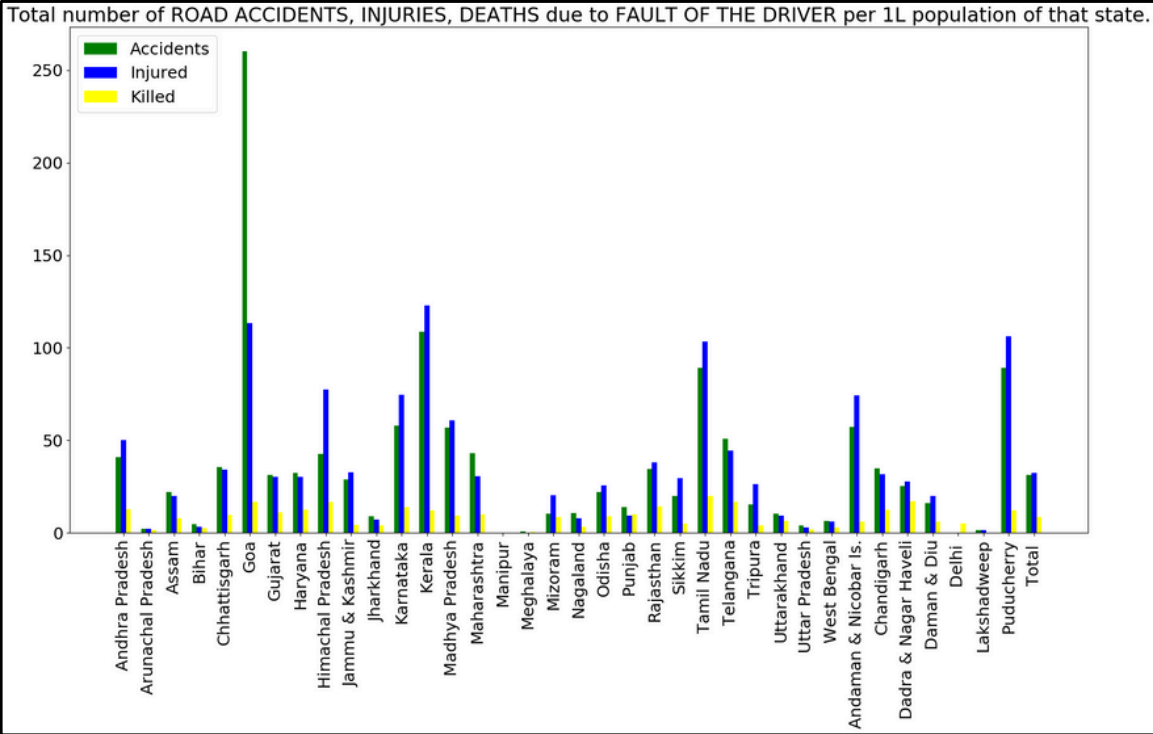
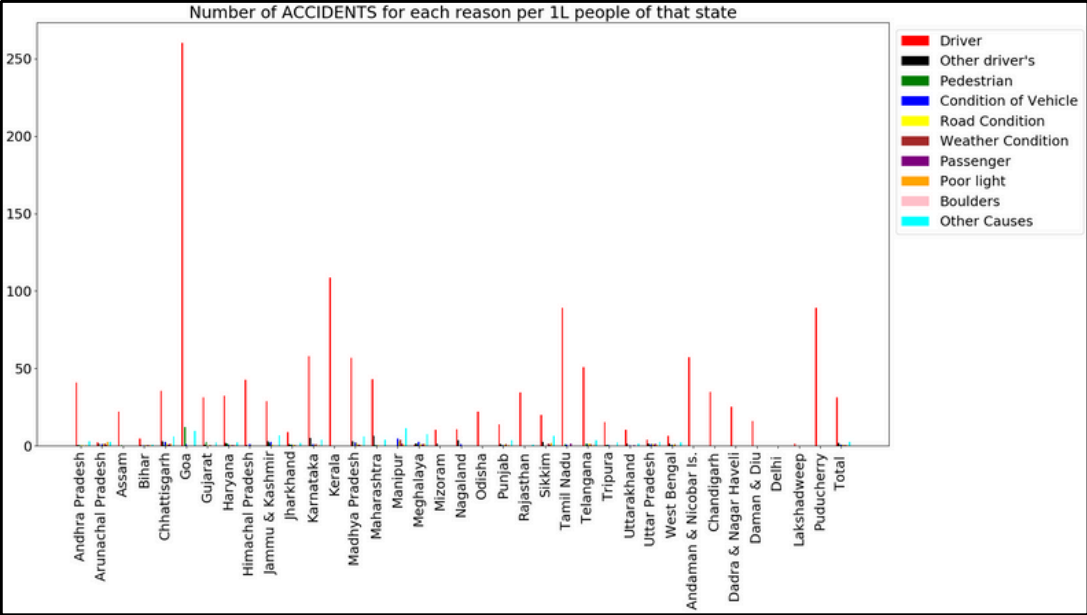


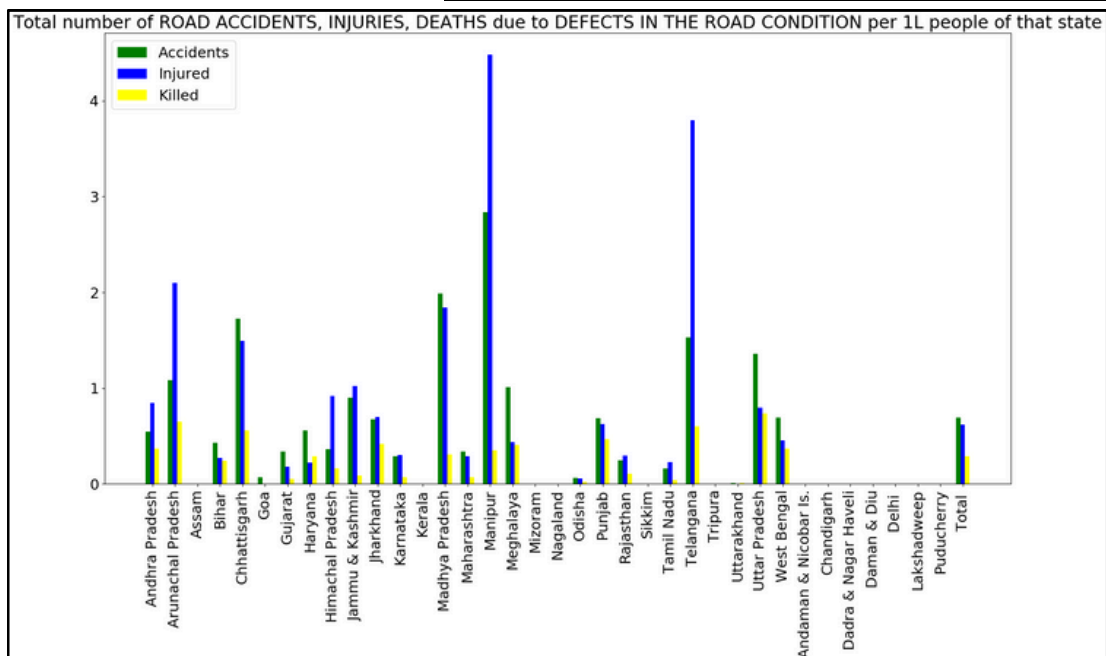
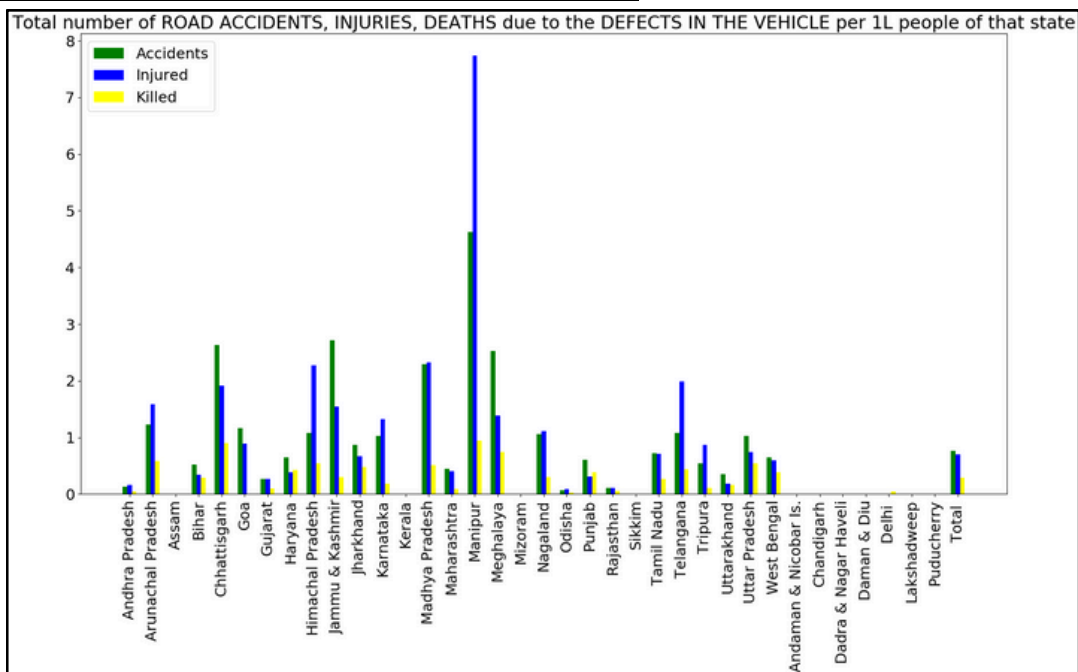
