Victoria Road Accident Analysis

Introduction:

Road traffic accidents significantly impact public health and community well-being. This analysis reviews 54,258 road accident records in Victoria, Australia, to identify patterns in the occurrence, types, and conditions of accidents. The insights aim to support the creation of more effective road safety strategies.

Analysis Overview:

The analysis examines accident trends across different times of the day, week, and year, shedding light on the frequency and severity of incidents to guide safety improvements.

Accidents by Time of Day (Line Chart):

Morning (8 AM) and evening (3 PM) commutes show a spike in accidents, highlighting rush hour risks. Late-night to early-morning hours see fewer incidents, suggesting lower traffic volumes as a factor.

Accidents by Day of Week (Line Chart):

A weekly tends shows an increase in accidents from Monday, reaching a peak on Friday, followed by a decrease over the weekend. This trend points to the potential impact of weekly commuting habits and suggests optimal times for the implementation of safety measures.

Accidents by Month (Line Chart):

The monthly trends, represented in a line chart, reveal a peak in accidents from February to March, with a significant decrease leading to the lowest point in September, followed by a decrease again in November and December. The fluctuations correspond to seasonal activities, public holidays, and varying road conditions, providing insights for targeted safety campaigns throughout the year.

Accident Occurrence Patterns Across Time of Day and Day of Week (Heat map):

Explores the relationship between the time of day and the day of the week, identifying peak periods for accidents. Accident rates spike at certain times, often during rush hours when more cars are on the road, raising the chance of incidents. Weekdays have a consistent pattern due to regular commutes, unlike weekends which see more leisure travel. The heatmaps red areas show when and where accidents mostly happen, pointing out key times for safety focus. Blue areas spots suggest fewer accidents, likely during late nights or quieter weekend hours.

The Accident Occurrence on Different Accident Type (Table Summary):

A categorization of accidents by type and their counts and percentages of total accidents, deaths, and injuries. Collision with vehicle is identified as the most frequent, constituting the majority of injuries and deaths. This is contrasted by other accident types, which show minimal occurrences. To reduce these collisions, we need to better understand why they happen through more data. Meanwhile, teaching drivers safe habits, doing regular drug tests, and applying strict road rules can lower accident numbers.

Accident Occurrence Patterns by Type and Light Condition (Heat Map):

A heatmap reveals that most accidents happen during daylight, mainly because there are more cars on the road, especially during busy rush hours. At night, when its dark and street lights are off, accidents decrease, likely due to less traffic outside of rush hours. To address the high accident rates during the day, managing traffic flow and enhancing driver awareness during peak times can help reduce the number of incidents.

Road Geometry and Accident Frequency (Table Summary):

Analyzing accidents by road type shows most happen away from intersections, where cars often speed up, increasing the chance of serious crashes. Distractions and risky overtaking add to the danger since these stretches usually have fewer controls compared to intersections. However, "T" intersections still see many accidents due to the challenge of navigating multiple signals and traffic directions. Places like dead ends or areas with "Y" intersections have fewer accidents, thanks to less traffic and slower speeds. Solutions could include adding more traffic controls on open roads, improving signage at intersections, and promoting safer driving habits to reduce accidents.

Local Government Areas and Accident Rates (Bar Chart):

It shows the Local Government Areas(Councils) with the highest number of accidents, revealing that urban centers like Melbourne and Casey are significant hotspots, mainly because of heavy traffic and more people living there. It suggests that making specific safety plans for each place could really help. For example, putting better crosswalks in busy city parts and clearer signs in country areas could make roads safer for everyone.

Speed Zones and Accident Frequency (Bar Chart):

The distribution of accidents across different speed zones, illustrating how the frequency of accidents varies with the posted speed limit. The data on road accidents shows that most happen in areas with speed limits of 60 km/h and 50km/h. These areas, usually in towns or suburbs, see a lot of accidents because there are many cars, people, and bikes around. Speed zones of 80 km/h and 100 km/h also see many accidents, likely on highways or country roads where cars go fast, making crashes more dangerous. Very slow (30 km/h) and very fast (110 km/h) areas have fewer accidents, possibly because drivers are more careful or there's less traffic. This tells us that

setting the right speed limits and making sure they're followed can help prevent accidents, especially in places where lots of people are moving around or cars are going fast.

