Spatio-Temporal Clustering of Road Accidents in Kelantan, Malaysia

Wan Fairos Wan Yaacob, Shahirah Binti Ibrahim, Ainin Sorfina Afizan, Nur Azreen Faizul Azran, Syerina Azlin Md Nasir, Norazlina Che Harun

In-Text Citation: (Yaacob et al., 2021)
To Cite this Article: Yaacob, W. F. W., Ibrahim, S. B., Afizan, A. S., Azran, N. A. F., Nasir, S. A. M., & Harun, N. C. (2021). Spatio-Temporal Clustering of Road Accidents in Kelantan, Malaysia. *International Journal of Academic Research in Business and Social Sciences*, *11*(9), 522–534.

Full Terms & Conditions of access and use can be found at
http://hrmars.com/index.php/pages/detail/publication-ethics

Spatio-Temporal Clustering of Road Accidents in Kelantan, Malaysia

Wan Fairos Wan Yaacob[1], Shahirah Binti Ibrahim[2], Ainin Sorfina Afizan[3], Nur Azreen Faizul Azran[4], Syerina Azlin Md Nasir[5], Norazlina Che Harun[6]

[1,2,3,4]Universiti Teknologi MARA Cawangan Kelantan, Kampus Kota Bharu, Lembah Sireh, 15050 Kota Bharu, Kelantan, Malaysia, [5]Universiti Teknologi MARA Cawangan Kelantan, Bukit Ilmu, 18500 Machang, Kelantan, Malaysia, [5]Ibu Pejabat Polis Kontigen Kelantan, Malaysia
Email: wnfairos@uitm.edu.my

Abstract
Road accidents have become a global issue concern. Accidents may occur in different places with different incidents that can make it difficult to determine which areas are prone to accidents. This information is needed by the community and the respective authority for law enforcement. This study utilized spatio-temporal clustering to analyze the high-risk area of road accidents in the state of Kelantan, Malaysia. It aimed to identify the hotspot area of accident location in Kelantan using spatio-temporal analysis and cluster the road accident locations according to the geographical area in Kelantan using cluster analysis. Analysis of spatio-temporal is utilized to identify the hotspot areas of high-risk road accidents by mapping spatio-temporal heterogeneity road accidents' cases of ten districts in Kelantan by day. The results indicated that the area of Kota Bharu was identified as the hotspot of road accident location in Kelantan. By using K-means clustering, four different clusters were formed. The first cluster was Kota Bharu which represented a very high-risk accident area. The second cluster of high-risk accident areas were Gua Musang, Pasir Mas and Tanah Merah, while the third cluster which was a moderate-risk accident areas consisted of Machang, Kuala Krai, Tumpat, Pasir Puteh and Bachok. Lastly, the fourth cluster of low-risk accident area was Jeli. The findings from this study can be used by the authorities in preventing and reducing the statistics of road accident cases in Kelantan and can be further utilized by the other states in Malaysia.
Keywords: Clustering, K-Means, Road Accident, Hot Spot, Spatial Mapping

Introduction
Road Traffic accidents have become a major public health problem in most developing countries. According to the World Health Organization (2018), road injury has been reported as the eighth of the top tenth global cause of deaths in 2016. From this ranking, it shows that road accident has given great impacts on society and directly impacted the economy and the country as a whole. Malaysia is one of the developing countries that is rapidly building its

highways and road networks that indirectly has increased occupancy of the vehicle. Malaysia was recorded the third highest fatality rate for road accidents in Asia and Asian, behind Thailand and Vietnam (WHO, 2018). This alarming issue is worrying and need to be addressed. The need for the development of efficient road network and safety is very important to improve and reduce statistics of traffic accidents especially in the east coast Malaysia, particularly Kelantan state where there are abundant of cases involving traffic accidents. Highlights of road traffic accidents in Kelantan depicted that on average Kelantan recorded 609 statistics of road accidents during 'Op Selamat' and 11 of them were fatal accidents involving 13 deaths (Sun Daily, 2019). Accidents occur mainly during the festive seasons due to the increasing number of vehicles on the road. According to the State Investigations and Enforcement Department, accidents regularly happen in the areas that have been classified as the 'black area'. When this happens, it will not only burden the healthcare delivery system due to limited hospital beds and resource utilization, but it will also result in loss of productivity and income, with social and economic consequences. The use of data in making an informed decision by understanding the accident risk area in the cities is important as a key to traffic safety management. Over the years, studies related to road accidents were carried out in most countries such as Iran (Soltani & Askari, 2014; Soltani & Askari, 2017; Zangeneh et al., 2018), China (Cheng et al., 2018; Fan et al., 2018), United Kingdom (Anderson, 2009; Aljofey & Alwagih, 2018; Kazmi et al., 2020), Indonesia (Sitohang & Merni, 2020), Vietnam (Le et al., 2020), Malaysia (Musdholifah & Hadhim, 2013; Shahid et al., 2015) and many more. Thus, many researchers have contributed to road accident studies in identifying the accident hotspot sites using analysis such as spatial-temporal clustering of Geographical Information System (GIS) (Kavgisiv et al., 2015; Zangeneh et al., 2018; Lakshmi et al., 2019; Kazmi, 2020), Kernel Density Estimation (Musdholifah & Hadhim, 2013; Fan et al., 2018), Triangular Kernel Clustering (Asadi & Pegun, 2019), Deep Embedded Clustering (Shahido et al., 2015), K-Means Clustering (Ghadi & Thorok, 2020; Sitohang & Merni, 2020) and many others.

