Spatio-Temporal Analysis and Hotspot Identification of Road Traffic Accidents in Lahore Using GIS and Data Analytics

Muhammad Mukarram Muneer[1], Khalid Mahmood[2] and Muhammad Arham Javed[3]
https://doi.org/10.62345/jads.2025.14.3.183

Abstract
In recent years, there has been a serious concern about the increasing rate of road Traffic accidents (RTAs) that cause different types of injuries, including disabilities, fatalities, etc. To overcome these losses, this research aims to identify hotspot regions and the time intervals of most occurrence in Lahore District using various Spatio-temporal techniques, including Kernel Density Estimation (KDE), Co-map analysis, and Python, Tableau, and Excel-based approaches. In addition, those hot zones are found in the district's busiest regions in the north, notably in Interior Lahore, with notable clusters in Ravi, Data Gunj Baksh, and Shalamar Towns. The most populous towns seem to have the highest rate of accidents. Temporal analysis shows an average of approximately 10 accidents per hour, peaking on Mondays and weekends, and in busy hours. The study provides actionable insights for traffic authorities, offering a replaceable framework for mitigating high-risk zones, with the potential to significantly reduce accident rates and improve road safety in Lahore and similar urban areas.
Keywords: Road Traffic Accidents, Hotspots, Rescue 1122, and Spatio-temporal Analysis.

Introduction
RTAs have become a common and pathetic issue in the modern era. Every year, approximately 1.35 million individuals are killed, and nearly 50 million individuals are injured or become disabled on the roadway, and every day, almost 3,700 people are killed globally in RTAs (WHO, 2018). Male youngsters who are under 25 seemed to be more involved in RTAs as compared to females, with 73% about 3 quarters. (WHO, 2018). Globally, the eighth leading cause of fatalities is road Accidents (RAs) in all groups of people around 5-29 years. Every year, the crash death rates (CDRs) are three times lower in developed countries than in undeveloped countries, but no elimination has been made in the number of CDRs in the undeveloped countries from 2013 to 2016. The maximum ratio of deaths in RTAs is due to motorcyclists, cyclists, and pedestrians. (WHO, 2018).
It has been predicted that MLICs (Middle and Low-income Countries) may face economic losses of approximately 834 billion dollars (USD, 2010) from 2015 to 2030 due to deaths and injuries. The economy of every country is directly associated with RTAs, which consumes the GDP (Gross Domestic Product) at around 3% cost (Satria, 2020). Road Traffic Injuries (RTIs) have also developed as the main fitness problems globally, resulting in more than a million fatalities yearly (Lam et al., 2019).
The Global Burden of Disease (GBD) study reports that road traffic injuries (RTIs) cause about 10% of global fatalities and rank as the third leading cause of disability in the past two decades. Nearly two-thirds of road traffic accidents (RTAs) occur in developing countries due to poor road safety, inadequate infrastructure, and untrained drivers (Abbasi et al., 2019; Kent et al.,

[1]*Remote Sensing, GIS, Climatic Research Lab (National Center of GIS and Space Applications), University of the Punjab, and AIFA Education, Training and Research Centre, Pakistan. Email: mmukarrammunir@786gmail.com*
[2]*Institute of Space Science, University of the Punjab, Lahore, Pakistan*
[3]*Emergency Services Academy, Lahore, Pakistan.*

2020). Globally, RTAs cause more deaths than HIV/AIDS, with high mortality rates in Thailand (23.3) and Myanmar (12.9) per 100,000 population (WHO, 2018; Satria, 2020). In Pakistan, WHO (2020) reports 28,170 RTA fatalities, accounting for 1.93% of total deaths, with an age-adjusted rate of 15.18 per 100,000.

Geo-spatial technology, like biotechnology and nanotechnology, is an emerging field encompassing sub-disciplines such as Remote Sensing (RS), Surveying, Photogrammetry, and Cartography. By integrating data from these disciplines, a Geographic Information System (GIS) is formed. Originating from geography, cartography, and computer science, GIS is widely used for spatial statistics, RS, surveying, computer cartography, and photogrammetry. It involves specialized hardware and software to collect, store, process, analyze, and visualize real-world geographic data (Chang, 2015). In accident research, GIS is beneficial for examining the geographic distribution of accidents, identifying their spatial patterns, and analyzing their local spatial statistical properties (Bucsuházy et al., 2020).

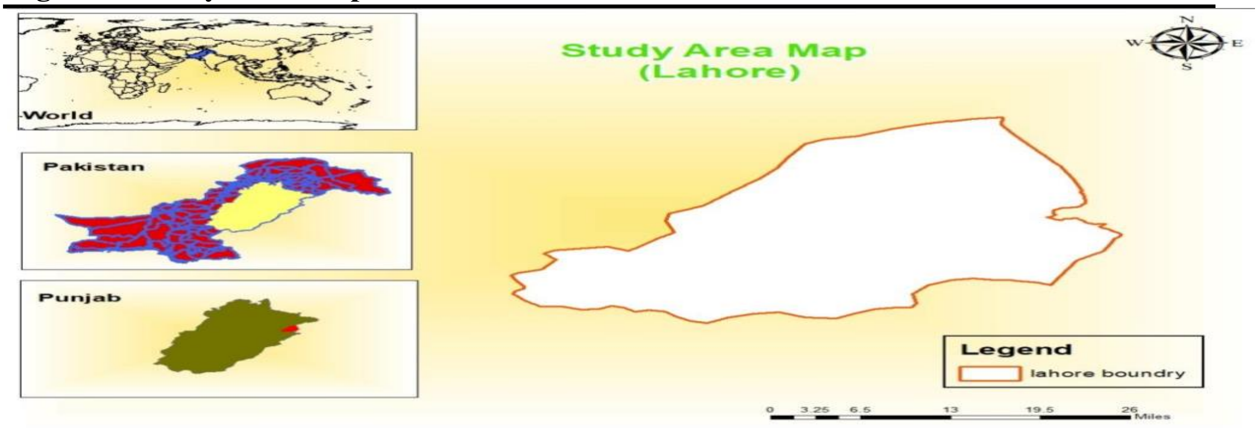
The particular objectives of my study are to find out the risky time periods of the day as well as the week, to determine the clusters in existing data on the maps, and to create some important recommendations and precautionary measures for the future.

Study Area and Data

Study Area

My Study Area was the "Lahore" District of the Punjab province, Pakistan. Metropolitan Lahore (31°15'-31°45'N, 74°01'-74°39' E) is a major city in Pakistan and the capital of the Punjab province with a flat alluvial plain. According to the results of the 6th Population Census Report 2017, Lahore is the second-most populous city in Pakistan (Ahmad et al., 2010), and it is concentrated after Karachi, with a population of 11 million and an area of 1,172 square kilometers. Lahore is the 16th most extensive and fastest growing worldwide (http://www.citymayors.com/statistics/largest-cities-population-125.html). It is found at a height of 217m above mean sea level (AMSL). It is bounded on the north and west by the Sheikhupura district, on the south by the Kasur district, and on the east by India (Wagah border). It has expanded in size over the last 12 to 15 years and is now nearly double what it was (Basharat, 2014).

Figure 1: Study Area Map of Lahore District



Lahore is divided administratively into 10 Towns which are Ravi Town (31km^2), Iqbal Town (516km^2), Wagha Town (438km^2), Shalamar Town (24km^2), Data Gunj Baksh Town (30km^2), Samanabad Town (37km^2), Gulberg Town (44km^2), Nishter Town (495km^2)), Aziz Bhatti Town (69km^2), & Cantonment (97km^2) (Mazhar and Jamal. 2009: Gabriel. 2010) and 274 Union Councils (UC) (politicpk.com, n/d).

It has different types of roads like Primary Roads, secondary Roads, & tertiary Roads, Motorways, Trunk Roads, Residential Roads, Unclassified Roads, Service Roads and some others small roads including link Roads and streets, etc (MM Pakistan Pvt Ltd).

Data Collection

Accurate data collection is essential for ensuring the credibility and reliability of any research. In this study, road traffic accident (RTA) location coordinates serve as the primary data source. Traffic accident datasets (TAD), often maintained in Excel format, provide valuable temporal and historical information but are limited in conducting detailed spatial analyses. For Kernel Density and spatio-temporal analysis, both accident datasets and shape-files are required.

The TAD used in this study was obtained from the Rescue 1122 department in Punjab, Pakistan, which compiles accident records in Excel format. The data-set includes key variables such as longitude, latitude, incident address, date and time, response time, number of victims, number of fatalities, types of injuries, dispatched vehicle type, and a Google Maps link for each accident location.

