International Journal of Information Technology and Computer Science Applications (IJITCSA)
p-ISSN: 2964-3139 e-ISSN: 2985-5330
Vol. 1, No. 2, page 86 – 95
Submitted 05/04/2023; Accepted 11/05/2023; Published 02/06/2023

Business Intelligence (BI) Approach for Traffic Accidents Analysis

[1*]Shujana Madnira, [2]Carlos Sandez, [3]Aaron Abby
[1,2,3]School of Information Technology, Mapúa Institute of Technology, PHILIPPINES
e-mail: guyanaa931@gmail.com

Corresponding Autor: Shujana Madnira

Abstract
'FLCRASH' is the data source used for this report that consists of crash data from the year 2008 to 2009. The purpose of this report is to find ways to reduce the number of fatal accidents on the road by identifying the main culprits on why these accidents happen so frequently. Business intelligence is used in the process of finding the relationships and trends in the major causes of the crashes. The technologies related to Business Intelligence that will help in aiding the organization to their goal are also discussed along with the supporting factors for the use of Business Intelligence for the organization. A dashboard has been created from various graphs and charts for easy viewing and understanding of the situation so logical decisions can be made based on it. This report seeks to help us better understand the nature of why traffic accidents happen in the first place and how to prevent them from happening in the future.

Keywords— Business Intelligence, accident analysis, traffic accidents, data visualization

1 Introduction

It's a fact that driving cars or riding motorcycles is more dangerous that other forms of transportation like Trains. It's also a surprise that even flying in an airplane 30,000 feet above the ground is safer than using motor vehicles. It is stated that the odds of dying in a vehicle to be 1 in 98 in a lifetime [1]. This is a worrying statistic when compared to only 1 in 7178 in a lifetime when flying in an airplane. Therefore, it is important for us to find ways to reduce the mortality rate of driving motor vehicles when on the road.

Traffic accidents not only affect the lives of the victim, but they also affect the economy as a whole. Traffic accidents causes a large percentage of casualties and long-term injuries therefore the productivity of the working-class citizens Is greatly reduced resulting in the regression of the economy [2],[3]. Florida Transportations is a company which analyses crash data from traffic accidents from the state of Florida in the United States. Their goal as an organization is to reduce the number of accidents occurring in the Country by finding out the cause of the crashes and analyzing existing accidents [4],[1]. A lot of the accidents occurring are accidents that can be prevented if proper actions are taken beforehand. They hope that their efforts will be beneficial to the state of Florida and the country.

These accidents occur due to many different contributing factors. Florida Transportation's uses a Business Intelligence approach to analyses and identify why these accidents occur therefore we can find ways to prevent future tragedies from ever happening. Traffic accidents not only affect human lives but also stunt economic growth as the aftermath of these accidents can lead to long term disabilities and workers are forced to retire early[5],[6],[3]. They are also one of the major causes of deaths that happen in the country. Traffic accidents effects millions of people in the world so it is important to find a way to improve the situation at hand. This data source has 1,871,649 observations and 31 columns. It also consists of 31 variables.

2 Business Issues

The first issue is to find ways to reduce the total number of crashes or accidents that could end in fatalities in all the counties in Florida. These accidents are one of the leading causes of death in Florida, therefore steps should be taken to reduce the amount especially since traffic accidents like these can be easily preventable with the right precautions [7]. The county, Miami-Dade, has the greatest number of casualties in Florida. This county is also the county with the highest amount of people in it therefore accidents happen very frequently. The total crash fatalities for the county based on the data source is 1269 deaths. This figure is considerably more than that of [7] the second highest county based on crash injuries which stands at 965 losses. Palm Beach and Hillsborough have almost the same amount of fatalities at 797 and 762 respectively.

Another issue that we need to solve is determining the main factors on why these accidents happen and how they impact the number of crash injuries and fatalities [8]. One of the factors is driver distraction or their carelessness while driving on the road. This has caused over 247,800 injuries to occur from traffic accidents. It has also resulted in 1337 deaths. When they fail to drive carefully and are not careful, they can also cause unnecessary danger to other people on the road and themselves [9]. Another factor why traffic accidents occur is the failure to Yield Right-of-way. This has contributed to 109,808 injuries and 665 fatalities from accidents. Drivers that fail to determine the right course of action and yield on the road can and will cause unnecessary accidents [10].

The next issue then is finding out what part of the vehicle these crashes occur at. Determining how and where these accidents occur can help us to find ways to find ways to reduce the severity of the crashes [11]. For example, 515,011 injuries and fatalities caused by traffic accidents occur at the rear-end or rear-end collisions. This is almost half the number of the third worst crash area, angle collisions, that is at 294,737 crashes.

3 Proposed Solution

The solution that I propose for the first issue, finding ways to reduce the total number of crashes or accidents that could end in fatalities in all the counties in Florida, is using business intelligence tools to study the data from all the traffic accidents [12]. After analysing all the information, we will use visual representations like charts to show the severity of the casualties caused by traffic accidents in Florida. Our company will use various Business Intelligence software that can help us prevent these accidents from happening in the first place. We can predict where accidents are more likely to occur and act accordingly and determine where our resources should be used at. The state can use this information and decide to run a safety campaign at where most of the accidents happen which is at Miami-Dade and Broward for instance [13].

