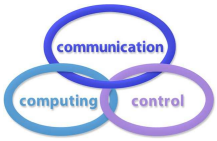
INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 20, Issue: 2, Month: April, Year: 2025
Article Number: 7017, https://doi.org/10.15837/ijccc.2025.2.7017



CCC Publications

Model for Analysis and Evaluation of road crashes resulting in fatalities using Business Intelligent Systems Approach

F. Córdova, C. Montt, N. Lagos

Felisa Córdova*
Facultad de Ingeniería, Arquitectura y Diseño
Universidad San Sebastián
Bellavista 7, Recoleta, 8420524 Santiago, Chile
felisa.cordova@uss.cl

Cecilia Montt
Facultad de Ingeniería
Universidad de Santiago de Chile
Ecuador 3769, Estación Central, 9160000
Santiago, Chile

Nicolás Lagos
Facultad de Ingeniería, Arquitectura y Diseño
Universidad San Sebastián
Bellavista 7, Recoleta 8420524
Santiago, Chile
*Corresponding author: felisa.cordova@uss.cl

Abstract

Road traffic accidents are the leading cause of death for young people in today's world. Given the proliferation of pedestrians, cyclists and motorcyclists in the city, it is necessary to find the most frequent causes to prevent these accidents. To resolve this situation problem an exploratory analysis of the databases is carried out to identify the main variables, their attributes and relationships. A Data Model is designed where the most frequent underlying causes of these accidents are identified, paying special attention to the actors involved, such as pedestrians, cyclists and motorcyclists. The integrated database is analyzed with a business intelligence approach to obtain behavioral patterns of the actors and determine the relevance of the variables involved in these events. The results obtained show that the main accidents with causes of death are collisions between vehicles and the crash, by not maintaining the prudent and reasonable distance caused by the physical conditions of the driver, whether fatigue and/or drunkenness. Since most accidents occur in the early morning or late afternoon, the use of illuminated beacons is recommended for cyclists and motorcyclists. Also, accident indicators are identified, and measures are recommended to improve the system. This work explores a method of Business Intelligence, with applications in the subject of study.

Keywords: traffic accidents, exploratory analysis, business intelligence approach.

1 Introduction

There are several ways to address traffic accidents, depending on the amount and complexity of the data available. Different authors propose a data mining framework for analyzing accident data identifying some of the factors that affect traffic accidents, using techniques data mining to predict the occurrence and severity of accidents [1]. Some explore the use of big data and data mining techniques for data analysis. It focuses on the visualization and prediction of accidents, highlighting the importance of integrating large volumes of data to improve the road safety [2],[3]. Also, there are other authors who analyze injury data in traffic accidents using data mining and machine learning. It focuses on identifying risk factors and patterns associated with accidents and serious injuries, proposing preventive measures to reduce accidents [4]. Advanced analysis is used to predict the severity of collisions, considering environmental and road infrastructure factors. Predictive models based on historical data are proposed to improve traffic risk assessment and management [5]. Other approaches discuss the differences in the reduction of traffic fatalities between countries for example between the U.S. and Germany, and provide an overview of international accident data collections [6]. Another study done in the USA uses decision trees and neural networks to analyze the severity of injuries in traffic accidents. It uses a real data set of police accident reports from a number of years seeking to identify patterns and determinants in the severity of accidents [7],[8]. An article was found that presents a spatial analysis model for road classification accident prone. It uses spatial data and artificial intelligence methods to identify high-risk areas and propose interventions to improve road safety [9].

Some articles offer valuable perspectives on the assessment and analysis of accidents in transport systems, such as a fuzzy logic-based approach to assessing the risk of accidents in transport systems, providing a tool to improve road safety [10]. Also, a study that applies machine learning techniques to analyze traffic accident data, identifying patterns and factors that contribute to the occurrence of accidents [11]. However, in the literature examined, no studies of road accidents are presented that use business intelligence tools to perform big data analysis and search for user behavior patterns, opening a good opportunity to explore the use of these advanced tools. In this context, Power BI tool allows to collect, manage, and analyze data from a variety of data sources present in sheets or databases, allowing to see not only what happened in the past and what is happening in the present, but also what might happen in the future, through predictions using time series techniques and exponential smoothing algorithms.

2 Methodology proposed

The methodology proposed in this study consist in the following four stages:

2.1 Data Acquisition

In this research, databases from 2017 to 2022 of traffic accidents were used, provided by the National Commission of Traffic Accidents of Chile CONASET [12],[13],[14] which are obtained in the field by units of Chilean Police. The databases are made up of four different Excel spreadsheets, these are: claims database, people, vehicles, and people – vehicles. These databases contain: type of accident, cause, whether the person is a pedestrian, driver or passenger, the type of vehicle, age of those involved, result, whether they are fatal, serious, less serious or minor, as well as other data that help identify the accident. To work with a single Excel database, the Power BI business intelligence tool was used to unify the data from the different databases.

