Analysis of Road Accidents by using Power BI

ITIS414 Project Report

Fatema Hussain Habib
Information Systems
University of Bahrain
Sakhir, Bahrain
20192408@stu.uob.edu.bh

Walaa Hussain Ali
Information Systems
University of Bahrain
Sakhir, Bahrain
202104424@stu.uob.edu.bh

Manar Eyad Ahmed
Information Systems
University of Bahrain
Sakhir, Bahrain
202006306@stu.uob.edu.bh

Fatema Jaafar Ahmed
Information Systems
University of Bahrain
Sakhir, Bahrain
202002065@stu.uob.edu.bh

Alanood Abdulaziz Alhammadi
Information Systems
University of Bahrain Sakhir, Bahrain
202009426@stu.uob.edu.bh

Supervised by: Dr. Ahmed Zeki

Abstract—*This project focuses on applying business intelligence practices to a road accidents dataset using Power BI. Business intelligence involves the collection, storage, and analysis of data generated by a company's operations. By leveraging Power BI, this project aims to gain insights from the road accidents dataset and apply business intelligence techniques to enhance understanding and decision-making in accident prevention and road safety.*

Keywords— *Road Accidents, Power BI, analysis*

i. INTRODUCTION

Every year, road accidents claim millions of lives worldwide, making them a serious problem. These occurrences have significant negative effects on the economy and public health. They are among the main causes of death and material loss. The World Health Organization (WHO) estimates that road traffic crashes claim the lives of 1.35 million people annually, with millions more suffering non-fatal injuries that frequently result in permanent disability.
The financial impact consists of direct costs to families, medical bills, and lost production as a result of injuries and fatalities. Numerous factors, such as human behavior, vehicle conditions, road infrastructure, and environmental impacts, contribute to the complexity of traffic accidents.

It is essential to analyze these complex components to create preventative methods that work. The development of business intelligence (BI) and data analytics technologies has made it feasible to examine accident data in greater detail, spot trends, and extract useful information that can guide intervention and policy decisions. The goal of this project is to do a thorough analysis of road accident data by utilizing the capabilities of Power Bl, a top BI tool. The project aims to identify high-risk indicators, explain the underlying causes of road accidents, and suggest data-driven solutions to reduce the frequency of these unfortunate occurrences by utilizing Power Bl's sophisticated analytical skills. The ultimate objective is to improve road safety, which will save lives and lessen the financial burden brought on by traffic accidents.

ii. BACKGROUND

Since the invention of cars, road safety has been a worry, but as more and more vehicles are driven on the road, the problem has become much worse. Historically, legislation requiring seat belt use, speed limits, and penalties for drunk driving have all been used to reduce traffic accidents. Although the effectiveness of these initiatives has varied, the development of digital technology and data analytics opens up new possibilities for tackling issues related to road safety. Large datasets may be integrated, analyzed, and visualized with the use of business intelligence (BI) solutions like Power Bl, which promote a deeper comprehension of challenging problems. With the help of Power BI's dynamic dashboards, real-time data processing, and sophisticated analytical features, stakeholders can keep an eye on trends, pinpoint hotspots, and forecast the likelihood of future traffic incidents.
The development of focused treatments and policies requires the use of this data-driven methodology.

iii. PROBLEM DESCRIPTION

The necessity for a thorough examination of the causes and contributing variables of road accidents is highlighted by the persistently high incidence of these incidents worldwide.
Conventional accident investigation techniques frequently fail to offer the detailed insights required for successful preventative measures. The complexity of road accidents, which include interactions between human, vehicular, infrastructure, and environmental elements, is a problem.
Important concerns consist of:
Human factors include speeding, driving while intoxicated, breaking traffic laws, and distracted driving.
Infrastructure factors include bad road conditions, inadequate signage, and unsafe intersections. Vehicle factors include mechanical problems, poor vehicle

maintenance, and design defects. Environmental factors include the state of the weather, problems with visibility, and natural hazards. Finding patterns and correlations in extensive accident data analysis is crucial to resolving these problems. The goal of this project is to analyze this kind using Power BI, to produce insights that can be used to tailor actions and policies to increase road safety.