The available TAD is based on 125682 RTAs for the 16-month period from February 2021 to May 2022. Excel data has been divided into five categories: (1) Total Victims, (2) Minor injuries, (3) Major injuries, (4) Deaths and (5) Not Mentioned. In Europe, a severe injury is characterized as a person who has been hospitalized for more than 24 hours, whereas a minor injury is defined as less than 24 hours (Eurostat, 2020). Based to my data, non-fatal victims who get transported to a hospital by rescue-1122 have a severe injury, but those who return home after receiving first aid from rescue 1122 have a minor injury. The total number of Death, Major, and Minor Injury victims are 543, 46262 and 58440 respectively which are calculated by the use of Tableau (Data Visualization Software). In terms of the severity of the injuries, there are three primary categories of RTAs, but the fourth is "not mentioned" in my data. When one or more persons are killed in an accident, it is referred to as a fatality accident. Severe injury accidents occur when a severely injured individual is involved in a crash, but no deaths are reported, whereas minor injury accidents result in just minor hurt individuals. "Not mentioned" suggests that Rescue 1122 was unaware of the victims' condition. It might just be a property damage accident. Victims are sometimes transferred to hospitals or sent home before rescuers arrive, if they are stable or dead, with the assistance of someone or a family member. MM Pakistan (Pvt) Ltd also provided shape-files of Lahore's districts, Tehsils, towns, union councils, and roads. These shape-files are utilized in ArcGIS 10.4.1 for various analysis.

Data Organization or Preparation

Data organization is also an essential component of every research project. Data were first pre-processed before being utilized for further analysis. For cleaning the data, I utilized the Python language's "Pandas" library and MS Excel tools. TAD is read out in the Jupyter notebook (a web-based Python platform), and the insufficient columns and empty rows and columns are removed using the following programs:

Figure 2: Data Cleaning

```
import numpy as np
import pandas as pd
data=pd.read_excel("Updated RTCs final file for thesis with extra Columns.xlsx")
data1=data.drop("Emergency Type", axis=1)
df=data1.drop("District Name", axis=1)
df=df.drop_duplicates()
df=df.dropna(axis=0,how="all")
df=df.dropna(axis=0,how="all")
df
```

	Incident Location	Dispatched Vehicles Type	Date	Month	Year	Day	Day Division	Hours Status	Time	Season	...	Reason for late response	Total Victims	First Aid	Shifted	Dead	Not Mentioned
0	Faiz Rd, Muslim Town, Lahore, Punjab 54000, Pa...	Ambulance	1	Feb	2021	Monday	NIGHT	Non-Busy Hours	01:20:00	winter	...	Call to calar	0	0	0	0	0
1	Office No 49 G, PMA Trade Center,, Ferozepur R...	Ambulance	1	Feb	2021	Monday	NIGHT	Non-Busy Hours	01:52:00	winter	...	NaN	1	1	0	0	0
2	Qartaba Chowk Clock Tower, Ferozepur	Ambulance	1	Feb	2021	Monday	NIGHT	Non-Busy Hours	02:39:00	winter	...	yes	0	0	0	0	0

As we all know, an excel file cannot be utilized for creating a shape file in ArcGIS (10.4.1). For that purpose, the compiled excel data had been converted into a CSV file for the creation of point shape files of TAD of Lahore district for further GIS based analysis. I used excel formulas to pre-process the TAD and spread it in additional columns, Python programs, and ArcGIS software tools which are date, month & year separately, day division (day & night), day (Mon, Tue, etc), time of accident, hours status (busy, non-busy & normal), season, and towns.

Before pre-processing steps, date, month, year, and time were combined in one column but now I distributed them in separate columns for each through excel formulas and short tools. New columns of "day" (Mon, Tue, etc) and "Day division" are also created using excel formulas and python programs. In "day division", I divide the time into two parts day and night by using this scheme:

i. April to Sept (Day = 5:31am to 18:30pm and Night = 18:31pm to 5:30am)
ii. Oct to March (Day = 6:16am to 17:45pm and Night = 17:46pm to 6:15am)

New category of Season as per the following scheme has been created. The year has been divided into four seasons based on specific date ranges: **winter** extends from December to February, **spring** from March to May, **summer** from June 5 to September 10, and **autumn** from September 5 to November 10. Traffic hours are categorized into three types. **Busy hours** run from 9:00 AM to 4:59 PM. **Normal hours** include 7:00 AM to 8:59 AM and 5:00 PM to 10:59 PM. **Non-busy hours** occur from 11:00 PM to 6:59 AM.

New Column "Town" was created with the help of ArcGIS 10.4.1 software. For this purpose, I used the shape files of the Towns of Lahore and the Point shape file of TAD. Firstly, I separated each town with the help of clipping with point shape files and added the extra field (Town). In the end, collect all towns in a separate single-shape file.

Methodology

Various approaches have been studied to enhance the understanding of RTAs, with complexity increasing as more data becomes available, leading to more detailed outcomes. Based on the available data-set, an evaluation of different methodologies was conducted, considering their outcomes and respective advantages and limitations. The most suitable methods for achieving this study's objectives were identified as: (a) Kernel Density Estimation (KDE), (b) hotspots detection using Python, Tableau, or Excel, (c) temporal analysis using Python, Tableau, or Excel, and (d) Co-map analysis.

Hot Spots Determination

Python, Tableau, and Excel-Based Analysis

Data analysis involves examining raw data to extract meaningful insights, while data science is a multidisciplinary field that applies logical, statistical, and computational techniques to structured and unstructured data using tools such as data mining and big data analysis (Khadka, 2019). In this study, Python, Excel, and Tableau were employed to analyze the accident data-set and conduct a comparative assessment of their functionalities. Python was primarily used for data cleaning and pre-processing, utilizing the Pandas and Matplotlib libraries. Pandas, an open-source library under the BSD license, provides data structures and functions for manipulating and analyzing tabular and time-series data (Aurora, 2020). Matplotlib, a widely used Python plotting library, supports data visualization through a range of plots and charts and can be integrated with graphical tool kits such as wxPython, Tkinter, and Qt. It also offers an object-oriented API for embedding visualizations into applications (Aurora, 2020).

In Excel, I split the required data by using the Power-pivot data manipulation extension, add a slicer, and categorize into quarters based a sales date. The slicer provides the ability to the executive to determine sales based on quarter.

This research will also shows how to dissect data by using Python along with the libraries of Python especially Pandas to analyze different types of data. This research also demonstrates Tableau and Excel-based data analysis and will also conclude the preparation of data analysis, performing simple analysis based on statistics and creating meaningful data visualization.

Temporal Analysis

The temporal analysis affords valuable insight into hazardous material release management because this analysis can create a baseline of activity and expose new trends. Temporal data could be shown using a simple line chart or spider plots because they can demonstrate continuity and the chronological order of Accidents (Boba, el 2005). Therefore, the particular interest of this research is a line graph. This research used four different dimensions: spatial, non-spatial, temporal, and spatio-temporal analyses. After all, data can be presented by the hours, days, months, seasons, and yearly (Gerber et al., 2014). Temporal data could be revealed using a simple line chart (Boba et al., 2005)

Python, Tableau, and Excel-based analysis

Details of this approach I already discussed in 3.1.1. These three software's are again utilized for the temporal analysis. Data-set was integrated at seasonal, monthly, weekly, day and hourly intervals and these procedures were performed to identify changes over time by plotting RTAs against time. This has been complete to calculate changes by time.

Table 1: Accidents on Different Temporal Intervals

Time Interval	Number of Intervals	Number of Accidents per Interval/125682
Months	16	7855.12
Weeks	69.3	1813.59
Days	485	259.13
Hours	11640	10.79
Minutes	698400	0.1799570
Seconds	41904000	0.0029992

The Co-Map Method

Spatio-temporal methodologies are applied when data possess both spatial and temporal attributes. An event in a spatio-temporal database represents a phenomenon occurring at a specific location *x* and time *t*. In this study, spatio-temporal analysis was performed in ArcGIS 10.4.1 using the Co-map method, which effectively integrates spatial and temporal information. The Co-map approach subdivides crash data into selected time intervals—such as hours of the day, day/night, days of the week, or months—followed by spatial analysis, such as Kernel Density Estimation (KDE). The results are then presented sequentially to illustrate spatial distribution changes over time. Previous studies (Plug et al., 2011; Asgary et al., 2010; Plug et al., 2012) have successfully applied Co-map to visualize accident variations across different temporal scales. In this research, maps were generated based on hourly and day divisions within specific towns, identifying death and major injury hotspots at corresponding times.

Kernel Density Estimation (KDE)

Hotspots are locations with high rates of accidents, fatality, injury, and property damages. This particular identification of RTA applications is the supreme essential and effective technique in transportation systems (Loo & Anderson, 2015). We argue that the detection of incident hotspots locations mainly follows the incident-based and characteristic technique.

To calculate the amount according to unit area of an object, a point or poly-line characteristic can create an easy flat surface by way of the use of a KF. In theory, there is a single smooth surface separately at every point on the basis of density factor of point or poly-line. The value of the surface is maximum at the location of the point feature and decreases as the distance from the point feature rises. At the threshold disassembling from the point feature, it approaches to “0”. If the population field is set to "NONE," the value is allocated as "1" at any point, and the value entered in a "Population field" is equal to the density. Summing the values of all Kernel surfaces that are located in the cell's center determines the allocation of each raster cell of output in KDE. The planar KDE mathematical formulation is presented by Xie & Yan (2008).