2.2 Extraction of relevant data. Analysis of attributes

Once the databases are integrated, a data mining framework supported by Power BI allows the process of extracting, transforming and cleaning data to be carried out, for a region of the city of Santiago, which has several communes and has the highest number of accidents with serious and fatal results. The relevant actors involved in the accident rate are identified. Next, an analysis of the

relevance of variables is applied to show the attributes considered most influential and the incidence of the causes in the fatality and serious injuries of accidents of motorcyclists, cyclists and pedestrians. All this, considering the traffic accident data for 6 years (2017 to 2022), which account for the historical situation of accidents with fatal results [12], [13].

2.3 Obtaining patterns of behavior and indicators

Next, the data is exploited to obtain behavior patterns of the different actors involved in the accidents using the Power BI tool. Fatal accidents are characterized, highlighting rush hour hours, age ranges, type of user, type of road and cause. The most relevant indicators to assess accident rate are identified.

2.4 Measures for reducing causes

Different measures for trying to reduce the causes of these accidents are proposed.

3 Exploratory Analysis

3.1 Relevant variables and Attributes

Based on the exploratory analysis of the databases of the 6 years available, the following relevant variables and their associated attributes are identified and presented in Table 1.

Table 1: Relevant variables and its attributes

Category	**Detail and Attributes**
Driver – Passenger	Attributes: Sex, Age, Nationality, Deceased, Serious
People-Vehicles	**Detail: User Type** Attributes: Pedestrian, Cyclist, Motorist, Passenger. **Detail: Vehicle Type** Attributes: Automobile, Bicycle, Motorcycle, Other Vehicles. **Detail: Ages** Attributes: Under 15 years, 15 to 29 years, 30 to 39 years, 40 to 49 years, from 50 to 55 years, over 55
Accidents	**Detail: Actors involved in accidents** Attributes: Motor vehicle drivers, pedestrians,. bicycle and motorcycle drivers
Causes of the accident	**Detail: Environmental causes** Attributes: Overtaking, Speeding, Direction of traffic. **Detail: Human causes in passengers, driver, pedestrians** Attributes: Poor physical conditions, drowsiness, fatigue, under the influence of alcohol, in a state of drunkenness, drug use time of the accident. **Detail: Area where it is produced** Attributes: North, South, East, West, urban, rural.

3.2 Data Model

Once the data has been ordered, the data is modeled to obtain relationships between the causes of accidents involving bicycles, motorcycles and pedestrians, and other antecedents of those involved.

On the other hand, the location of accidents in the city is modeled, indicating the places where accidents with bicycles, motorcycles and pedestrians occur and their causes, which will allow road studies to be carried out and precautions to be taken to prevent the causes.

The Data Model implemented in Power BI considers four driver files:

- T_Persona (People) describing the characteristics of people with their attributes present in Table 1.
- T_RU that considers year, commune, day, season of the year, date, time, time, identification of the accident, month.
- T_Veh considers the type of vehicles involved in accidents, consequences and maneuvers with their attributes present in Table 1.
- T_Per_Veh considers damage to people, deceased, serious, minor, the time of the accident, day, time and month.

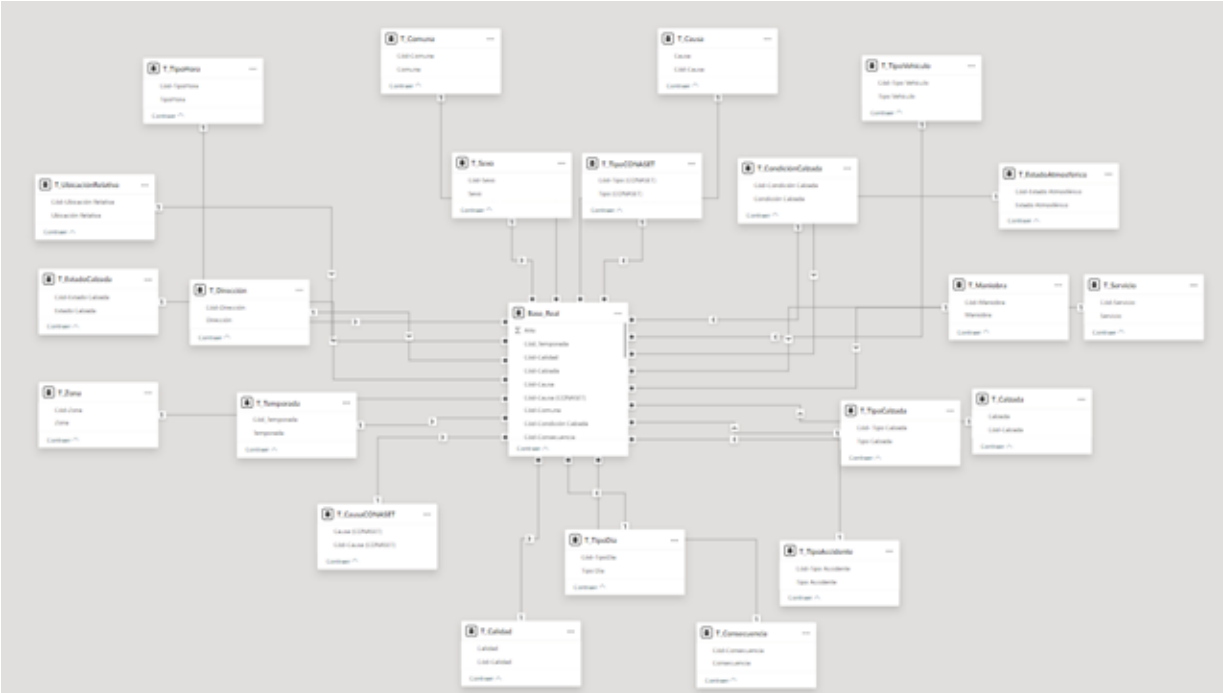


Figure 1: Data Model (Power BI)