The solution for the second problem, determining main factors on why these accidents happen and how they affect the number of crash injuries, we will use Data Mining techniques and analyse the previous data of traffic accidents to find out the main culprits of why these accidents happen. By finding out which factor is the most problematic, we can suggest to the Florida County Commission ways to improve the situation [14]. They can launch road safety programs all around Florida that can reduce the traffic accident rates hence reducing fatality rates.

For the next problem, we can use the information gathered from the analysis of the graphs to determine which part of the vehicle should be better engineered or be designed in a smarter way. This can improve the outcome that might occur when these traffic accidents happen so the number of fatalities can be reduced. Rear-end collisions are the most common crashes that happens in Florida so extra work should be done at that specific part of the vehicle [15].

4 Support Factors

Organizations need the applications of Business Intelligence in their company. There are several relevant factors in supporting the proposed solutions. The first factor is the increase of data volumes. The advantage of Business Intelligence is being able to process large amounts of data in a single time. Florida Transportations uses Business Intelligence tools to analyze and process past crash data and reports, so they can figure out the factors that influence these traffic accidents. This will allow them to find ways to prevent these accidents from occurring in the first place thus reducing the number of injuries and casualties that can happen because of vehicle crashes [16].

The next factor to support these solutions is the increased complexity of the decision-making process. It is to decide when you have a very large amount of structured and unstructured data. Business Intelligence solutions allows Florida Transportations to easily make precise decisions. Florida Transportations visualizes the data given. For example, when determining the human factors that cause traffic accidents, you can visualize this data and easily understand the data better. Companies can then make wiser decisions based on these visual representations.

The need for quick reflexes is another factor where business intelligence is needed in organizations like Florida Transportations. In today's world, opportunities that your organization receives can close very quickly before you have the chance to take it. The people who make these decisions need to react fast and accordingly. Traffic accidents happen every day by the millions. They don't wait around for us to slowly figure out how to manage them. There are many delays which will hinder a company's decision making, the first is the delay in converting data from many sources into information that can be used. Doing this the traditional way will surely slow the progress of a company like Florida Transportations as the volume of data they work with is huge as the number of accidents that occur is a lot. Next is to integrate said data across many different sources and eventually produce the information needed for the decision makers to use. Effective use of these Business Intelligence solutions can greatly help reduce the duration needed for a good decision to be made.

The final factor is the technological advancements or progress that have been made that can benefit organizations today both hardware-wise and software-wise. The development of Decision Support Systems (DSS) has enabled organizations to make wiser decisions that can help a company like Florida Transportations in the long run. Other advancements include data warehousing and data mining. Florida Transportations have all the right materials to make effective business intelligence tools, so the most effective business intelligence solutions can be used [17].

5 Required Technologies

There many capabilities of Business Intelligence that can benefit organizations today. The first is organizational memory. This is known as a way to store information and knowledge. Unstructured data and structured data or data resources are accumulated and used to help organizations like Florida Transportations to make important decisions that will shape the future of the company. These Business Intelligence tools are used to reduce the amount of unwise decisions that are being made, an example of technologies that are important to organizational memory is online transaction processing (OLTP). It is a type of online processing system. It stores all the data regarding transactions from an organization. A company like Florida transportations needs technologies like OLTP so they can avoid overspending is certain places by storing all the data regarding transactions.

Another type of this technology is Enterprise Resource Planning or "ERP". It stores and processes the data in terms of the company's raw materials, human resources and manufacturing processes. ERP can be used to provide better organize and manage the employees of Florida Transportations on which area's should they be focusing on most.

The third technology related to organizational memory is Data Warehouse. This collects a large volume of data from multiple different systems which is later transformed to create consistency and finally be loaded in format to be used for an analysis by the company. Data warehousing is important to the organization as this allows them to analyse large amount of data regarding traffic accidents crashes that is the main purpose of the company. ERP on the other hand can be used to provide better organize and manage the employees of Florida Transportations on which area's should they be focusing on most.

The next capability is Information Integration. This is the ability to link together structured and unstructured data together to form relevant information to facilitate insight creation. The first is Environmental Scanning. This is where you check the events and trends that are outside or external from the organization but are important to the future company decisions.

Another type of technology is Text mining or Web mining. This means mining the contents of an unstructured data to find trends in it. Florida transportations uses this to check for relevant trends that can help determine why these vehicle crashes occur. Next is Web Scraping where we scourge the internet for published data. The employees at Florida transportations can use this method to look for other trends regarding traffic accidents in other counties in America to have a better understanding of the situation.

The Internet of Things devices are devices that are connected via the Internet, so they are able to both send and receive data. Insight Creation is the act or a way to develop insights based on the information you have to ultimately make wiser decisions. Patterns and trends are both identified along with the relationships to help in the decision-making process and make it more streamlined [18].

Data mining is one of the types of technologies that is important in this area. Hidden patterns are found in large data sets that is stored electronically so we can generate new information from it to help the organization. Florida Transportations uses this method to find trends that are otherwise missed by traditional methods like which part of the vehicles are most effected during accidents.

The next piece of technology used is Business Analytics. This largely refers to previous skills, technologies, and practices and focuses on creating new insights from the data that has been collected. This is used by Florida transportations, so they can possess superior knowledge to create new insights from the data they have. Another one of these technologies is called Real Time Support System. It uses data mining or business analytics to help them to make operational decisions according to the data they have mined previously. A company like Florida Transportation uses this method to support the higher ups that make the decisions, so they don't make the wrong one and make more ones that will allow them to reach the goals they have set.