Formulas of KDL for calculations in given in equation 4.10:
The density at the new spot (x, y)

$$Density = \frac{1}{(radius)^2} \sum_{i=1}^{n} \left[\frac{3}{\pi} \cdot pop_i \left(1 - \left(\frac{dis_i}{radius} \right)^2 \right)^2 \right] \quad (1)$$

$$For\ dis_i < radius \quad (2)$$

Where,

i=number of points

pop_i = value of the population field

dis_i = distance from the point to (x, y) point

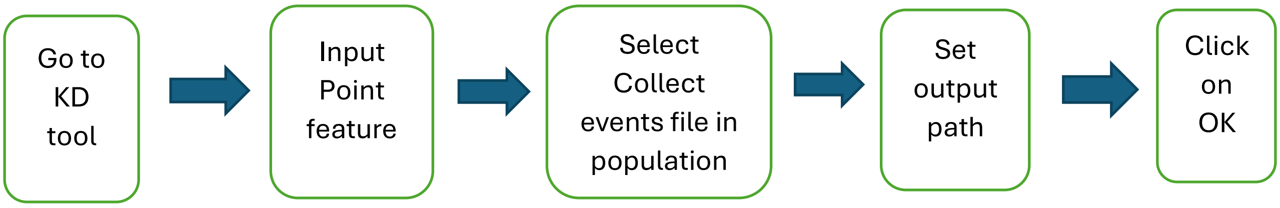
“The formula will require to be calculated for each point. Then a raster is formed. These calculations are applied to every center of the cell in the output raster (Silverman, 1986)."

Default search radius (Bandwidth):

$$SearchRadius = 0.9 \times \min\left(SD, \sqrt{\frac{1}{1n(2)}} \times D_m\right) \times n^{-0.2} \quad (3)$$

Where.
SD = Standard distance.
D_m = weighted median distance between (w) and mean center.
n = sum of the values of the populated field.
Customarily, the threshold distance is dependent on the input units (linear) but in any other cases it relies upon the units for output projection. The unit zone which is selected for KDE is expanded by an appropriate component then it is written into raster data of output. This tool has implemented to the data to notice the density of spots or points at distinct places. The following steps are given in flowchart;

Figure 3: Steps for KDE



Results and Discussion

Using the three techniques described earlier, the results of data analysis and visualizations are presented in this section. Each map, table, and graph illustrates traffic safety patterns in Lahore District, providing valuable insights for decision-makers.

Python, Tableau & Excel Based Data Analysis

By using the Excel pivot table, I performed different analyses of data. First I found the all death spots and then sorted & filtered the top 30 locations where maximum deaths occur as shown in Figures 4 (a) and 20 (a). Similarly, found all major & minor injury spots and then also sorted & filtered the top 30 locations where maximum major & minor injuries happened respectively and mapped which are shown in figures 4 (b), (c), and 5 (b) & (c).

Figure 4 (a): Death Spots

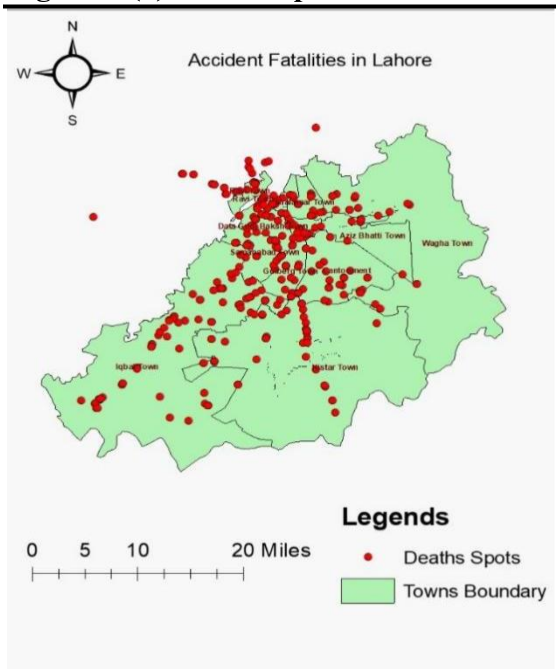


Figure 4 (b): Major Injury Spots

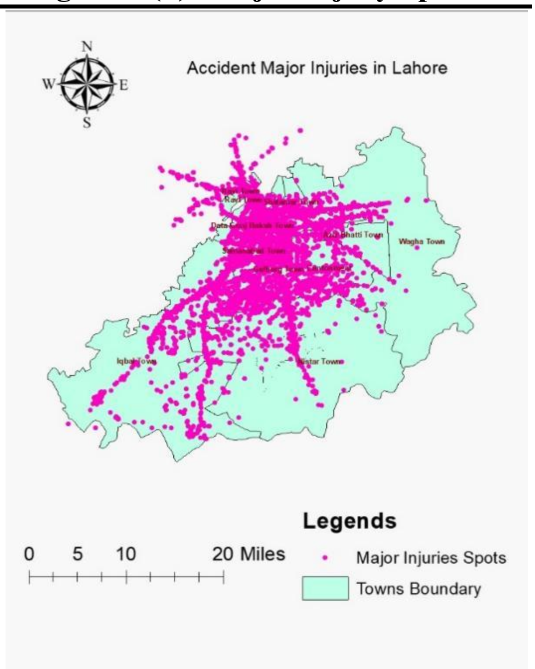


Figure 4 (c): Minor Injury Spots

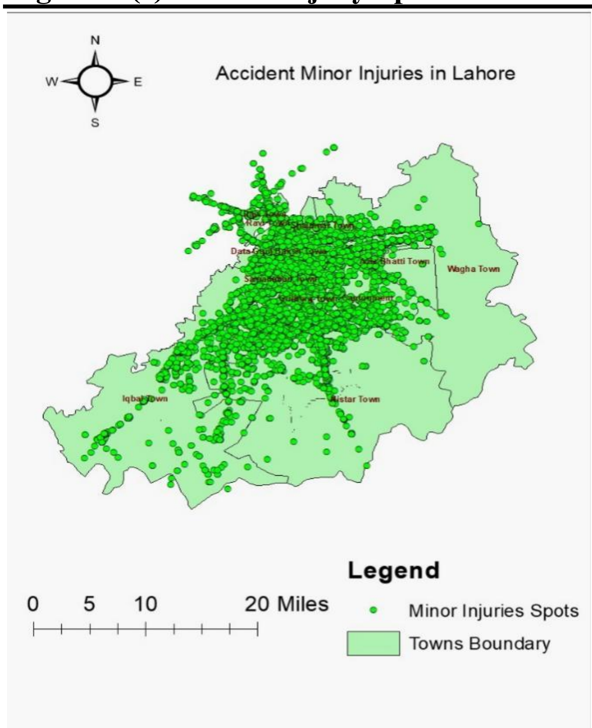


Figure 5 (a): Top 30 Death Spots

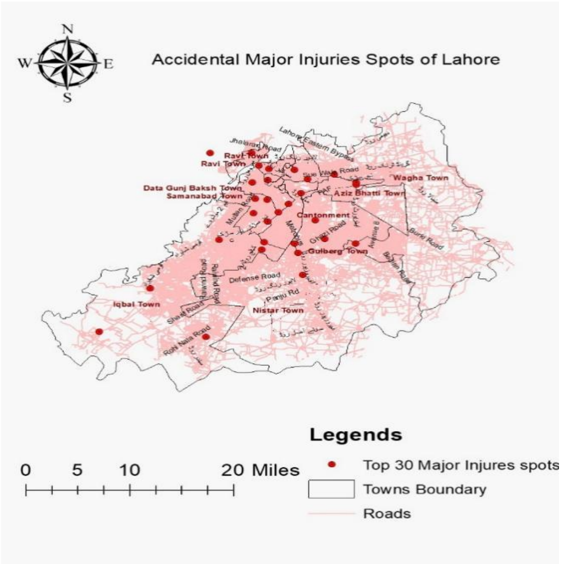


Figure 5 (b): Top 30 Major Injury Spots

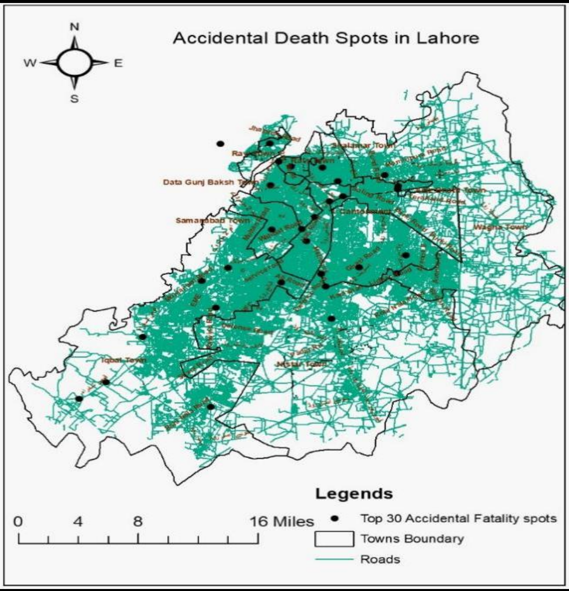
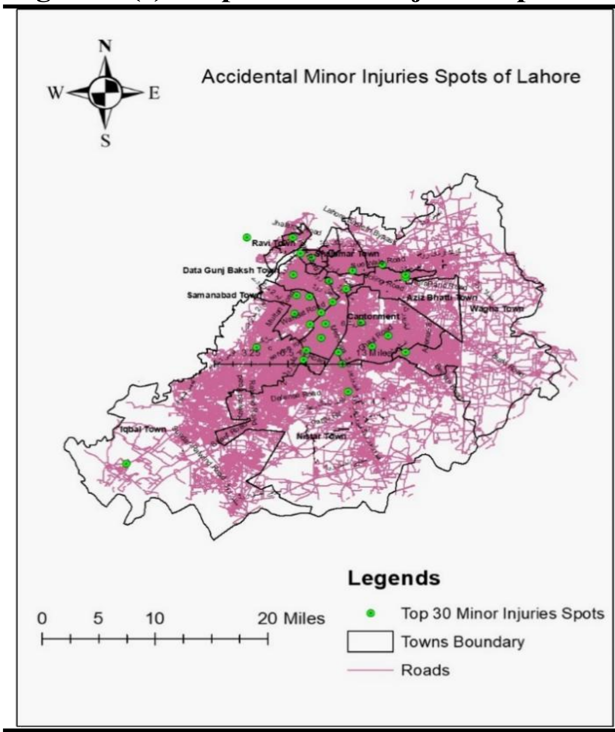


Figure 5 (c): Top 30 Minor Injuries Spots



Results can be visualized after the data analysis by using the pandas' library of python programming in Jupyter notebook. Figure 6 represents the highest accident-based analysis on each longitude and latitude. In this analysis, every spot represents a specific number of accidents, this was done after a few lines of programming code.