4 Results

4.1 Analysis of accidents by zone

Figure 2 shows the number of accidents that occurred in the city during the 6-year sample present in the CONASET databases, from a cumulative total of 184,001 between 2017 and 2023. In the examination of accidents that occurred in the 5 zones of the city (North, South, West, East, Center) it was determined that the most affected area was the South Zone every year as shown in Figure 3 with 58,109 accidents per year. These communities are populous and have medium-low socioeconomic levels. Figure 4 presents the two communes Santiago and Puente Alto with the highest number of accidents in the Metropolitan Region (MR). The communes with high accidents are the most populated

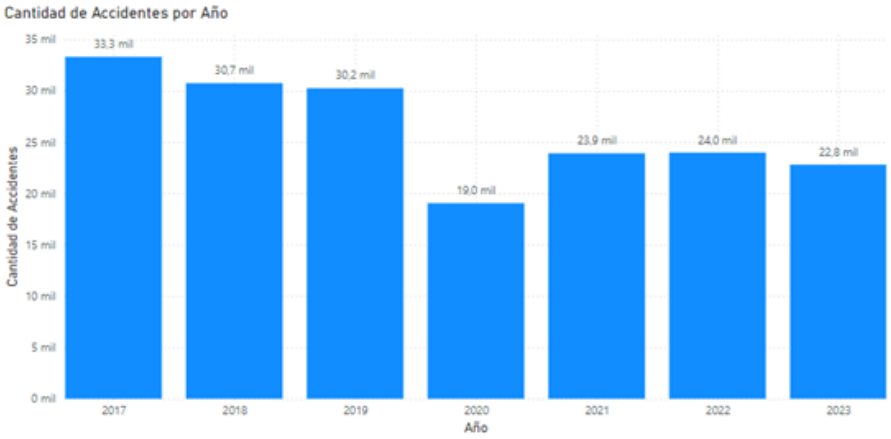


Figure 2: Number of accidents per year (2017-2023)

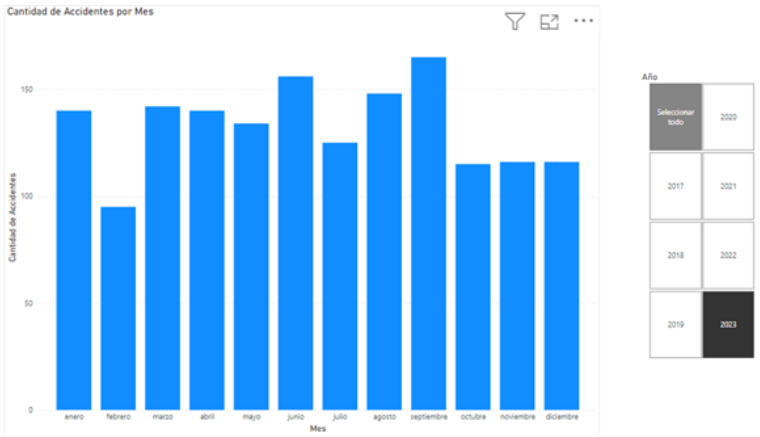


Figure 3: Accidents of the South Zone

in the MR and represent almost 40% of the population. That is the main reason for the increased occurrence of accidents. On the other hand, they are poor communes in bicycle lane infrastructure, it is common in non-working hours to reduce accidents because there is very little movement of vehicles and pedestrians.

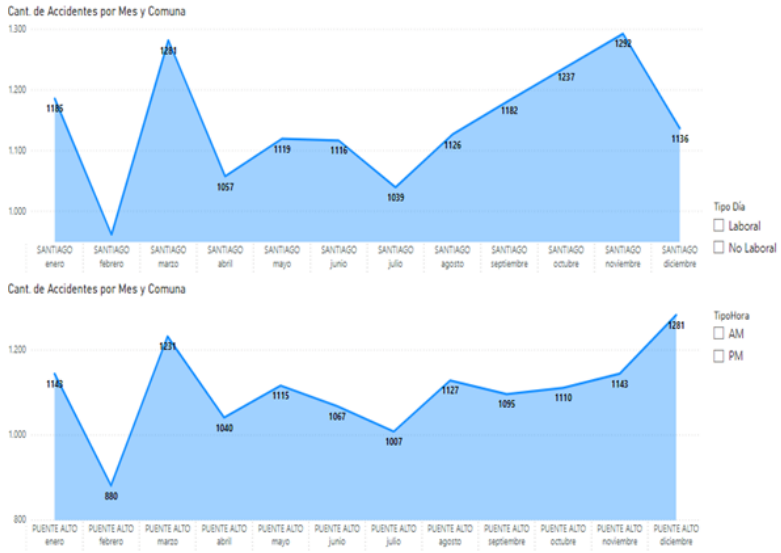


Figure 4: Communes of Santiago and Puente Alto with the highest number of accidents)

Figure 5 shows the accidents with the second highest number of accidents in Commune of Maipu.