The final capability of BI solution is the Presentation. The information that we have gathered should be presented in a way that is user-friendly and easy to understand for other people or the people who you are making it for. They present all the findings that we previously got by insight creation and can help the decision makers to make the best possible decision. Online Analytic Processing tool (OLAP) is an analytical processing system where different dimensions of data is extracted to later be analysed and later be used to help in the making of decisions in an organization [17], [19]. Florida Transportation's uses this to reduce the time it takes to analyse these data to save time and resources. Visualization is the act of portraying complex information in the form of graphs or other graphics. This is used by many organizations including Florida Transportations to display all their findings to the decisions makers to help them make understand the whole picture first before making a decision. They are usually conveyed as graph's, table's, tree map's and many more.

A dashboard is a customized interface that is navigable and is displayed using charts, tables, graphs and such. The users can acquire the information they need just by interacting with the dashboard. Similarly like with visualizations, dashboards help companies to make more informed decisions. Scorecards are also like dashboards, but they show and monitor performance by paying attention to certain metrics and comparing to a single target. Florida Transportation often uses this to compare the number of fatalities from traffic accidents between two different years to see if they have increased or decreased.

Business Performance Management (BPM) both monitors and manages organizational performance so that the organization goals can be reached effectively. It allows company to track what goes on in the company is real time in any given time. Florida Transportations uses this to check the progress towards the company's goals and whether they are heading in the right direction. For example. have the findings of the company resulted in less fatalities from car crashes from occurring?

6 Result and Discussion

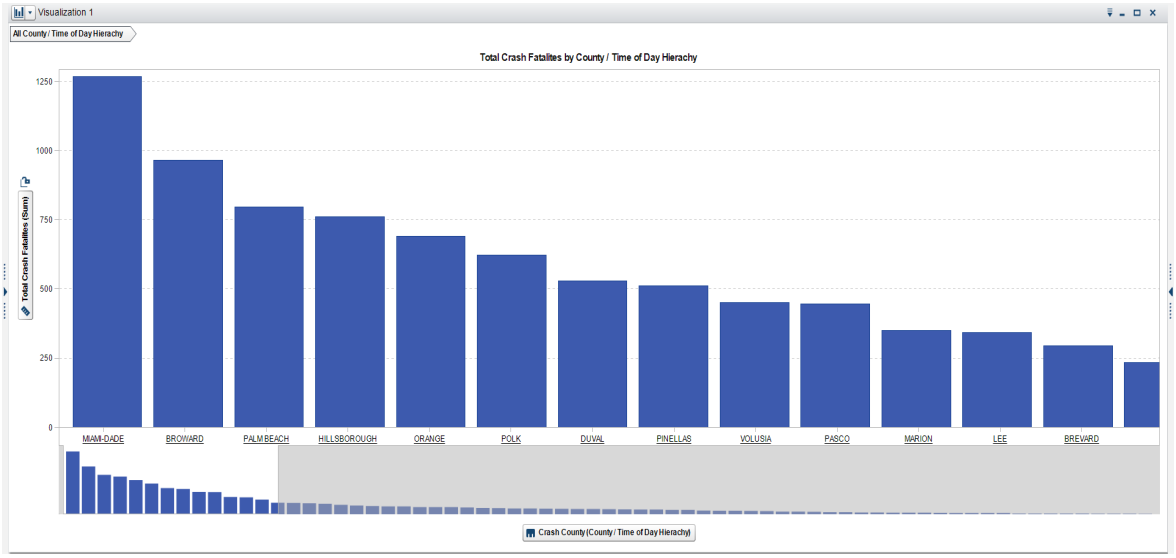

Fig. 1 Total Number of Fatalities

All of the data shown are the traffic accident report from 2008 ,2009 and 2010. Figure 1 shows the total number of fatalities that have occurred in the County of Florida. From this bar chart, it's clear that Miami-Dade has the highest number of casualties that is 1269 deaths from vehicle related accidents. Therefore, it is important for us to use more of our resources and efforts towards this specific county. This might have been caused by the fact that Miami-Dade is the County with the greatest number of people living in it therefore more people have motor

vehicles [5], [8], [10], [19]. Broward also has an alarming number of casualties at 965 deaths followed closely by Palm Beach with 797 casualties. Florida Transportations will use this information to organize more Road Accident Awareness Campaigns in the county's that are most affected by it.