A study done by Zamzuri et al (2019) identified accident variables such as weather, light condition, traffic system, lane marking, road geometry, collision type and accident severity when investigating the interrelationship between these variables to further understand the cause of accidents. Furthermore, many studies reported on the causes or factors that are closely related to accident severity specifically on accidents occurrence and their locations. Ghadi and Thorok (2020) highlighted that the main difficulty in analyzing accident data was its heterogeneity and found that few studies have been carried out on spatial dependency factors. In most cases, the data mining approach has been applied to explore the issue further. Data mining is a well-known technique for summarizing data by finding patterns that can be understood and useful for data owners. Specifically, transforming continuous data using discretization in the preprocessing techniques is found in many research works for spatial data mining (Ahmad-Azani et al., 2018). Hence many researchers use data mining in analyzing traffic accident datasets such as association rule, K-Means clustering (Kumar et al., 2016) and machine learning algorithm of the decision tree, naive Bayes, and K-nearest neighbour classifiers to predict accident severity (Beshah & Hill, 2010).

According to Aljofey and Alwagih (2018), identifying the most frequent accident frequencies is not enough and they suggested the need to cluster accident frequencies times for location instead of clustering accident frequencies locations. Clustering the road accident "black area" or known as hotspots are generally based on the available data associated with the accident itself (namely, time of day, type of victim, type of vehicle). Bil et al. (2019)

applied a clustering method for the identification of traffic crash hotspots at the rural parts of primary roads in the Czech road network, while Sitohang and Merni (2020) applied K-Means clustering for the analysis of accident-prone areas in Indonesia. In Hungary, the accident data were manipulated to identify through spatial segmentation, black spots identification, and decision analysis using K-Means clustering and Empirical Bayesian (EB) method (Ghadi & Thorok, 2020). Work by Cheng et al. (2018) utilised spatio-temporal analysis in traffic crash study to identify the hotspots of accident locations in Wujiang, China. They used spatial join analysis to visualize the crash number at each intersection and display the traffic crashes over time and area using space-time cube analysis. From the spatial autocorrelation, the hotspots area was identified and the trends of crash density in the time-space cube which denoted as cold and hot spots were reported.

Furthermore, a study by Ivan and Haidu (2012) applied Spatio-temporal analysis on road traffic accidents in Cluj-Napoca and highlighted the importance of Geographic Information System (GIS) in displaying the spatial distribution of accidents along with the road network. The study included time attributes which indicated the moment of the road accidents occurrence and area attributes. The results were able to identify the high number of accidents that tend to happen at the East-West back-bone road, Traian Vuia-Calea Floresti which is the exit-access points in the city. A study done by Shahid et al. (2015) adopted GIS to determine the spatial and temporal variation in the incidence of road traffic accidents and fatalities across the states of Peninsular Malaysia in Malaysia. The findings showed more accidents but lower fatalities in more urbanized and developed states. Meanwhile, the less urbanized and developed states such as Kelantan and Perlis showed lower accident levels but more serious fatalities. Most of the accidents were found to happen in the festival month (Hari Raya Puasa) followed by the month of mid-year school holidays.

Anderson (2009) adopted Kernel Density Estimation (KDE) and K-means clustering to profile road accident hotspots in the United Kingdom (Anderson, 2009). The work was further improved by Kazmi et al (2020) who proposed the identification of accident hotspots by exploiting GIS technology coupled with KDE. This was carried out in the study for intra-urban traffic accident data in metropolitan Shiraz, Iran which aimed to identify accident prone zones and sensitive hours (Soltani & Askari, 2014). The findings provided information about traffic conditions in that metropolitan, the most potentially accident-prone roads and some causes of the accident. KDE method was also applied to generate maps in determining Road Traffic Accidents (RTAs) hotspots in Shiraz and the results indicated that the highest RTAs incident was found in North-western parts of the metropolitan. A network-based KDE was applied by Kaygisiv et al. (2015) to examine the hotspots of pedestrian accidents and their variation on the Eskisehir Motorway from the year 2005 to 2010. The findings were further tested using Nearest Neighbor Distance and K-function methods which allow to investigate the detected hotspots and the reasons that may have caused the changes over time. Apart from that, KDE was also applied to investigate patterns of single-vehicle crashes (SVCs) in Western Australia between 1999 and 2008 (Plug et al., 2011). The hotspots of SVCs were measured in three locations of West Australia (WA), Metropolitan area and Perth Local Government Area (LGA) through Spatial Analyst KDE tool in ESRI's ArcGIS 9.3 software. The results of this study showed that the SVC distribution over time of the day on weekdays and weekends appeared to be more frequent in the afternoon until night (3 pm- midnight) and early in the morning (10 pm - 2 am) respectively. Meanwhile, spatial analysis showed that the area around Perth Metropolitan was clustered as a hotspot region. In China, a study was carried out to analyze the traffic collision data in the Jianghan District of Wuhan using the network Kernel Density

Estimation (KDE) and K-function. The results of this study showed these methods generated different perspectives of spatio-temporal clustering patterns. However, research related to data clustering as a basis for identifying areas prone to traffic accidents is still limited.

