Regulatory Affairs of Road Accident Data – India 2020

Patterns, Risks, and Policy-Relevant Insights for Million-Plus Cities

Author: Dharana Singh
Date: August 13, 2025
Tableau Dashboard

Objective:

The objective of this project is to analyse road accidents across major Indian cities in the year 2020, with a focus on identifying patterns, causes, and severity levels. The analysis aims to:

- Quantify the scale of accidents in terms of total incidents, injuries, and fatalities.
- Calculate key safety metrics, including Fatality Rate, Injury Rate, and Severity Score.
- Examine the distribution of accidents by city type (e.g., metro, tourist, industrial, tier-2) and cause group (behavioural, infrastructure, environmental).
- Highlight high-risk areas and accident patterns to support policy recommendations and urban safety planning.

Dataset & Methodology:

Data Source:

- Road Accidents in India 2020 – Data.gov.in
- Road Accident Data 2020 – Kaggle

Data Cleaning & Transformation (ETL):

The raw dataset required significant restructuring and quality checks before it could be analysed effectively. The following steps were performed using Excel Power Query and Tableau:

1. Restructuring the Data

- Pivoted the outcome columns so that each subcategory appears only once, with associated counts for minor injuries, grievous injuries, persons killed, and total accidents.

2. Data Profiling & Validation

- Classified each row into one of four categories:
  i) Valid Data – All key metrics are non-zero and logically consistent.
  ii) No Impact – Accidents > 0 but injuries and deaths are zero.
  iii) Empty Rows – All values are zero.
  iv) Data Error – Accidents recorded as zero but injuries or deaths > 0.

3. Categorization for Analysis

- Grouped cities into Metro, Tourist, Industrial, and Tier-2 categories.
- Grouped accident causes into Behavioural, Infrastructure, and Environmental categories.

4. Calculated Fields in Tableau

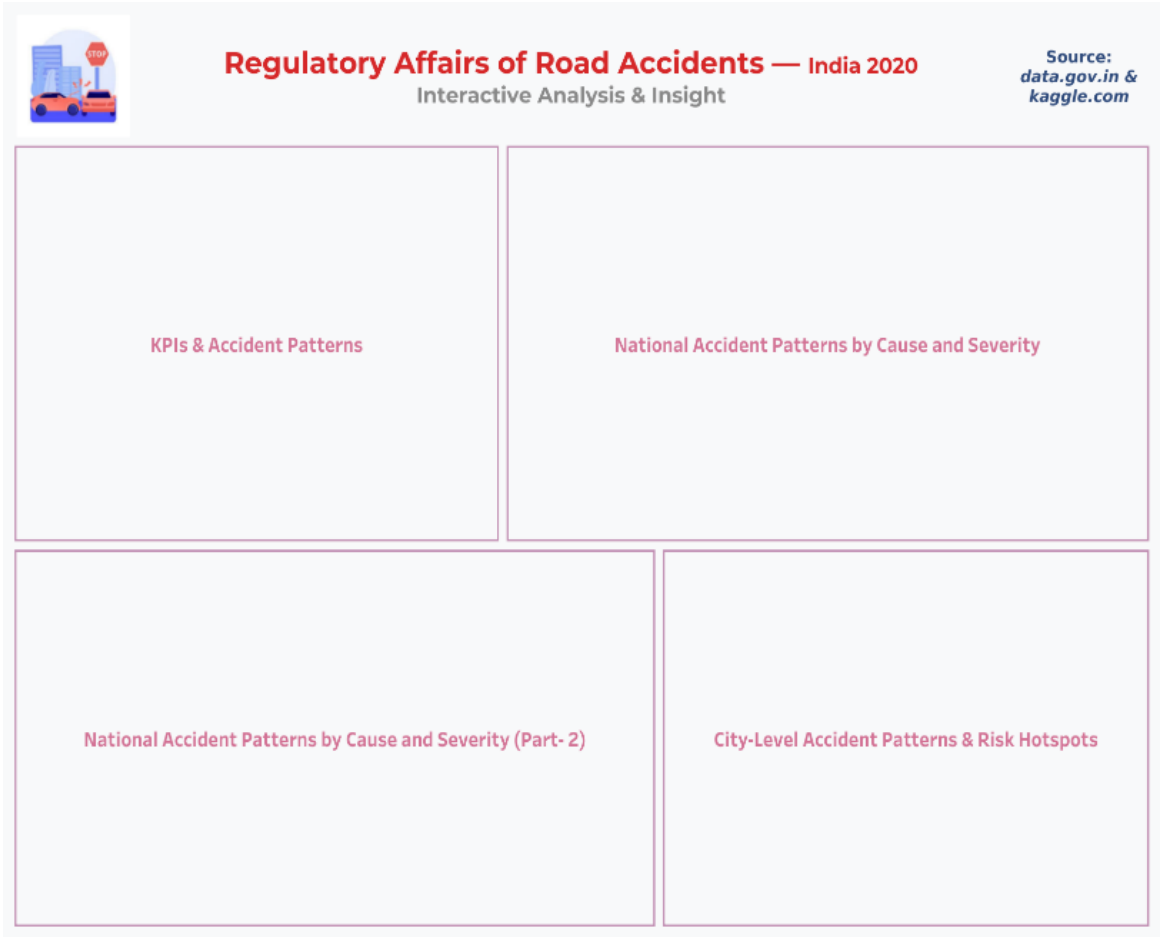
- Total Injured = Minor Injured + Grievously Injured
- Fatality rate (%) = (Persons Killed / Total accidents) * 100
- Injury Rate (%) = (Total Injured / Total accidents) * 100
- Severity Score (Raw) = (1 * Minor Injuries) + (3 * Grievous Injuries) + (5 * Deaths)