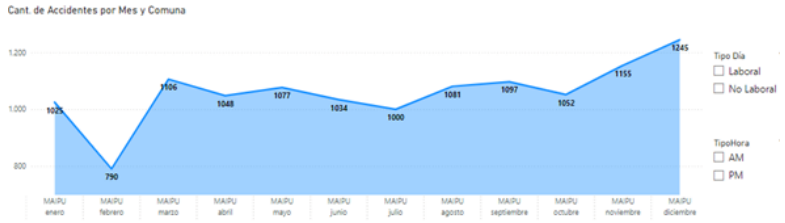


Figure 5: Commune of Maipu with the second highest number of accidents

A summary of months and work and non-work hours of a commune with a high accident rate is shown in Figure 6.

Comuna	SANTIAGO							Total
Tipo Día	Laboral			No Laboral			Total	
Mes	AM	PM	Total	AM	PM	Total		
enero	274	477	751	143	291	434	1.185	1.185
febrero	190	373	563	138	260	398	961	961
marzo	282	489	771	197	313	510	1.281	1.281
abril	253	405	658	158	241	399	1.057	1.057
mayo	293	385	678	171	270	441	1.119	1.119
junio	247	418	665	167	284	451	1.116	1.116
julio	225	397	622	165	252	417	1.039	1.039
agosto	289	432	721	132	273	405	1.126	1.126
septiembre	255	451	706	165	311	476	1.182	1.182
octubre	275	478	753	184	300	484	1.237	1.237
noviembre	335	485	820	185	287	472	1.292	1.292
diciembre	270	447	717	140	279	419	1.136	1.136
Total	3.188	5.237	8.425	1.945	3.361	5.306	13.731	13.731

Comuna	PUENTE ALTO							Total
Tipo Día	Laboral			No Laboral			Total	
Mes	AM	PM	Total	AM	PM	Total		
enero	208	452	660	163	320	483	1.143	1.143
febrero	143	357	500	121	259	380	880	880
marzo	251	481	732	172	327	499	1.231	1.231
abril	220	395	615	142	283	425	1.040	1.040
mayo	231	422	653	158	304	462	1.115	1.115
junio	221	377	598	150	319	469	1.067	1.067
julio	193	350	543	138	326	464	1.007	1.007
agosto	222	411	633	141	353	494	1.127	1.127
septiembre	224	416	640	145	310	455	1.095	1.095
octubre	214	424	638	142	330	472	1.110	1.110
noviembre	234	428	662	167	314	481	1.143	1.143
diciembre	221	479	700	186	395	581	1.281	1.281
Total	2.582	4.992	7.574	1.825	3.840	5.665	13.239	13.239

Comuna	MAIPU						Total	Total
Tipo Día	Laboral			No Laboral				
Mes	AM	PM	Total	AM	PM	Total		
enero	200	425	625	147	253	400	1.025	1.025
febrero	144	304	448	108	234	342	790	790
marzo	265	383	648	172	286	458	1.106	1.106
abril	242	372	614	151	283	434	1.048	1.048
mayo	293	366	659	141	277	418	1.077	1.077
junio	215	350	565	156	313	469	1.034	1.034
julio	183	392	575	154	271	425	1.000	1.000
agosto	226	390	616	160	305	465	1.081	1.081
septiembre	204	402	606	164	327	491	1.097	1.097
octubre	233	383	616	149	287	436	1.052	1.052
noviembre	231	449	680	134	341	475	1.155	1.155
diciembre	227	480	707	176	362	538	1.245	1.245
Total	2.663	4.696	7.359	1.812	3.539	5.351	12.710	12.710

Figure 6: Summary of months and work and non-work hours of a commune with a high accident rate.

The accident trends per year considering quarterly information is presented in Figure 7 showing a decrease only during the pandemic year and remained stable in previous years over 30,000 accidents.

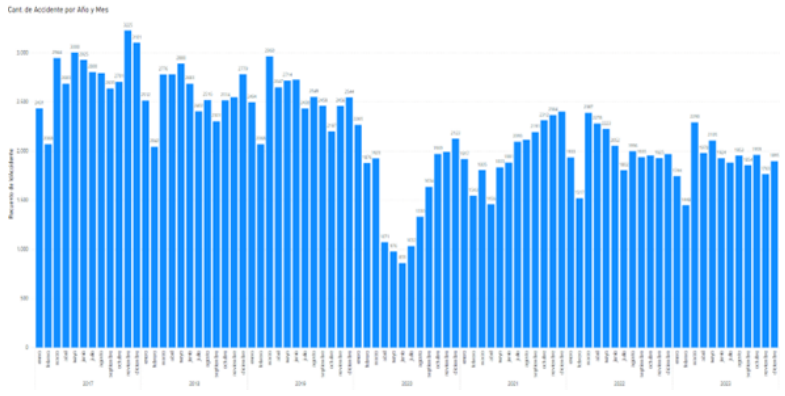


Figure 7: Total of Acident Trends by Years (2017-2023)

4.2 Analysis of People Vehicles accidents

Figure 8 presents the people vehicles accident (2023) in the Commune of Santiago, Figure 9 in the commune of Puente Alto and Figure 10 in the commune of Maipu.