Accidents may occur in different places with different incidents which make it difficult to determine which areas are prone to accidents classified as the 'black area'. Hence, there is a need to identify the hot spot areas along the road with high accidents to minimize the risk and use the information for road planning and design decision. Thus, this study aims this study aim to cluster the traffic accident area and identify the hotspots area of accident locations in Kelantan using spatio-temporal analysis. Findings from this research can provide a better understanding on safety research field especially in the state of Kelantan.

Method

This research used road accident data obtained from Kelantan Contingent Police Headquarters, Kota Bharu, Kelantan, Malaysia. The data set consisted of 1,139 road accidents from January 1st, 2019 to February 11th, 2019 in 10 districts of Kelantan State. After pre-processing, five variables were identified to be used for this study. The data set comprised the day of the accident, district of the accident site, longitude and latitude of the accident location for each accident occurrence. The accident occurrence used in this study was inclusive of minor injury, severe injury and fatal accident. Both longitude and latitude were used in spatio-temporal analysis to produce a spatial map of districts in Kelantan. The day represents the accident occurrence ranging from Sunday until Saturday. The areas of accident occurrence cover ten districts in the state of Kelantan. The brief information about this data is given in Table 1.

TABLE 1: Road Accident Data Attributes

Number	Attribute name	Values
1	Accident	Accident occurrence inclusive of minor injury, severe injury and fatal accident.
2	Day of Accident	Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday
3	District of Accident	Kota Bharu, Pasir Mas, Jeli, Tumpat, Bachok, Kuala Krai, Tanah Merah, Machang, Gua Musang, Pasir Puteh
4	Longitude	Degree, Minutes, Seconds
5	Latitude	Degree, Minutes, Seconds

Data pre-processing was performed before conducting the analysis which involved transforming the raw data into an understandable format. The process includes several steps such as data cleaning, data integration, data transformation and data reduction to ensure the reliability of the data. According to the data set obtained from the Kelantan Contingent Police Headquarters, there was a total of 1285 cases of road accidents in Kelantan from the period of January 1st, 2019 to February 11th, 2019. The data set focused on three categories of road accident which were minor injury, severe and fatal. Category of road accidents such as Rempit and Seksyen 48 were not included in this study. Thus, after the cleaning process, a total of 1139 cases were used for analysis in this study.

In order to identify the pattern of hotspots accident location distribution across the state based on districts, Spatio-temporal analysis was used. Spatio-temporal analysis was used to visualize the distribution of data across space and time. In this study, the spatial map was plotted for ten districts in Kelantan with different colour tones using R software to display the hotspot area of road accident location in Kelantan. The temporal analysis was also used to investigate the day effect which could relate the road accident pattern across time that might lead to other sources of evidence.

K-means clustering algorithm is one of the most frequently used clustering algorithms. It clusters the observations into k groups. Aljofey et al. (2018) used clustering and classification to analyse the accident times based on accident frequencies for highway locations which they proposed to do clustering technique and classification trees. The 'k' in k-means algorithm refers to the fixed number of clusters that will be chosen for the algorithm. The 'k' clusters for this study refer to the cluster road accident locations. This study used Orange 2.7 software to produce the k-means clustering. The k-means workflow of analysis done in Orange can be seen in Fig.1 below.
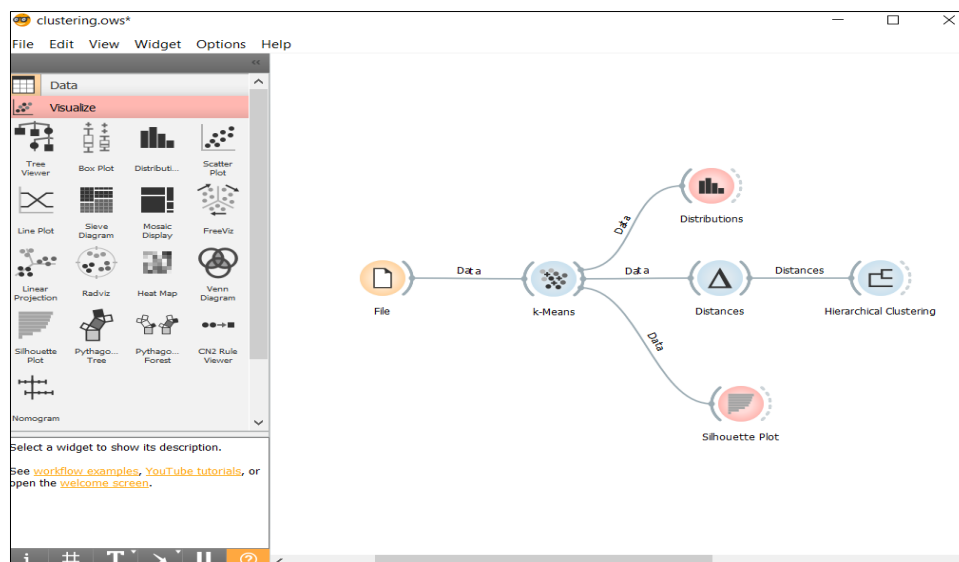

