

# Spatiotemporal Analysis of Traffic Accidents and Traffic Congestion Along Commonwealth Avenue Using Crowdsourced Waze Traffic Data

Simon Reinier D. Castro<sup>1</sup>, Claress Anne C. Tadle<sup>1</sup>, Bienvenido G. Carcellar III<sup>1</sup>

<sup>1</sup> Department of Geodetic Engineering, University of the Philippines - Diliman, Quezon City -  
simonreiniercastro@gmail.com, annetadle1501@gmail.com, bgcarcellar@up.edu.ph

**Keywords:** Waze, crowdsourced data, traffic accidents, traffic congestion, KDE

## Abstract:

With the growing demand for transportation and traffic management in Metro Manila, innovation in these industries is essential. Crowdsourced traffic data, particularly from smartphones, has emerged as an accurate and reliable source. This study utilized crowdsourced traffic data from Waze, through the Waze for Cities (WFC) program, to explore the relationship between traffic accidents and congestion. Focusing on Commonwealth Avenue, an accident-prone highway, and by applying Kernel Density Estimation (KDE), six areas of interest (AOIs) were identified. These areas align with more general accident-prone zones identified by the Metro Manila Development Authority (MMDA) through the Metro Manila Accident Recording and Analysis System (MMARAS) Annual Report. The study examined the behavior of traffic jams in these areas, considering variables such as congestion extent, additional time delay, and travel speed. Results showed that the presence of accidents significantly increased the speed difference between regular and jammed conditions, as indicated by the Mann-Whitney U test at a 5% significance level, except in two areas of interest (AOI) near Quezon Memorial Circle, which confirms the potential of using Waze data in assessing traffic conditions. From the findings, it is suggested that Waze data be integrated with traditional data sources used by government agencies for more effective traffic management, enhancing the accuracy and responsiveness of traffic assessments, interventions, and policy implementations.

## 1. Introduction

### 1.1. Background of the Study

Plagued by the negative effects of centralized employment, severe traffic congestion in Metro Manila has been a complex, long-standing problem caused by inadequate road infrastructure and lackluster implementation of traffic mitigation policies leading to an inefficient public transport system. Unsurprisingly, it has consistently been described to be one of the worst cities in terms of urban mobility readiness and traffic indices (Thibault et al., 2022). As such, an increasing number of commuters opt for more convenient options such as ride-hailing services or even investing in personal vehicles themselves, which only increases the volume of traffic and, proportionately, traffic accidents (World Health Organization, 2023). There lies much to be desired from the way traffic data was being handled, leading to the complexity of the problem, but fortunately, recent innovations such as the Waze For Cities Program (WFC) provide supplementary data which are both accurate and precise. In utilizing a crowdsourced database, traffic accidents on the road may be monitored in real time through its users and their reports. Commonwealth Avenue, one of the largest and busiest major thoroughfares in Metro Manila, was deemed a suitable area for the study due to its large daily traffic volume, increasing the probability of the presence of Waze users and user-submitted accident reports.

### 1.2. Research Objective and Significance

The main objective of the study was to establish the relationship between the presence of accident reports and the detected jams from Waze Traffic Data reported within Commonwealth Avenue throughout the year of 2022. Accident-prone road segments were then identified as areas of interest (AOI) in order to observe the distribution of these reports segregated into each hour of the day (HoD) per day of

the week (DoW). The effect of jams in relation to traffic accidents were then observed based on the average length of the jams, delay in seconds, and average speed, all of which have been tested for significance using statistical analysis. In doing so, these tangible parameters pertaining to the relationship of traffic accidents with traffic congestion may be further explored.

## 2. Review of Related Literature

### 2.1. Current Traffic Monitoring Practices

Recent innovations in technology have expanded the means of collecting road traffic data with greater accuracy and precision (Shahgolian & Gharavian, 2018). However, traffic monitoring in the Philippines has yet to keep up. Particularly, on Commonwealth Avenue, Metro Manila Development Authority (MMDA) personnel deployment and closed-circuit television (CCTV) surveillance cameras were still the primary means of traffic monitoring, both limited by access, reliability, and cost (Zhang et al., 2021). With a car-centric and road-based transportation system, there lies a great need for study of efficient means of traffic management other than the conventional approaches.

### 2.2. Benefits of Crowdsourced Data and Waze

The abundance of smartphones among Filipinos, consequently boasting one of the highest number of active Waze users, suggests the plausibility of utilizing its citizens as sensors as a proactive means to foster a communal awareness for traffic management through engagement (Mukheja et al., 2017). Despite being limited to Waze users, it allows for accurate and precise traffic monitoring, assessing traffic jams for their lengths and the caused delay in time, comparing it to free-flowing conditions. In fact, the number of reported traffic accidents from Waze matches the Metro Manila Accident

Recording and Analysis System (MMARAS) statistics of Commonwealth Avenue (CA) for the year of 2022 recorded by the MMDA (MMDA, 2022). Although the submitted reports came from users with their own biases, the similarity with MMARAS statistics suggests that these adequately represent the traffic situation of the study area.

