Road Accident Analysis of Addis Ababa Sub cities

Data Visualization and Exploration (DS6151)

Final Project

2024 G.C

Teshager Admasu

Natnael Chala

Rediet Tenaw

Contents

Abstract

This study explores patterns, trends, and factors contributing to road accidents in Addis Ababa based on a comprehensive analysis of accident data. Using exploratory data analysis (EDA) techniques, the study identifies critical factors influencing accident severity, including time of occurrence, road conditions, driver demographics, and collision types. The findings highlight significant patterns such as peak accident times, demographic risks, and the impact of environmental and behavioral factors, offering actionable insights for road safety improvements.

Introduction

Road accidents are a major concern globally, particularly in urban areas with high traffic volumes. Addis Ababa, as the capital city of Ethiopia, experiences a significant number of road accidents, leading to fatalities, injuries, and property damage. Understanding the underlying factors contributing to these accidents is essential for implementing effective safety measures. This study utilizes accident data to identify trends, risk factors, and potential interventions to reduce accident severity.

Methodology

This research is based on a detailed analysis of accident data from Addis Ababa. The methodology includes:

1. **Data Collection:** A dataset containing records of road accidents in Addis Ababa was used.
2. **Exploratory Data Analysis (EDA):** Various techniques, including univariate, bivariate, and multivariate analysis, were employed to examine the relationships between accident severity and contributing factors.
3. **Data Preprocessing:** The data was cleaned by handling missing values, encoding categorical variables, and removing duplicates. Outliers were addressed, and relevant transformations were applied.
4. **Visualization:** Graphs and charts were used to highlight key trends and patterns in the data.

Steps

Step 1 | Import Libraries

Step 2 | Loading the Dataset

Step 3 | EDA

- Step 3.1 | Dataset basic Information
- Step 3.2 | Missing Values
- Step 3.3 | Duplicated Values
- Step 3.4 | Univariate Analysis
- Step 3.5 | Bivariate Analysis
- Step 3.6 | Multivariate Analysis

Step 4 | Data Preprocessing

- Step 4.1 | Features Encoding
- Step 4.2 | Missing Value Handling
- Step 4.3 | Data Transformation
- Step 4.4 | Data Reduction
- Step 4.5 | Outlier Treatment
- Step 4.5 | Insights and Recommendations

This report is part of a GitHub repository containing: the dataset, a notebook and an interactive tableau dashboard.

Data Description

The data set is collected from Addis Ababa Sub city police departments for Master's research work. The data set has been prepared from manual records of road traffic accident of the year **2017-20.** All the sensitive information has been excluded during data encoding and finally, it has **32** features and **12316** instances of the accident

Here is the definition of the features listed, based on their names and example values:

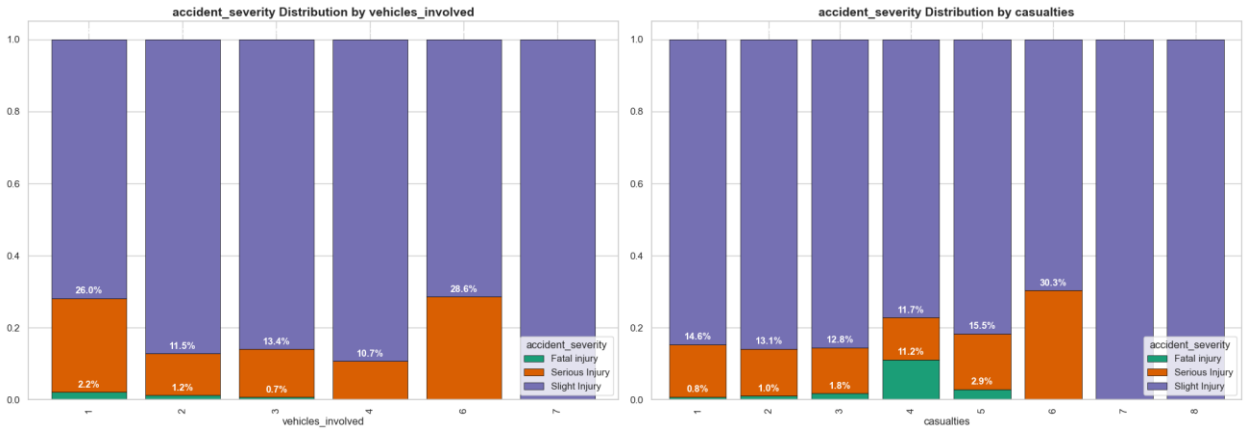
1. **time**: The time of the incident, recorded in HH:MM:SS format.
2. **day_of_week**: The day of the week on which the event occurred.
3. **driver_age**: The age group of the driver, categorized into ranges such as '18-30', '31-50', etc.
4. **driver_sex**: The gender of the driver ('Male', 'Female', or 'Unknown').
5. **educational_level**: The highest educational qualification of the driver.
6. **vehicle_driver_relation**: The relationship between the vehicle owner and the driver (e.g., 'Owner', 'Employee', 'Other').
7. **driving_experience**: The driver's experience level, categorized by years or specific conditions like 'No Licence'.
8. **vehicle_type**: The type of vehicle involved in the incident (e.g., 'Automobile', 'Taxi', 'Motorcycle').
9. **vehicle_owner**: The owner of the vehicle, such as 'Owner', 'Governmental', or 'Organization'.
10. **service_year**: The service duration of the vehicle, categorized by years.
11. **vehicle_defect**: Whether the vehicle had a defect at the time of the accident (e.g., 'No defect', '7', '5').
12. **accident_area**: The type of area where the accident occurred (e.g., 'Residential areas', 'Office areas').
13. **lanes**: The type of road lanes where the accident occurred (e.g., 'Undivided Two way', 'One way').
14. **road_allignment**: The road's alignment, describing its geometry (e.g., 'Tangent road with flat terrain').
15. **junction_type**: The type of road junction at the accident location (e.g., 'T Shape', 'Crossing').
16. **surface_type**: The type of road surface (e.g., 'Asphalt roads', 'Gravel roads').
17. **road_surface_conditions**: The condition of the road surface during the accident (e.g., 'Dry', 'Wet or damp').
18. **light_condition**: The lighting condition at the time of the accident (e.g., 'Daylight', 'Darkness - no lighting').

19. **weather_condition**: The weather condition at the time of the accident (e.g., 'Normal', 'Raining').
20. **collision_type**: The type of collision that occurred (e.g., 'Vehicle with vehicle collision', 'Collision with pedestrians').
21. **vehicles_involved**: The number of vehicles involved in the accident.
22. **casualties**: The number of individuals injured or killed in the accident.
23. **vehicle_movement**: The movement status of the vehicle at the time of the accident (e.g., 'Going straight', 'Parked').
24. **casualty_class**: The role of the casualty in the accident (e.g., 'Driver or rider', 'Pedestrian').
25. **casualty_sex**: The gender of the casualty.
26. **casualty_age**: The age or age group of the casualty.
27. **casualty_severity**: The severity of the casualty's injury ('1', '2', '3', or 'na').
28. **casualty_work**: The employment status or occupation of the casualty.
29. **casualty_fitness**: The fitness or health condition of the casualty (e.g., 'Normal', 'Deaf').
30. **pedestrian_movement**: The movement or behavior of the pedestrian at the time of the accident.
31. **accident_cause**: The primary cause of the accident (e.g., 'Overspeed', 'Drunk driving').
32. **accident_severity**: The severity of the accident's outcome ('Slight Injury', 'Serious Injury', 'Fatal injury').