From the presented maps, it is evident that accident-related fatalities, major injuries, and minor injuries in Lahore are highly concentrated along the central urban corridors and major road networks, particularly in areas with high traffic density. The spatial clustering of top 30 accident spots highlights that most severe incidents (deaths and major injuries) occur along primary transport routes and intersections, suggesting a strong link between traffic volume, road design, and accident severity. In contrast, minor injuries appear more dispersed across the city, indicating that less severe accidents are widespread and not limited to specific hotspots. Overall, these spatial patterns justify the need for targeted traffic management, road safety interventions, and urban planning measures focused on high-risk corridors to reduce fatalities and severe injuries.

Figure 6: Python-Based top 30 Hazardous Spots

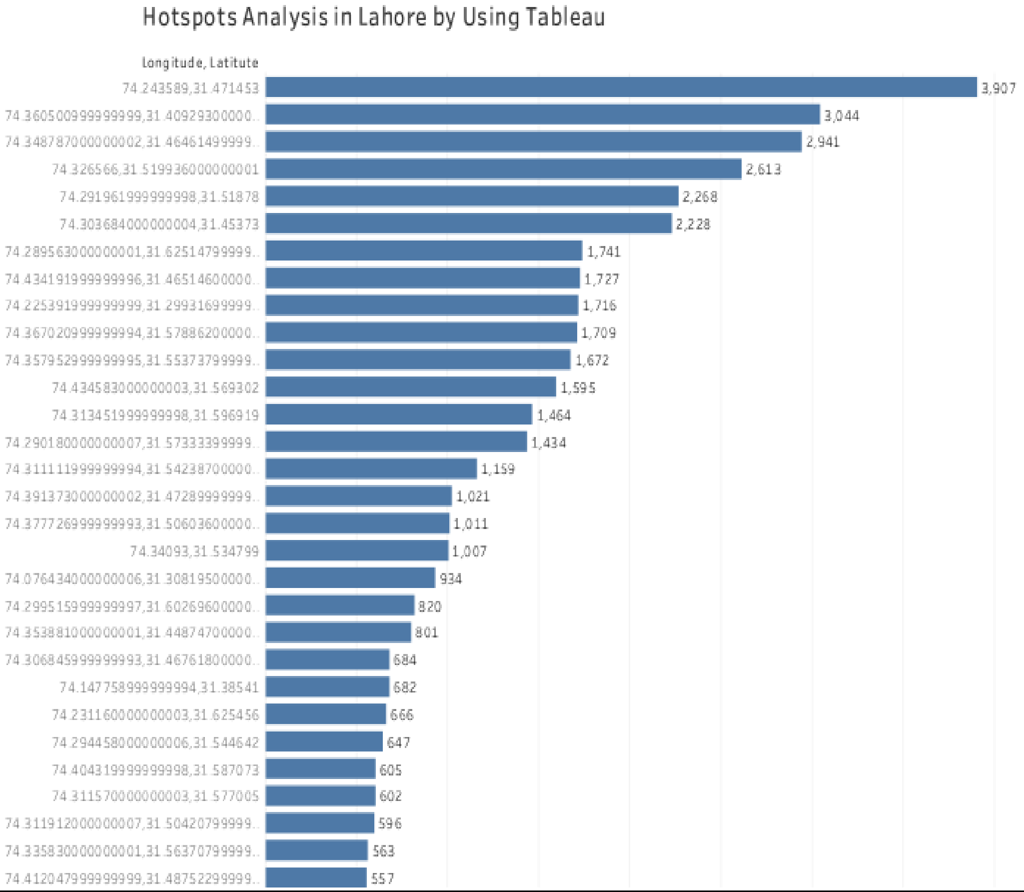
In [127]:
```
df1=df["Latitude, Longitude"].value_counts().head(30)
df2 = df1.to_frame(name="Number of Accidents").reset_index()
df2.index.name = 'index'
df2.rename(columns = {'index':'Latitude, Longitude'}, inplace = True)
df2
```

Out[127]:

index	Latitude, Longitude	Number of Accidents
0	31.471453,74.243589	3907
1	31.409293,74.380501	3044
2	31.484815,74.348787	2941
3	31.519936,74.328588	2813
4	31.51876,74.291982	2288
5	31.45373,74.303884	2228
6	31.625148,74.289583	1741
7	31.485148,74.434192	1727
8	31.299317,74.225392	1718
9	31.578882,74.387021	1709
10	31.553738,74.357953	1672
11	31.589302,74.434583	1595
12	31.598919,74.313452	1484
13	31.573334,74.29018	1434
14	31.542387,74.311112	1159
15	31.4729,74.391373	1021
16	31.508038,74.377727	1011
17	31.534799,74.34093	1007
18	31.308195,74.078434	934
19	31.602898,74.299516	820
20	31.448747,74.353881	801
21	31.487818,74.308848	684
22	31.38541,74.147759	682
23	31.625458,74.23118	668
24	31.544842,74.294458	647
25	31.587073,74.40432	605
26	31.577005,74.31157	602
27	31.504208,74.311912	598
28	31.583708,74.33583	583
29	31.487523,74.412048	557

By using Tableau software, results can be visualized after the data analysis. Figure 9 represents the simple highest crash-based analysis on each longitude and latitude. This analysis is performed for the results verification and comparison with the other approaches. In this analysis, every spot represents a specific number of accidents with a bar chart, this was also created after sorting and filtering the data.

Figure 7: Tableau-based Top 30 spots w.r.t no. of Accidents



In my large data-set, there are only three spots exited where all types of injuries occurred (death, major and minor injuries) as shown in figure 8. The top two spots out of 3 are also part of the top 30 hotspots which were determined in the approach of "hotspots analysis by crash SI". This result is also obtained by using the python programming in Jupyter notebook.

Figure 8: Spots with all types of injuries

In [141]:

```
df1=df.loc[df["Dead"]>=1]
df2=df1.loc[df["Shifted"]>=1]
df3=df2.loc[df["First Aid"]>=1]
df4=df3.sort_values(by="Total Victims", ascending=False)
df4[["Latitude, Longitude","Total Victims","Dead","Shifted","First Aid"]]
```

Out[141]:

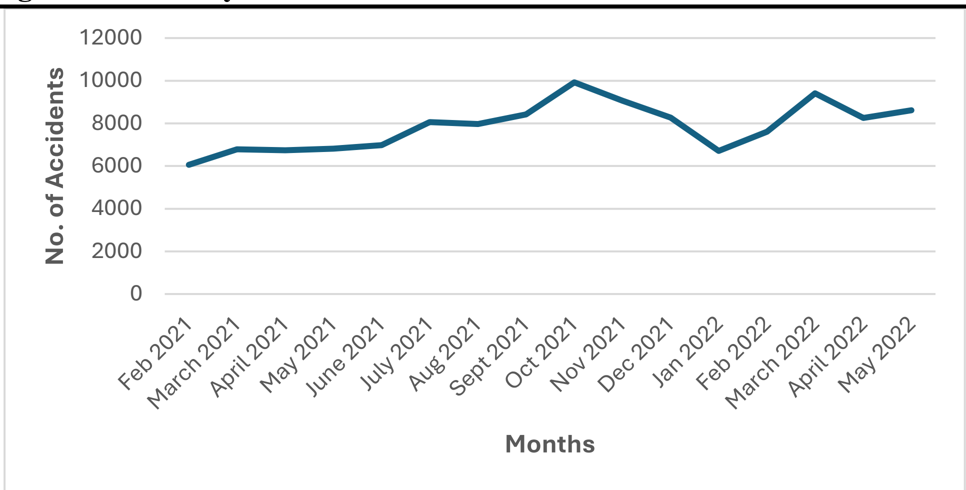
	Latitude, Longitude	Total Victims	Dead	Shifted	First Aid
104650	31.5047752,74.3317265	8	1	5	2
121147	31.38541,74.147759	7	1	5	1
67619	31.406268,74.360619	3	1	1	1