- Severity Rate (%) = (Severity Score / Total Accidents) * 100
- Weighted Minor Injuries = SUM (Minor Injury) * 1
- Weighted Grievous Injuries = SUM (Grievously Injured) * 3
- Weighted Fatalities = SUM (Persons Killed) * 5

Analysis & Visuals:

Navigation & Dashboard Structure

- To make the analysis interactive and easy to explore, a navigation dashboard was created as the entry point.
- It contains four buttons that link to the main dashboards:
  a) KPI & Accidents Patterns: National-level accident statistics
  b) National Accidents Patterns by Cause & Severity: Accident causes and distributions
  c) National Accidents Patterns by Cause & Severity – 2: Cause severity and injury patterns
  d) City Level Accidents Patterns & Risk Hotspots: City-specific risks and hotspots
- This ensures the end user can jump directly to the area of interest without scrolling through unrelated visuals.



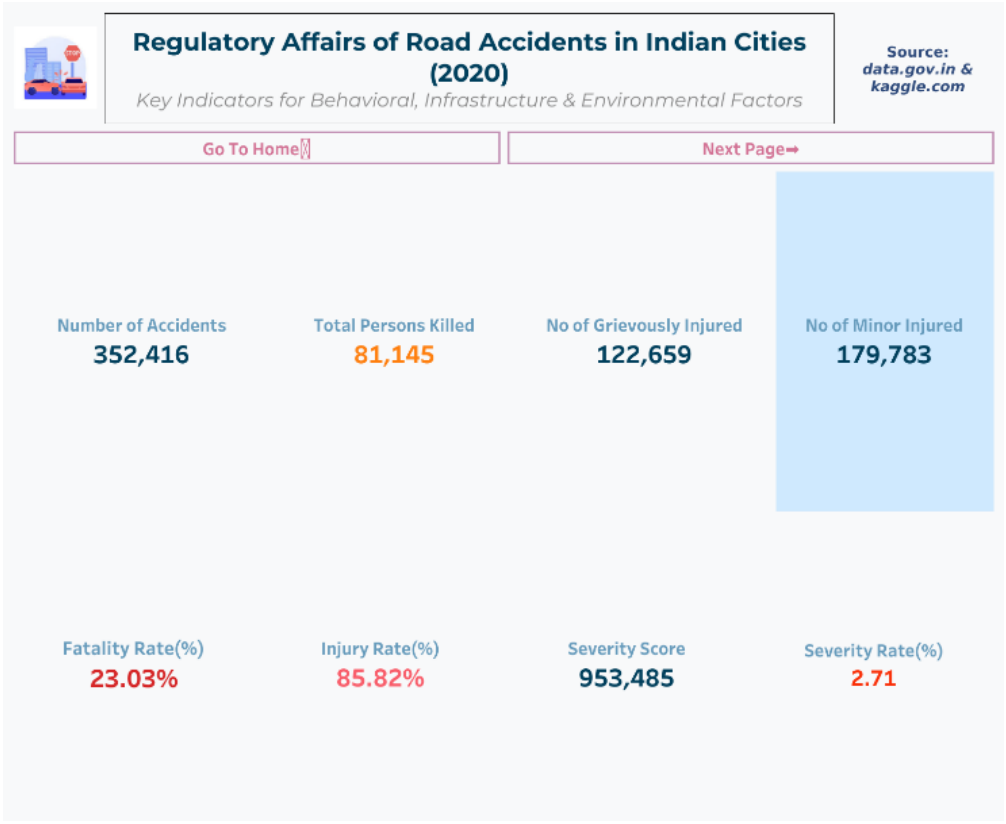
National Level Key Indicators:

The dashboard begins with a set of KPI cards that provide an at-a-glance view of the accident situation in India for 2020.

- Total Number of Accidents: 352,416 accidents occurred across various cities.
- Total Persons Killed: 81,145 fatalities were recorded in these accidents
- Grievously Injured: 122,659 people suffered serious, life-altering injuries.
- Minor Injured: 179,783 people sustained less severe injuries.
- Fatality rate: On average, 1 person is killed for every 23 accidents out of 100. This highlights the deadly nature of certain accidents.
- Injury Rate: In 85 out of every 100 accidents, at least one person sustained injuries (either grievous or minor).
- Severity Score: This score is a composite measure that reflects both the number and seriousness of accident outcomes. Different weights are assigned to each injury type to account for their impact:
  a) Minor injuries (weight: 1)
  b) Grievous injuries (weight: 3)
  c) Fatalities (weight: 5)

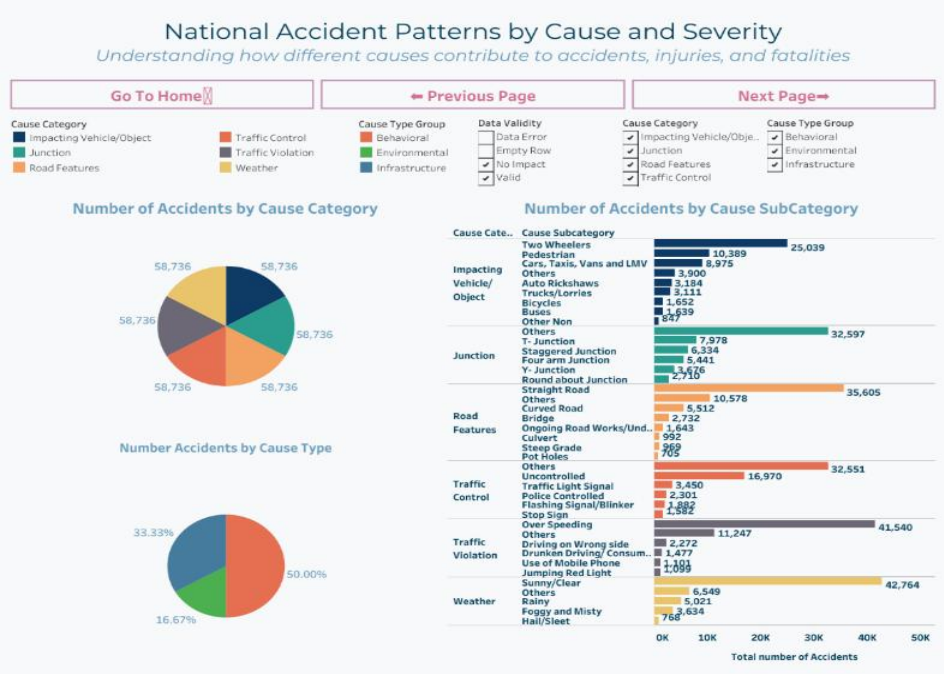
  By summing these weighted values for all accidents, the Severity Score provides a single figure that captures overall accident seriousness.