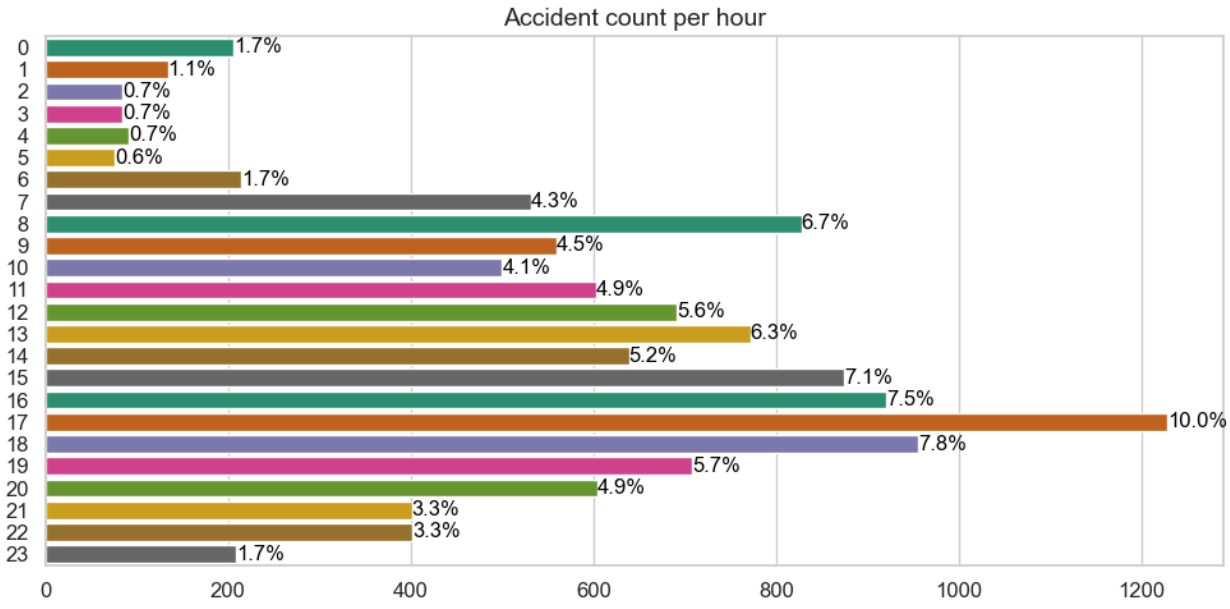
Results and Analysis

1. Accident Trends

- **Number of Vehicles:** As the number of vehicles involved in an accident increase, the fatality rate decreases. No fatalities were observed for accidents involving four or more vehicles.

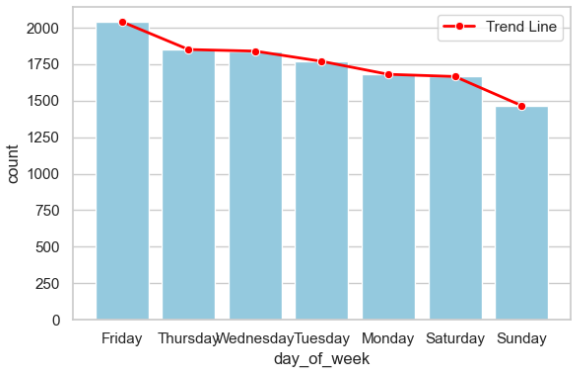


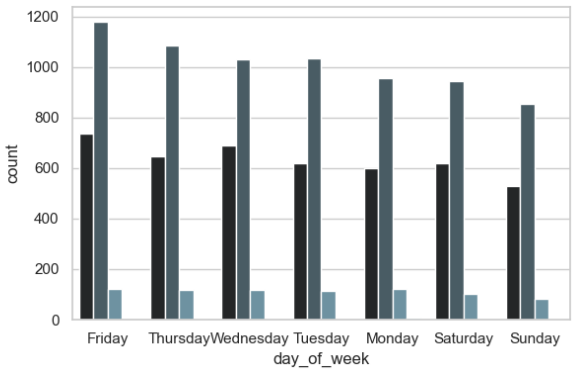
- **Time of Day:** Most accidents occur during three critical periods:
    - **8:00 AM**: Start of work hours.
    - **1:00 PM**: Lunch hour.
    - **5:00 PM**: End of work hours.