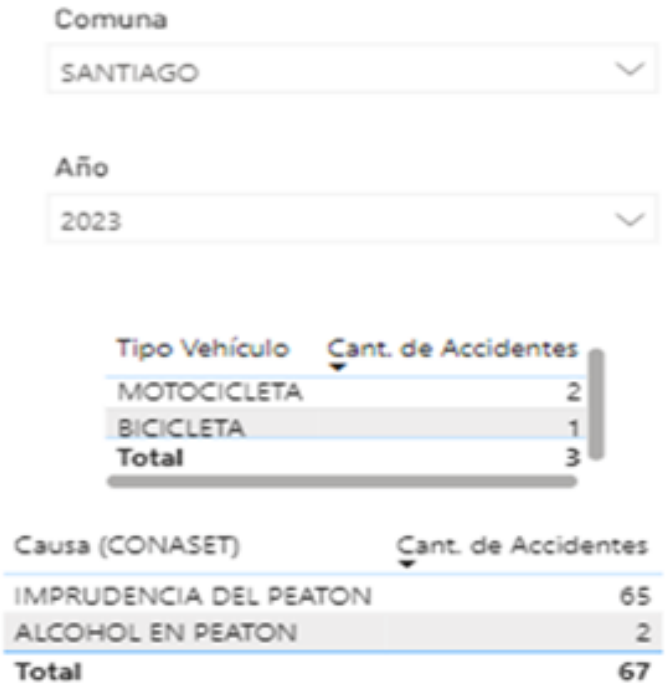


Figure 8: Total of Accident Trends by Years (2017-2023)

In these cases, pedestrian accidents are greater because they are the most unprotected, they involve cars since these are a high percentage greater than bicycles, motorcycles and pedestrians.

4.3 Analysis of Causes of accidents

Figures 11, 12 and 13 presents the most important causes of fatality and serious injury in female and male from the Communes of Santiago, Puente Alto and Maipu which are caused by different physical conditions of the drivers, such as the influence of alcohol and drugs, drunkenness, fatigue, in a normal state and in case of fleeing. The vehicles considered in the accidents are truck, jeep, motorcycle, bicycle, motorized skating, scooter/ Bike Motorcycle, car and taxi. The main causes correspond to outrage, fall, crash, collision, other types and rollover.

The largest number of accidents occur due to outrage, crash and collision, the latter being the source of serious injury. Death results are also caused by outrage and collision.

4.4 Accidents rate indicators

Performance Indicators are considered in Governance Solutions [15]. Relevant indicators were detected in this study to evaluate the accident rate of pedestrians and bicycle and motorcycle drivers, such as:

- Injury rate: The number of injuries sustained by pedestrians, bicyclists, and drivers in traffic accidents.
- Fatality rate: the number of deaths caused by road accidents in relation to the population or number of vehicles in circulation.
- Factors contributing to accidents: such as speeding, alcohol, driver distraction, lack of visibility, or poor road infrastructure design, can help identify areas for improvement.

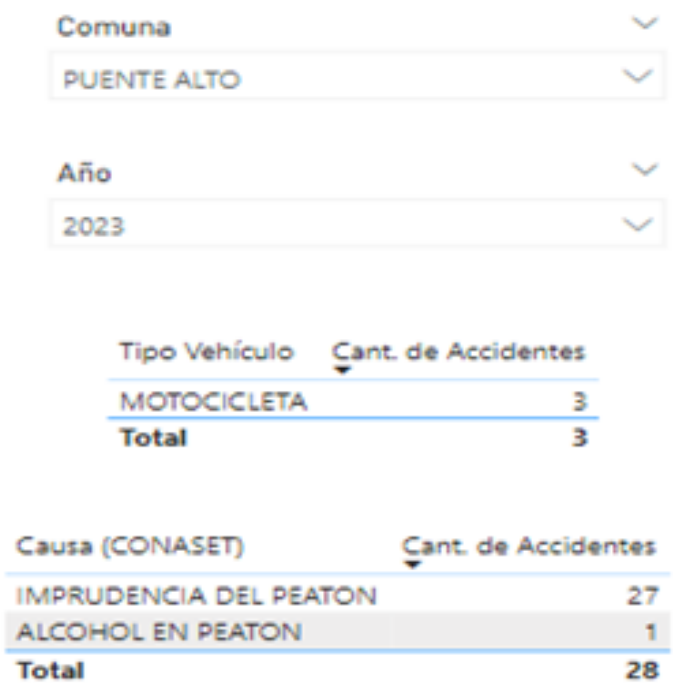


Figure 9: Causes of fatality and serious injury in female and male from the Commune of Santiago