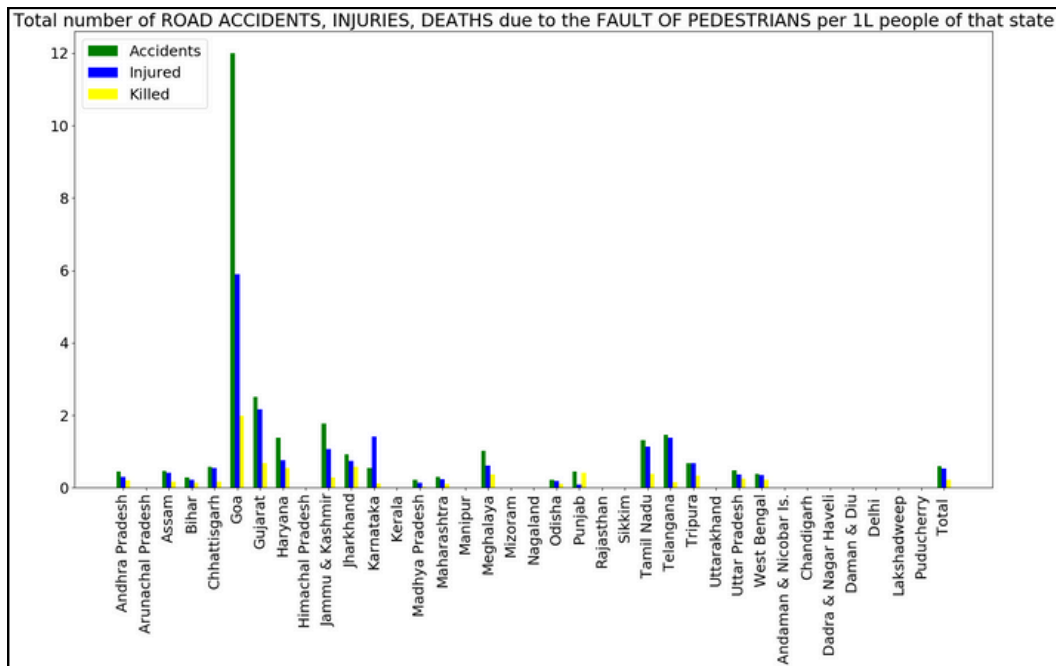
Total ROAD ACCIDENTS, INJURIES, DEATHS in general for each state per 1L population of that state

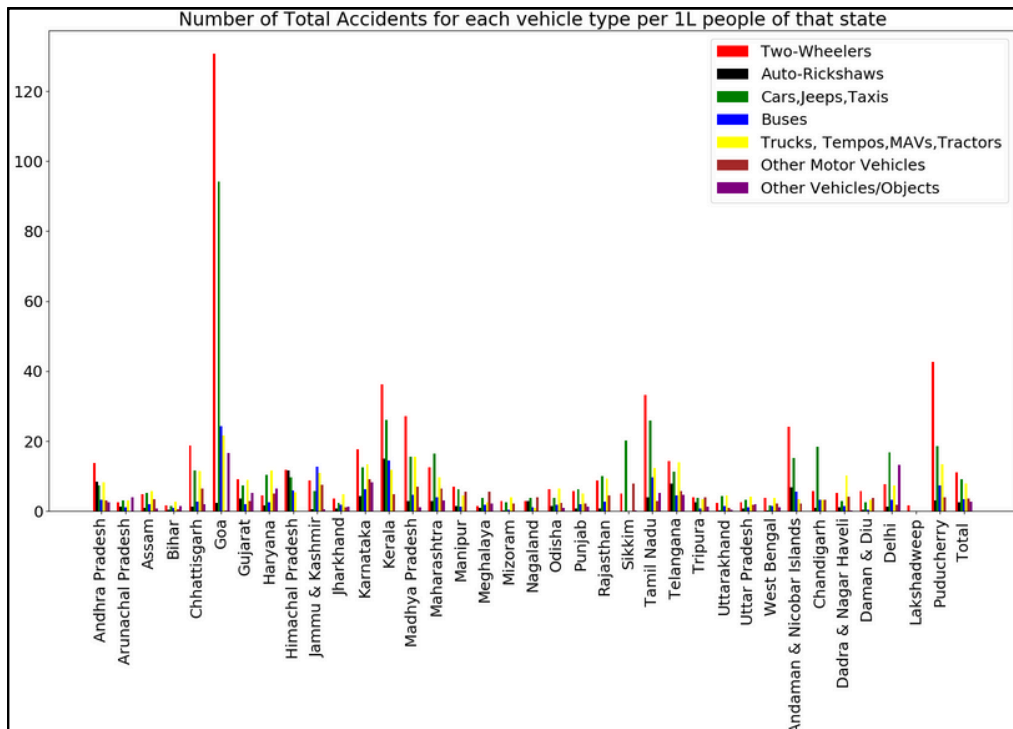