- **Daylight:** A majority (71.4%) of accidents happen during daylight.

- **Day of the Week:** Fridays are the most accident-prone. The accident decreases from Friday to Thursday then Wednesday right up to Sunday in an orderly fashion

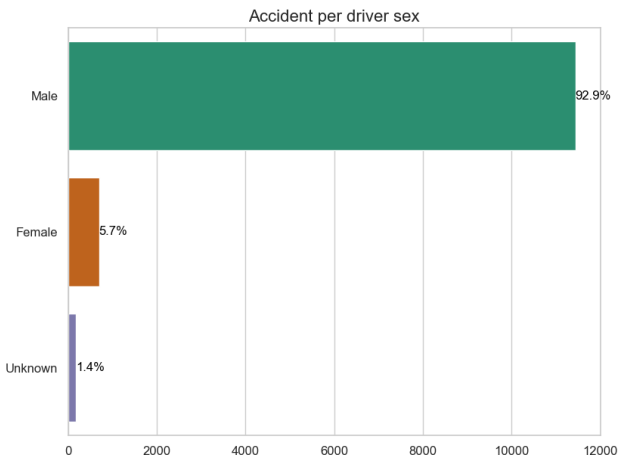


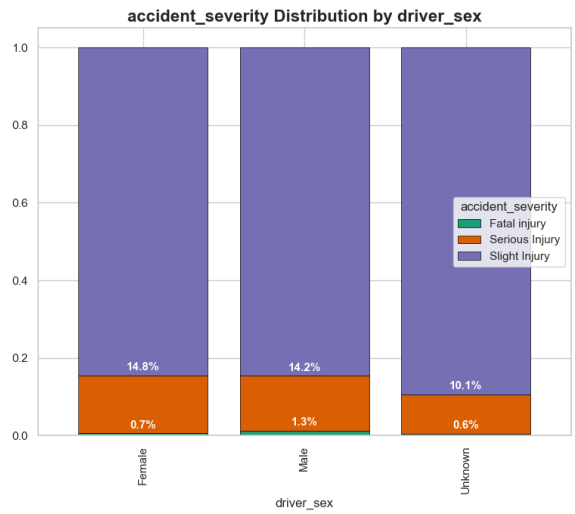


-
-
-

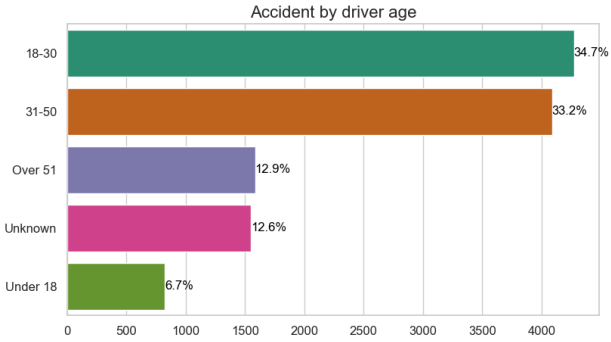
2. Demographics and Severity

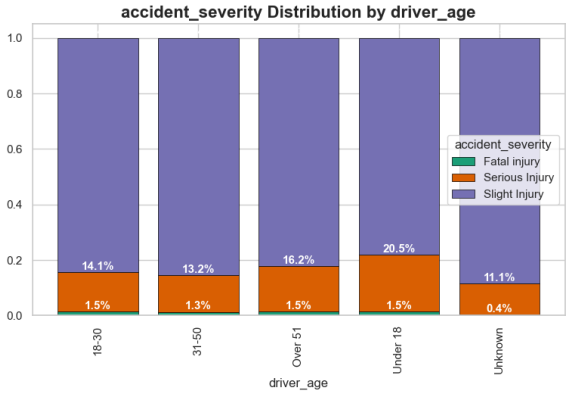
- **Gender:** Male drivers are involved in more accidents, though the severity of accidents involving females is comparable.



