DASHBOARD CREATION OF VIRTUAL INDIAN COLLISION ANALYSIS

REPORT OF VIRTUAL INDIAN COLLISIONS

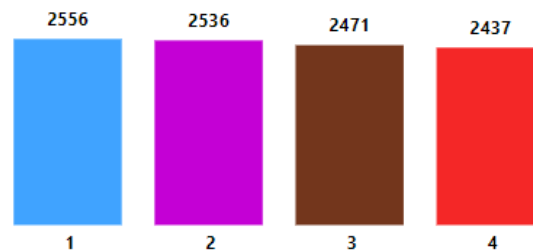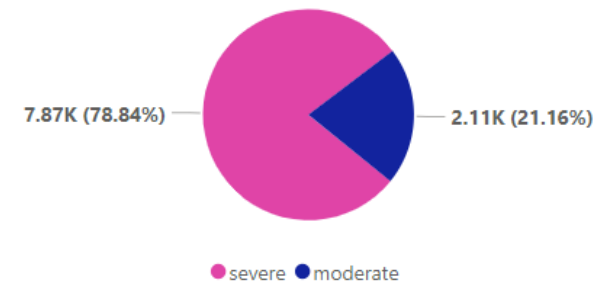| Honda | Hyundai | Mahindra | Maruti Suzuki | Tata Motors |

NUMBER OF ACCIDENTS

10K

AVERAGE AGE OF THE DRIVER

51

IMPACT OF SAFETY FEATURES

ABS	TCS	ESC	Accidents ▲
unknown	unknown	unknown	5
unknown	unknown	False	7
unknown	False	unknown	19
False	unknown	unknown	19
False	unknown	False	28
False	False	unknown	35
unknown	False	False	38
unknown	True	unknown	44
True	unknown	unknown	45
False	False	False	46
unknown	unknown	True	56
False	True	unknown	97
Total			**10000**

NO. OF ACCIDENTS BY SAFETY RATINGS

	1	2	3	4
	2556	2536	2471	2437

NO. OF ACCIDENTS BY VEHICLE TYPE

- sedan — 3.0K
- hatchback — 2.9K
- pickup — 2.0K
- SUV — 2.0K

BREAKUP OF CRASH SEVERITY

7.87K (78.84%) — 2.11K (21.16%)

● severe ● moderate
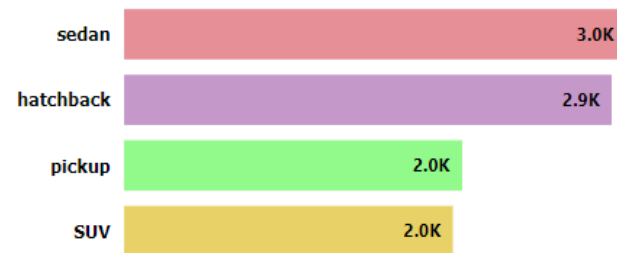
INFLUENCE OF WEATHER CONDITIONS AND TIME OF THE DAY

	clear	fog	rain	unknown	Total
afternoon	1779	309	629	299	3016
evening	1734	303	595	312	2944
night	1227	206	402	205	2040
morning	1180	195	415	210	2000

● clear ● fog ● rain ● unknown
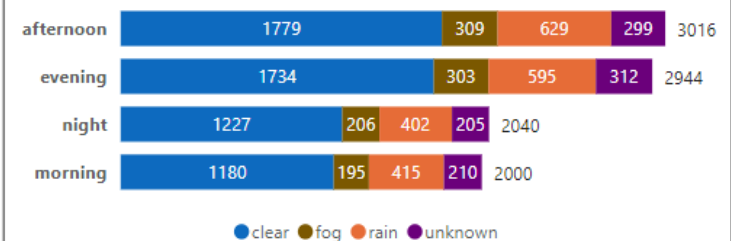
This is my 2nd attempt in trying to publish the dashboard in my profile while still honing my in Power BI. I chose the synthetic data provided by the Kaggle user Swayam Patil for creating this dashboard.