Conclusion:

This comprehensive analysis of road accidents in Victoria underscores the importance of strategic road safety interventions tailored to specific times, locations, and accident types. By focusing on areas with the highest risk, implementing safety campaigns at critical times, and improving infrastructure, we can aim to reduce the incidence and severity of road accidents.

Recommendations:

- Optimize traffic flow during peak hours to reduce congestion-related accidents.
- Strictly enforce speed limits in high-risk zones.
- Conduct safety awareness campaigns at critical times, particularly during peak accident periods.
- Enhance pedestrian safety through improved crosswalks and signage.
- Reevaluate road designs, particularly at problematic intersections, to minimize risks.
- Increase nighttime visibility with better street lighting.
- Implement continuous road safety education targeting all road users.

Appendix:

Tools:

- **MS Excel:** Used for initial data cleaning and analysis
  Data Cleaning: Remove Duplicates, Filtering, Formatting, Spell Check, Change Case, Find and Replace
  Formula & Functions: Trim, Substitute, Xlookup, Proper, Ifs, Aggregate, Date, Time, Index, Match, Left, Right, Middle, Len
  Pivot Table and Charts

- **Tableau:** Enabled advanced data visualization and dashboard creation for interactive analysis.
  Data Preparation: Joining Table, Parameters, Filters, Calculated Fields, Aggregate Functions
  Data Visualization: Line Chart, Bar Chart, Table, Heatmaps, Sliders, Dropdown Menu
  Dashboard Creation for Business Intelligence

- **MS SQL:** Employed to verify the accuracy of the findings derived from Tableau analyses

  ### List of Queries:
  Table name: accident, nodeA

- **KPI: Total Number of Accident, Deaths and Injuries**

  ```
  SELECT
  ```

```
        COUNT(a.accidentNo) AS  TotalAccident,
        SUM(a.personsKilled) AS TotalDeaths,
        SUM(a.PersonsInjured) AS TotalInjured
FROM
        accident a
```

- **Accidents by Time of Day**

```
SELECT
        DATEPART(HOUR,a.AccidentTime) AS byHour,
        COUNT(a.AccidentNo) AS totalAccidents,
        SUM(a.PersonsInjured + a.PersonsKilled)  As TotalInjuriesDeaths
FROM
        accident a
GROUP BY
        DATEPART(HOUR,a.AccidentTime)
ORDER BY
        byHour ASC;
```

- **Accidents by Day of Week**

```
SET DATEFIRST 1

SELECT
        DATENAME(dw,a.accidentDate) AS dayOfWeek,
        COUNT(a.accidentNo) AS totalAccident,
        SUM(a.PersonsInjured + a.PersonsKilled) AS injuriesAndDeaths
FROM
        accident a
GROUP BY
        DATENAME(w,a.accidentDate),DATEPART(w,a.accidentDate)
ORDER BY
        DATEPART(w,a.accidentDate) ASC;

/*
SET DATEFIRST 1: Setting Monday the start of the week for any queries that
come after it in the same session.

DATEPART() function in GROUP BY clause: allows it to be later utilized in the
ORDER BY clause to sort the days of the week in calendar order.
*/
```

- **Accidents by Month**

```
SELECT
        DATENAME(Month,a.accidentDate) AS byMonth,
        COUNT(a.accidentNo) AS totalAccident,
        SUM(a.PersonsInjured + a.PersonsKilled) AS injuriesAndDeaths
FROM
        accident a
GROUP BY
        DATENAME(Month,a.accidentDate)
ORDER BY
        MIN(DATEPART(Month,a.accidentDate)) ASC;

/*
Even a function isn't listed in the GROUP BY or SELECT clause, it can
still be used in the ORDER BY clause by wrapping it with an aggregate
function. This works because aggregate functions like MIN() sort the data by
```

```
the smallest value in each group defined by the GROUP BY clause, without
breaking the grouping rules.
*/
```