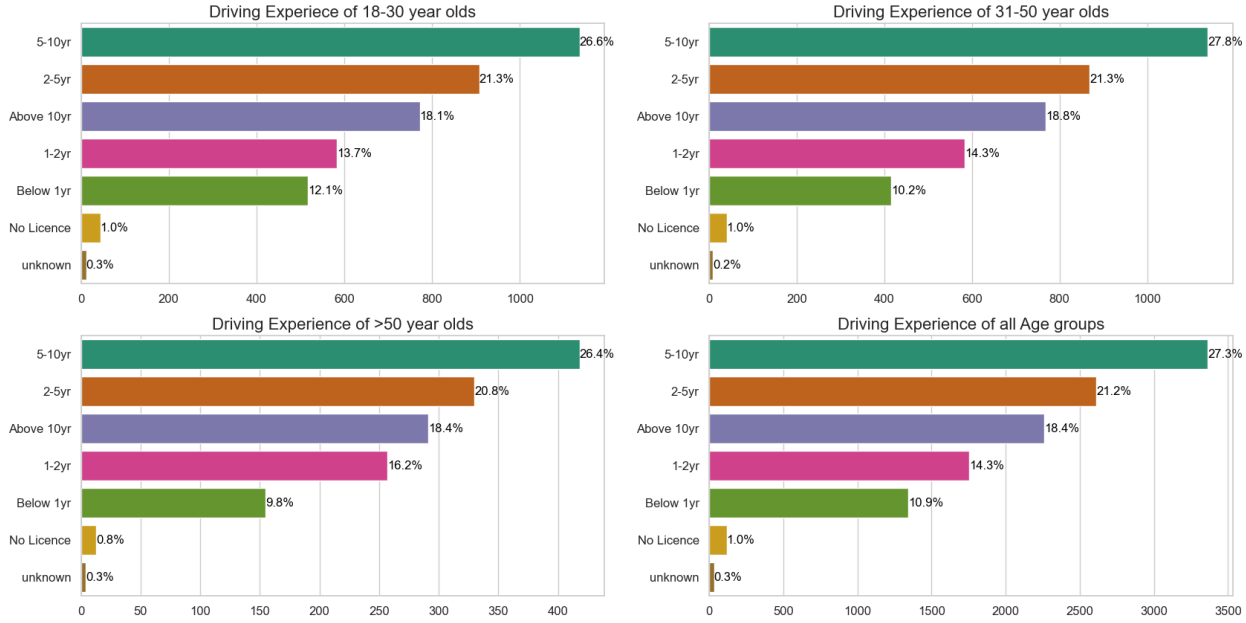


- **Age:** Underage drivers have the highest percentage of serious accidents (20.5%). Drivers aged 31–50 have the lowest serious accident rate (13.2%) and fatalities (1.3%).