iv. DATA COLLECTION

Data for this analysis is sourced from the well-known data science and machine learning portal Kaggle. A range of datasets, including comprehensive records of traffic occurrences, meteorological data, and vehicle information, are available on Kaggle and can be used in the analysis of road accidents. The dataset chosen for this project includes extensive details on road accidents, such as:

• Date and time of the accident

• Location

• Weather conditions at the time of the accident

• Road conditions (e.g., wet, dry, icy)

• Type of vehicles involved

• Number of casualties (fatal and non-fatal)

• Causes and contributing factors (e.g., speeding, distracted driving)

By importing this data into Power BI, the project will leverage its powerful data integration, cleaning, and visualization features to explore the data thoroughly. This will involve creating interactive dashboards that highlight key metrics, identify accident hotspots, and reveal trends over time.

v. SIMILAR WORK

Similar to what Power BI might give, various studies have been undertaken in the field of road safety and accident prevention that provide significant information. As an example, a lot of study has been done on the relationship between bad weather and highway events. The results show some alarming figures: bad weather accounts for around 24% of all crashes, resulting in about 7,400 fatalities and 673,000 injuries annually [1]. Similarly, they have examined age and gender in their study on pedestrian accidents. Research examining collision data from Melbourne from 2004 to 2013 showed that different age groups and genders display distinct crash patterns. For example, younger and older pedestrians experience crashes at various periods, while men and women experience crashes at different times [2].

Some innovative approaches are being studied, such as the use of smartphone applications to track driver behavior and prevent accidents. These applications, which have shown potential in reducing accidents in early tests, employ the sensors and cameras in smartphones to identify tiredness and distractions [3]. Additionally, to determine the locations and times of accidents that are most likely to happen, researchers have examined localized data, such as that from Oyo State, Nigeria. They have identified times and places with greater accident rates by examining data from previous years, which has given them important information for focused safety measures [4].

Together, these studies underscore the complexity of road accidents and highlight the importance of data-driven strategies in preventing them.

vi. APPLYING BUSINESS INTELLIGENCE ON DATASET

1. The Most Day of Week:

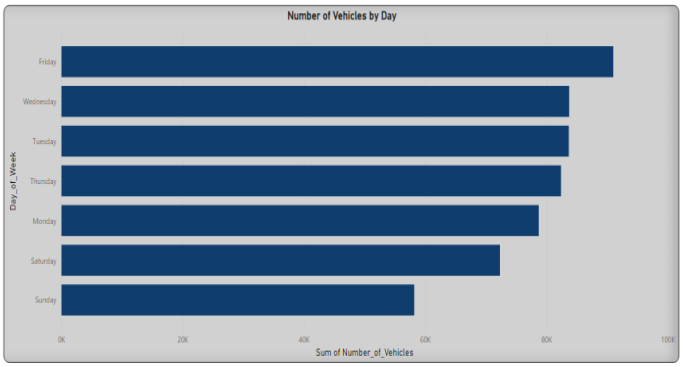

Figure 1. Number of Vehicles by Day of Week

According to the analysis, there were the most vehicle accidents on Friday (91 K), followed by Wednesday (83.7 K), and Tuesday (83.6 K). We observe that because Friday is the day before the weekend, there are more accidents and more people on the streets. People disregard traffic safety during the day as they hurry home. Also, the majority of people go out at night on Friday since they view it as the start of the weekend.

2. The Most Type of Vehicle:

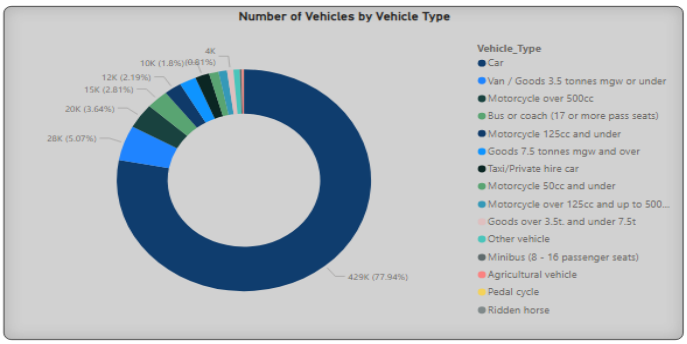

Figure 2. Number of Vehicles by Vehicles

In this figure, the type of vehicle most involved in accidents is cars (77.94%), as most people use cars to travel home or to work, so they have become an essential factor for living, especially in civilized cities, as they are safer and faster. It also requires some work on driving knowledge. Then the van (5.07%), and the third motorcycle (3.64%).