Fig.1: The K-Means workflow in Orange Canvas

The implementation of k-means clustering was based on the following algorithm:
   i.   Choose the number of clusters, k.
   ii.   Choose a set of k initial centroids.
   iii. Assign each instances in the data set to the closest centroid using Euclidean distance.
   iv.   Then for each cluster, a new centroid was computed as a centre of clustered data instances in this update step.
   v.   Repeat the previous two steps of assignment and update until the cluster assignment does not change.

In the Euclidean space, the mean of cluster j is computed with:

$$m_j = \frac{1}{|C_j|}\sum_{x_i \in C_j} x_i \qquad\qquad (1)$$

where $|C_j|$ is the number of data points in cluster C_j. The distance from one data point x_i to a mean (centroid) m_j was computed as:

$$dist(x_i, m_j) = |x_i, m_j| \qquad (2)$$
$$= \sqrt{\left(x_{i1} - m_{j1}\right)^2 + \left(x_{i2} - m_{j2}\right)^2 + \cdots + \left(x_{ir} - m_{jr}\right)^2}$$

To evaluate the k-means and decide on the number of k cluster to be formulated from the data, this study used the Silhouette Score method. Silhouette score is a measurement calculated to decide the number of clusters that suits best for the data. The value varies between -1 and 1. The value closest to 1 indicates the best number of cluster whereas the value that is closest to -1 indicates the worst cluster. The highest silhouette score will indicate the best cluster to be formed. The formula for the silhouette score is:

$$Silhouette\ Score = (x - y)/max(x, y) \qquad (3)$$

where:

y is the mean intra cluster distance (mean distance to other instances in the same cluster) and x is the mean nearest cluster distance (mean distance to the instances of the next closest cluster). According to Kauffman and Rousseuw (1990), the bigger the silhouette coefficient that is close to 1, the better the quality of the formed cluster. A silhouette value of range between 0.5 to 0.7 is considered good cluster and more than 0.7 is considered strong quality number of cluster formed.

Results and Discussion
The frequency of accident by district in Kelantan during the period of study is described in Fig.2. The districts in the bar chart were ranked according to the number of accident cases for each district. It can be seen that Kota Bharu ranked as the highest number of road accident cases (437), followed by Gua Musang, Pasir Mas and Tanah Merah which recorded 105, 98 and 90 cases, respectively. The number of road accident was lowest in Jeli which intuitively indicated the safest district among all the districts in Kelantan. Several reasons may cause the high accident level in Kota Bharu as compared to Jeli. One of the main reasons is the high population density in the district. Kota Bharu is the district with the highest population (608.6 thousand) and the smallest district area of 115.6km2 compared to Jeli with a population of 51.9 thousand and district area of 1,330km2. Thus, this makes Kota Bharu a highly condensed area with rapid urban development, infrastructure and focus on employment sector. This would increase the number of vehicles and traffic flow which is also the risk of road accident. Meanwhile, Jeli's economy is dominated by agriculture and plantation. The urban population is much less than Kota Bharu and the traffic situation is much better than Kota Bharu. Another possible reason for the high case of road accidents in Kota Bharu is uneven condition of road structure such as sharp turns and uneven roads that might contribute to accident.
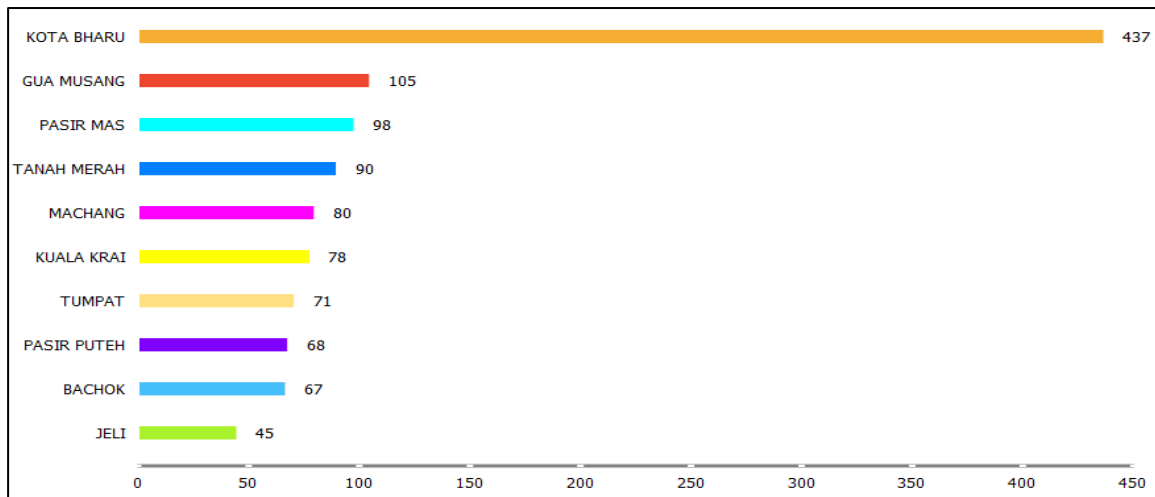
Fig.2: The Frequency of Road Accident Cases by District in Kelantan