### 3. Methodology

As shown in Figure 1, the study encompassed several pivotal stages. The data collection, cleaning, and aggregation were processed through a Python script, while the processing of the spatial data further was done in GIS software. Results were analyzed and visually presented through maps.

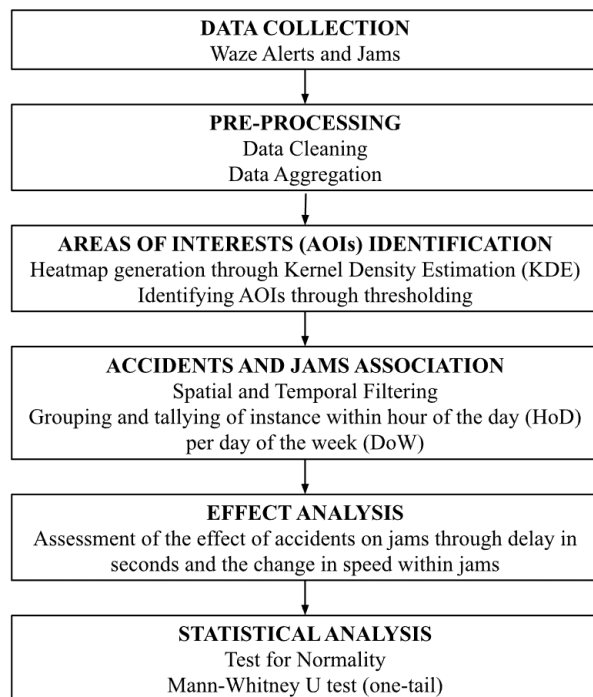


Figure 1. General workflow of methodology.

#### 3.1. Data Collection

Waze traffic data was composed of two main data sets: (1) *Alerts* and (2) *Jams* which were updated every two (2) minutes (Waze, n.d.) and accessed through an official Waze Application Program Interface (API) link provided by the University of the Philippines - National College of Public Administration and Governance (NCPAG) through their Waze for Cities affiliation access via Big Query. These provide relevant information, describing where and when these instances take place based on anonymous user reports.

For the purposes of this study, these were limited to those within the northbound and southbound sections of CA throughout 2022.

Alerts are a point data set (44,918 points) containing multiple attributes, notably the time and place from which the alert was recorded while Jams (53,845 polylines) represent a dynamic deviation from the average speed within certain local areas (Gonzalez et al., 2021).

#### 3.2. Pre-Processing

A total of 3,700 accident-related alerts were extracted and subjected to the established criteria for duplicates from Gonzalez et al. (2021) of being within 20 meters and minutes apart from the time reported, leaving 3,478 unique accident reports. Multiple overlapping jam polylines representing jams were consolidated according to shared ID, date and time of report with essential attributes such as segment length and time delay caused, measured in terms of deviation from free-flow conditions, were aggregated through averaging, noting their variance and minimum and maximum values.

#### 3.3. Areas of Interest (AOI) Identification

Kernel Density Estimation Algorithm (KDE), a nonparametric technique commonly used for point pattern analysis that estimates density from a dataset, was used to identify local hotspots of accidents along Commonwealth Avenue using the Waze traffic accident alerts to serve as AOIs. Its accuracy depends on the bandwidth used, which in this study were set to be 20 m, 100 m, and 250 m for processing in QGIS (Harirforoush & Bellalite, 2019; Hashimoto et al., 2016; Jia et al., 2018; Thakali et al., 2015; Xie & Yan, 2008).

#### 3.4. Accidents and Jams Association

A Python script was then run to identify accidents occurring within these established AOIs, following the same spatial (20 meters) and temporal (20 minutes) thresholds to establish their probable causal relationship, forming three distinct cases: jams, accidents, and jams with accidents. Instances for each case were tallied across each HoD and each DoW.

#### 3.5. Effects of Accident to Jams

After determining whether a certain jam is associated with an accident, the characterization of the jams were analyzed. Jams associated with accidents were then further analyzed based on their average length, time delay cause, and the average speed within. The mean and standard deviation are computed for each of these descriptors to summarize the central tendency and dispersion of the data, providing a detailed descriptive analysis showing the nature and severity of the jams with or without the associated accidents.

#### 3.6. Statistical Analysis

Mann-Whitney U test, a non-parametric test for two ordinal groups not normally distributed with varying sample sizes, was used to compare the two primary datasets utilized in traffic congestion and accidents (Laerd Statistics, 2013). Additionally, the average length of the jam and the time delay caused were tested for the significance of each to the other through tangible parameters such as distance and time. The mean and SD are computed for each of these descriptors to summarize the central tendency and dispersion of the data, providing a detailed descriptive analysis showing the nature and severity of the jams with or without the associated accidents.