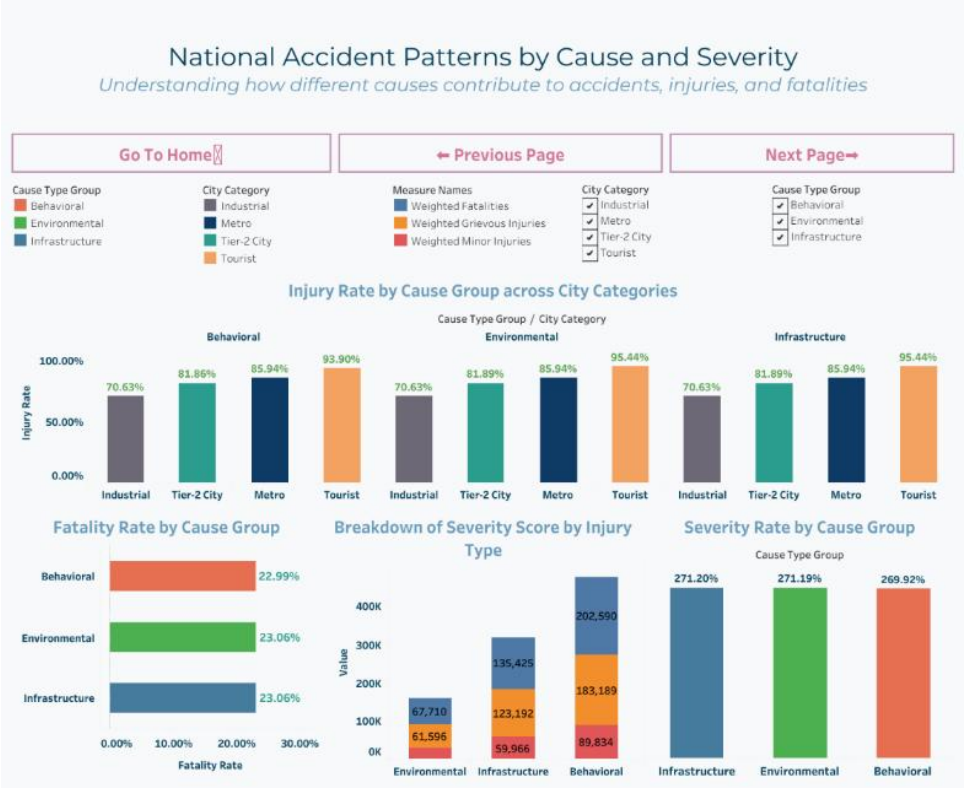
- Severity Rate: When converted into a rate, it shows the average weighted severity per 100 accidents. For 2020, the severity rate is 2.71, meaning that on average, every 100 accidents in Indian cities resulted in a combined injury–fatality severity equivalent to 2.71 weighted points.