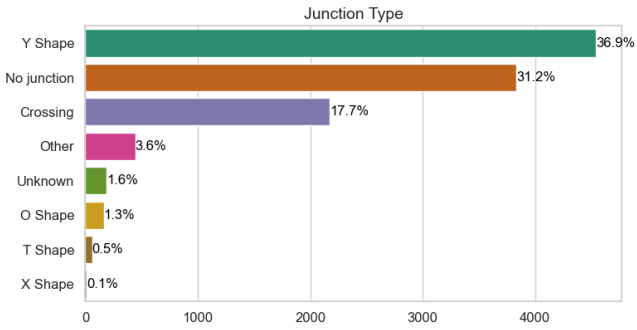


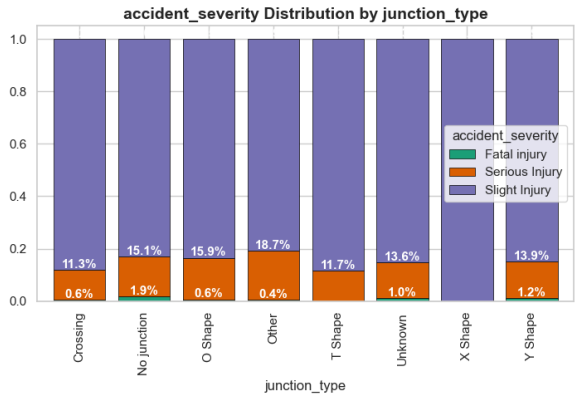
- **Experience:** Accident frequency increases with driving experience up to 10 years and declines thereafter.



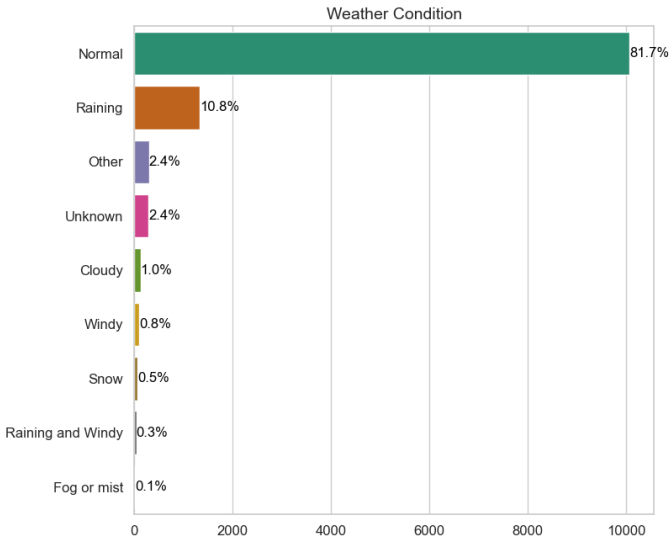
3. Road and Environmental Conditions

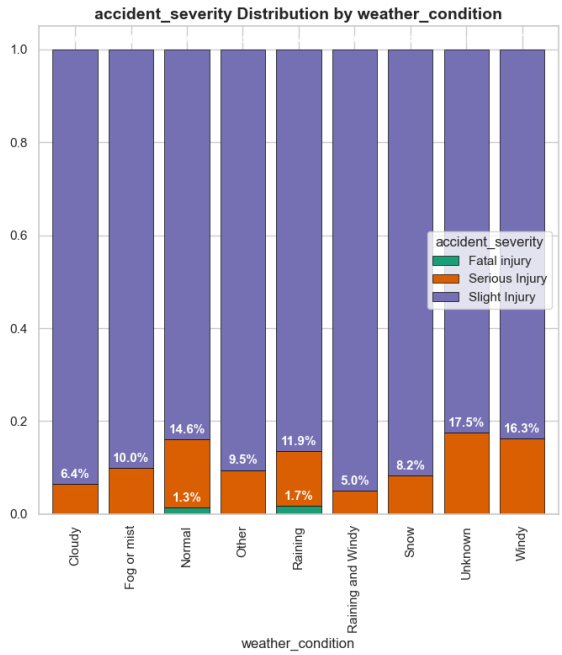
- **Junctions:** “No Junction” points have the highest fatality proportion (1.9%). Serious accidents are most common in “O Shape” and “No Junction” areas.