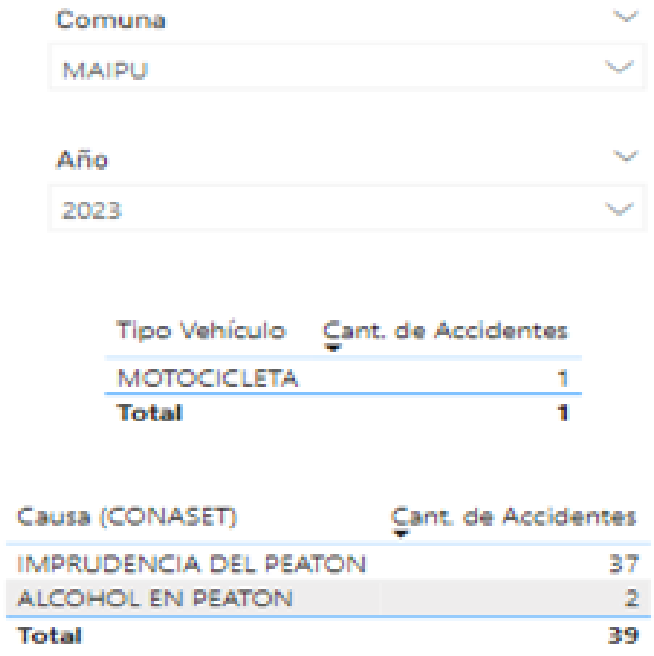
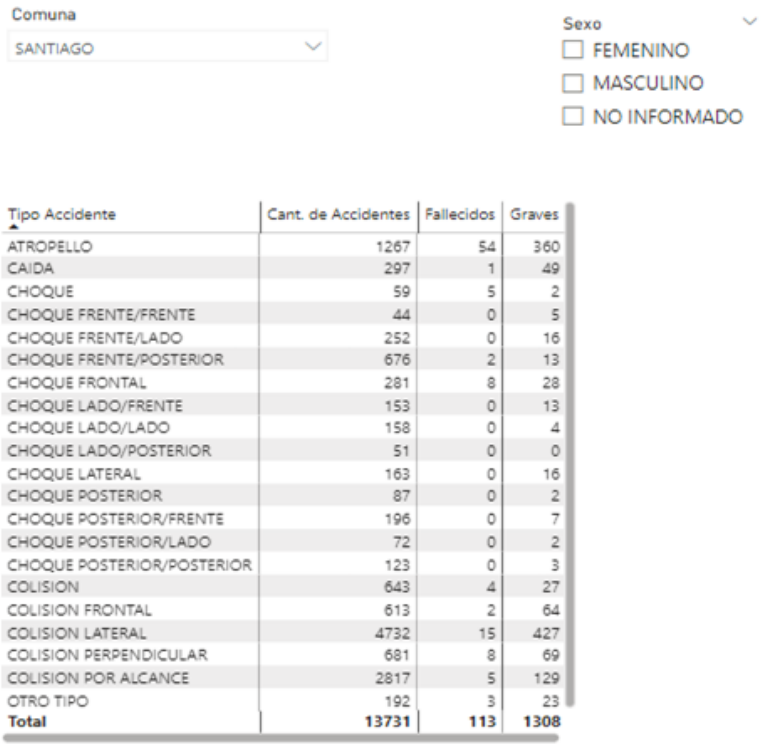


Figure 10: Causes of fatality and serious injury in female and male from the Commune of Puente Alto

- Temporal trends: Analyzing trends over time can provide insight into whether the safety measures in place are having a positive impact on reducing accidents.

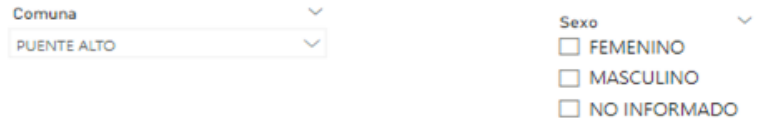
4.5 Recommendations for improving the analysis

It is recommended to propose the following data to be recorded in future investigations on causes:



Tipo Accidente	Cant. de Accidentes	Fallecidos	Graves
ATROPELLO	1267	54	360
CAIDA	297	1	49
CHOQUE	59	5	2
CHOQUE FRENTE/FRENTE	44	0	5
CHOQUE FRENTE/LADO	252	0	16
CHOQUE FRENTE/POSTERIOR	676	2	13
CHOQUE FRONTAL	281	8	28
CHOQUE LADO/FRENTE	153	0	13
CHOQUE LADO/LADO	158	0	4
CHOQUE LADO/POSTERIOR	51	0	0
CHOQUE LATERAL	163	0	16
CHOQUE POSTERIOR	87	0	2
CHOQUE POSTERIOR/FRENTE	196	0	7
CHOQUE POSTERIOR/LADO	72	0	2
CHOQUE POSTERIOR/POSTERIOR	123	0	3
COLISION	643	4	27
COLISION FRONTAL	613	2	64
COLISION LATERAL	4732	15	427
COLISION PERPENDICULAR	681	8	69
COLISION POR ALCANCE	2817	5	129
OTRO TIPO	192	3	23
Total	**13731**	**113**	**1308**