- **Accident Occurrence Patterns Across Time of Day and Day of Week**

```
SELECT
	DATEPART(HOUR,a.AccidentTime) AS byHour,
	DATENAME(w,a.accidentDate) AS dayOfWeek,
	COUNT(a.accidentNo) AS totalAccident
FROM
	accident a
GROUP BY
	DATEPART(HOUR,a.AccidentTime), DATENAME(w,a.accidentDate)
ORDER BY
	byHour ASC, MIN(DATEPART(w,a.accidentDate)) ASC;
```

- **The Accident Occurrence on Different Accident Type**

```
SELECT
	a.accidentType,
	COUNT(a.accidentNo) AS totalAccident,
	CAST((COUNT(a.accidentNo) * 100.0) / SUM(COUNT(a.accidentNo)) OVER() AS
DECIMAL(10,2)) AS totalAccidentPerc,
	SUM(a.PersonsKilled) AS totalDeaths,
	CAST((SUM(a.PersonsKilled) * 100.0) / SUM(SUM(a.PersonsKilled)) OVER()
AS DECIMAL(10,2)) AS totalDeathsPerc,
	SUM(a.PersonsInjured) AS totalInjuries,
	CAST((SUM(a.PersonsInjured) * 100.0) / SUM(SUM(a.PersonsInjured)) OVER()
	AS DECIMAL(10,2)) AS totalInjuriesPerc
FROM
	accident a
GROUP BY
	a.accidentType
ORDER BY
	2 DESC;

/*
CAST(... AS DECIMAL(10,2)): To ensure percentages values are rounded to no
more than two decimal places for precision control.

100.0: to ensure floating-point division, which provides a more accurate
percentage value rather than integer division which could truncate the
result.

OVER(): A window function applied to compute aggregate sums of accidents,
deaths, injuries separately across the entire dataset, enabling percentage
calculations based on each accident type.
*/
```

- **Accident Occurrence Patterns by Type and Light Condition**

```
SELECT
	a.lightConditionDesc,
	a.accidentType,
	COUNT(a.accidentNo) AS totalAccident
FROM
```

```
        accident a
GROUP BY
        a.lightConditionDesc, a.accidentType
ORDER BY
        1 ASC, 2 ASC;
```

- **Road Geometry and Accident Frequency**

```
SELECT
        a.RoadGeometryDesc,
        COUNT(a.accidentNo) AS totalAccident,
        CAST((COUNT(a.accidentNo) * 100.0) / SUM(COUNT(a.accidentNo)) OVER() AS
        DECIMAL(10,2)) AS totalAccidentPerc,
        SUM(a.PersonsKilled) AS totalDeaths,
        CAST((SUM(a.PersonsKilled) * 100.0) / SUM(SUM(a.PersonsKilled)) OVER()
        AS DECIMAL(10,2)) AS totalDeathsPerc,
        SUM(a.PersonsInjured) AS totalInjuries,
        CAST((SUM(a.PersonsInjured) * 100.0) / SUM(SUM(a.PersonsInjured)) OVER()
        AS DECIMAL(10,2)) AS totalInjuriesPerc
FROM
        accident a
GROUP BY
        a.RoadGeometryDesc
ORDER BY
        2 DESC;
```

- **Local Government Areas and Accident Rates**

```
SELECT
        n.lgaNameAll,
        COUNT(a.accidentNo) AS totalAccident
FROM
        accident a
JOIN nodeA n ON a.accidentNo = n.accidentNO
GROUP BY
        n.lgaNameAll
Order By
        totalAccident DESC;
```

- **Speed Zones and Accident Frequency**

```
SELECT
        a.speedZone,
        COUNT(a.accidentNo) AS totalAccident
FROM
        accident a
GROUP BY a.speedZone
Order By totalAccident DESC;
```