THAILAND ROAD ACCIDENTS

BY KYLIE TAN, MENTORED BY DR KAM TIN SEONG

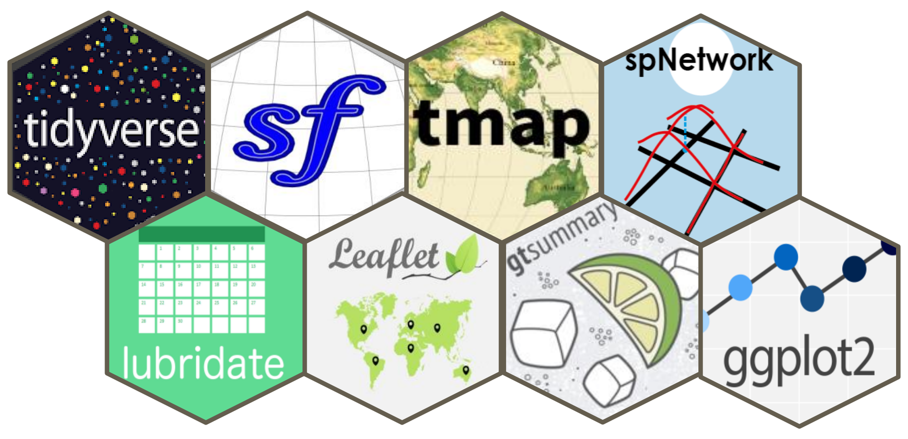
Problem + Motivation

Road traffic injuries represent a significant public health challenge in Thailand, leading to approximately 20,000 fatalities annually, or 56 deaths per day according to WHO. Despite numerous government initiatives, the situation remains critical, highlighting the urgent need for a comprehensive analysis of contributing factors.

This study investigates road traffic accidents in the **Bangkok Metropolitan Region (BMR)** using **spatial and spatio-temporal point patterns analysis**. It specifically focuses on the influence of temporal factors such as season, day of the week, and time of day.

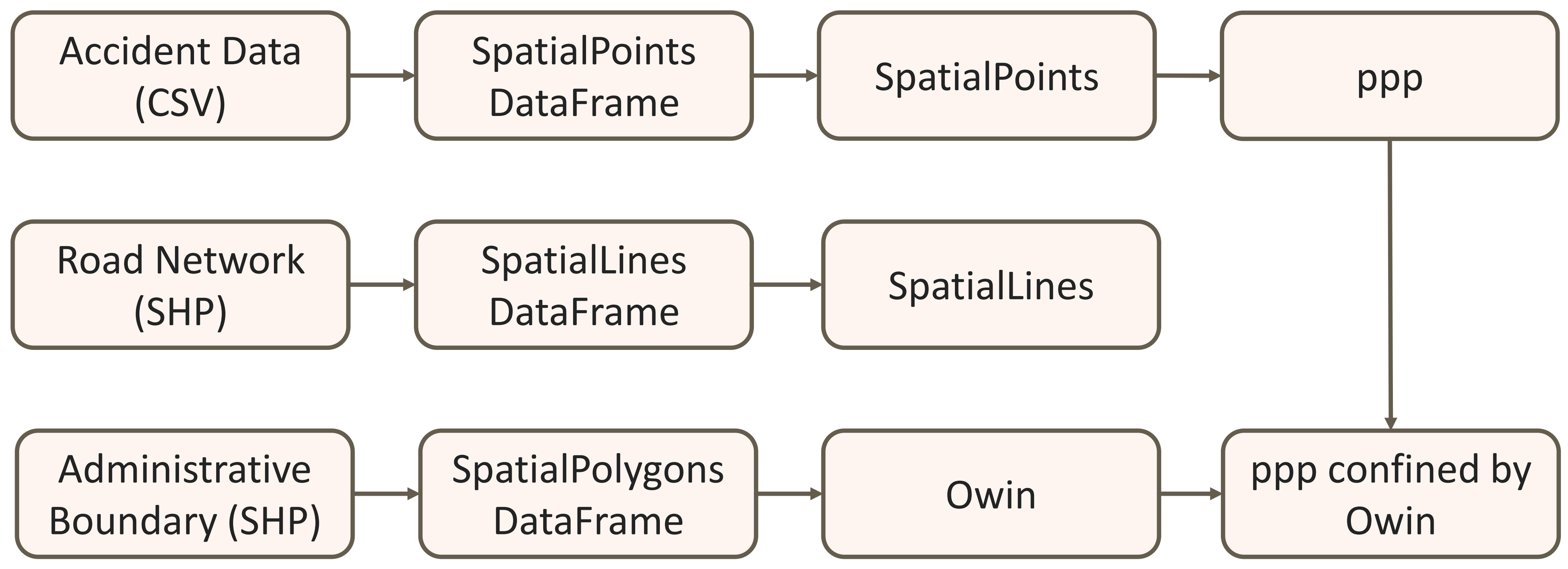
Our analysis utilises the Thailand Road Accidents (2019-2022) data in CSV format from Kaggle, alongside Thailand Roads and Subnational Administrative Boundaries in SHP file formats from the Humanitarian Data Exchange website.