Spatio-temporal analysis is one of the best ways to determine hotspot area of road accident locations in Kelantan. The results of the analysis are presented through a spatial map. Using the frequency of road accident cases that occur in the respective districts of the state, spatio-temporal analysis was used to identify the hotspot areas of accident occurrence. Figure 3 shows the district involved with road accidents in Kelantan from January 1st, 2019 to February 11th, 2019. The data points in the figure represent the 10 districts of road accident locations. Fig.3 (a) to (g) below illustrates the spatial maps that consist the number of concurrences road accident cases in 10 districts of Kelantan. The pattern of accident occurrence was described according to seven days in a week for the whole duration of road accident cases shown by a series of figures from Sunday until Saturday. The overview of the pattern of road accidents occurrence in Kelantan can clearly be seen according to the days. All spatial maps below showed different shades of red based on the distribution of road accident cases according to the districts. Based on the spatial map below, Kota Bharu was identified as the hotspot or very high road accident area in Kelantan for almost every day in a week since it showed the darkest shades of red colour in all spatial maps. Meanwhile, other districts exhibited low occurrence of road accident cases except for Gua Musang and Kuala Krai which indicated a medium-high accident distribution. An interesting finding from this study illustrated the third darker shades for road accident cases which occurred on Sunday at Gua Musang and Saturday at Kuala Krai.