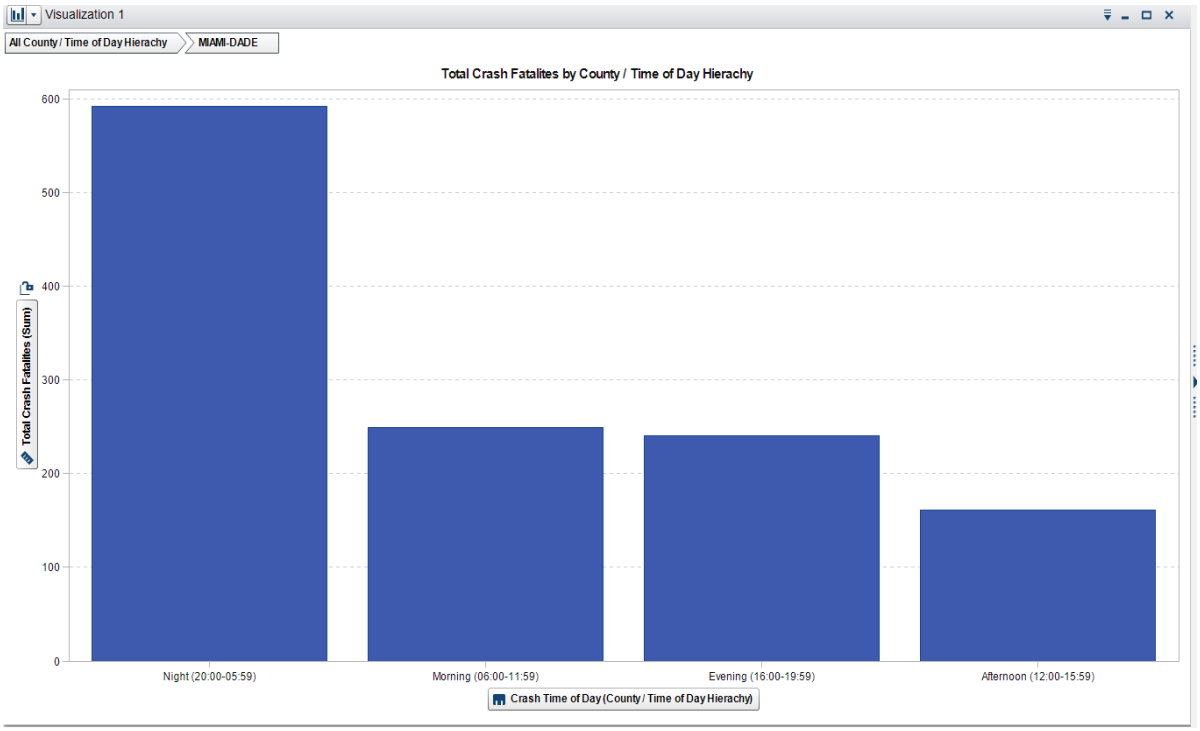

Fig. 2 Number of Crash Time of Day

We applied hierarchy to the first visualization to find out what time of day these traffic accidents occur the most. We can see from the chart that fatal accidents happen considerably more during night with 592 deaths occurring in the county, Miami-Dade. This might occur due to the fact that drivers that works long hours are often too fatigued to properly handle a vehicle or drivers not using their headlights properly or not even using them at all. The second deadliest time to be on the road is during the morning from 6.00 A.M to 11.59 AM with 249 deaths occurring. Most individuals with jobs start their shift at 9AM so there are a lot of vehicles on the road the time so the likelihood of accidents occurring will increase. The number of casualties is very similar on the evening at 240 casualties. Most people leave their work at 6 P.M so it can be assumed that there are also a lot of cars on the road at this time as well.

Decision makers can use this information to safely assume that the majority of accidents occur at night therefore employers should avoid letting their employees work overtime for too long and should place designated areas for their employees to rest so they can drive home safely after a long day at work [20], [21].