## 4. Results and Discussion

#### 4.1. Identified Areas of Interests (AOIs)

In performing the KDE to identify AOIs, the 250-meter bandwidth was chosen as it produced the smoothest, most pronounced, and contiguous hotspots. Six (6) areas of interest within Commonwealth Avenue (CA) were identified

independently for both lanes—four (4) along the northbound side and two (2) along the southbound side. Table 1 indicates the list of accident-prone areas identified.

AOI	Description
1	Quezon City Memorial Circle (QMC) towards UP Diliman - <i>Northbound</i>
2	Petron near Zuzuarregui Street - <i>Northbound</i>
3	Around Puregold Jr. Tandang Balara - <i>Northbound</i>
4	Opposite side of Commonwealth Marker, before INC - Lokal ng Capitol - <i>Northbound</i>
5	Luzon Avenue - <i>Southbound</i>
6	Quezon City Memorial Circle (QMC) towards Philcoa Terminal - <i>Southbound</i>

Table 1. AOIs identified and their corresponding description.

Remarkably, the established AOIs coincide with the accident-prone areas reported in MMARAS 2022, which include: (1) Commonwealth Market to Litex Area, (2) Elliptical Road (Philcoa) to Technohub Area, (3) IBP Road Sandigan (COA) to BF Road Meralco, (4) St. Peter Church and Don Antonio Area, and (5) Tandang Sora Avenue and Luzon Avenue Area.

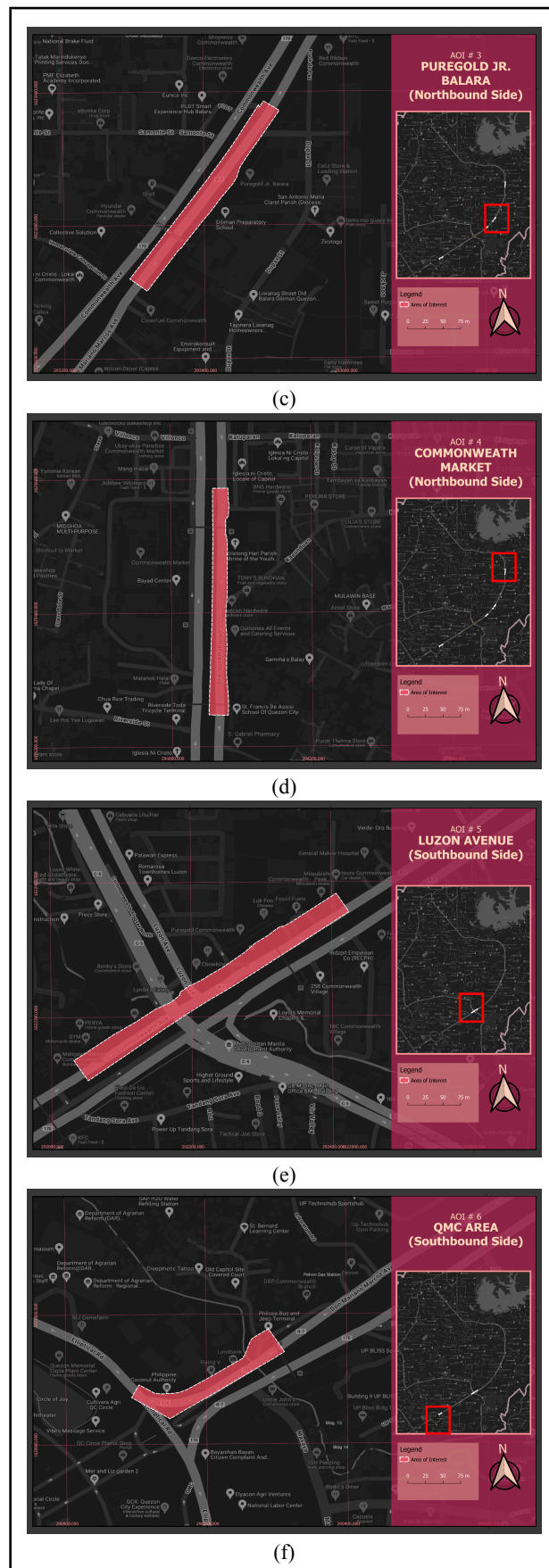
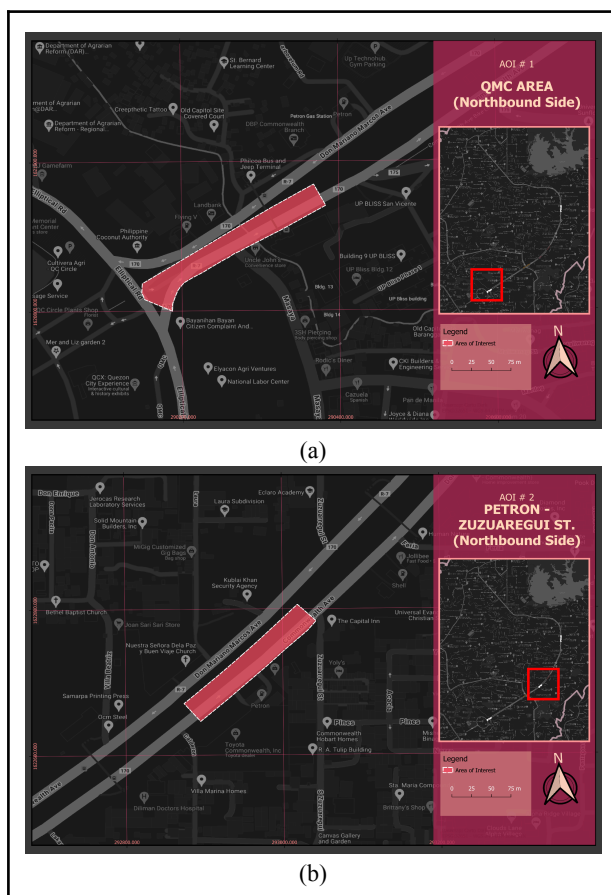