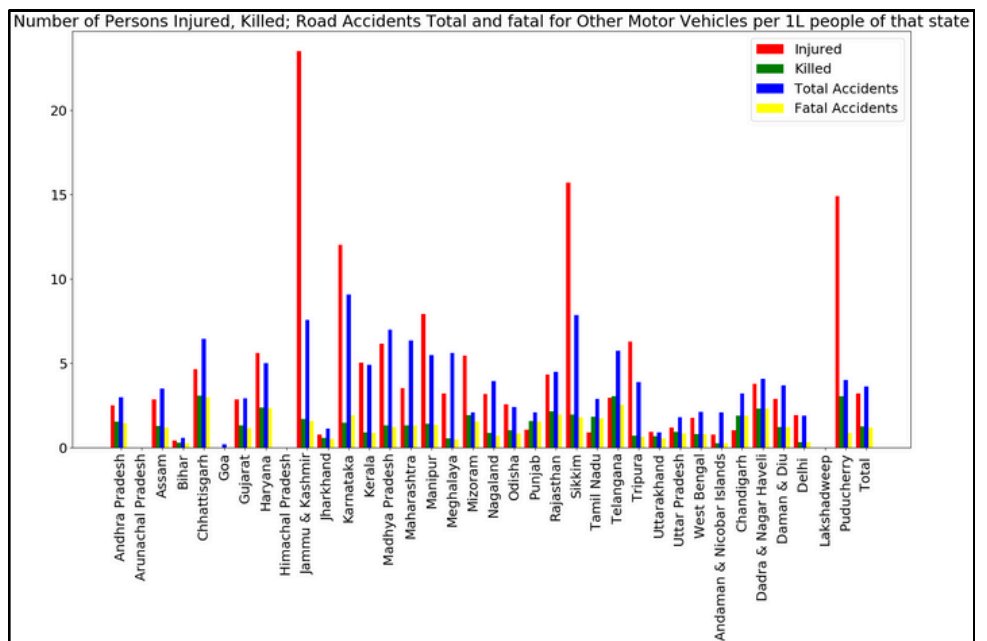
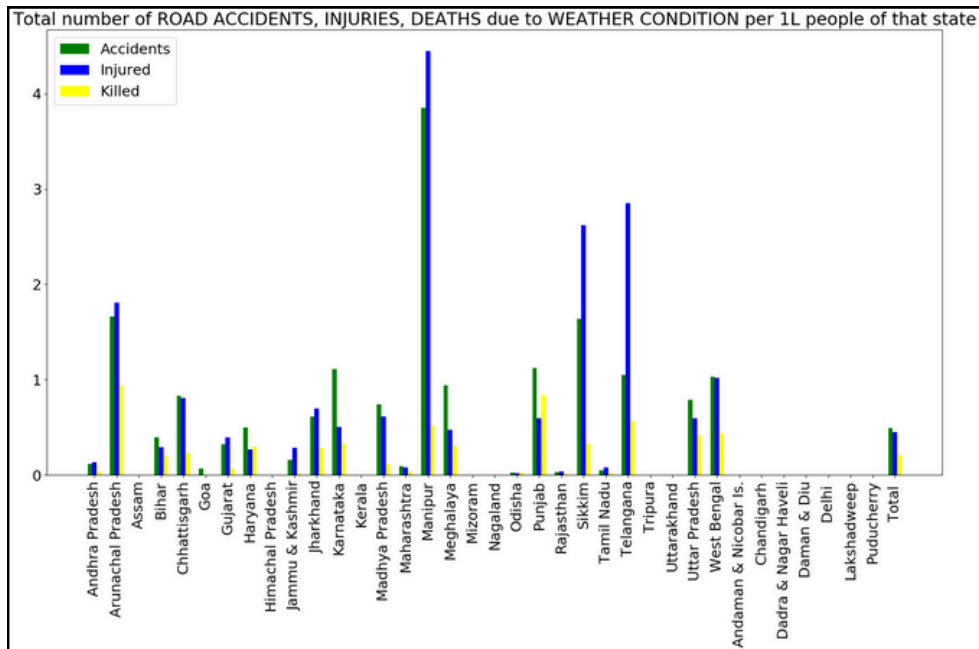


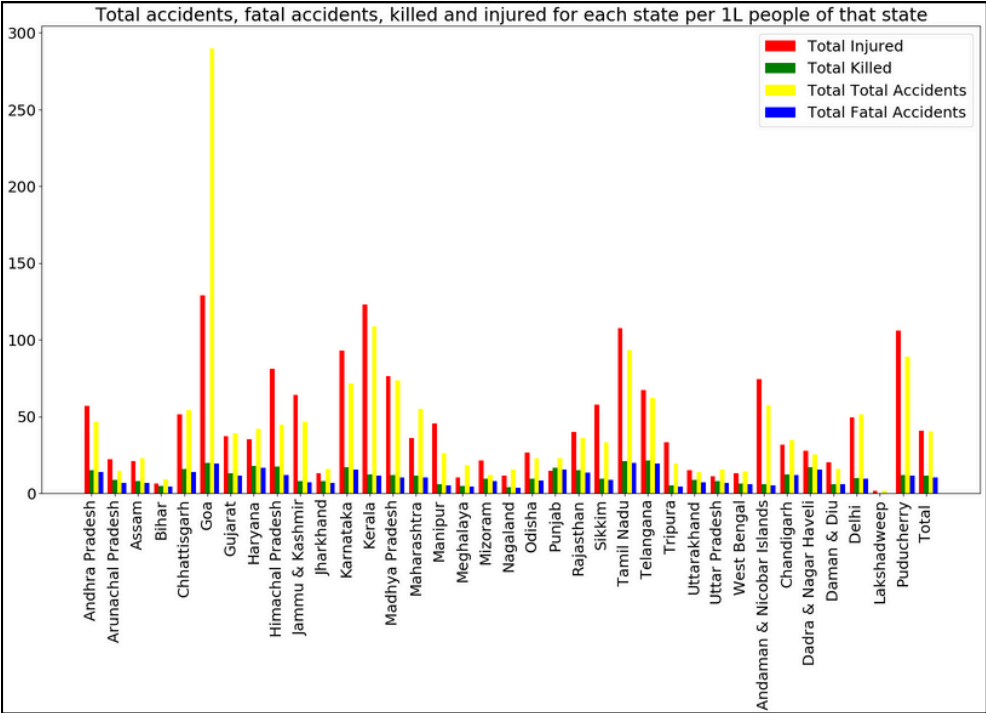
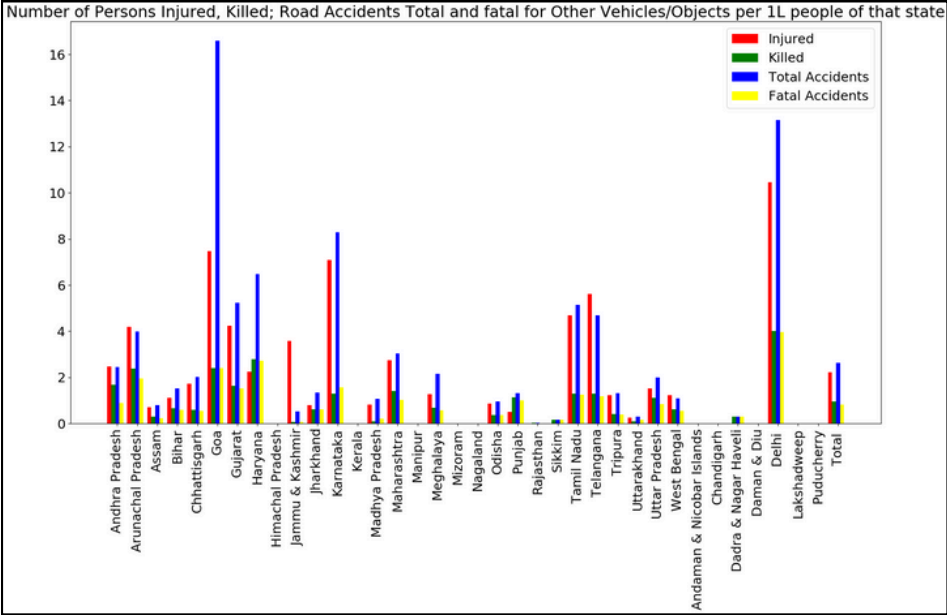