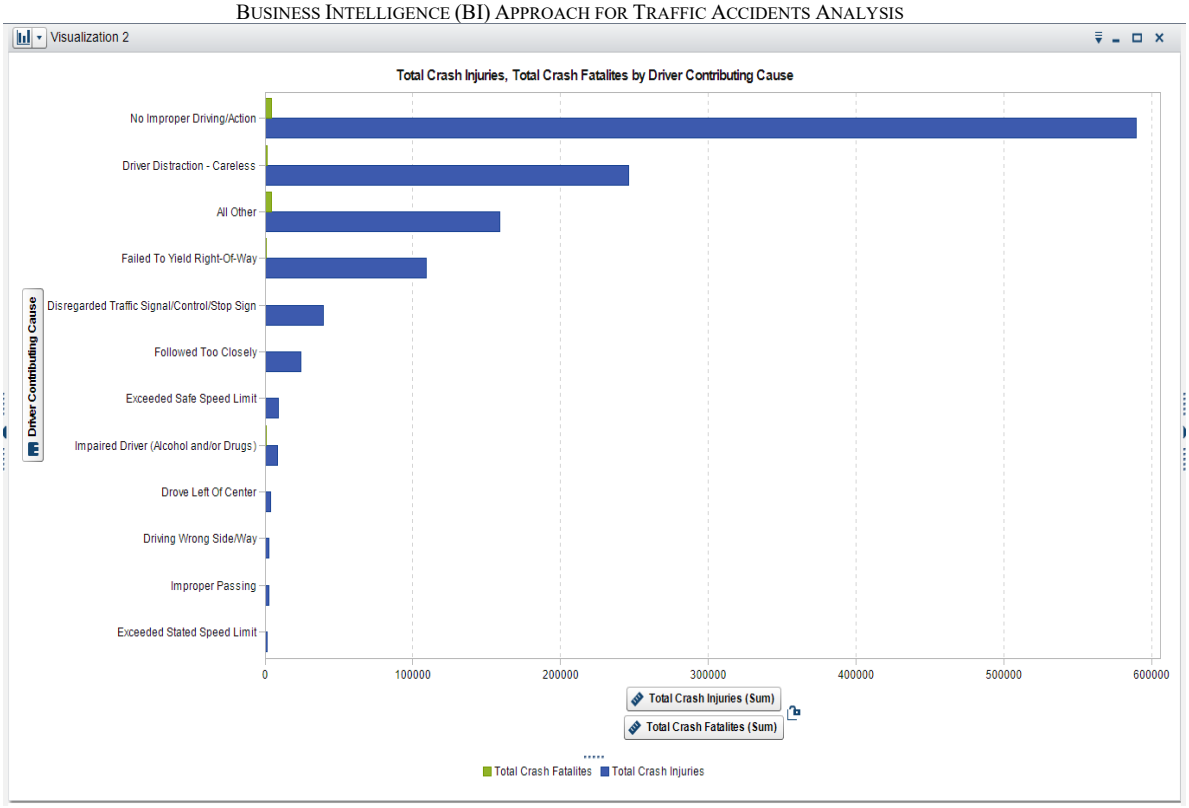

Fig. 3 The Main Accident Cause Contributor

Figure 3 shows what are the main contributing cause for these road accidents. Many accidents still occur even if the drivers don't show signs of improper driving or action on the road with the number being at 593,897 injuries and 4298. Driver distraction or driver carelessness is the second most important cause of traffic accidents. It has caused 247,800 injuries and 1337 fatalities. This cause is one of the most preventable causes of accidents, but many people die as a result of this. Therefore, it is most important that proper education is given to drivers by the organization in the hopes of it reducing the number of deaths from this cause. Failure to yield right-of-way on the other hand has caused 109,808 injuries and 665 deaths. It is no surprise that driving while under the influence of substances is very deadly. This has caused over 8228 injuries and 976 deaths. This is also an easily preventable situation therefore based this information, the organization can spread the awareness of not drinking while driving and work closely with the Police Department to catch these drunk drivers before they cause any unnecessary harm to themselves and others [21], [19].

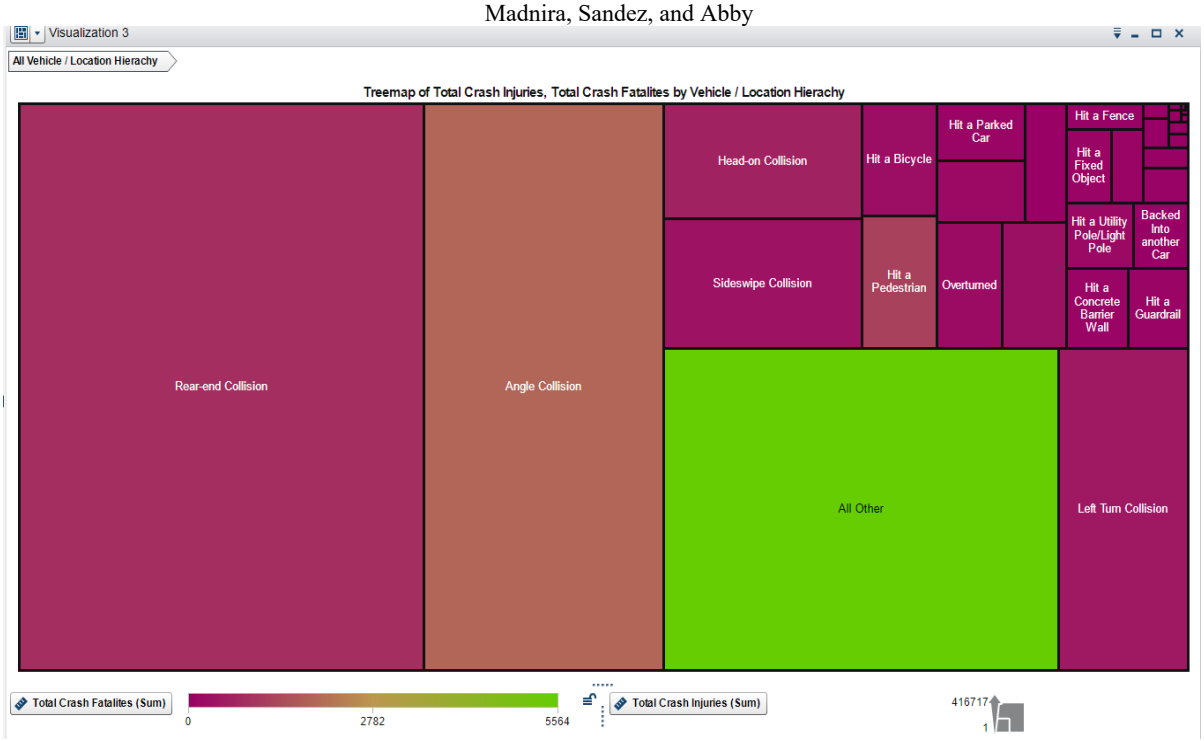

Fig. 4 Crash Fatalities

Figure 5 shows as the colour of the tree map goes towards the colour green, more deaths occur. This figure tells us which part of the vehicle is most affected when traffic accidents do occur. From the tree map above, we can conclude that rear-end collisions are the most common form of accidents that happen on the road with the numbers being 347,169 injuries and 1885. An angle collision is also very deadly with 416,717 injuries and 1885 casualties' deaths which is much more that rear end collisions. A head on collision is also very fatal with a death count of 634. This information can be used car manufacturers to figure out which part of the vehicles needs the most improvement in terms of safety features. The decision made by these companies based on the following data can definitely help reduce the number of casualties.