a) Sunday

b) Monday

c) Tuesday

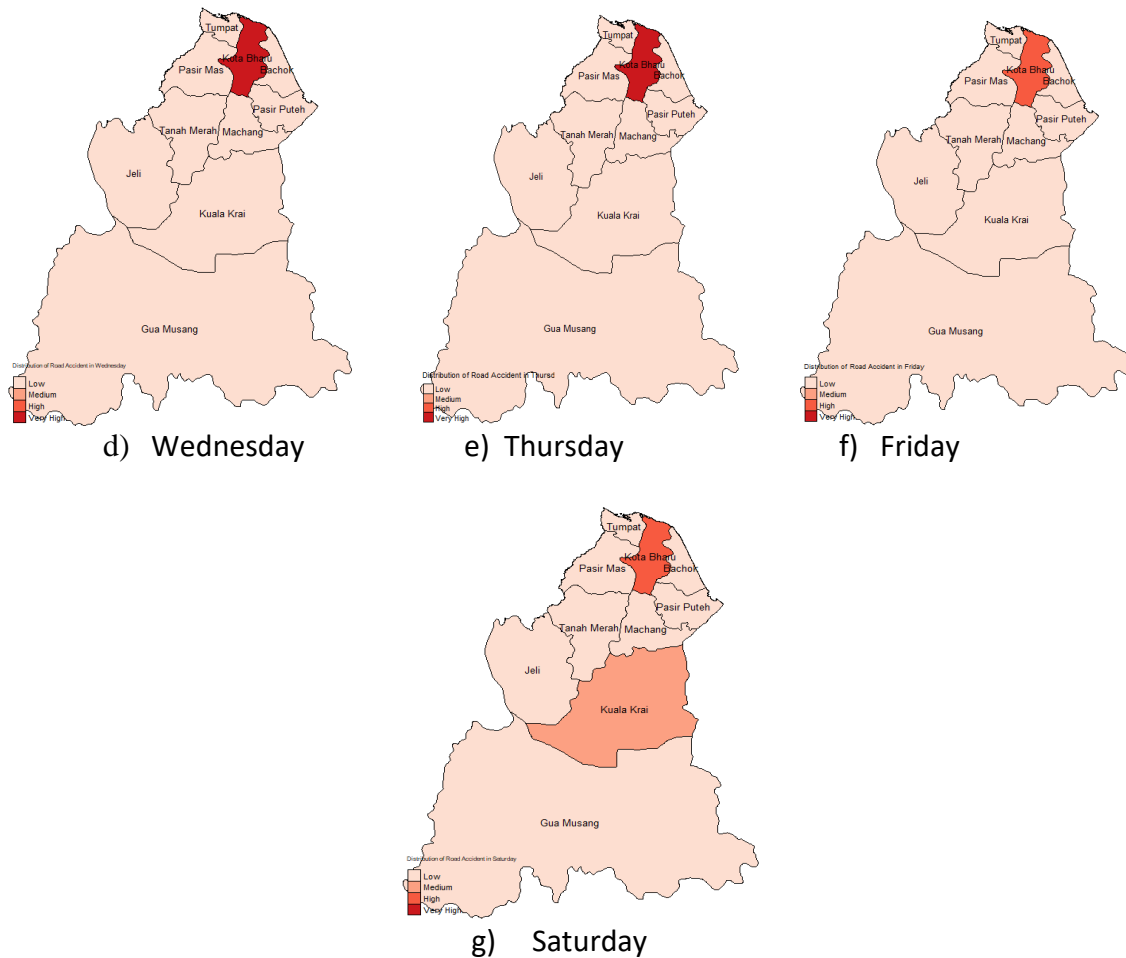d) Wednesday   e) Thursday   f) Friday

g) Saturday
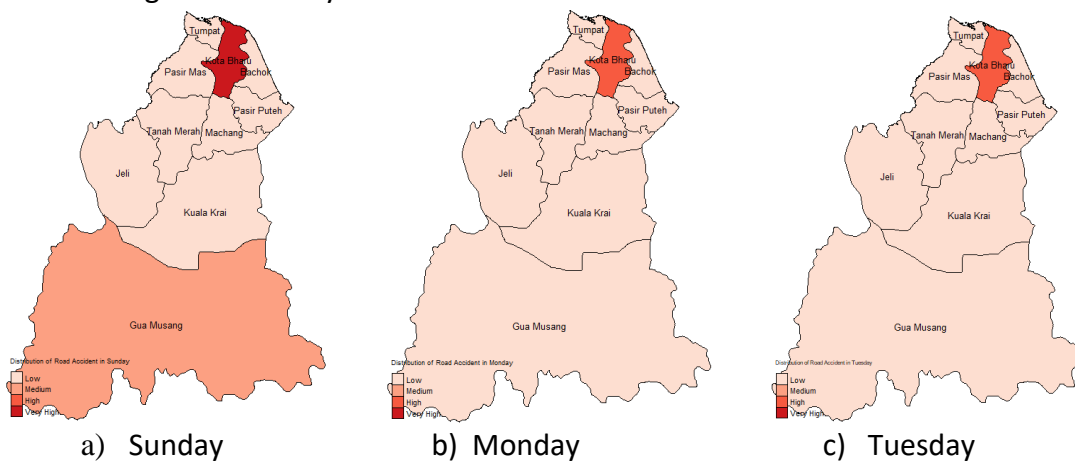
Fig.3: Spatial Distribution of Accident Cases in Kelantan State According to Days

The results from k-means clustering algorithm associated to the silhouette score obtained for cluster 2, 3, 4, 5, and 6 in Table 2 showed that the best number of cluster chosen was 4 since the silhouette score for cluster 4 was the highest value with 0.50 as compared to other number of clusters performed. Silhouette score is a measurement calculated to decide the number of clusters that suits best for the data.

TABLE 2: Silhouette Score for Optimal k

Number of Cluster (k)	Silhouette Score
3	0.49
4	**0.50**
5	0.44
6	0.44
7	0.37

Fig.4 displays the k-means hierarchical dendrogram of road accident cases according to the identified group of accident occurrence. Based on Fig.4, the dendrogram shows four different spatial patterns for the distribution of road accident in Kelantan using K-Means Clustering. Cluster 1 consisted only one district which was Kota Bharu. Cluster 2 consisted of three districts namely Gua Musang, Pasir Mas and Tanah Merah. Cluster 3 consisted of four districts including Machang, Kuala Krai, Tumpat, Pasir Puteh and Bachok, while Cluster 4

consisted only one district which was Jeli. These four spatial road accident distribution patterns were determined due to different levels of road accident number.
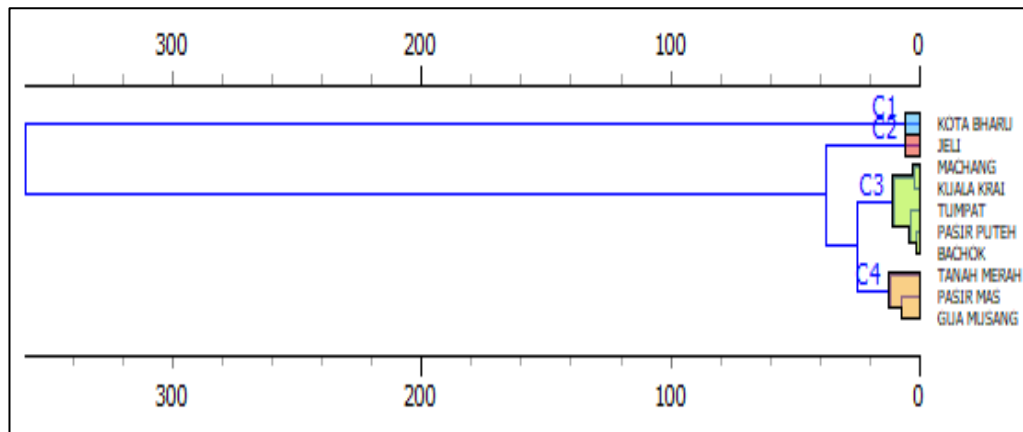