Temporal Analysis

Python, Tableau, and Excel-Based Data Analysis:

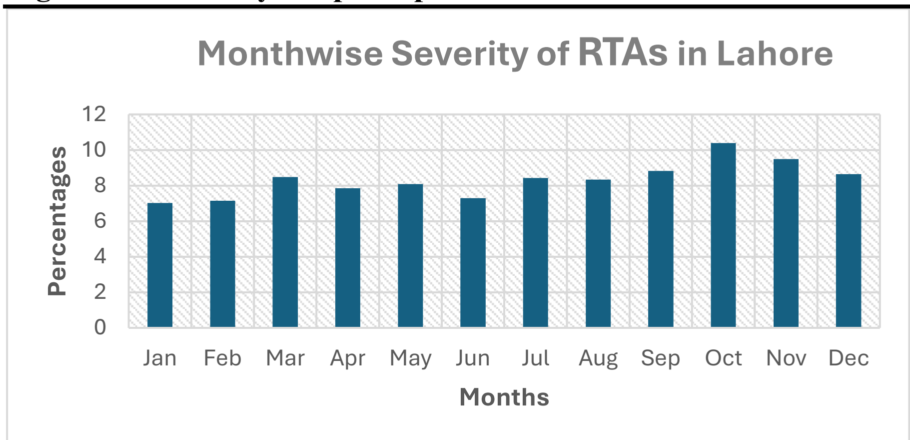
The entire data-set of 125682 RTAs from Feb 2021 to May 2022 in the spreadsheet, was utilized for all explanatory graphs in this portion. To study how the location of recognized clusters varies over time, the data were split into monthly, seasonal, day-by-day, and hourly components. Temporally, data analysis was done by using Python Programming, Excel Pivot Tables, and Tableau-based data analysis and visualization.

Figure 9: Monthly Variation of RTAs in Lahore



Monthly variation of RTC victims in Lahore shows a slight increase in March 2021, followed by a steady trend until June, then a sharp rise in July that persisted until August. The numbers rose again, peaking in October 2021, declined from November 2021 to January 2022, and reached a second peak in March 2022, after which fluctuations continued until the study's end. The two peaks occurred in October 2021 and March 2022, both months with moderate weather. Notably, accident numbers increased from mid-2021 onward, likely due to the easing of COVID-19 restrictions and the removal of smart lockdown measures in Lahore.

Figure 10: Monthly temporal patterns of RTAs



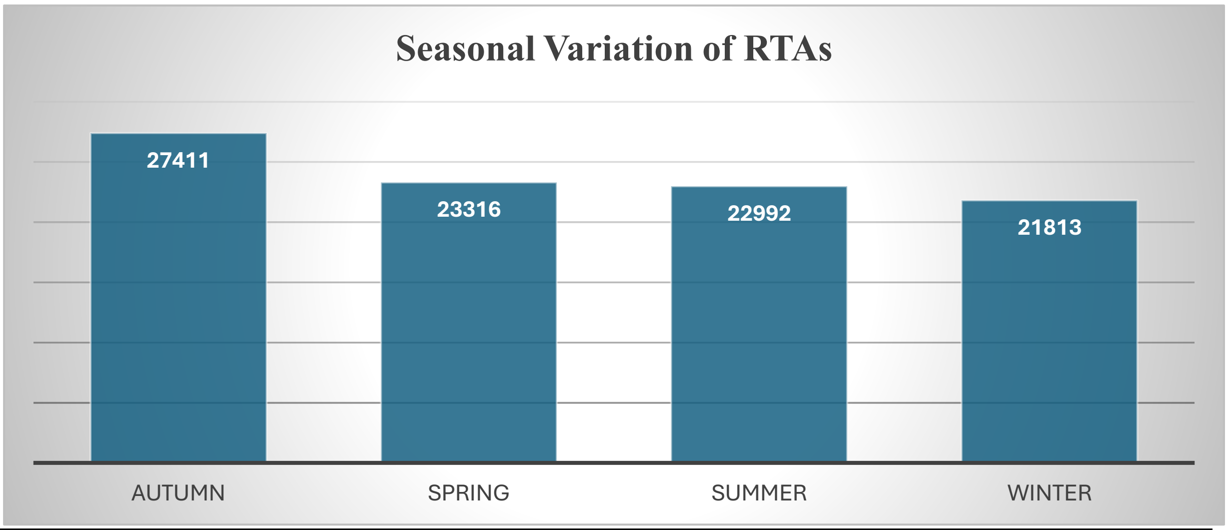
The lowest value of RTC victims was reported in Jan as shown in figure 10, it may be because individuals become more cautious in winter, and bike riders minimize the speed of their automobiles due to reason of the impact of low temperatures on them thus less number of automobiles are on the roads and mobility of this school going age groups has reduced of some extent. The use of helmets is practiced more in the winter season as compared to the summer which automatically reduced the RTAs. Dataset is divided into six classes i.e. (1) Number of accidents, (2) Number of Total Victims, (3) Minor Injuries, (4) Major injuries, (5) dead, and (6) Not Mentioned from Feb 2021 to May 2022. By using excel and python programming, a summary of RAs in all months which shown below in table 2.

Table 2: Summary of RTAs

Sr.	Months	No. of Accidents	No. of Total Victims	Dead	Major Injuries	Minor Injuries	Not Mentioned
1	Feb 2021	6058	3666	11	1454	1461	740
2	March 2021	6784	5915	18	2401	2614	882
3	April 2021	6733	6084	20	2251	2877	936
4	May 2021	6820	6492	16	2491	2902	1083
5	June 2021	6972	6636	33	2703	2957	943
6	July 2021	8057	7394	28	2917	3295	1154
7	Aug 2021	7963	7573	35	2949	3322	1267
8	Sept 2021	8421	8242	40	3218	3860	1124
9	Oct 2021	9924	9090	42	3411	4558	1079
10	Nov 2021	9066	8463	41	3161	4383	878
11	Dec 2021	8264	7855	41	3119	4108	587
12	Jan 2022	6716	6653	55	2757	3586	255
13	Feb 2022	7609	7160	34	2839	4252	4
14	March 2022	9414	8859	45	3613	5162	0
15	April 2022	8259	7917	43	3362	4482	0
16	May 2022	8621	8289	41	3616	4621	0
17	Feb 2021 to May 2022	125681	116288	543	44808	56979	10192

One important thing is noticed that, after august 2021, rescue 1122 gradually tried to decrease the number of "not mentioned" and in the last three months of my data, there were 0 "not mentioned" victims. It means that they are improving their services day by day because "not mentioned" are those accidents in which rescue don't know about the condition (dead, major or minor injury) of the victims, it seems like the rescue team reached late at the accident spot and victims go to hospitals with the help of others or by own or go home after the minor injury or dead body sent to the home to the victims.

Figure 11: Seasonal Variation of RTAs in Lahore



Normally, the ratio of RTAs is higher in moderate weather which exists are spring and autumn. After the data analysis shown in figure 11, it was also found that accidents are high in autumn and spring as compared to winter and summer. The same reason which I already discussed in the previous monthly results.
Day-wise analysis of TAD leads to determinate the peak of the accidents that took place during particular days. A research of the condition related with that specific day helpful in determining the causes of reoccurring accidents on the similar days.