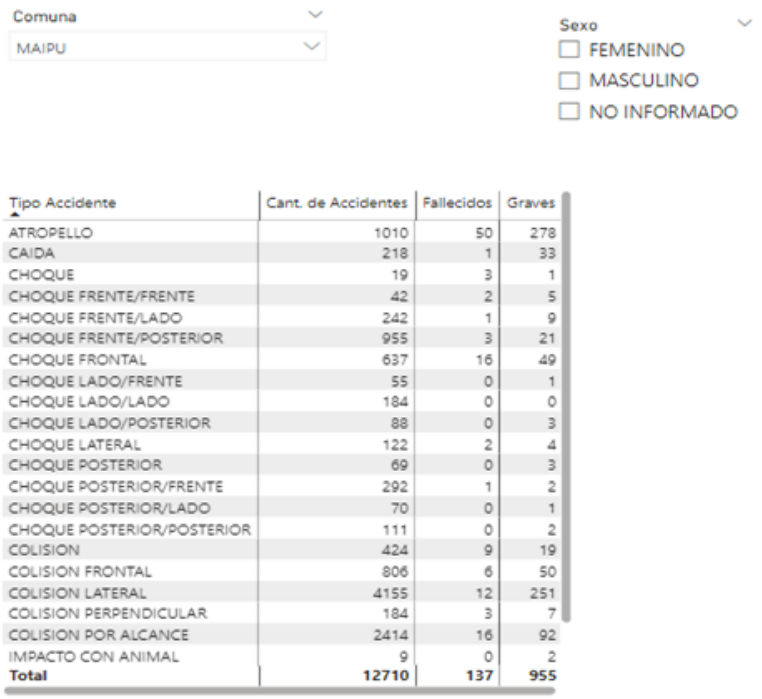
Figure 11: Causes of fatality and serious injury in female and male from the Commune of Maipu



Tipo Accidente	Cant. de Accidentes	Fallecidos	Graves
ATROPELLO	1199	40	291
CAIDA	298	0	44
CHOQUE	22	2	1
CHOQUE FRENTE/FRENTE	109	0	4
CHOQUE FRENTE/LADO	549	0	8
CHOQUE FRENTE/POSTERIOR	1603	1	31
CHOQUE FRONTAL	829	20	61
CHOQUE LADO/FRENTE	54	0	1
CHOQUE LADO/LADO	181	0	1
CHOQUE LADO/POSTERIOR	25	0	2
CHOQUE LATERAL	98	1	2
CHOQUE POSTERIOR	52	0	0
CHOQUE POSTERIOR/FRENTE	258	0	3
CHOQUE POSTERIOR/LADO	63	0	2
CHOQUE POSTERIOR/POSTERIOR	49	0	0
COLISION	187	3	14
COLISION FRONTAL	753	9	58
COLISION LATERAL	3281	10	178
COLISION PERPENDICULAR	1389	12	121
COLISION POR ALCANCE	1699	8	52
IMPACTO CON ANIMAL	17	0	3
Total	**13239**	**114**	**928**

Figure 12: Causes of fatality and serious injury in female and male from the Commune of Puente Alto

- Accurately track and height of all events.
- Pedestrian status at the time of the accident (sober, drunk, intoxicated).
- Driver use of cell phone.
- Speeding of the vehicle involved in a hit-and-run.
- Use of high-visibility clothing.



Tipo Accidente	Cant. de Accidentes	Fallecidos	Graves
ATROPELLO	1010	50	278
CAIDA	218	1	33
CHOQUE	19	3	1
CHOQUE FRENTE/FRENTE	42	2	5
CHOQUE FRENTE/LADO	242	1	9
CHOQUE FRENTE/POSTERIOR	955	3	21
CHOQUE FRONTAL	637	16	49
CHOQUE LADO/FRENTE	55	0	1
CHOQUE LADO/LADO	184	0	0
CHOQUE LADO/POSTERIOR	88	0	3
CHOQUE LATERAL	122	2	4
CHOQUE POSTERIOR	69	0	3
CHOQUE POSTERIOR/FRENTE	292	1	2
CHOQUE POSTERIOR/LADO	70	0	1
CHOQUE POSTERIOR/POSTERIOR	111	0	2
COLISION	424	9	19
COLISION FRONTAL	806	6	50
COLISION LATERAL	4155	12	251
COLISION PERPENDICULAR	184	3	7
COLISION POR ALCANCE	2414	16	92
IMPACTO CON ANIMAL	9	0	2
Total	**12710**	**137**	**955**

Figure 13: Causes of fatality and serious injury in female and male from the Commune of Maipu

A study of the accident rate of motorcyclists is recommended, analyzing maneuver performed, result of this, area of it (commercial and residential), speed limit allowed level of experience of the driver, relevant driver's license, use of safety implements and type of motorcycle (scooter, scooter, etc.). On the other hand, a virtual reality awareness campaign on driving with alcohol or speeding is recommended.