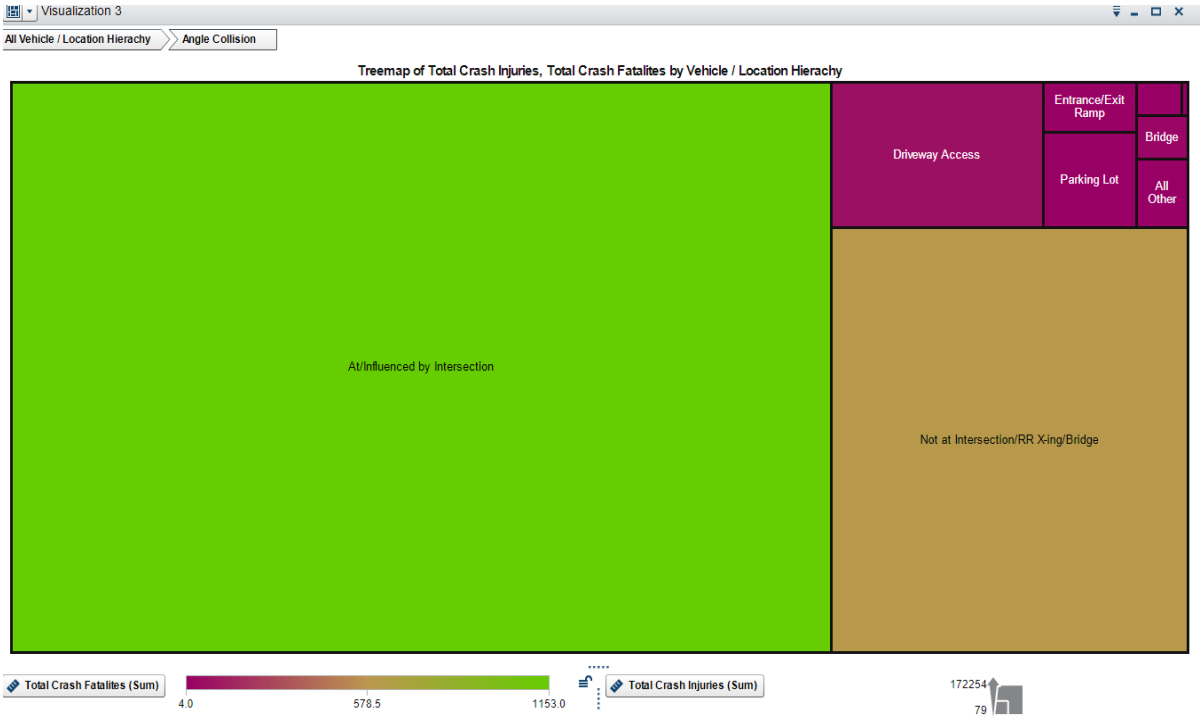

Fig. 5 Total Crash Fatalities

We have applied hierarchy to see the locations on which these accidents happen on the road. For angle collisions, they mostly happen at or are influenced by an intersection. 1153 deaths occurred at this location. The second deadliest location at 607 deaths are at other places besides an intersection that is at a railroad crossing or a bridge. A parking lot is mostly safe with only 4 casualties occurring ass most cars are stationary or are moving very slowly. The organization can use this information to help make important decisions like suggesting to the local government to put more traffic police at the locations that deaths happen the most. Doing this will most likely reduce the number of casualties from happening on these locations [17].