The dataset consists of the following information

1. **Vehicle_make**: The make or manufacturer of the vehicle involved in the crash, such as Maruti Suzuki, Hyundai, Tata Motors, Mahindra, or Honda.
2. **Vehicle_type**: The type of vehicle involved in the crash, including classifications like hatchback, sedan, SUV, or pickup.
3. **Vehicle_year**: The year of manufacture for the vehicle.
4. **Engine_type**: The type of engine powering the vehicle, such as petrol, diesel, CNG (compressed natural gas), or electric.
5. **Engine_displacement**: The displacement of the engine in cubic centimeters (cc), indicating its size and power output.
6. **Transmission_type**: Indicates whether the vehicle has a manual or automatic transmission.
7. **Number_of_cylinders**: The number of cylinders in the vehicle's engine, influencing its performance and efficiency.
8. **Vehicle_weight**: The weight of the vehicle in kilograms (kg), affecting its stability and crash dynamics.
9. **Vehicle_length**: The length of the vehicle in millimeters (mm).
10. **Vehicle_width**: The width of the vehicle in millimeters (mm).
11. **Vehicle_height**: The height of the vehicle in millimeters (mm).
12. **Safety_rating**: Assigns a safety rating to the vehicle based on predefined criteria, indicating its crashworthiness and occupant protection capabilities.
13. **Number_of_airbags**: Specifies the number of airbags installed in the vehicle, a crucial safety feature.
14. **ABS_presence**: Indicates whether the vehicle is equipped with an Anti-lock Braking System (ABS).
15. **ESC_presence**: Indicates whether the vehicle is equipped with Electronic Stability Control (ESC).
16. **TCS_presence**: Indicates whether the vehicle is equipped with Traction Control System (TCS).
17. **TPMS_presence**: Indicates whether the vehicle is equipped with a Tire Pressure Monitoring System (TPMS).
18. **Crash_location**: Specifies the location of the crash as urban or rural.
19. **Weather_conditions**: Describes prevailing weather conditions during the crash, including categories like clear, rain, or fog.
20. **Road_surface_conditions**: Indicates the condition of the road surface at the time of the crash, such as dry, wet, or muddy.
21. **Time_of_day**: Specifies the time of day when the crash occurred, categorized into morning, afternoon, evening, or night.
22. **Day_of_week**: Indicates the day of the week when the crash occurred, facilitating analysis of temporal patterns and trends.
23. **Driver_age**: The age of the driver involved in the crash.
24. **Driver_gender**: The gender of the driver involved in the crash (Male, Female).
25. **Crash_severity**: Classifies the severity of the crash as minor, moderate, or severe.

Data Cleaning

This dataset had 10K rows. There were some mismatched values which I had removed using the replace option present in the Power BI.

1. In real life scenarios, electric vehicles do not have IC engines, but here in the 2nd and 10th row a numerical value is present for electric vehicle category indicating discrepancy. Consequently, such rows were eliminated.

2. Similarly, there were many rows with missing and mismatched values. These accrued to 2.5k rows approximately. Deleting them would have a significant impact in analysis and dashboard. Hence, the value "unknown" was substituted in the required columns to replace them.

3. Following the previous step, adjustments were made to certain columns with mismatched data types. Upon completion, the data was loaded into Power BI.

Dashboard Creation

- **Filter View**

  To enable users to access individual manufacturers' numbers, a new version of filter was utilized and adjusted to ensure readability and ease of access all while not compromising the aesthetic integrity of the dashboard.

 | Honda | Hyundai | Mahindra | Maruti Suzuki | Tata Motors |
  |-------|---------|----------|---------------|-------------|

- **Adding Cards**

  To reveal the total number of accidents and the average age of the driver a basic DAX function involving count and average to create the measure which were used for the respective field. These were then implemented in the cards to share the required information.

 | NUMBER OF ACCIDENTS | AVERAGE AGE OF THE DRIVER |
  |---------------------|----------------------------|
  | 10K | 51 |
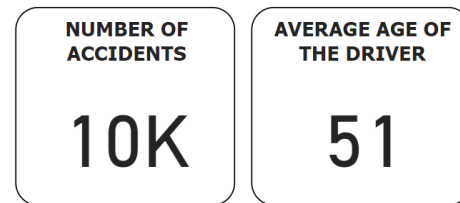
- **Pie Chart**

  Three types of accidents were identified: minor, moderate, and severe. To analyze the proportion between these types, a pie chart was employed for its simplicity and ease of understanding. The "minor" type accounted for only 0.01% of the total number of accidents and thus it was removed from the chart due to its negligible contribution. The design was kept minimal for better clarity.