Figure 2. Identified AOIs based on KDE; (a) AOI 1, (b) AOI 2, (c) AOI 3, (d) AOI 4, (e) AOI 5, and (f) AOI 6. (Refer to Table 1 for details corresponding to each AOI)

While these locations correspond to those identified in MMARAS, the integration of Waze, through its accurate geographic data, allows for the precise identification of both the location and directionality of incidents (i.e., northbound or southbound), in contrast to the relatively vague descriptions provided by MMARAS.

The AOIs generated, in turn, enhanced the specificity and contextual accuracy of the identified accident-prone areas. For instance, while MMARAS vaguely describes the stretch from IBP Road Sandigan (COA) to BF Road Meralco as accident-prone, the heat map presented in Figure 2.d identifies a specific segment within this area where accidents occur more frequently.

Through the detection of AOIs, a deeper spatial contextualization was conducted by analyzing key characteristics of the surrounding environment. This included assessing the condition of roads—such as the number of lanes, lane dedication, and clarity of road markings—the presence and density of structures, and the location of high-activity establishments. These factors helped highlight the complex traffic dynamics along Commonwealth Avenue (CA) and the specific influences shaping each AOI. As illustrated in Figure 3, which shows the conditions in AOI 4 during 2022, similar insights were drawn for other AOIs using imagery and data from Google Maps.

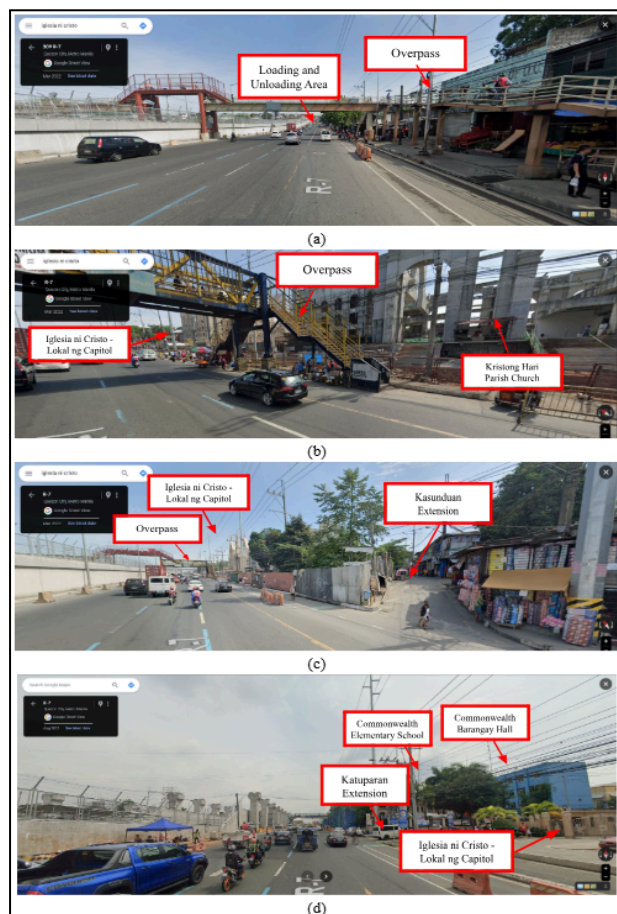


Figure 3. Images of AOI # 4 showing the locations of (a) overpass 1, loading and unloading terminal, (b) Iglesia ni Cristo - Capitol, overpass 2, (c) Kasunduan extension, (d) Katuparan extension, and Commonwealth Barangay Hall. © Google Maps Street View Imagery Copyright (2022).

Despite the distinct features of each AOI, several recurring issues emerged. These included inconsistent lane counts across segments, poorly defined road markings, and the presence of high-traffic establishments such as malls, churches, markets, and pedestrian overpasses. Additionally, the widespread absence of designated or official loading and unloading zones was a common concern across most AOIs.

## 4.2. Temporal Distribution of Jams and Accidents

After identifying the AOI boundaries, each case—accidents only, traffic jams with accidents, and traffic jams without accidents—was organized by hour of day (HoD) per day of week (DoW) to describe its temporal distribution.