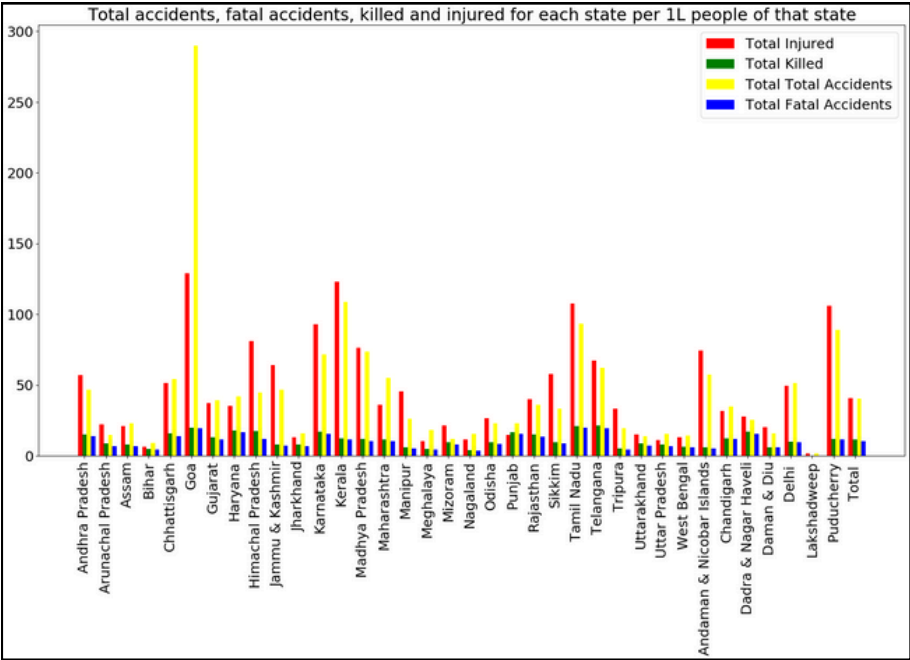
Number of people INJURED for each reason per 1L people of that state

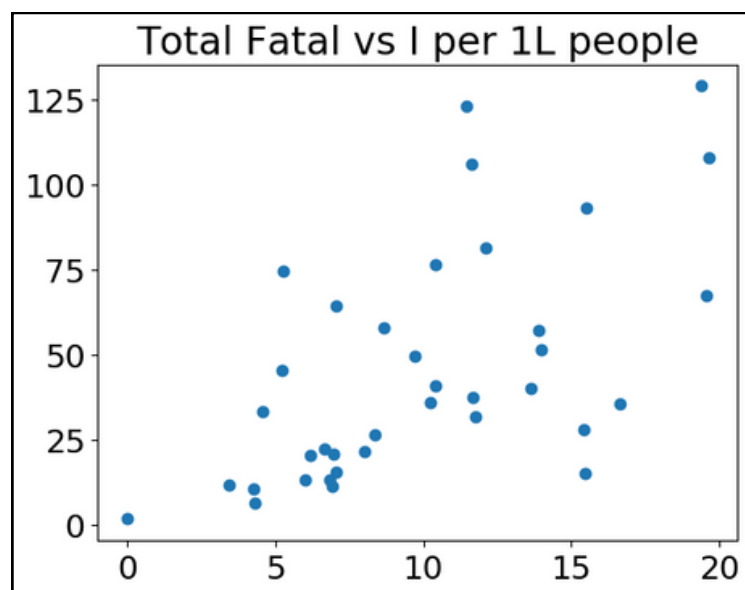