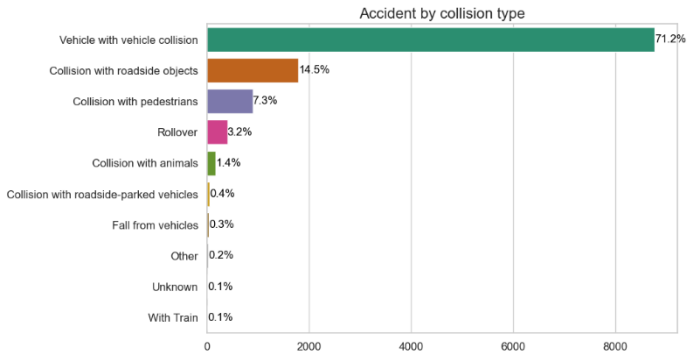
- **Weather:** Rain accounts for the highest fatality percentage (1.7%), despite 81.7% of accidents occurring in clear weather. Windy conditions lead to the highest serious injury rate (16.3%).

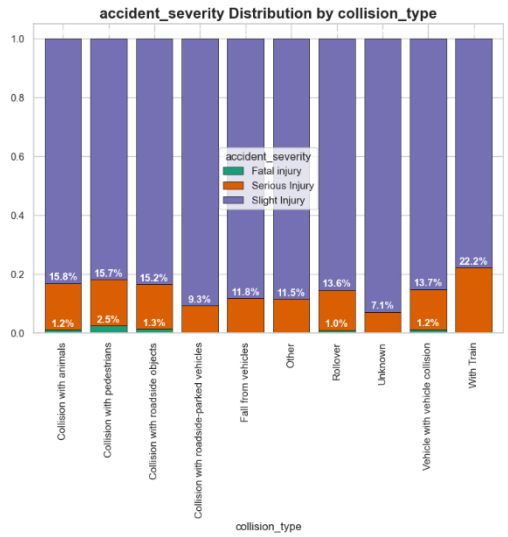




4. Collision Types

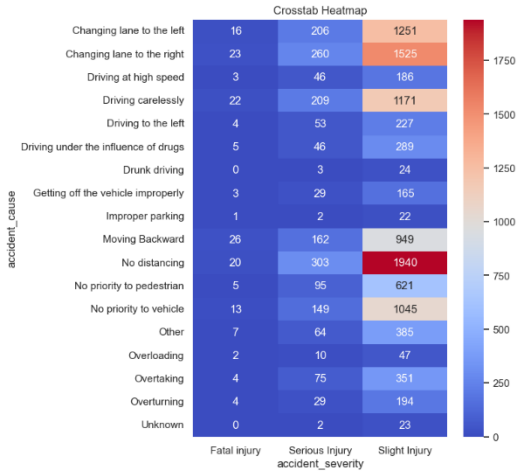
- **Vehicle-to-Vehicle Collisions:** Represent 71.2% of all accidents.
- **Train Collisions:** Lead to the highest percentage of serious injuries (22.3%).
- **Pedestrian Collisions:** Account for the highest fatality percentage (2.5%).