### 4.2.1. Traffic Jams without Accidents



Figure 4. Tallied Jams without accidents occurring across AOIs within HoD per DoW.

As shown in Figure 4, traffic jams without associated accidents typically peak on Tuesdays, followed closely by Mondays. This trend is consistent across all AOIs except AOI 3, where the highest congestion occurs on Fridays and Saturdays. This deviation may be attributed to increased weekend activity, as AOI 3 is near several commercial establishments, including Ever Gotesco Commonwealth. In contrast, Sundays consistently record the fewest traffic jams, likely due to reduced weekday-related activity and its designation as a rest day.

Upon closer examination of Figure 5a, it becomes evident that Areas of Interest (AOIs) 1 through 4 experience dual peaks throughout the day: one occurring from approximately 8 AM to 9 AM, and a more pronounced peak from late afternoon to evening, around 5 PM to 8 PM. Interestingly, all of these AOIs are situated along the northbound side of Commonwealth Avenue. In contrast, Figure 5b shows that the southbound AOIs 5 and 6 exhibit a distinct common peak during the early hours, from 5 AM to 8 AM, with a minor discernible peak at night.



This behavior is notably contrasting, indicating that mobility peaks on the northbound side occur in the morning, while on the southbound side, peaks are more prominent in the evening.

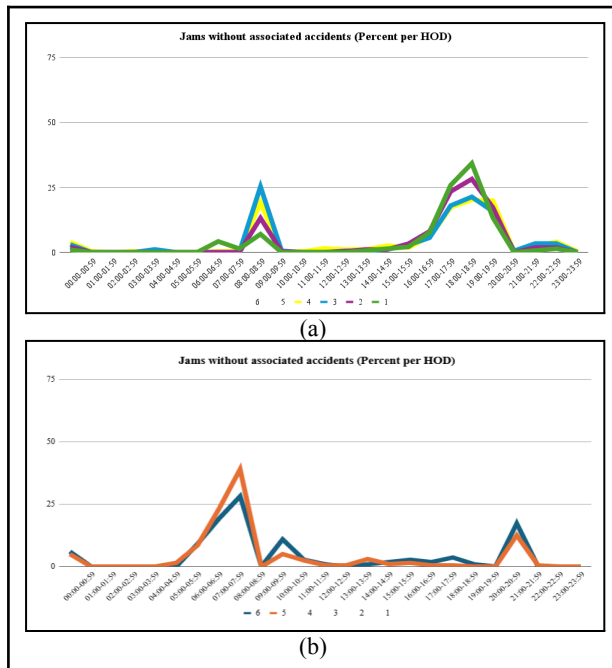


Figure 5. Graph of the Jams without associated accidents occurring across northbound (a) and southbound (b) AOIs within HoD per DoW.

#### 4.2.2. Traffic Jams with Associated Accidents



Figure 6. Tallied Jams with accidents occurring across AOIs within HoD per DoW.

Upon closer inspection, the hourly distribution reveals that such events occur more distinctively in the late afternoon to

evening in AOIs located on the northbound side, while jams with accidents appear more prominently in the mornings on the southbound side, as shown in Figure 7.

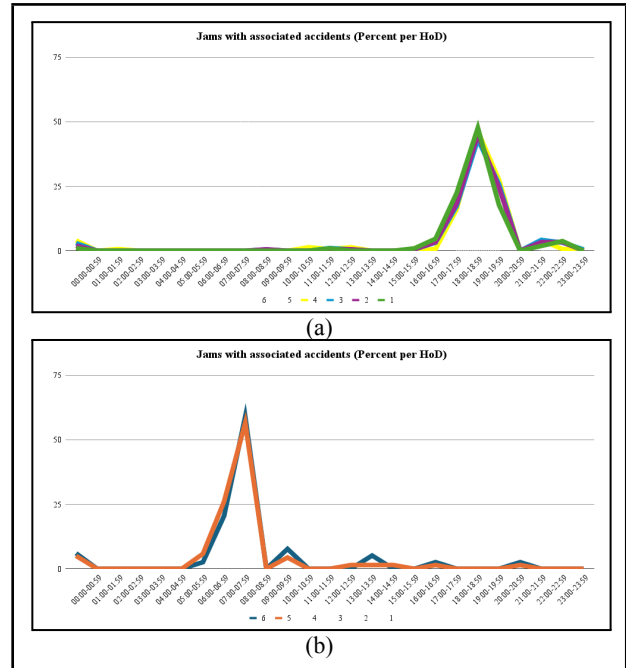


Figure 7. Graph of the Jams with accidents occurring across northbound (a) and southbound (b) AOIs within HoD per DoW.

#### 4.2.3. Distribution of Accidents

Unlike the previous events, accident-related alerts without associated jams do not exhibit a consistent pattern across the HoD and DoW. Their distribution is relatively erratic compared to the first two instances. The most notable observation is that the highest number of accidents occurs on Tuesdays in AOIs 1 and 2, and also on Tuesdays in AOIs 3, 4, and 6. Meanwhile, AOI 5 experienced the most accidents on Saturdays. This suggests that each AOI has its own distinct accident-prone timeframe unique to themselves.