  BREAKUP OF CRASH SEVERITY

  7.87K (78.84%)   2.11K (21.16%)

  ● severe ● moderate
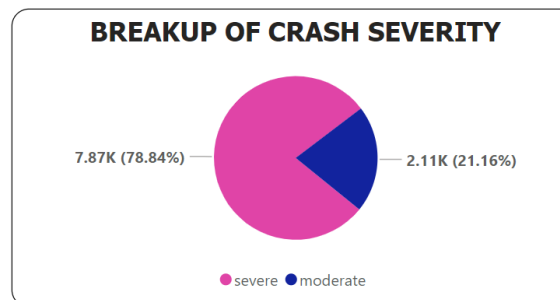
- **Clustered Column Chart**
  This type of chart is particularly effective for comparing categorical data types simultaneously. To calculate the number of cars involved in low to high safety ratings, the previously used measure involving the count function along with categorical safety ratings data were applied.
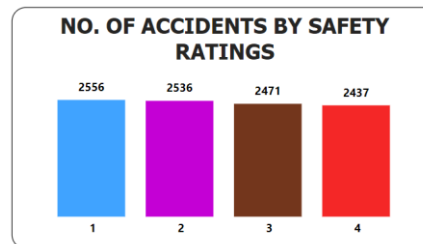
- **Stacked Bar Chart**
  To visualize the impact of both weather conditions and the time of day on crashes, a stacked bar chart was employed. It helps us understand the number of accidents that occurred under different weather conditions and at different times of the day.
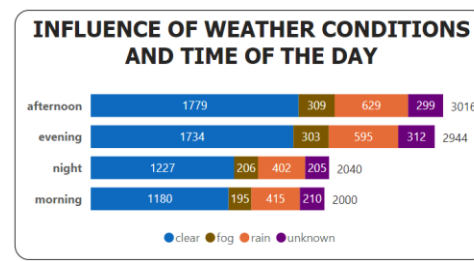
- **Bar Chart**
  A simple bar chart displays the total information about the types of vehicles involved in the crash. To achieve this, the Count DAX function was utilized along with the vehicle type to create the required visual representation.
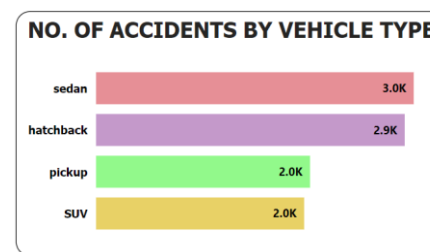
- **Table creation**

   In the analysis, the impact of safety features was examined. Three safety features in cars involved in accidents were considered: Anti-Lock Braking System (ABS), Traction Control System (TCS), and Electronic Stability Control (ESC). These columns had many missing values, accounting for approximately 25% of the total rows.

   Removing such a large portion of data would significantly impact the analysis. To address this, the tables option was effectively utilized to combine all these columns and obtain a summarized result. These columns had three possible values: True, False, or Unknown. The tables visual was used to identify which vehicle had which safety feature installed when involved in an accident. There are 27 different combinations representing this data in a concise way.

IMPACT OF SAFETY FEATURES

ABS	TCS	ESC	Accidents ▲
True	False	unknown	123
unknown	False	True	123
True	False	False	185
False	True	False	202
False	False	True	226
unknown	True	True	333
True	unknown	True	337
True	True	unknown	365
True	True	False	715
False	True	True	763
True	False	True	791
True	True	True	5054
Total			**10000**

Editing Interactions

Power BI has a feature that highlights parts of all visuals when selecting a specific data point. However, this may pose readability and comprehension challenges for many users. To address this, all interactions pertaining to this dashboard were adjusted. Now, Power BI revisualizes the exact chart without highlighting when selecting a certain data point, ensuring a smoother user experience

Comments

In my opinion, after going through some certain posts in social media, I realized that most of the consultants generally prefer their dashboards to be simple and effective when it comes to understanding. Hence, I chose to go with the white background and remove the redundant titles, legends etc. This is subjective and the style might differ based on organizations' approach and practices.

Finally, I would like to thank you for reading till the end. I am eagerly looking forward to your opinions and feedback. Thank you very much.