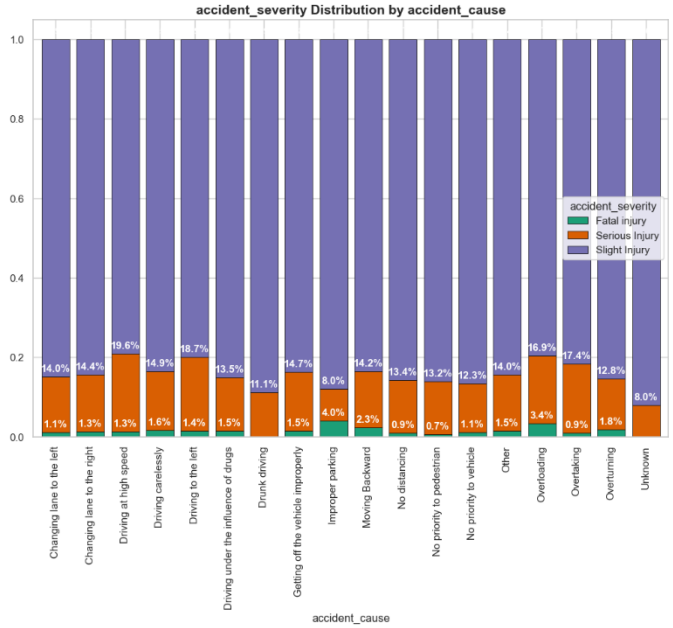




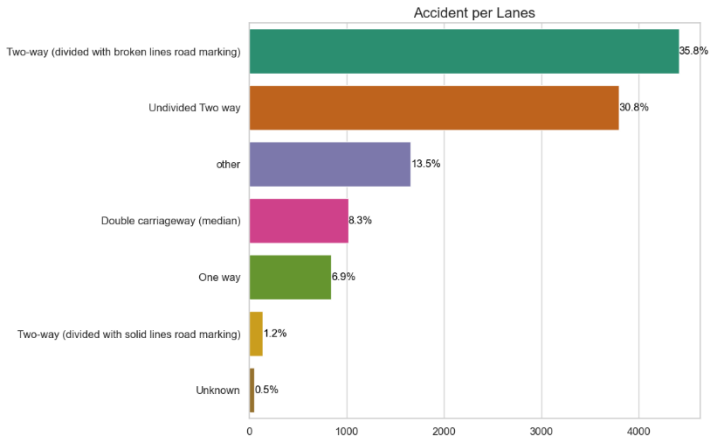
5. Causes of Accidents

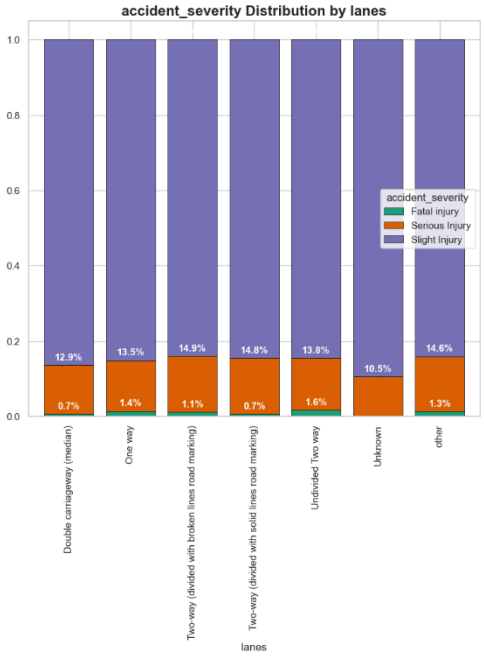
- **Over speeding:** While less frequent, it results in the highest proportion of serious accidents.
- **Improper Parking:** Causes the highest proportion of fatalities (4%).





Changing Lanes: The most prominent cause, followed by no distancing and careless driving.





6. Feature Importance

- **Mutual Information Scores:** Features with MI scores below 0.01, such as “Casualty Sex,” were deemed insignificant and recommended for exclusion.

Discussion

The analysis highlights significant patterns in road accidents in Addis Ababa. Key insights include the influence of driver behavior, time, and environmental conditions on accident severity. The findings suggest targeted interventions, such as:

- **Time-Based Measures:** Enhanced traffic management during peak hours.
- **Driver Training:** Focused programs for underage and inexperienced drivers.
- **Infrastructure Improvements:** Upgrades at high-risk junctions and better drainage systems for rainy conditions.
- **Policy Interventions:** Stricter enforcement of speed limits and parking regulations.

Conclusion and Recommendations

This study provides a data-driven foundation for improving road safety in Addis Ababa. Key recommendations include:

1. **Education Campaigns:** Targeting drivers about risks associated with overspeeding and improper parking.
2. **Improved Traffic Management:** Addressing peak-hour congestion and enhancing road infrastructure.
3. **Technology Integration:** Using AI and real-time monitoring to identify and address high-risk scenarios.
4. **Policy Reforms:** Strengthening laws on underage driving and enforcing penalties for violations.

By implementing these measures, Addis Ababa can reduce accident severity and enhance road safety for all users.
