

# Capstone Project Data Analyst: Road Accident Analysis

**Project description:** India has the highest number of road fatalities in the world.

In 2016, for which global figures are available, India accounted for more than a third of global road accident deaths. The World Health Organization says such deaths are under-reported and estimated that in 2016, the figure for India was likely twice as big as that reported by the government.

In this project, perform a detailed statistical data analysis on Road Accidents of India over a period of 2003 to 2016 years using various Python Libraries.

The relevant data supporting for the analysis is available as listed below,

DF - Analysing Accidents per Lakh Population-State-Year.

DF1 - Analysing Offender and Victim Deaths per Gender, State.

DF2 - Analysing Deaths occurred due to improper use of Safety Accessories.

DF3 - Analysing rate of accidents from the year 2003 to 2016 for each state.

DF4 - Analysing accidents/injuries/deaths occurring as per number of lanes.

DF5 - Analysing accidents/injuries/deaths occurring due to various faults/reasons.

DF6 - Analysing accidents/injuries/deaths as per the types of vehicles.

DF7 - Analysing number of accidents which take place as per time of occurrence.

and the exact names of the files for the above data description are,

df = roadAccStats13-16.csv

df1 = Details\_of\_road\_accident\_deaths\_by\_situation\_state\_2014.csv

df2 = Persons\_killed\_due\_to\_Non-use\_of\_Safety\_Device\_2016.csv

df3 = datafile.xls - total number of accidents from 2003 to 16 per state.

df4 = laneAccidents.csv

df5 = reasonOfAccident.csv

df6 = typeOfVehicle.csv

df7 = timeOfOccurence.csv

These data files are available in a Data Base folder, which will be shared with you.

[https://drive.google.com/drive/folders/1AUGLM2uPwT5EbF\\_2kBWcqitZ1fOaljD2?usp=sharing](https://drive.google.com/drive/folders/1AUGLM2uPwT5EbF_2kBWcqitZ1fOaljD2?usp=sharing)

With your analysis, you must answer to the below questions either in a statistical analysis or in a graphical representation using python code snippets.

- 1. The percentage of road accidents during all the years.**
- 2. Mean Accidents per 1L population for each year.**
- 3. The highest number of accident states and least number of accident states.**
- 4. Offenders and victims who died according to gender as well the as the total deaths.**
- 5. Percentage of Deaths occurring due to non-wearing of helmets between male and female.**
- 6. The number of accidents happening per state from the year 2003 to 2016.**
- 7. Number of ACCIDENTS for 1,2,3,4 LANE per 1L population of resp. state.**
- 8. Number of people INJURED for 1,2,3,4 type of lane per 1L population of resp. State.**
- 9. Number of people KILLED for 1,2,3,4 LANES per 1L population of resp. States.**
- 10. Number of Accidents, people KILLED, INJURED on SINGLE LANE per 1L population.**
- 11. Number of accidents, people INJURED, KILLED on DOUBLE LANE per 1L population.**
- 12. Number of accidents, people INJURED, KILLED on THREE LANE per 1L population.**
- 13. Number of accidents, people INJURED, KILLED on FOUR LANE per 1L population.**
- 14. Total Number of INJURED, KILLED, ROAD ACCIDENTS irrespective of lanes per 1L population of resp. State.**
- 15. Number of people KILLED for each different REASON per 1L population of that state.**
- 16. Number of people INJURED for each reason per 1L people of that state.**
- 17. Number of ACCIDENTS for each reason per 1L people of that state.**
- 18. Total number of ROAD ACCIDENTS, INJURIES, DEATHS due to FAULT OF THE DRIVER per 1L population of that state.**
- 19. Total number of ROAD ACCIDENTS, INJURIES, DEATHS due to the FAULT OF DRIVER'S FROM OTHER VEHICLES per 1L people of that state.**
- 20. Total number of ROAD ACCIDENTS, INJURIES, DEATHS due to the FAULT OF PEDESTRIANS per 1L people of that state**
- 21. Total number of ROAD ACCIDENTS, INJURIES, DEATHS due to the DEFECTS IN THE VEHICLE per 1L people of that state.**

- 22. Total number of ROAD ACCIDENTS, INJURIES, DEATHS due to DEFECTS IN THE ROAD CONDITION per 1L people of that state.**
- 23. Total number of ROAD ACCIDENTS, INJURIES, DEATHS due to WEATHER CONDITION per 1L people of that state.**
- 24. Number of Total Accidents for each vehicle type per 1L people of that state.**
- 25. Number of Persons Killed for each vehicle type per 1L people of that state.**
- 26. Total accidents, fatal accidents, killed and injured for each state per 1L people of that state.**
- 27. Number of Accidents happening in DAY and NIGHT TIME for 2014 and 2016.**

## **Evaluation Scheme:**

**Total marks: 100**

### **Deliverables [Total marks - 95]:**

1. Uploading all the data files into python environment – 5 marks
2. Creating the specific variables for each analysis – 5 marks
3. For completing the above stated 27 analyses, each carries 3 marks – 81 marks.
4. For writing clean and pep8 standard python syntaxes – 4 marks

### **Project Submission [Total marks - 5]:**

1. Once the project has been created, upload all the files on GitHub & commit (save) all the changes, make sure you add a readme file containing detailed description of your thoughts during the project creation. **[3 marks]**
2. Once done, kindly copy the GitHub link of your project & submit the same using your dashboard. **[2 mark]**