All analyses will be conducted in R to identify key accident hotspots and inform the development of effective road safety policies and interventions. Packages used includes:



Data Wrangling

Overview



Key Steps

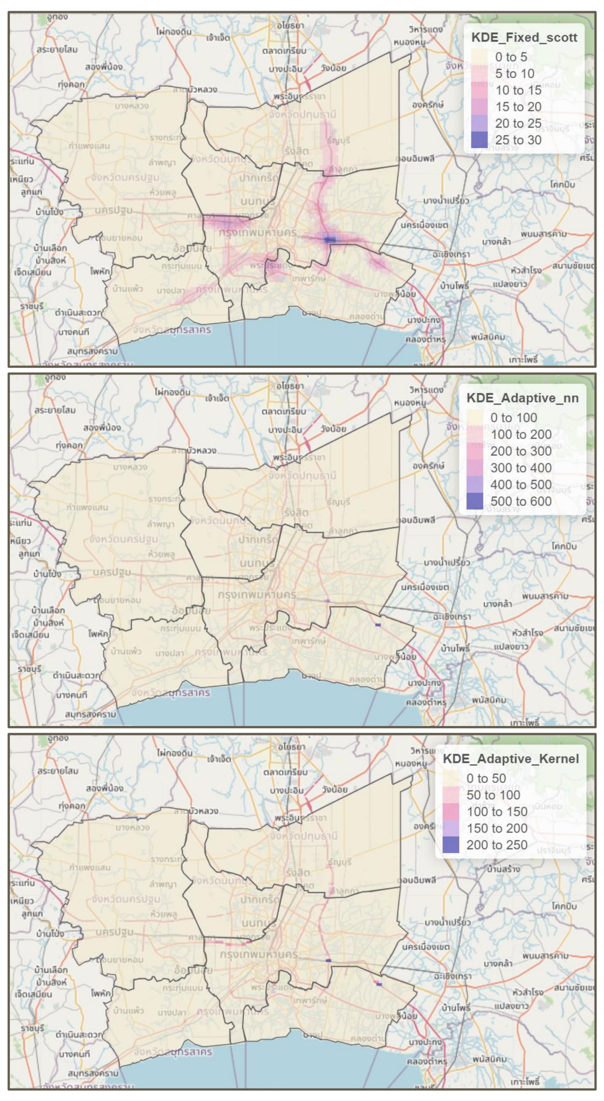
- Transformed dataset into an *sf* object using longitude and latitude coordinates.
- Ensured spatial data is reprojected to **EGSP:32647** (local UTM projection used in TH).
- Changed the roads data from multi-linestrings to linestring for further analysis.
- Filtered for roads within BMR using *st_intersection()*.
- Selected specific highway types based on road classifications from OpenStreetMap.
- Applied jittering to ensure accident points do not overlap to enhance visualisation of intensity and spatial distribution of traffic accidents.
- Confined the data to the study area through an *owin* object.

Analysis + Insights

Test of Clustering

The **Clark-Evans test** yielded an **R-value of 0.22579**, suggesting a clustered distribution of accident points. With a p-value of less than 2.2e-16, well below the significance level of 0.05, we reject the null hypothesis (H_0) in favor of the alternative hypothesis (H_1). This implies that the accident points are **not randomly distributed, but rather clustered**, suggesting potential underlying factors influencing their spatial distribution.

First-Order Spatial Point Pattern Analysis



The analysis shows that the highest concentration of accidents occurs in **Bangkok province**, especially near the Bangkok–Ban Chang Motorway and the Bangkok Outer Ring Road (Motorway Route 7 and Route 9). Borommaratchachonnani Road (Highway 338) along the western edge of Bangkok also shows high accident rates, as do other major highways within BMR.