Figure 12: Severity of RTAs w.r.t to the type of Day

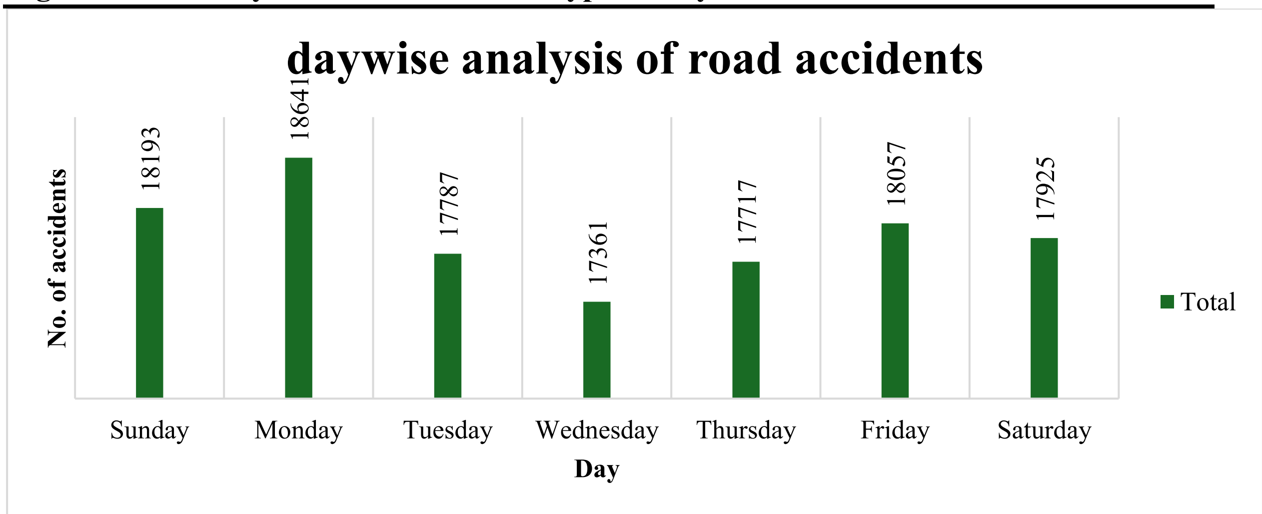
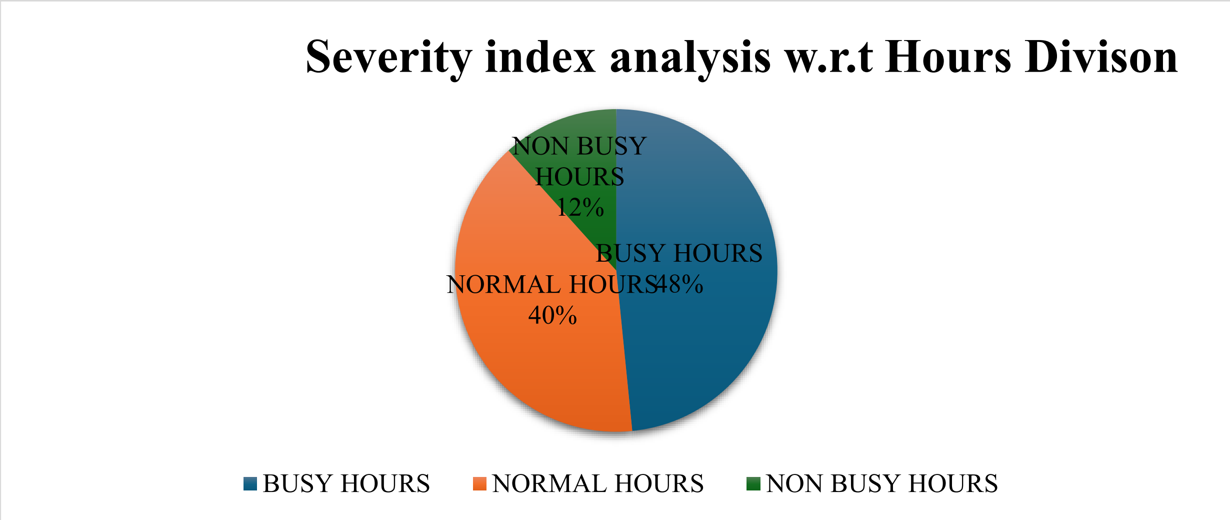
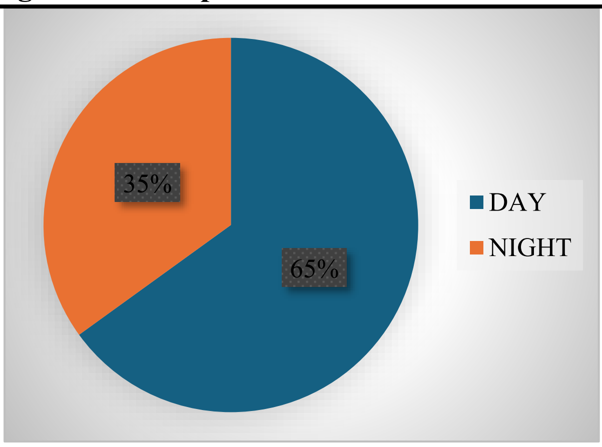


Figure 13: RTA Analysis w.r.t Hours Division



The highest number of accidents in Lahore occur on Mondays and weekends. As the capital of Pakistan's most populous province, Lahore attracts people from neighboring districts for its diverse facilities, including educational institutions, hospitals, courts, industries, and recreational sites. Many office workers and students travel to Lahore on Mondays or Sundays, while major appointments are often scheduled on Mondays, and tourists visit on weekends. Conversely, workers and students often return to their hometowns on Fridays or Saturdays. As noted in Section 2, 24 hours were divided into three categories for analysis, revealing that 48% of accidents occurred during busy hours—coinciding with office, educational, court, and industrial timings—while 40% happened during normal hours. Overall, 88% of accidents occurred from morning until night, underscoring Lahore's continuous high activity levels.

Figure 14: Temporal Variation of RTAs during Day & Night



In Lahore, a huge variation is shown in the graph of day and night RTAs because accidents are usually reported from 9 am to 7 pm. Lahore is the capital as well as the largest city of the Punjab Province therefore many other residential visits Lahore during day time.

Co-map Method

The Spatio-temporal analysis is performed with the usage of the Co-map method which is the most suitable regarding my data-set. All results are shown below:

The Co-map method was applied to assess whether hazardous locations in Lahore's RTAs are influenced by both spatial and temporal factors. The analysis incorporated time-based categories day/night and busy, non-busy, and normal hours—alongside town-wise zones and injury types (deaths and major injuries). Figures 15 and 16 depict fatality patterns, while Figures 17 and 18 address major injury distributions, revealing distinct spatio-temporal variations in RTCs.

As shown in Figure 15(a), during busy hours (09:00–14:59), 242 of 543 fatalities occurred, coinciding with high traffic volumes. The highest counts were in Iqbal Town (~70 deaths), followed by Shalamar (~30), Ravi (~25), Data Gunj Buksh and Gulberg (~20 each), Nishter (16), Aziz Bhatti (15), Cantonment (14), Wagah (7), and Samanabad (5). Multiple small clusters appeared in Nishter, Iqbal, Data Gunj, Aziz Bhatti, Shalamar, and Wagah Towns. In Figure 15(b), during non-busy hours (late night/early morning), only 45 fatalities were recorded, over half of them in Nishter Town (21). Ravi recorded 5 deaths, Data Gunj Buksh 4, and no deaths occurred in Aziz Bhatti. Clusters were identified in Ravi, Data Gunj, and Shalamar, often linked to employee travel, heavy truck operations, poor lighting, and driver fatigue. Figure 15(c) shows that in normal hours (07:00–08:59 & 17:00–22:59), 225 fatalities occurred, with the highest in Iqbal Town (>80), followed by Data Gunj Buksh and Ravi (~30 each), Nishter (~25), Cantonment and Gulberg (~20 each), Shalamar (~15), Wagah (>10), and Samanabad/Aziz Bhatti (~10 each).

In Figure 16(a), daytime accounted for 354 fatalities, reflecting high vehicular activity. Iqbal Town recorded ~130 deaths, followed by Gulberg (~35), Cantonment (~30), Samnabad (~30), Data Gunj Buksh (~30), Ravi (~40), Aziz Bhatti (~20), Nishtar (>25), Shalamar (~20), and Wagah (~10). Conversely, Figure 16(b) shows that nighttime fatalities totaled 158, with Iqbal (35), Nishter (30), Data Gunj Buksh (28), Gulberg (25), Ravi (23), Shalamar (15), Wagah (7), Samnabad (6), and Aziz Bhatti (4) reporting lower but still significant clusters. Similar to non-busy hours, nighttime patterns are influenced by worker travel, heavy truck movements, poor lighting, and fatigue-related crashes involving trucks and motorcycles.