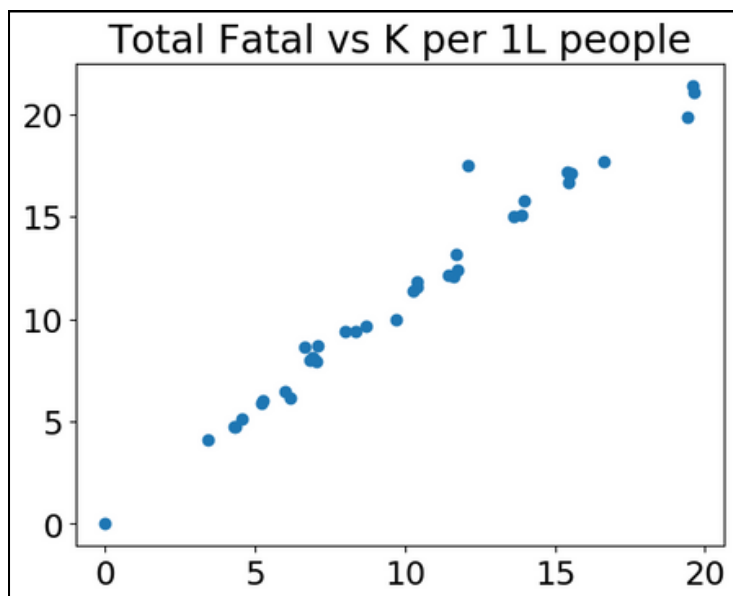
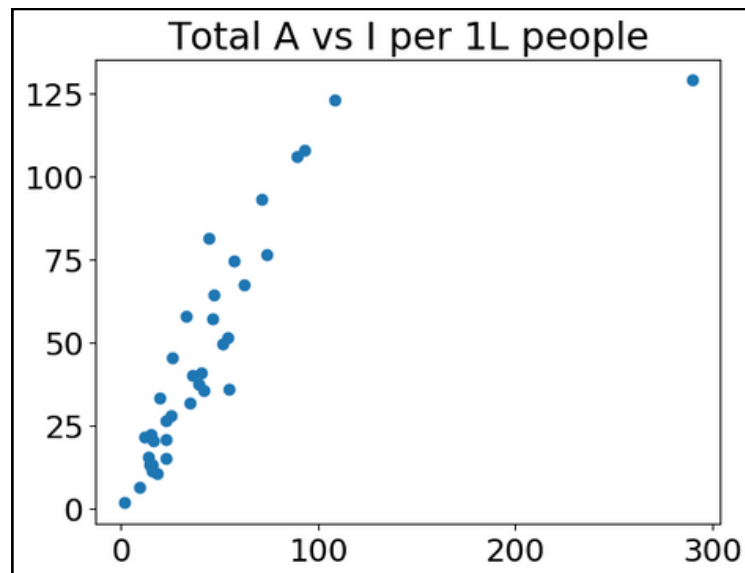
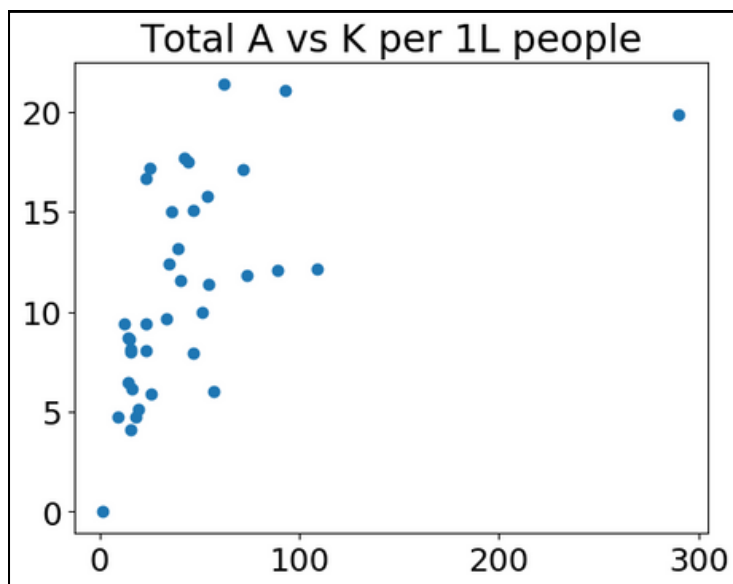


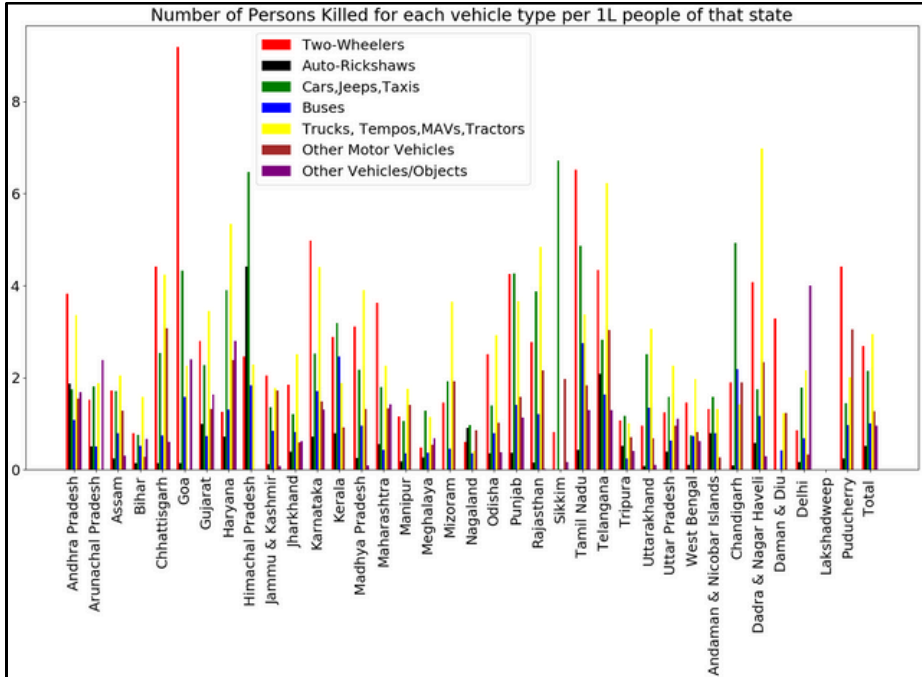