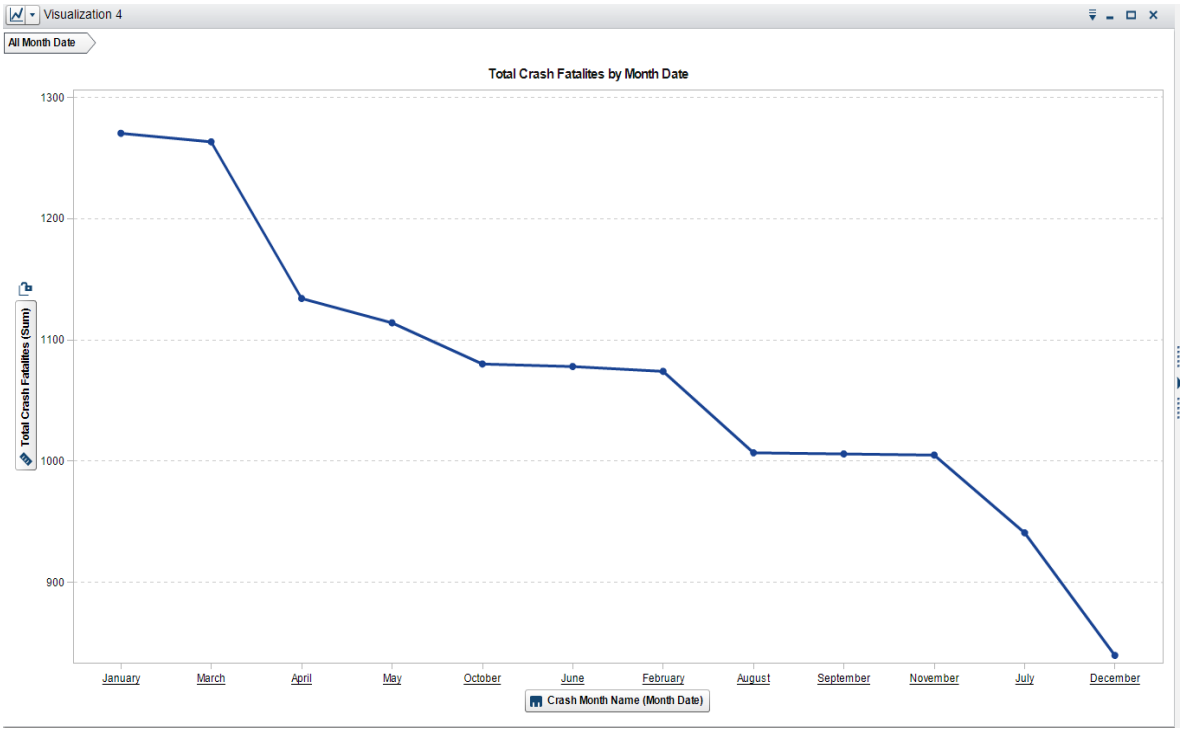

Fig. 6 Monthly Crash

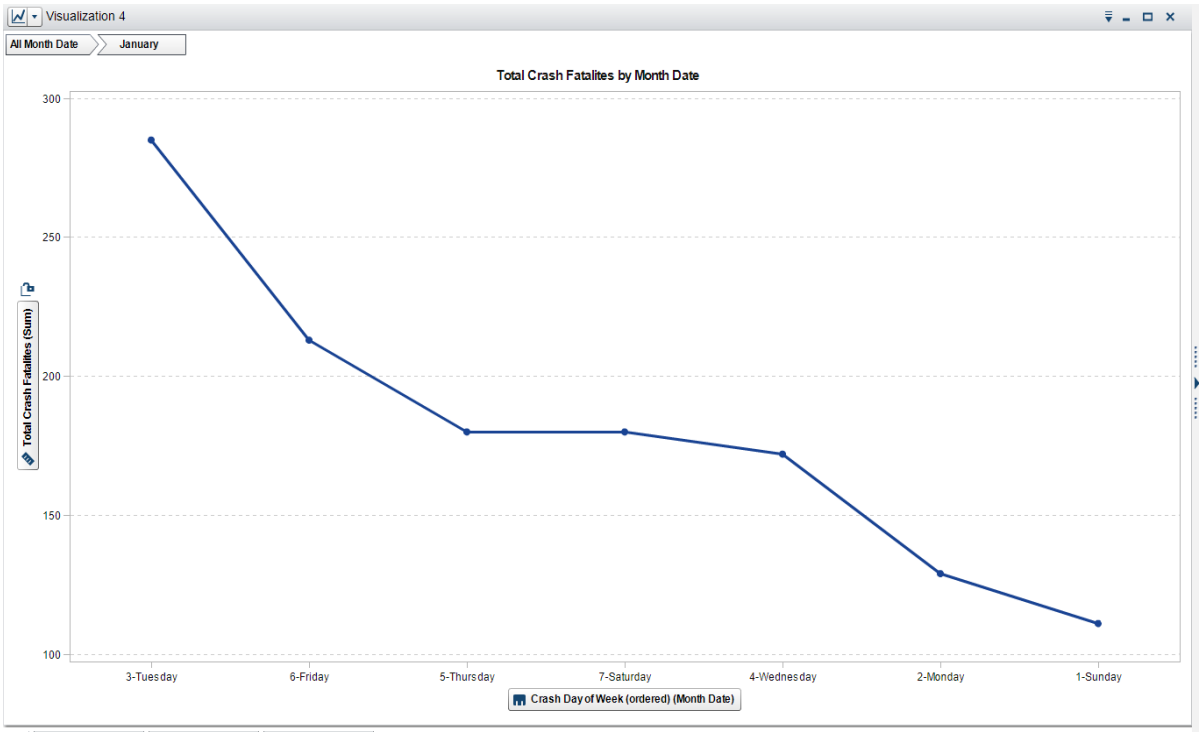

Fig. 7 Weekly Crash

Based on figure 6 and 7, January and March are the month with the most fatalities out of all the months with around 1260 deaths each. It can be assumed that more death from traffic accidents occur at the start of the year as December only has 829 deaths. After applying hierarchy, we can see that the day with the greatest number of fatalities in the month of January is a Tuesday followed closely by Friday. The reason for Friday having the second highest death toll could be explained by the fact that people might get too excited to go back home to start their weekend that they neglect their safety when driving back home [18]. The organization can use this information to predict which months and which day of the week these accidents occur most therefore being able to make the right decisions on when to use their resources most and act accordingly.

7 Challenges in Implementation

When wanting to implement a business intelligence solution, many problems or challenges would occur. In this organization for example, the first challenge is that the organization might not want to use business intelligence solutions as it is costly [16], [17], [19]. This is especially true for smaller businesses as they have limited funds to use in the first place. The next challenge is the lack of expertise in the field of business intelligence. If you want to benefit the most from using business intelligence solutions, then you need the right people. It might be hard for some companies to find these specific people that suits the organization's needs. Difficulties in driving user adoption is also a major challenge [20]. This organization might find it hard to change their old ways of analytics as there are so used to it and changing the way they run things since decades ago might seem daunting to them.

8 Best Practices

The solution we propose for the first challenge of organizations finding Business Intelligence solutions to be costly is having an expert in the field to justify why it is beneficial to the organization as a whole to implement Business Intelligence and that the cost to do so is worth it in the end. For organizations lacking in expertise, one solution is to have Business Intelligence to be accessible to everyone regardless of their level. This way more people will be familiar with Business Intelligence therefore it becomes more common within organizations [21]. For the next problem, to increase user adoption for Business Intelligence, the decision makers within the organization must be involved during the implementation process therefore they can understand better the importance of using Business Intelligence and how it will save them precious time that is otherwise wasted using traditional methods for example.