Fig 15 (a): Death Spots during Busy Hours

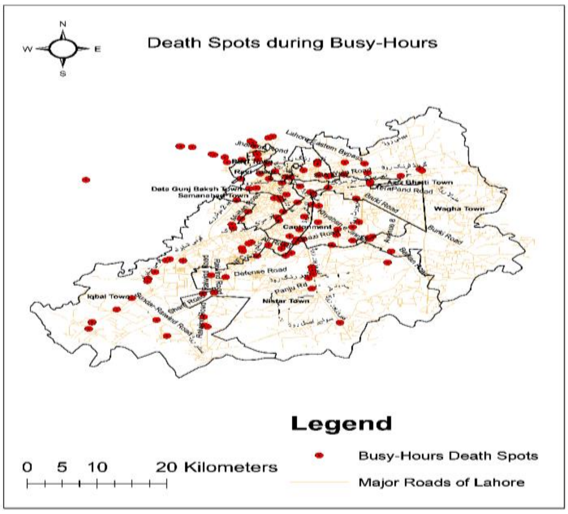


Fig 15 (b): Death Spots during Busy Non Hours

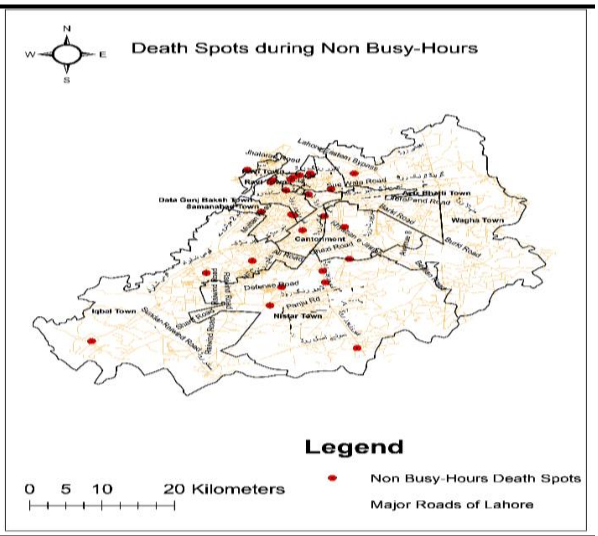


Figure 15 (c): Death Spots during Normal Hours

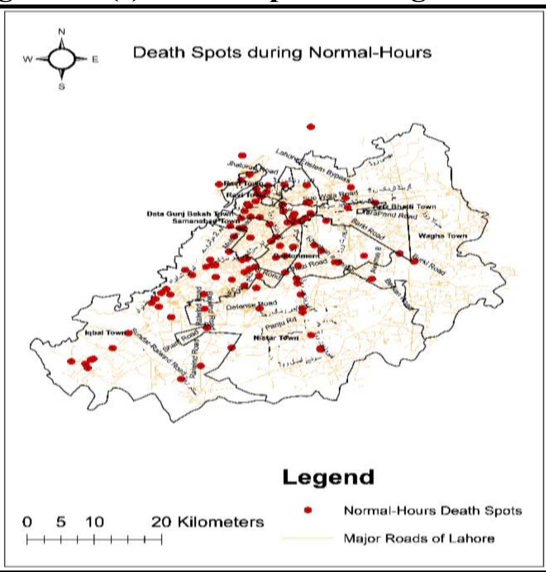


Figure 16 (a): Death Spots during Day Time

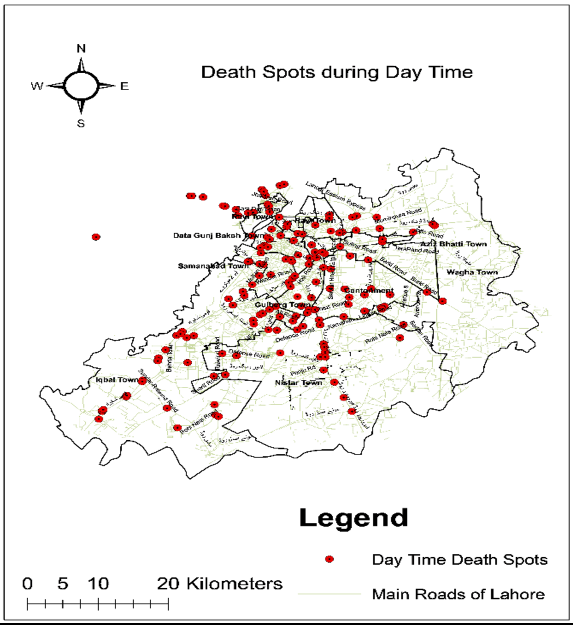
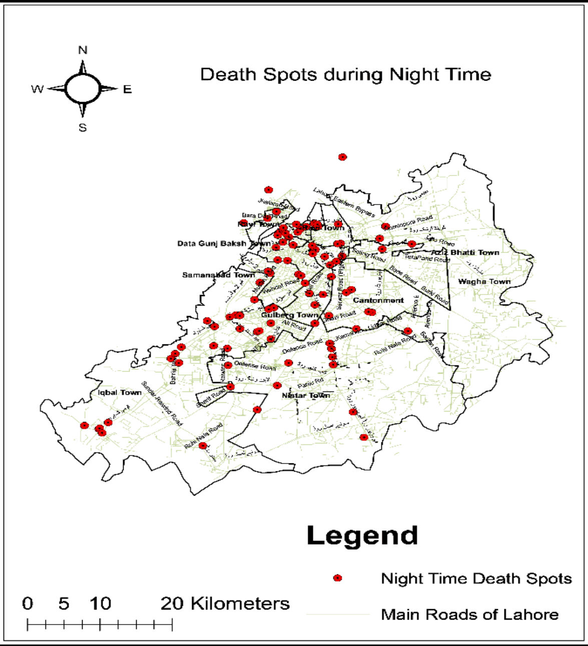


Fig 16 (b): Death Spots during Night Time



In **Figure (17a-c)**, major injury patterns demonstrate clear temporal and spatial variations. During busy hours (09:00–14:59), 19,029 major injuries were recorded, with the highest concentrations along Multan and Ferozpur Roads. Significant clusters were observed in Data Ganj Baksh, Ravi, Shalamar, Samanabad, and Gulberg Towns, with smaller clusters in Cantonment, Iqbal, and Nishter Towns. In non-busy hours (late night/early morning), major injuries dropped to 3,548, with Ferozpur Road as the most affected corridor. High clusters persisted in Data Ganj Baksh, Ravi, and Shalamar, while Gulberg, Cantonment, Iqbal, and Nishter showed smaller clusters. During normal hours (07:00–08:59 & 17:00–22:59), 16,636 major injuries were reported, again concentrated along Multan and Ferozpur Roads, with spatial patterns mirroring busy hours—large clusters in Data Ganj Baksh, Ravi, Shalamar, Samnabad, and Gulberg, and smaller clusters in Cantonment, Iqbal, and Nishter Towns. These patterns indicate that major injury hotspots in Lahore are heavily associated with high-traffic corridors and central urban towns, with reduced but persistent clustering during low-traffic periods.

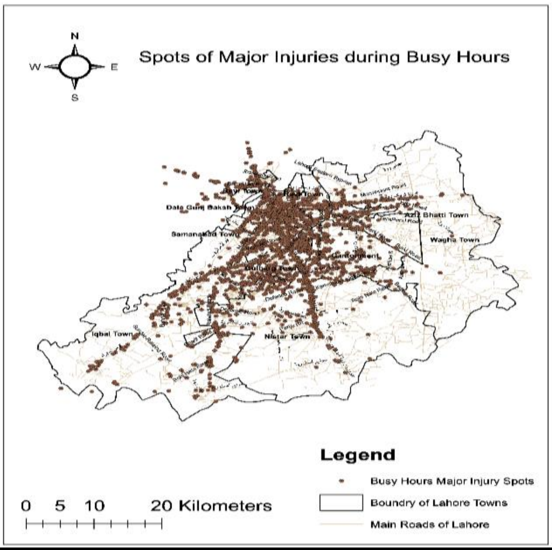
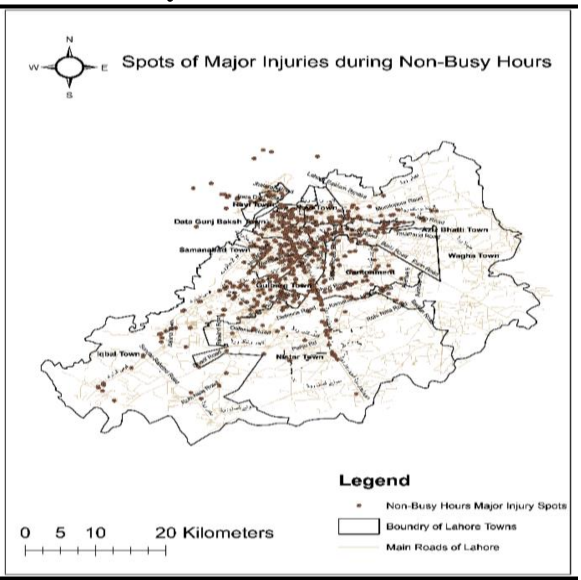
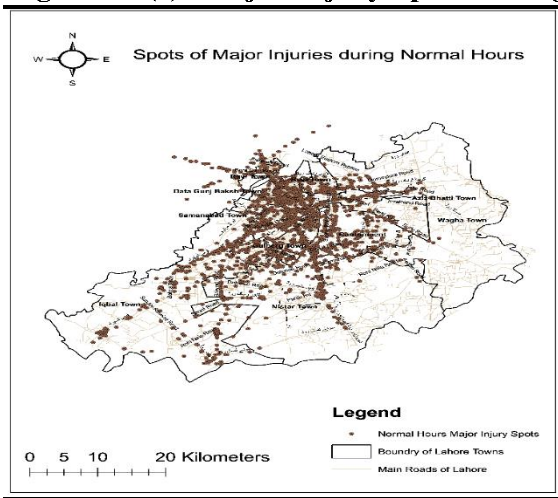
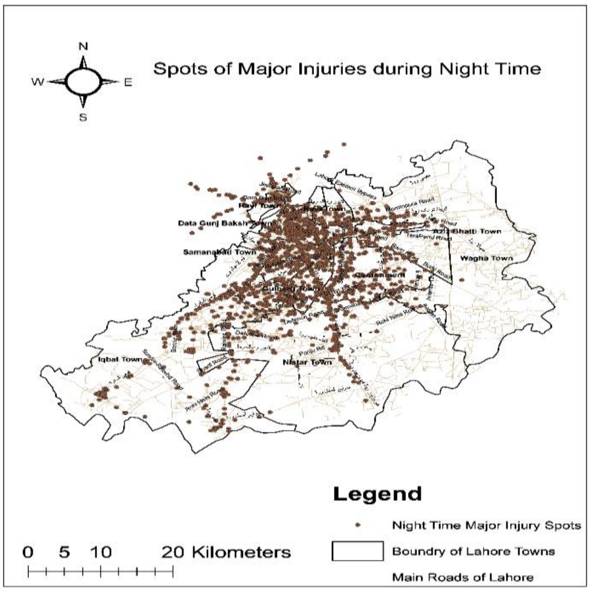
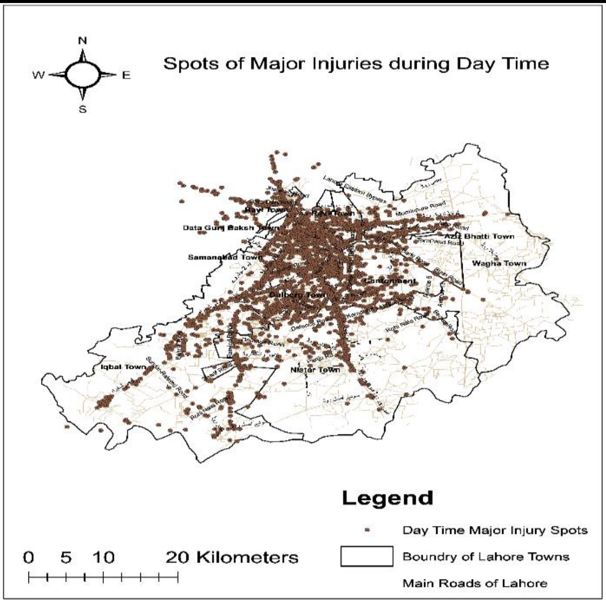
Figure 17(a): Major Injury Spots during Busy Hours	**Figure 17(b): Major Injury Spots during Non-busy Hours**
	