#### 4.3. Characterization of Jams within the AOIs

The characterization of jams—whether associated with accidents (JA) or not (JO)—revealed interesting patterns and commonalities among the identified AOIs. The results of this characterization are discussed in the following sections.

##### 4.3.1. Characterization of Jam Extents and Delay

Based on Table 2, traffic jams associated with accidents (JA) generally have longer average lengths than those without accidents (JO), with the exception of AOIs 1 and 6, which are highlighted in red. This suggests that accidents increase jam lengths, as the Waze data reflects. However, AOIs 1 and 6—both situated near Elliptical Avenue at the entrance to Commonwealth Avenue—exhibit distinct traffic behavior. Their proximity to this major intersection may contribute to atypical congestion patterns, where traffic buildup likely extends beyond the boundaries of Commonwealth Avenue and spills over into Elliptical Avenue. Furthermore, comparing the standard deviations (SD) of both scenarios reveals that JO cases exhibit greater variability, with an SD of 2481.863 meters, compared to 2018.215 meters for JA cases.

Length of Jams (m)			SD (m)	
JO	JA	AOI	JO	JA
4444.437	4234.628	1	2481.663	2018.215
3738.996	3902.682	2	2141.893	1945.421
3478.915	3723.259	3	2134.284	1957.030
3051.458	3436.645	4	2489.562	2037.524
2622.012	2727.183	5	1261.166	1367.514
2452.159	2312.967	6	1364.638	1205.145

Table 2. Average length of jams with (JA) and without (JO) associated accidents. Red texts highlight higher JO length than JA.

Additionally, delays in time needed to traverse such segments measured in seconds are analyzed and represented through violin graphs, as shown in Figure 8.

Generally, the time delay across all AOIs was longer, by 2 to 3 minutes, for jams associated with accidents compared to those without. Notably, AOI 1 presents an interesting case: although the jam extent was shorter for accident-related jams, they resulted in significantly greater delays—requiring an additional four (4) minutes to traverse the area compared to jams without accidents.

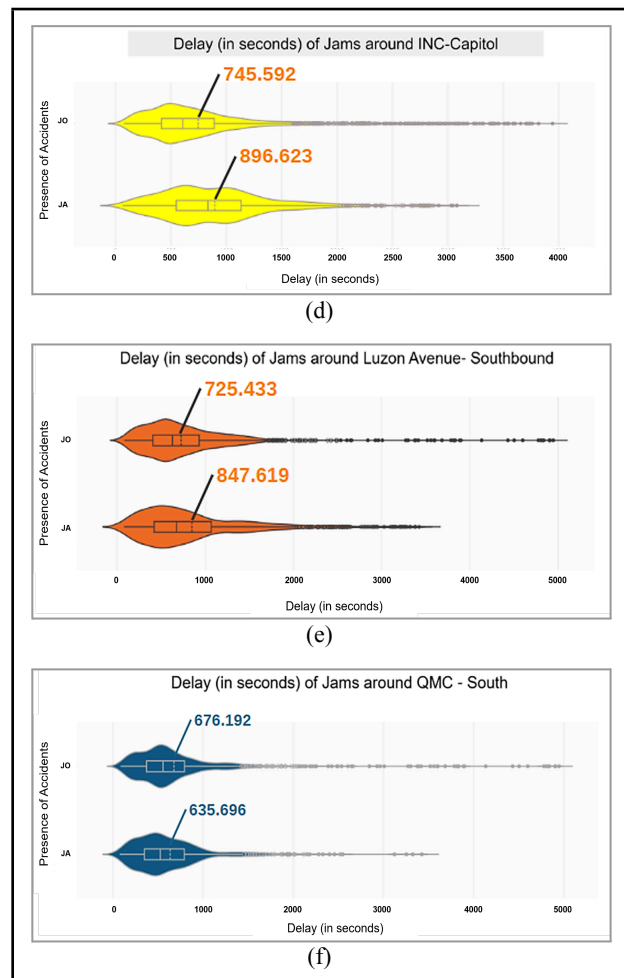
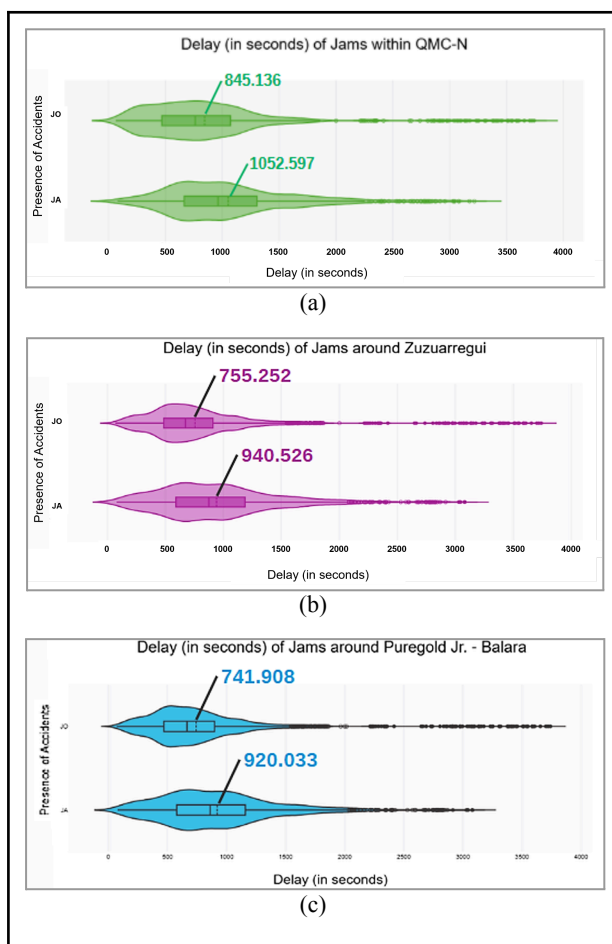