Fig.4: Hierarchical Dendrogram using K-Means Clustering of Road Accident Cases in Kelantan based on Districts

In Table 3, Cluster 1 clustered Kota Bharu district as a very high-risk accident area, followed by Cluster 2 as the high accident risk area that was represented by Pasir Mas, Gua Musang and Tanah Merah. Next, Cluster 3 which described the moderate accident risk area included the districts of Machang, Kuala Krai, Tumpat, Pasir Puteh and Bachok. The last cluster of Jeli was clustered in a low-risk accident area.

TABLE 3: Road Accident Clusters Occurrence using K-Means Hierarchical Clustering

Cluster	District Of Incident	Cluster Name
1	Kota Bharu	Very High Accident Risk Area
2	Pasir Mas	High Accident Risk Area
	Gua Musang	
	Tanah Merah	
3	Machang	Moderate Accident Risk Area
	Kuala Krai	
	Tumpat	
	Pasir Puteh	
	Bachok	
4	Jeli	Low Risk Accident Area

Conclusion

Spatial and temporal analysis of road accidents in Kelantan showed that Kota Bharu is the district with the highest record of total road accident cases compared to the other districts in Kelantan. The highest case of accidents occurred in Kota Bharu based on both spatial mapping and cluster analysis since Kota Bharu is the focus place for working with more compact residents as compared to the other districts. Meanwhile, Jeli recorded the lowest accident occurrence among all the districts. The intensity of the accident tends to vary by location and day. Based on the daily spatial and temporal pattern of the accident from spatial mapping, high accident occurrence was spotted in Kota Bharu. The medium-high distribution of road accident cases also occurred in Gua Musang on Sunday. This might be due to the increased

number of people who were traveling back to Lembah Kelang or other states. People who commute long distance over the weekend by going in or out of Kelantan will travel through Gua Musang. Hence, there is an influx of vehicles crossing Gua Musang since Gua Musang is the place where people stop for rest. Those people travelling to Kelantan will contribute to the congestion of vehicles that will happen on Sunday in Gua Musang. Thus, it increases the traffic flow and risk of accident occurrences. For the physiological aspect of driver, Gua Musang will be the place where people need to stop for rest after driving halfway of the journey. This condition might also contribute to the risk of accident occurrence. Meanwhile, the medium-high increase of accident cases at Kuala Krai on Saturday might be due to the heavy traffic along Kuala Krai roadway. The increased number of traveling on Saturday is because it is the day for weekend to end for people who start working on Sunday. Saturday is rush hour for people to go back to the places they work. Hence, there is a rush vehicle transportation going in and out of Kuala Krai. There will be a tendency of people to rush in their driving mode because they want to reach their destination early to prepare for work on the next day as Saturday is the day that weekend ends. That would increase the number of transportation getting into Kuala Krai or going out of it. Further investigation can be done to study the data of travelers/workers who travel from Kota Bharu to Kuala Krai over the weekend. This is because most people who stay in Kuala Krai during working days travel back to their hometown in Kota Bharu over the weekend. Based on the clustering finding that identified the cluster with very high accident risk location in Kota Bharu, and high accident risk locations at Gua Musang, Pasir Mas and Tanah Merah, monitoring from the authorities should be emphasized especially places like hilly roads, twin lines and traffic light areas. This effort might help to resolve and reduce the road accident rates.

Acknowledgements

References

Ahmad-Azani, N. I., Yusoff, N., & Ku-Mahamud, K. R. (2018). Fuzzy Discretization Technique for Bayesian Flood Disaster Model. Journal of Information and Communication Technology, 18(2), 167-189.

Aljofey, A. M., & Alwagih, K. (2018). Analysis of Accident Times for Highway Locations Using K-Means Clustering and Decision Rules Extracted from Decision Trees. International Journal of Computer Applications Technology and Research, 7(1), 1–11. https://link.springer.com/article/10.1007/s40534-016-0095-5

Anderson, T. K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. Accident Analysis and Prevention, 41(3), 359–364. https://doi.org/10.1016/j.aap.2008.12.014

Asadi, R., & Regan, A. (2019). Spatio temporal clustering of traffic data with deep embedded clustering. Proceedings of the 3rd ACM SIG SPATIAL International Workshop on Prediction of Human Mobility, PredictGIS 2019, 45–52. https://doi.org/10.1145/3356995.3364537

Beshah, T., & Hill, S. (2010, March). Mining road traffic accident data to improve safety: Role of road-related factors on accident severity in Ethiopia. In AAAI Spring Symposium: Artificial Intelligence for Development (Vol. 24, pp. 1173-1181).

Bíl, M., Andrášik, R., & Sedoník, J. (2019). A detailed spatiotemporal analysis of traffic crash hotspots. Applied Geography, 107, 82-90.