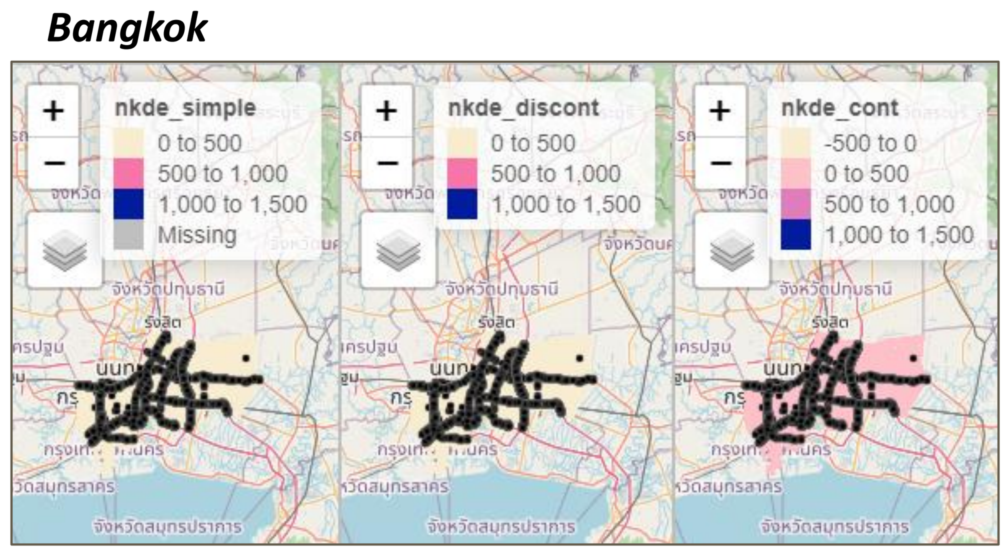
Both the Adaptive Nearest Neighbour and Adaptive Kernel KDE methods highlight a significant accident concentration in the Khlong Chang Tai area along Highway 3701 in **Samut Prakan province**, which was not captured by the Fixed Bandwidth KDE method.

When applied at the province-level, KDE maps demonstrate limitations when analysing smaller regions, as fixed KDE maps can over-smooth the data while adaptive KDE maps may under-smooth, complicating the extraction of meaningful insights. Furthermore, the reliance on grid pixels and Euclidean distance restricts the ability to identify finer patterns within localised areas, reducing the effectiveness of the analysis.

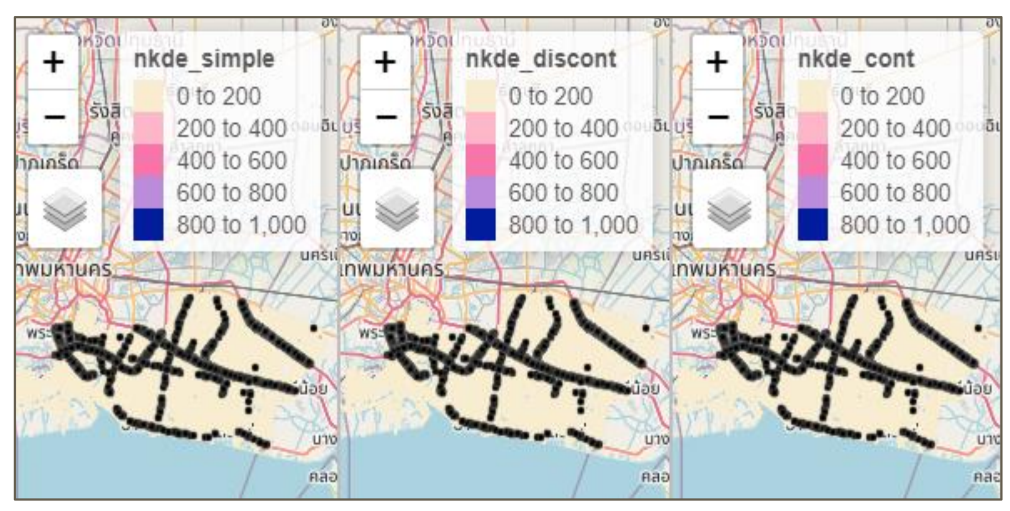
Network Constrained Kernel Density Estimation (NKDE)

In real-world scenarios, events such as accidents tend to follow specific networks, like road systems, rather than being randomly distributed. Traditional KDE operates under the assumption that events occur across an open, two-dimensional space, which fails to accurately represent network-based patterns like road traffic. To address this, NKDE **extends spatial KDE by estimating event density strictly along the network**.