National Accident Patterns by Cause and Severity

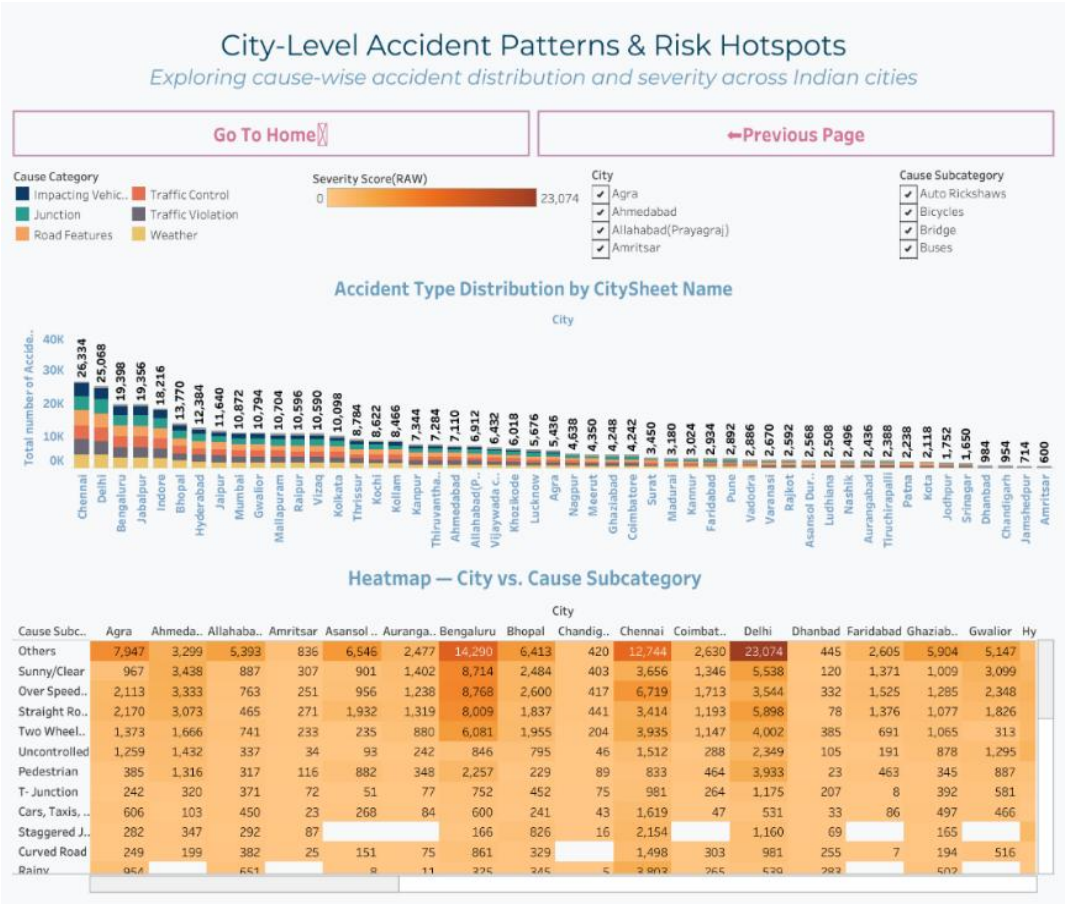
1. Number of accidents per Cause Category:
   - A pie chart compares the total accidents across six categories: Impacting Vehicle, Junction, Road Features, Traffic Control, Traffic Violation, and Weather.
   - Observation: Accidents are fairly evenly distributed across cause categories, with no single category exceeding ~17% share.
2. Number of Accidents by Cause Sub Category:
   - A bar chart breaks down each cause category into subcategories.
   - Observation: Sunny/Clear weather accounts for the highest accidents within Weather causes, while Over Speeding is the top contributor in Behavioural causes.
3. Number of Accidents by Cause Type (Environmental, Behavioural, Infrastructure):
   - A pie chart groups causes into three high-level types.
   - Observation: Behavioural factors (Traffic Control, Traffic Violation, Impacting Vehicles) account for ~50% of total accidents, making them the largest contributor.
4. Fatality Rate by Cause Group:
   - A bar chart shows the fatality rate for each cause group.
   - Observation: All groups have a fatality rate of ~23%, which matches the national average, suggesting that fatality likelihood is consistent across cause types.
5. Breakdown of severity score by Injury type:
   - A stacked column chart shows the weighted composition of severity scores (Minor, Grievous, Fatalities) for each cause group.
   - Observation: Fatalities and grievous injuries occur at nearly the same levels, while minor injuries are less frequent.
6. Severity Rate by Cause Group:
   - A bar chart displays severity rates per cause group, with values exceeding 100%.
   - Interpretation: A rate >100% means multiple injuries and/or fatalities occur per accident on average, indicating high severity.