5 Conclusion

This work introduces a business intelligence approach and tools to perform big data analysis and search for user behavior patterns in road crushes resulting in fatalities and serious injuries. In the exploratory analysis relevant variables and their associated attributes were identified. The Data Model implemented in Power BI made it possible to obtain the relationships between actors, causes and effects of accidents in the most affected city communes. In the examination of accidents that occurred in the 5 zones of the city, it was determined that the most affected area was the South Zone. This work, unlike others that appear in the literature, makes use of combined Big Data analysis and Business Intelligence tools, not frequently used in this type of study. This work shows that, as the number of bicycles and motorcycles in the country has increased, accidents have also increased. It has been shown that the majority of pedestrian accidents result in death and serious injury and that the most important cause is pedestrian recklessness. Given the above, it is recommended to increase the regulations for the use of them and thus avoid accidents. In addition, vehicle lights should be turned on before sunset to prevent pedestrian deaths. In the next studies, the causes and conditions of the roads will be analysed in greater detail to provide municipalities with road safety support.

Funding

This work was supported by Vicerrectoría de Investigación y Doctorados de la Universidad San Sebastián – Fondo USS-FIN-25-APCS-06

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Das, A.; Sun, X. (2014). A Data Mining Framework for Analyzing Road Accident Data. *Journal of Transportation Safety & Security* 6(3), 180-203. doi:10.1080/19439962.2013.839929

[2] Anbaroglu, B.; Heydecker, B. G.; & Cheng, T. (2014). Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. *Transportation Research Part C: Emerging Technologies*, 48, 47-65. doi:10.1016/j.trc.2014.08.009

[3] Aggarwal, C. C. (2015). Data Mining: The Textbook. Londres: Springer. doi:10.1007/978-3-319-14142-8

[4] Yoon, Y.; Bai, Y. (2016). Injury Severity in Traffic Crashes with Light Commercial Vehicles. *Journal of Safety Research*, 59, 23-29. doi:10.1016/j.jsr.2016.09.006

[5] Xu, C.; Liu, X.; Jin, Z. (2018). Predicting Traffic Collision Severity: A Hybrid Deep Learning Framework. *Accident Analysis & Prevention*, 121, 338-347. doi:10.1016/j.aap.2018.06.014

[6] Broughton, J.; Baughan, C. J. (2002). The Effect of Seat Belt Usage on the Risk of Injury in Traffic Accidents: A Comparison of Data from the United States and Germany. *Accident Analysis & Prevention*, 34(2), 173-182. doi:10.1016/S0001-4575(01)00009-8

[7] Chen, C.; Jovanis, P. P. (2000). Method for Identifying Factors Contributing to Driver-Injury Severity in Traffic Accidents Involving Large Trucks. *Transportation Research Record: Journal of the Transportation Research Board*, 1717, 1-7. doi:10.3141/1717-01

[8] Boo, Y.; Choi, Y. (2021). Comparison of Prediction Model for Mortality Related to Injuries from Road Traffic Accidents after Correcting for Undersampling. *International Journal of Environmental Research and Public Health*, 18(5204). doi:https://doi.org/10.3390/

[9] Anderson, T. K. (2009). Kernel Density Estimation and K-means Clustering to Profile Road Accident Hotspots. *Accident Analysis & Prevention*, 41(3), 359-364. doi:10.1016/j.aap.2008.12.014

[10] Doe, J.; Smith, J. (2020). A Fuzzy Logic-Based Approach for Accident Risk Assessment in Transportation Systems. *International Journal of Computers, Communications and Control*, 15(3), 345-356.

[11] Brown, A.; Johnson, R. (2022). An Analysis of Traffic Accident Data Using Machine Learning Techniques. *International Journal of Computers, Communications and Control*, 17(2), 210-220.

[12] CONASET. (2019). Perfil de los fallecidos en siniestros de tránsito. Santiago de Chile. [Online]. Available: https://www.conaset.cl/wp-content/uploads/2020/08/Perfil-Fallecido-2019.pdf

[13] CONASET. (2020). Costo social de los siniestros de tránsito en Chile. Santiago de Chile. [Online]. Available: https://www.conaset.cl/wp-content/uploads/2020/11/Costos-accidentes-2019.pdf

[14] Montt, C.; Rubio, J. M.; Lanata, S. (2013). Análisis de accidentes de tránsito con inteligencia computacional. Congreso Chileno de Ingeniería de Transporte. [Online]. Available: https://congresotransporte.uchile.cl/index.php/CIT/article/view/28446/30168

[15] ITF OECD. (2019). Road Safety in European Cities. Performance Indicators and Governance Solutions. [Online]. Available: https://www.itf-oecd.org/sites/default/files/docs/road-safety-european-cities-performance-indicators.pdf

Journal's webpage: http://univagora.ro/jour/index.php/ijccc/

This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).
https://publicationethics.org/members/international-journal-computers-communications-and-control

Cite this paper as:

Córdova, F.; Montt, C.; Lagos, N. (2025). Model for Analysis and Evaluation of road crashes resulting in fatalities using Business Intelligent Systems Approach, *International Journal of Computers Communications & Control*, 20(2), 7017, 2025.
https://doi.org/10.15837/ijccc.2025.2.7017