Figure 17 (c): Major Injury Spots during Normal Hours



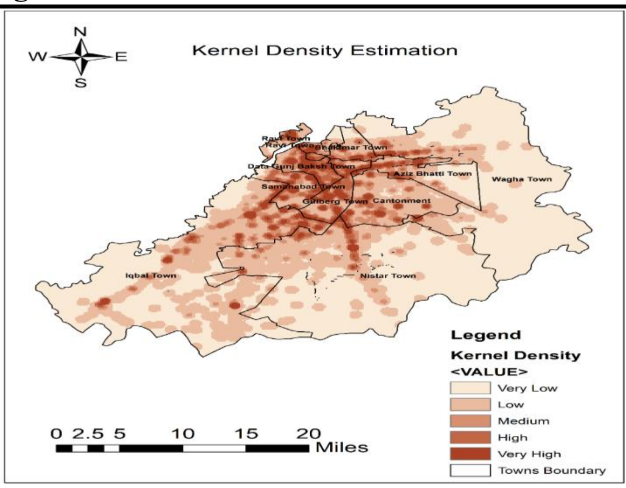
As shown in Figure 18(a), daytime accidents resulted in 26,352 major injuries, primarily due to the large number of road users during this period, with the highest occurrences on Multan and Ferozpur Roads. High-density clusters were observed in Data Ganj Baksh, Ravi, Shalamar, Samnabad, and Gulberg Towns, while smaller clusters appeared in Cantonment, Iqbal, and Nishter Towns. In contrast, Figure 18(b) shows that nighttime accidents caused 12,290 major injuries, mainly due to comparatively fewer traffic users. However, similar to daytime patterns, Ferozpur and Multan Roads remained the most affected, with high clusters in Data Ganj Baksh, Ravi, Shalamar, and Samanabad Towns, and smaller clusters in Gulberg and Iqbal Towns.

Fig 18 (a): Major Injury Spots during Day Time	Fig 18 (b): Major Injury Spots during Night Time
	

KDE

Figure 31 presented, clusters of very high and high destiny spots shown in Data Gunj Baksh, Samanabad, Ravi, Gulberg, & Shalamar Towns and also shown with the boundaries of Iqbal Town to Samanabad Town, Gulberg to Nishter Town, Aziz Bhatti to Gulberg & Shalamar to the Wagah Towns. The maximum part of Very low and Low densities are shown in Wagagh, Nishter, Aziz Bhatti, & Iqbal Towns and on Ferozpur Road.

Figure 19: Results of KDL



Conclusion

This study analyzed 125,682 road traffic accidents (RTAs) record from Rescue 1122 in Lahore District, covering February 2021 to May 2022, to identify spatial hotspots and hazardous time intervals. Using GIS-based techniques, KDE, Co-map spatio-temporal analysis, and tools such as Python, Tableau, and Excel, we identified high-risk zones including Shalamar, Data Gunj Baksh, Ravi Towns, Ferozpur Road, and Multan Road. Accident frequency was highest in interior Lahore due to dense daily traffic volumes. Temporal analysis revealed that RTAs averaged approximately 10 per hour, with peaks on Mondays and weekends, during moderate weather (March to October), and in busy hours although accident levels remained consistently high throughout most of the day due to Lahore's continues traffic activity. Factors contributing to accidents including untrained drivers, absence of pedestrian facilities, poor road infrastructure, irregular U-turns, lack of speed control, abrupt curves and nonfunctional traffic signals. The results provide actionable insights for traffic policymakers and authorities to apply targeted safety measures in high risk areas. The methodological analysis particularly the integrating of spatial-temporal and SI based analysis can be replicated in other districts to better understand and mitigate RTS patterns across Pakistan and beyond.

References

- Abbasi, E., & Hadji Hosseinlou, M. (2019). The importance of exercise and general mental health on prediction of property-damage-only accidents among taxi drivers in Tehran: A study using ANFIS-PSO and regression models. *Journal of Advanced Transportation, 2019,* 1–10. https://doi.org/10.1155/2019/4361607
- Ahmad, W., Ijaz, B., Javed, F. T., Jahan, S., Shahid, I., Khan, F. M., & Hassan, S. (2010). HCV genotype distribution and possible transmission risks in Lahore, Pakistan. *World Journal of Gastroenterology, 16*(34), 4321–4328. https://doi.org/10.3748/wjg.v16.i34.4321
- Asgary, A., Ghaffari, A., & Levy, J. (2010). Spatial and temporal analyses of structural fire incidents and their causes: A case of Toronto, Canada. *Fire Safety Journal, 45*(1), 44–57. https://doi.org/10.1016/j.firesaf.2009.09.003
- Aurora, S. (2020, January 2). *Big data and analytics: Top 5 Python libraries for data science*. Simplilearn. https://www.simplilearn.com/top-python-libraries-for-data-science-article
- Basharat, D. M. (2014). *Groundwater environment and evaluation of long-term sustainability of the aquifer under Lahore, Punjab, Pakistan* [Unpublished doctoral dissertation]. University of the Punjab.
- Boba, R. (2005). *Crime analysis and crime mapping*. Sage Publications.
- Bucsuházy, K., Matuchová, E., Zůvala, R., Moravcová, P., Kostíková, M., & Mikulec, R. (2020). Human factors contributing to the road traffic accident occurrence. *Transportation Research Procedia, 45,* 555–561. https://doi.org/10.1016/j.trpro.2020.03.057
- Gerber, M. (2014). Predicting crime using Twitter and kernel density estimation. *Decision Support Systems, 61,* 115–125. https://doi.org/10.1016/j.dss.2014.02.003
- International Transport Forum, United Nations Economic Commission for Europe, & Eurostat. (2020). *Glossary for transport statistics* (5th ed.). OECD Publishing; European Union. https://doi.org/10.1787/505b670b-en
- Kent, M., Cole, E., & Turner, J. (2020). Disability, mobility and transport in low- and middle-income countries: A thematic review. *Sustainability, 12*(2), 589. https://doi.org/10.3390/su12020589
- Khadka, B. (2019). *Data analysis theory and practice: Case Python and Excel tools* (pp. 4–9). Centria University of Applied Sciences.
- Lam, C., Pai, C. W., Chuang, C. C., Yen, Y. C., Wu, C. C., Yu, S. H., & Chiu, W. T. (2019). Rider factors associated with severe injury after a light motorcycle crash: A multicentre study in an emerging economy setting. *PLOS ONE, 14*(6), e0219132. https://doi.org/10.1371/journal.pone.0219132
- Loo, B. P., & Anderson, T. K. (2015). *Spatial analysis methods of road traffic collisions*. CRC Press.

- Ma, X., Chen, F., & Chen, S. (2015). Modeling crash rates for a mountainous highway by using refined-scale panel data. *Transportation Research Record, 2515,* 10–16. https://doi.org/10.3141/2515-02
- Plug, C., Xia, J. C., & Caulfield, C. (2011). Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis & Prevention, 43*(6), 1937–1946. https://doi.org/10.1016/j.aap.2011.05.007
- PoliticPK. (2024, July 5). *Lahore City UCs – Union Councils Complete List*. Political & Current Affair News, Election Results. Retrieved from https://www.politicpk.com/lahore-city-ucs-union-councils-complete-list
- Satria, R. (2020). *Spatial analysis of traffic accidents using GIS: The case of Banda Aceh, Indonesia* [Master's thesis]. University of Twente. https://doi.org/10.3990/1.9789036549819
- Satria, R., Tsoi, K. H., Castro, M., & Loo, B. P. (2020). A combined approach to address road traffic crashes beyond cities: Hot zone identification and countermeasures in Indonesia. *Sustainability, 12*(5), 1801. https://doi.org/10.3390/su12051801
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall.
- World Health Organization. (2018). *Global action plan on physical activity 2018–2030: More active people for a healthier world*. World Health Organization. https://apps.who.int/iris/handle/10665/272722
- Xie, Z., & Yan, J. (2008). Kernel density estimation of traffic accidents in a network space. *Computers, Environment and Urban Systems, 32*(5), 396–406. https://doi.org/10.1016/j.compenvurbsys.2008.05.001