3. The Most Road Surface Conditions:

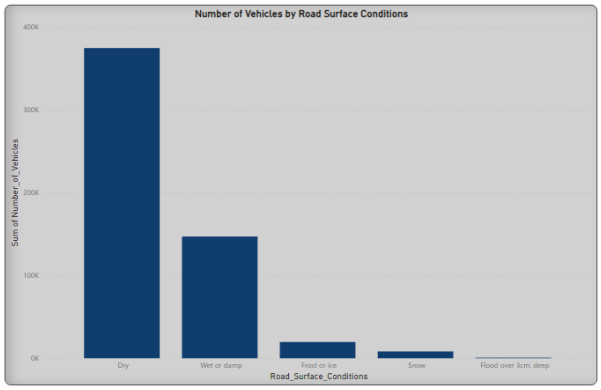

Figure 3. Number of Vehicles by Road Surface Conditions

Usually, after the weather, the road surface conditions are an important factor in many accidents between vehicles, but in this figure, we conclude road surface conditions did not have a significant impact, as most accidents occurred in dry road surface conditions, which were without any factors (374.5 K). This is because people believe that there are no changes in the road surface conditions, which allows them to drive without paying attention to road safety or speed restrictions. Secondly, it was wet or damp (146.9 K).

4. The Most Accident Severity:

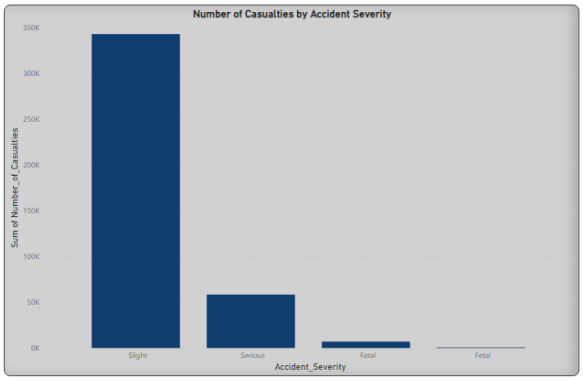

Figure 4. Number of Casualties by Accident Severity

We will notice in this figure the relationship between the number of casualties and the accident's severity. Accidents are often slight, but casualties are high by 343K. For example, this is due to not wearing a seat belt or other safety laws. Then there were serious accidents (58K), followed by fatal accidents (6.9K).

5. The Most Time:

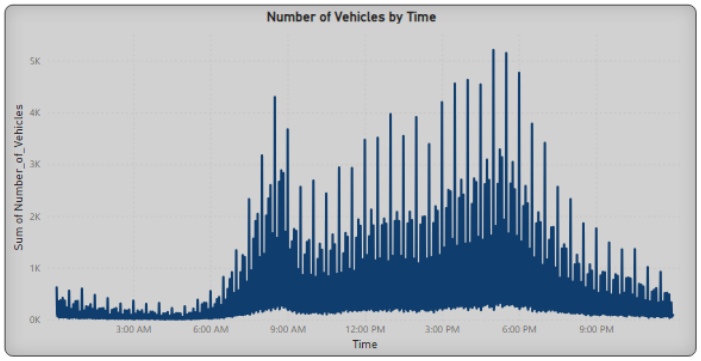

Figure 5. Number of Vehicles by Time

The figure shows the complex relationship of an accident between the number of vehicles and time. The highest time when an accident occurs among the number of vehicles is at five in the evening (5.2K), which is the peak time for people returning home from work. It was followed by 5:30 pm (5.1K).