Figure 8. Violin plot comparing travel delay (in seconds) with and without the presence of accidents across the six AOIs, labeled from a to f, respectively.

#### 4.3.2. Effect of Accident on Average Traffic Speed

Based on the Table 3, accidents cause observable differences to the flow of traffic, both in regular, free-flow conditions and jammed traffic, albeit sporadic. These results indicate that although accidents may generally impede the flow of traffic, it is essentially worse for areas already experiencing a traffic jam.

AOI	Regular Speed (kph)		Jammed Speed (kph)		Difference (kph)	
	JO	JA	JO	JA	JO	JA
1	27.527	29.603	13.920	11.930	13.607 (-49.31%)	17.674 (-59.70%)
2	30.877	29.013	13.334	12.029	17.544 (-56.82%)	16.984 (-58.54%)
3	31.146	29.436	12.713	11.762	18.433 (-59.18%)	17.674 (-60.04%)
4	31.316	30.180	11.844	11.193	19.472 (-62.18%)	18.987 (-62.91%)
5	28.839	29.585	11.384	10.931	17.455 (-60.53%)	18.654 (-63.05%)
6	29.889	31.822	11.531	11.718	13.358 (-44.69%)	20.104 (-63.18%)

Table 3. Average speed of jams with and without associated accidents.

#### 4.4. Statistical Analysis

Since the Mann-Whitney test was used for statistical analysis, a normality test was first performed. Using both the Kolmogorov-Smirnov and Shapiro-Wilk tests, the speed data for jams with associated accidents and those without were analyzed. At a significance level of  $\alpha = 0.05$ , the results showed that none of the data followed a normal distribution. Therefore, a non-parametric test was deemed appropriate for this study.

To determine whether the difference in speed is significantly larger in jams with associated accidents compared to jams without associated accidents, a one-tailed Mann-Whitney U test was employed to assess the directionality. The general format of the hypotheses for the one-tailed test is as follows:

- *Null Hypothesis (H<sub>0</sub>):* The difference in speed during traffic jams with associated accidents is equal to or less than the speed difference during traffic jams without associated accidents.
- *Alternative Hypothesis (H<sub>a</sub>):* The difference in speed during traffic jams with associated accidents is greater than the speed difference during traffic jams without associated accidents.

As shown in Table 4, with the p-values listed with a significance level of  $\alpha = 0.05$ , the null hypothesis is rejected for both AOI 1 and 6, implying that there exists insufficient evidence to reject the null hypothesis – suggesting that accidents within these areas do not exert a statistically significant influence on the variance in speed during traffic congestions. Notably, the uniformity in significance across both areas underscores their similarity in traffic conditions and dynamics.

AOI	U Statistics	p-value	Decision
1	31162144	1	fail to reject H <sub>0</sub>
2	21803605.5	0.024	reject H <sub>0</sub>
3	19228400	<0.001	reject H <sub>0</sub>
4	16443697.5	<0.001	reject H <sub>0</sub>
5	18977292	<0.001	reject H <sub>0</sub>
6	30171625	1	fail to reject H <sub>0</sub>

Table 4. Mann-Whitney U Test summary for each AOI.

On the other hand, AOIs 2 to 4 rejects the null hypothesis. This means that the presence of accidents significantly increases the difference in speed during traffic jams in these areas.

An intriguing observation emerges from the analysis: the sole areas where the presence of an accident fails to yield a significant impact on the speed difference are situated at the entrance of Commonwealth Avenue, proximate to the Elliptical Avenue which may be attributed to the atypical congestion patterns mentioned previously. Conversely, the findings robustly affirm that accidents substantially escalate the speed difference, consequently leading to a noteworthy decrease in speed along Commonwealth Avenue, irrespective of their occurrence on either the northbound or southbound side.