Bora, D. J., & Gupta, A. K. (2014). A new approach towards clustering based color image segmentation. International Journal of Computer Applications, 107(12).

Cheng, Z., Zu, Z., & Lu, J. (2018). Traffic crash evolution characteristic analysis and spatiotemporal hotspot identification of urban road intersections. Sustainability (Switzerland), 11(1). https://doi.org/10.3390/su11010160

Fan, Y., Zhu, X., She, B., Guo, W. & Guo, T. (2018). Network-constrained spatio-temporal clustering analysis of traffic collisions in Jianghan District of Wuhan, China.

Ghadi, M. Q., & Török, Á. (2020). Evaluation of the Impact of Spatial and Environmental Accident Factors on Severity Patterns of Road Segments. Periodica Polytechnica Transportation Engineering.

Ivan, K., & Haidu, I. (2012). The spatio-temporal distribution of road accidents in Cluj-Napoca. Geographia Technica, 2, 32–38.

Kauffman, L and Rousseauw, P.J. (1990). Finding Groups in Data: An Introduction to Cluster Analysis. USA: Wiley Series in Probability and Statistics.

Kaygisiz, Ö., Yildiz, A., & Düzgün, Ş. (2015). Spatio-temporal pedestrian accident analysis to improve urban pedestrian safety: The case of the eskişehir motorway. Gazi University Journal of Science, 28(4), 623–630.

Kaygiz, O., Duzgun, S., Yildiz, A. & Senbil, M. (2014). Spatio-temporal accident analysis for accident prevention in relation to behavioral factors in driving: The case of South Anatolian Motorway. Transportation Research Part F 33, 128-140.

Kazmi, S. S. A., Ahmed, M., Mumtaz, R., & Anwar, Z. (2020). Spatiotemporal Clustering and Analysis of Road Accident Hotspots by Exploiting GIS Technology and Kernel Density Estimation. The Computer Journal.

Kumar, S., & Toshniwal, D. (2016). A data mining approach to characterize road accident locations. Journal of Modern Transportation, 24(1), 62–72. https://doi.org/10.1007/s40534-016-0095-5

Kumar, S., & Toshniwal, D. (2016). A data mining framework to analyze road accident data. Journal of Big Data, 2(1). https://doi.org/10.1186/s40537-015-0035-y

Lakshmi, S., Srikanth, I., & Arockiasamy, M. (2019). Identification of Traffic Accident Hotspots using Geographical Information System (GIS). International Journal of Engineering and Advanced Technology, 9(2), 4429–4438. https://doi.org/10.35940/ijeat.b3848.129219

Le, K. G., Liu, P., & Lin, L. T. (2020). Determining the road traffic accident hotspots using GIS-based temporal-spatial statistical analytic techniques in Hanoi, Vietnam. Geo-Spatial Information Science, 23(2), 153–164. https://doi.org/10.1080/10095020.2019.1683437

Musdholifah, A., & Hadhim, S. Z. M. (2013). Mining Spatio-Temporal Data of Fatal Accident. International Journal of Computer Applications, 63(8).

Op Selamat: 13 fatalities recorded in Kelantan. (2019). The Sun Daily. https://www.thesundaily.my/local/op-selamat--13-fatalities-recorded-in-kelantan-GK950329

Plug, C., Xia, J. C., & Caulfield, C. (2011). Spatia; and temporal visualization techniques for crash analysis. Accident Analysis and Prevention 43, 1937-1946.

Shahid, S., Minhans, A., Che Puan, O., Hasan, S. A., & Ismail, T. (2015). Spatial and temporal pattern of road accidents and casualties in peninsular Malaysia. Jurnal Teknologi, 76(14), 57–65. https://doi.org/10.11113/jt.v76.5843

Sitohang, H., & Merni, S. (2020, June). K-Means method for analysis of accident-prone areas in Palangka Raya. In Journal of Physics: Conference Series (Vol. 1566, No. 1, p. 012008). IOP Publishing.

Soltani, A., & Askari, S. (2014). Analysis of intra-urban traffic accidents using spatiotemporal visualization techniques. Transport and Telecommunication Journal, 15(3), 227-232.

Soltani, A., & Askari, S. (2017). Exploring spatial autocorrelation of traffic crashes based on severity. Injury, 48(3), 637–647. https://doi.org/10.1016/j.injury.2017.01.032

World Health Organization (WHO). (2018). The Top 10 Causes of Death. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

Zamzuri, Z. H., Shabadin, A., & Ishak, S. Z. (2019). Bayesian Network of Traffic Accidents in Malaysia. Journal of Information and Communication Technology, 18(4), 473-484.

Zangeneh, A., Najafi, F., Karimi, S., Saeidi, S. & Izadi, N. (2018). Spatial-temporal Cluster Analysis of Mortality from Road Traffic Injuries Using Geographic Information Systems in West of Iran During 2009-2014. Journal of Forensics and Legal Medicine, 55, 15-22l..