Bangkok



Samut Prakan

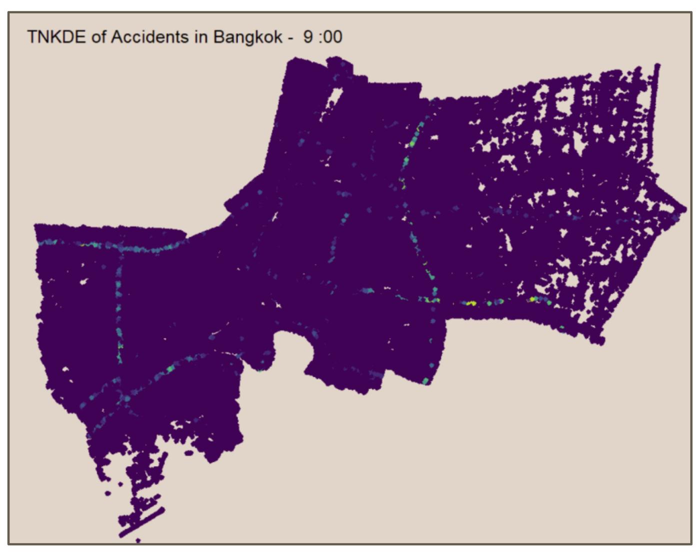


The resulting NKDE maps have not provided the expected insights, perhaps attributed to the high density of traffic accident points, which could obscure meaningful patterns. Analysing smaller subsets of events segmented by temporal patterns might offer deeper insights.

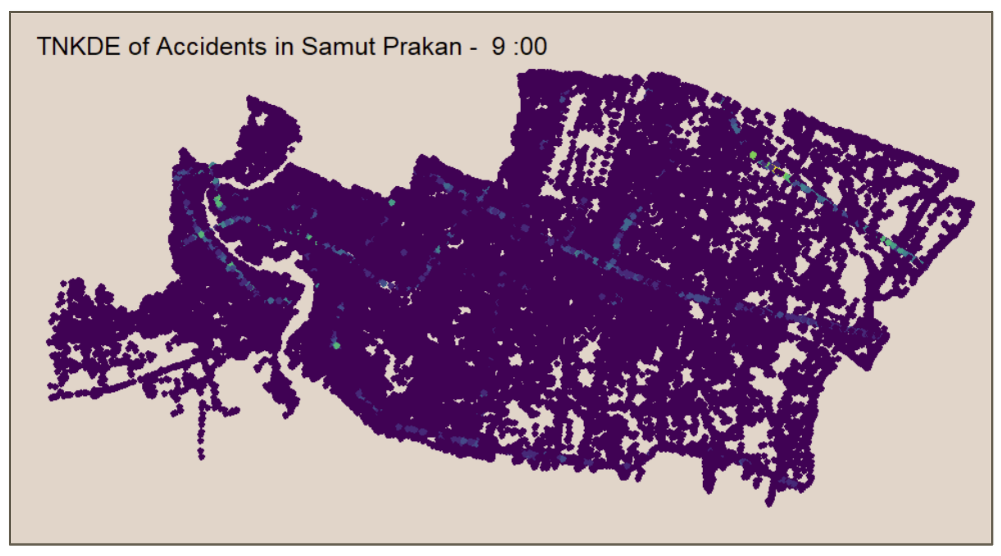
Temporal Network Kernel Density Estimation (TNKDE)

TNKDE **combines spatial and temporal dimensions** to enhance understanding of event distributions along networks. Unlike traditional NKDE, which focuses solely on spatial data, **TNKDE considers time-based variations**, such as time of day, and day of the week, allowing for a detailed analysis of how events like traffic accidents fluctuate over time and space.

Bangkok



Samut Prakan



In Bangkok, the TNKDE map shows increased accident hotspots during key time windows (7am-11am, 4pm, 7pm), particularly along major highways. Fewer hotspots are seen from late night through early morning, indicating that accident clusters are largely confined to high-traffic roads. Similarly, in Samut Prakan, the TNKDE map reveals consistent accident hotspots between 7am and 8pm, with fewer incidents during early morning hours. These hotspots remain concentrated in specific high-traffic areas throughout the day.

The TNKDE maps provide granular insights into how accident clusters evolve over time, highlighting how accidents are spatially and temporally concentrated along key routes. These findings are valuable for traffic safety planning, enabling the identification of critical accident-prone areas and trends over time.

Future Work

Future research could incorporate additional data from the original Thailand Road Accidents dataset, including weather conditions, road types, and traffic characteristics, to provide a more comprehensive view of the factors contributing to accidents. Integrating these variables would enhance the analysis, offering a deeper understanding of accident patterns.

SMU SINGAPORE MANAGEMENT UNIVERSITY
School of Computing and Information Systems