#### 5. Conclusion

In this study, Waze traffic data was utilized for traffic assessment, showcasing its potential to be accurate and precise. The results suggest a clear relationship between the presence of traffic accidents and jams, in that the former worsens the effects of the latter, as observed in the marked increase in time delay and length. The observed accident-prone areas from Waze traffic data mirrors what was observed from the MMARAS data from MMDA, a testament to its reliability. However, Waze alerts provide a more precise way of determining the directionality from which the reported traffic accident occurred, a feature often generalized in the MMARAS data. Also, the identified hotspots from the KDE were observed to be near high-activity zones such as churches, malls, schools, or even other highways. Given the temporal precision of the data set which includes date and time of the data set, the study has been able to determine precisely when accidents are at their peak, both in which day of the week and hour of the day. These highlight the potential of utilizing Waze data for the purposes of traffic assessment given proper support, integration, and upscaling into automation not only within the chosen study area but also well beyond given the conclusive results within its scope.

#### References

- González, S. S., Bedoya-Maya, F., & Calatayud, A. 2021. Understanding the Effect of Traffic Congestion on Accidents Using Big Data. *Sustainability*, 13(13), 7500. doi.org/10.3390/su13137500.
- Harirforoush, H., & Bellalite, L. 2019. A new integrated GIS-based analysis to detect hotspots: A case study of the city of Sherbrooke. *Accident Analysis & Prevention*, 130, 62–74. doi.org/10.1016/j.aap.2016.08.015.
- Hashimoto, S., Yoshiki, S., Saeki, R., Mimura, Y., Ando, R., & Nanba, S. 2016. Development and application of traffic accident density estimation models using kernel density estimation. *Journal of Traffic and Transportation Engineering (English Edition)*, 3(3), 262–270. doi.org/10.1016/j.jtte.2016.01.005.
- Jia, R., Khadka, A., & Kim, I. 2018. Traffic crash analysis with point-of-interest spatial clustering. *Accident Analysis & Prevention*, 121, 223–230. doi.org/10.1016/j.aap.2018.09.018.
- Laerd Statistics. 2013. *Mann-Whitney U Test in SPSS Statistics*. Laerd Statistics. <https://statistics.laerd.com/spss-tutorials/mann-whitney-u-test-using-spss-statistics.php> (29 October 2025).
- Metropolitan Manila Development Authority. 2022a. *Annual Average Daily Traffic*. Mmda.gov.ph. [https://mmda.gov.ph/images/Home/FOI/Annual-Average-Daily-Traffic-AADT/AADT\\_2022.pdf](https://mmda.gov.ph/images/Home/FOI/Annual-Average-Daily-Traffic-AADT/AADT_2022.pdf) (29 October 2025).
- Metropolitan Manila Development Authority. 2022b. *MMARAS Annual Report 2022*. [https://mmda.gov.ph/images/Home/FOI/MMARAS/MMARAS\\_Annual\\_Report\\_2022.pdf](https://mmda.gov.ph/images/Home/FOI/MMARAS/MMARAS_Annual_Report_2022.pdf) (29 October 2025).
- Mukheja, P., Kiran K, M., Velaga, N. R., & Sharmila, R. B. 2017. Smartphone-based crowdsourcing for position estimation of public transport vehicles. *IET Intelligent Transport Systems*, 11(9), 588–595. doi.org/10.1049/iet-its.2016.0247v.

Shahgholian, M., & Gharavian, D. 2018. Advanced Traffic Management Systems: An Overview and A Development Strategy. *ArXiv:1810.02530* [Eess]. doi.org/10.48550/arXiv.1810.02530.

Thakali, L., Kwon, T. J., & Fu, L. 2015. Identification of crash hotspots using kernel density estimation and kriging methods: a comparison. *Journal of Modern Transportation*, 23(2), 93–106. doi.org/10.1007/s40534-015-0068-0.

Thibault, G., de Clerq, M., Brandt, F., Nienhaus, A., & Bayen, A. 2022. Urban Mobility Readiness Index 2022 REPORT. In *OliverWyman Forum* (pp. 1–124). Berkeley | University of California. <https://www.oliverwymanforum.com/content/dam/oliver-wyman/ow-forum/template-scripts/urban-mobility-index/PDF/Mobility-Index-Report.pdf> (29 October 2025).

World Health Organization. 2023, May 31. *Department of Transportation, World Health Organization launch Philippine Road Safety Action Plan 2023-2028*. Wwww.who.int. <https://www.who.int/philippines/news/detail/31-05-2023-department-of-transportation--world-health-organization-launch-philippine-road-safety-action-plan-2023-2028#:~:text=In%20the%20Philippines%2C%20road%20> (29 October 2025).

Xie, Z., & Yan, J. 2008. Kernel Density Estimation of traffic accidents in a network space. *Computers, Environment and Urban Systems*, 32(5), 396–406. doi.org/10.1016/j.compenvurbsys.2008.05.001.

Zhang, Z., Han, L. D., & Liu, Y. 2021. Exploration and evaluation of crowdsourced probe-based Waze traffic speed. *Transportation Letters*, 14(5), 1–9. doi.org/10.1080/19427867.2021.1906477.