9 Conclusion

In conclusion, it is obvious that using Business Intelligence will greatly help Florida Transportations in reducing the amount of traffic accidents from occurring and as a result immensely reduce the number of deaths resulting from traffic accidents [17], [19], [22]. The decision-making process can be streamlined therefore decision makers can make better and wiser choices with the use of Business Intelligence in the organization. Moreover, with the use of the technologies in Business Intelligence, Florida Transportations will see great improvement in both their quality of work and the time taken to make important decisions will be reduced tremendously.

BIBLIOGRAPHY

[1]. Y. Su, "Game Analytics Research: Status and trends," Advances in E-Business Engineering for Ubiquitous Computing, pp. 572–589, 2019.

[2]. X. Wang, Y. Peng, S. Yi, H. Wang, and W. Yu, "Risky behaviors, psychological failures and kinematics in vehicle-to-powered two-wheeler accidents: Results from in-depth Chinese crash data," Accident Analysis & Prevention, vol. 156, p. 106150, 2021.

[3]. F. Basso, R. Pezoa, M. Varas, and M. Villalobos, "A deep learning approach for real-time crash prediction using vehicle-by-vehicle data," Accident Analysis & Prevention, vol. 162, p. 106409, 2021.

[4]. N. Veeramisti, A. Paz, and J. Baker, “A framework for corridor-level traffic safety network screening and its implementation using business intelligence,” Safety Science, vol. 121, pp. 100–110, 2020.
[5]. H. Surbakti and A. B. Ta'a, “Tacit Knowledge for Business Intelligence Framework: A Part of Unstructured Data?,” Journal of Theoretical and Applied Information Technology, vol. 96, no. 3, pp. 616–625, Feb. 2018.
[6]. B. Parashar and G. Rana, “The impact of artificial intelligence on Global Business Practices,” Reinventing Manufacturing and Business Processes Through Artificial Intelligence, pp. 95–113, 2021.
[7]. I. Fafaliou, M. Giaka, D. Konstantios, and M. Polemis, “Firms’ ESG reputational risk and market longevity: A firm-level analysis for the United States,” Journal of Business Research, vol. 149, pp. 161–177, 2022.
[8]. V. Mkrttchian, L. A. Gamidullaeva, and S. Panasenko, “Optimizing and enhancing digital marketing techniques in Intellectual Big Data Analytics,” Advances in Data Mining and Database Management, pp. 98–109, 2019.
[9]. A. A. Audu, O. F. Iyiola, A. A. Popoola, B. M. Adeleye, S. Medayese, C. Mosima, and N. Blamah, “The application of Geographic Information System as an intelligent system towards emergency responses in road traffic accident in Ibadan,” Journal of Transport and Supply Chain Management, vol. 15, 2021.
[10]. S. Das, X. Kong, and I. Tsapakis, “Hit and run crash analysis using association rules mining,” Journal of Transportation Safety & Security, vol. 13, no. 2, pp. 123–142, 2019.
[11]. K. Topuz and D. Delen, “A probabilistic bayesian inference model to investigate injury severity in automobile crashes,” Decision Support Systems, vol. 150, p. 113557, 2021.
[12]. M. Shahzad, “Review of road accident analysis using GIS technique,” International Journal of Injury Control and Safety Promotion, vol. 27, no. 4, pp. 472–481, 2020.
[13]. S. He, “Who is liable for the uber self-driving crash? analysis of the liability allocation and the regulatory model for Autonomous Vehicles,” Autonomous Vehicles, pp. 93–111, 2020.
[14]. M. Khder, “Web scraping or web crawling: State of Art, Techniques, approaches and application,” International Journal of Advances in Soft Computing and its Applications, vol. 13, no. 3, pp. 145–168, 2021.
[15]. A. Mohandu and M. Kubendiran, “Survey on big data techniques in intelligent transportation system (ITS),” Materials Today: Proceedings, vol. 47, pp. 8–17, 2021.
[16]. Y. Wang, W. Xu, W. Zhang, and J. L. Zhao, “SafeDrive: A new model for driving risk analysis based on Crash avoidance,” IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 3, pp. 2116–2129, 2022.
[17]. C. Lennerholt, J. Van Laere, and E. Söderström, “User-related challenges of self-service business intelligence,” Information Systems Management, vol. 38, no. 4, pp. 309–323, 2020.
[18]. M. K. Daradkeh, “Exploring the usefulness of user-generated content for business intelligence in Innovation,” International Journal of Enterprise Information Systems, vol. 17, no. 2, pp. 44–70, 2021.
[19]. M. D. Tamang, V. Kumar Shukla, S. Anwar, and R. Punhani, “Improving business intelligence through machine learning algorithms,” 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), 2021.
[20]. L. Wang, H. Zhong, W. Ma, M. Abdel-Aty, and J. Park, “How many crashes can connected vehicle and Automated Vehicle Technologies Prevent: A meta-analysis,” Accident Analysis & Prevention, vol. 136, p. 105299, 2020.
[21]. H. Ospina-Mateus, L. A. Quintana Jiménez, F. J. Lopez-Valdes, and K. Salas-Navarro, “Bibliometric analysis in Motorcycle Accident Research: A global overview,” Scientometrics, vol. 121, no. 2, pp. 793–815, 2019.
[22]. X. Zou, H. L. Vu, and H. Huang, “Fifty Years of Accident Analysis & Prevention: A Bibliometric and Scientometric Overview,” Accident Analysis & Prevention, vol. 144, p. 105568, 2020.