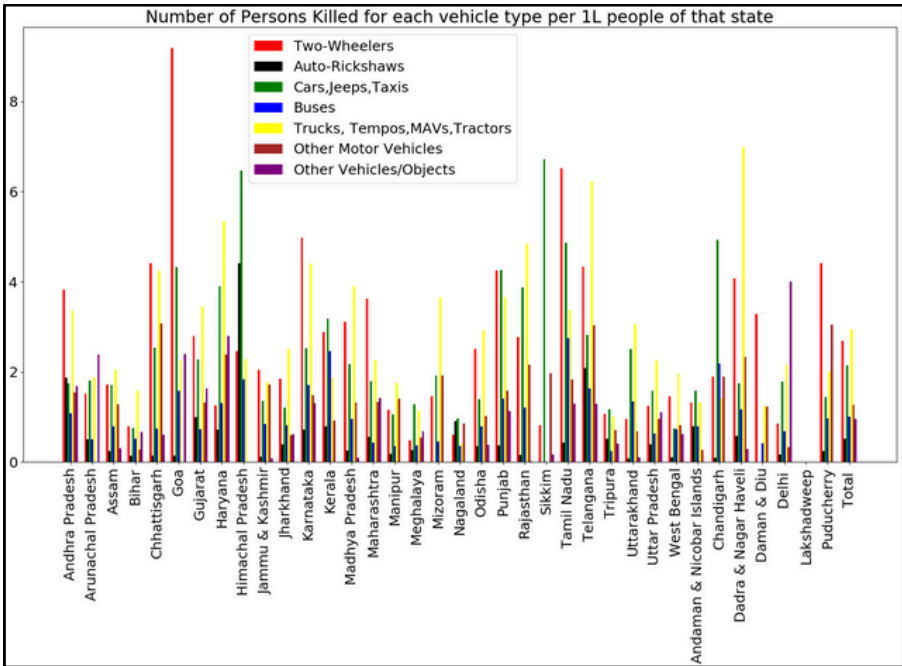
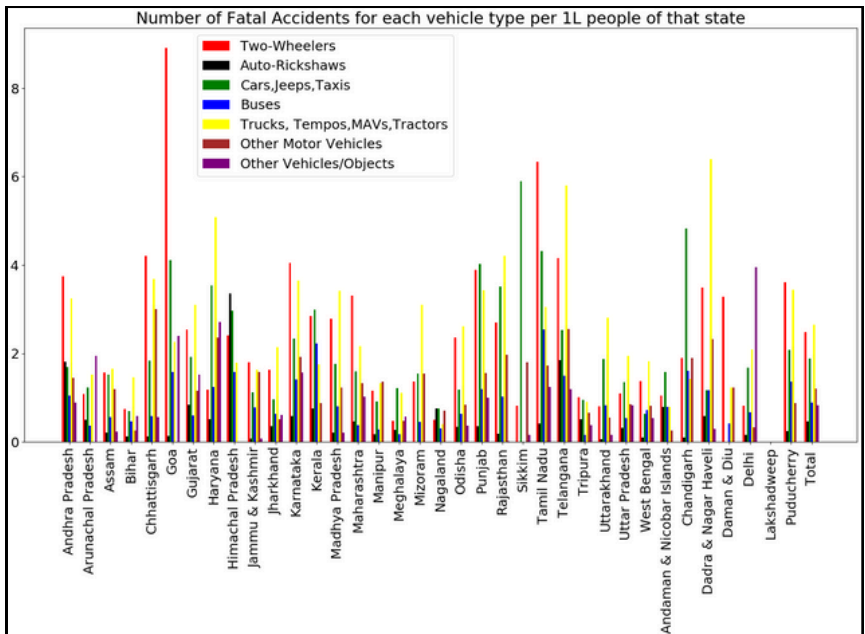






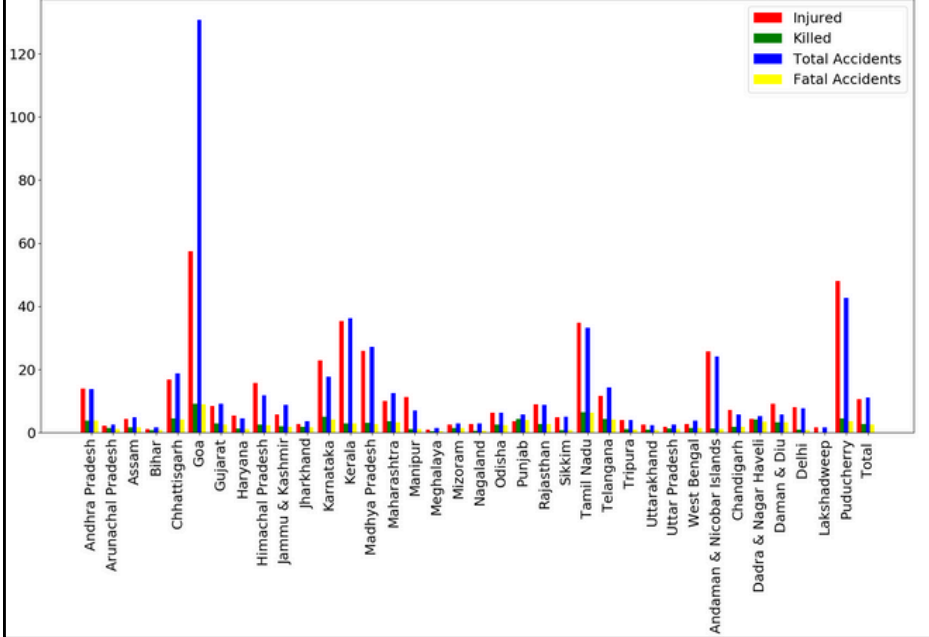




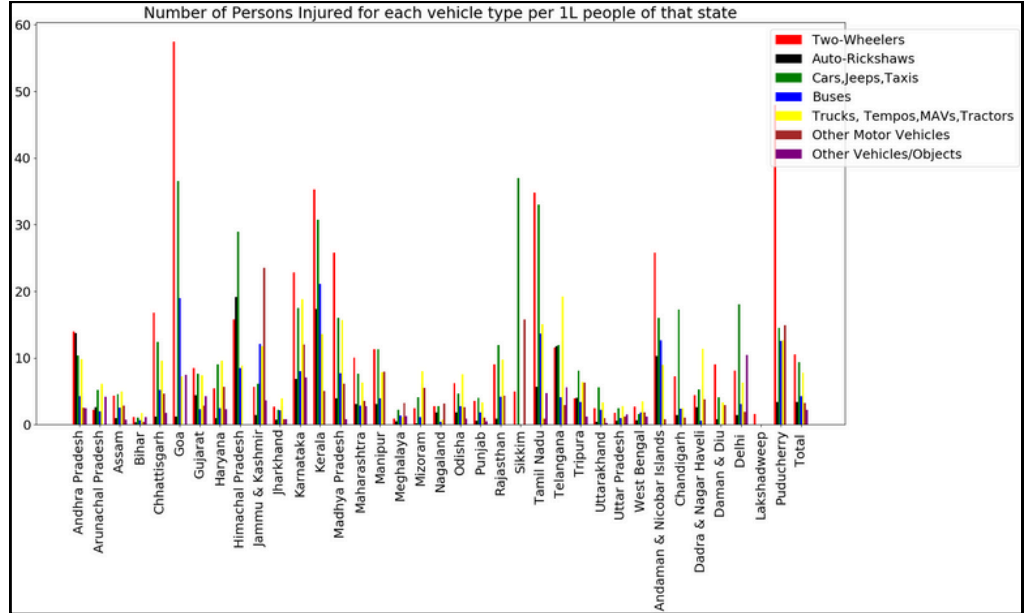




Number of Persons Injured, Killed; Road Accidents Total and fatal for 2 wheelers per 1L people of that state



Number of Persons Injured for each vehicle type per 1L people of that state



Number of Persons Injured, Killed; Road Accidents Total and fatal for Cars, Jeeps and Taxis per 1L people of that state