6. The Most Road Type:

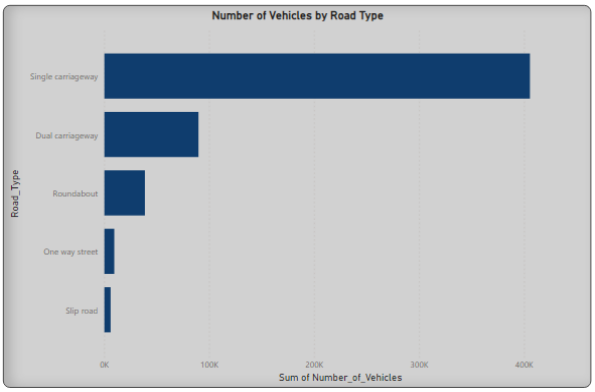

Figure 6. Number of Vehicles by Road Type

In the visualization analysis, single carriageway has the highest number of vehicles in accidents (405.6 k). The reason may be due to drivers losing control of their vehicles or driving recklessly. Second, the dual carriageway (89.9 k). It was followed by the roundabout (38.8 k).

7. Dashboard:

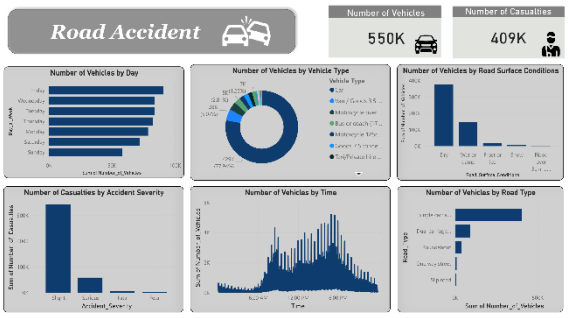

Figure 7. Accident Road Dashboard

In the dashboard, visualizations are created by Power BI. The primary data includes the number of casualties and the number of vehicles, and the secondary data includes the day of the week, accident severity, road surface conditions, road type, time, and vehicle type. These visualizations work to summarize the data analysis and help make decisions.

vii. DISCUSSION

After the issue is identified, the data is collected, modified, and then used by Power BI. Returning to the dashboard, we conclude that many accidents occur when employees return home at 5:00 pm, as in Figure 5, or on the weekend, on Friday, as in Figure 1, depending on the number of vehicles and the number of casualties.

viii. FINDINGS AND CONCLUSIONS

After conducting analysis on the given dataset, we found strong relationships between road accidents and the other factors like number of vehicles and their types, road surface conditions, date, time, number of casualties, accident's severity, speed, weather, and many others, by applying BI tools and techniques we successfully understand the reasons of occurring road accidents and its relationship with other factors. By utilizing BI, we can now take the best decision based on the analysis and try to minimize the road accidents by providing many solutions like: Raising the penalties for exceeding the speed limit and Fasten seat belts on the street, which will prevent many drivers from exceeding it, motivating drivers to adhere to traffic laws, Roads with snow, mud, or other debris that could endanger drivers and cause accidents should be closed, Place speed cameras in all streets in order to monitor violators.

The analysis conducted in this report highlights the importance of BI tools and techniques that may be applied to the critical task of preventing car accidents. By utilizing data analytics, dashboard and reporting, and visualization, we can obtain a significant understanding of the complex factors contributing to car accidents.

ix. REFERENCES

[1] P. A. Pisano, L. C. Goodwin, and M. A. Rossetti, "U.S. HIGHWAY CRASHES IN ADVERSE ROAD WEATHER CONDITIONS," 2023. [Online]. Available: https://media.syracuse.com/weather_impact/other/us%20highway%20crashes%20in%20adverse%20weather%20conditions.pdf.

[2] "Influence of pedestrian age and gender on spatial and temporal distribution of pedestrian crashes", Taylor and Francis Online. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/15389588.2017.1341630

[3] “Methodology and Mobile Application for Driver Behavior Analysis and Accident Prevention,” IEEE Journals & Magazine | IEEE Xplore, June. 04, 2019. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8730492

[4] F. J. Jegede, “Spatio-temporal analysis of road traffic accidents in Oyo State, Nigeria,” Accident Analysis and Prevention, June, 1988. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/0001457588900073

[5] “Car Accident Dataset,” Kaggle. [Online]. Available: https://www.kaggle.com/datasets/nextmillionaire/car-accident-dataset?resource=download

Fatema Hussain Habib	- Applying Business Intelligence on Dataset - Discussion
Walaa Hussain Ali	- Similar Work
Manar Eyad Ahmed	- Findings And Conclusions
Fatema Jaafar Ahmed	- Abstract
Alanood Abdulaziz Alhammadi	- Introduction - Background - Problem Description - Data Collection