City-Level Accident Patterns & Risk Hotspots

1) Injury Rate by Cause group as per City Category:
   - Chart Type: Grouped bar chart comparing injury rates for Behavioural, Infrastructure, and Environmental causes across four city types — *Industrial, Metro, Tourist, Tier 2*.
   - Observation: Injury rates range from ~70% to ~95%.
     - Industrial cities have the lowest injury rates (~70%).
     - Tourist cities report the highest injury rates (~94–95%), indicating a higher proportion of accidents resulting in injuries in tourism-heavy areas.
2) Number of accidents and its distribution across the cities:
   - Chart Type: Bar chart showing total accidents per city, with causes color-coded.
   - Observation: Chennai and Delhi record the highest number of accidents (~26k & ~25k respectively).
     - Within each city, accidents are evenly distributed across the six cause categories, indicating no dominant cause in any single city.
3) Heatmap- City Vs Cause Subcategory:
   - Chart Type: Heatmap showing the relationship between city, cause subcategory, and severity score.
   - Observation: The ‘Others’ subcategory skews severity scores upward, making it a potential data quality or classification gap to address.
     - Top cities with the highest severity scores include Delhi, Kanpur, Bengaluru, Jabalpur, and Chennai — highlighting key urban hotspots for severe accidents.



Key Insights:

1. Behavioural Causes Dominate Accident Share
   - Behavioural factors (e.g., traffic violations, poor traffic control) contribute to 50% of all recorded accidents, indicating the urgent need for stricter enforcement and road-user awareness programs. It is to be noted that Over Speeding alone accounts for a significant share within Behavioural causes.
2. Tourist Cities Face Higher Injury Risks
   - Tourist-heavy cities consistently show the highest injury rates across all three cause groups, suggesting that seasonal traffic surges and unfamiliar road users may be elevating accident risks.
3. Bengaluru Shows Elevated Risk in Specific Subcategories
   - Outside of the generic *"Others"* subcategory, Bengaluru records particularly high severity scores for Overspeeding and Sunny/Clear weather conditions — both of which point to preventable, human-driven risk factors.
4. Overspeeding & Sunny/Clear Weather Are Major Risk Drivers
   - These two subcategories alone account for the largest number of accidents nationally, underscoring the need for targeted enforcement and driver education even under seemingly safe conditions.

Policy Recommendations:

1) Stricter Enforcement on High-Risk Behaviours

- Introduce progressive penalties for overspeeding, with cumulative fines and licence point deductions for repeat offenders.
- Deploy speed cameras and automated enforcement systems in identified high-risk corridors.

2) Targeted Public Awareness Campaigns
   - Launch seasonal safety drives in tourist-heavy regions, focusing on safe driving practices, speed management, and defensive driving.
   - Use multilingual outreach to cater to both local drivers and visiting motorists.
3) Infrastructure Improvements in Accident Hotspots
   - Upgrade junction design, traffic light systems, and pedestrian crossings in high-tourist and high-traffic zones.
   - Implement smart traffic control technologies to reduce congestion and confusion at critical intersections.

Skills Demonstrated:

- **Data Cleaning & Transformation (Excel, Power Query)** – Standardized raw accident data, pivoted tables, and profiled records into valid/error/no-impact categories to ensure analytical accuracy.
- **Data Analysis (KPI Development & Aggregation)** – Designed and calculated key performance indicators such as Fatality Rate, Injury Rate, and Severity Rate to quantify accident impact.
- **Data Categorization & Grouping** – Created structured classifications for City Type and Cause Group to enable comparative analysis.
- **Data Visualization (Tableau)** – Built an interactive, multi-page dashboard with a navigation page, drill-down capabilities, and consistent thematic design.
- **Analytical Storytelling** – Translated complex accident statistics into a clear, policy-relevant narrative for decision-making.
- **Data Publishing & Portfolio Building** – Published analysis on Tableau Public and GitHub to demonstrate transparency, reproducibility, and professional presentation.

Executive Summary

This study analyses road accident patterns in India's million-plus cities during 2020, using official government datasets to uncover trends in accident causes, city categories, and severity levels. The analysis integrates calculated KPIs such as Fatality Rate, Injury Rate, and Severity Score to quantify accident impact.

Key findings reveal that **behavioural causes** (e.g., traffic violations, poor signal adherence) account for nearly half of all accidents, while **environmental conditions** exhibit the highest severity rates, indicating multiple injuries or fatalities per accident. Tourist cities record the highest injury rates, highlighting location-specific risks.

These insights provide a data-driven foundation for **policy interventions** in enforcement, public awareness, and infrastructure improvement, supporting safer urban mobility and targeted accident prevention strategies.