IS415 Geospatial Analytics and Applications

WhatTown User Guide

Done by:
Chia Yizheng, Chester
Matthew Ho Yi Wen

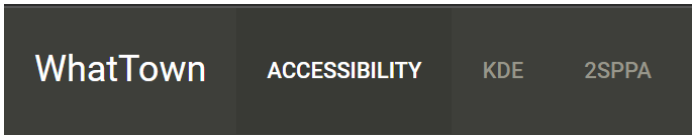
Overview

Welcome to WhatTown! WhatTown is a web application designed to transform how Singaporeans understand and interact with the geographical distribution of amenities such as supermarkets, wet markets, hawker centres, and more. Leveraging powerful tools like Geographical Accessibility Modelling, Kernel Density Estimation (KDE), and Spatial Point Pattern Analysis, WhatTown makes complex data analysis accessible to everyone—no technical skills required.

Getting Started

Upon accessing WhatTown, you will be greeted by a simple and intuitive interface designed to guide you through various exploratory and analytical processes. We can get started with the navigation bar:

Navigation Bar



The navigation bar located at the top of the webpage allows you to navigate between our 3 core features - accessibility modelling, Kernel Density Estimation, and 2nd Order Spatial Points Pattern Analysis.

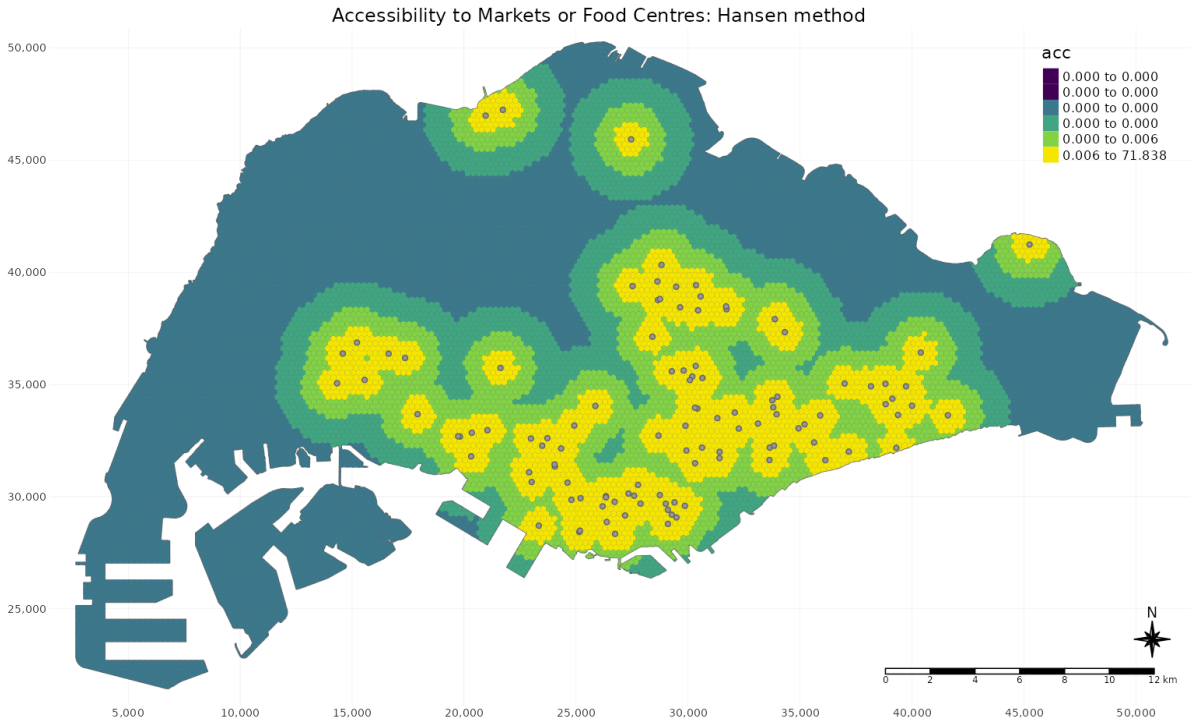
Features

Accessibilty Modelling (“Accessibility”)

In this feature, you will be able to select facilities to model accessibility for, choose your customization options for the generated map, and also select the modelling method of your choice. Here is a breakdown of each parameter in the left sidebar:

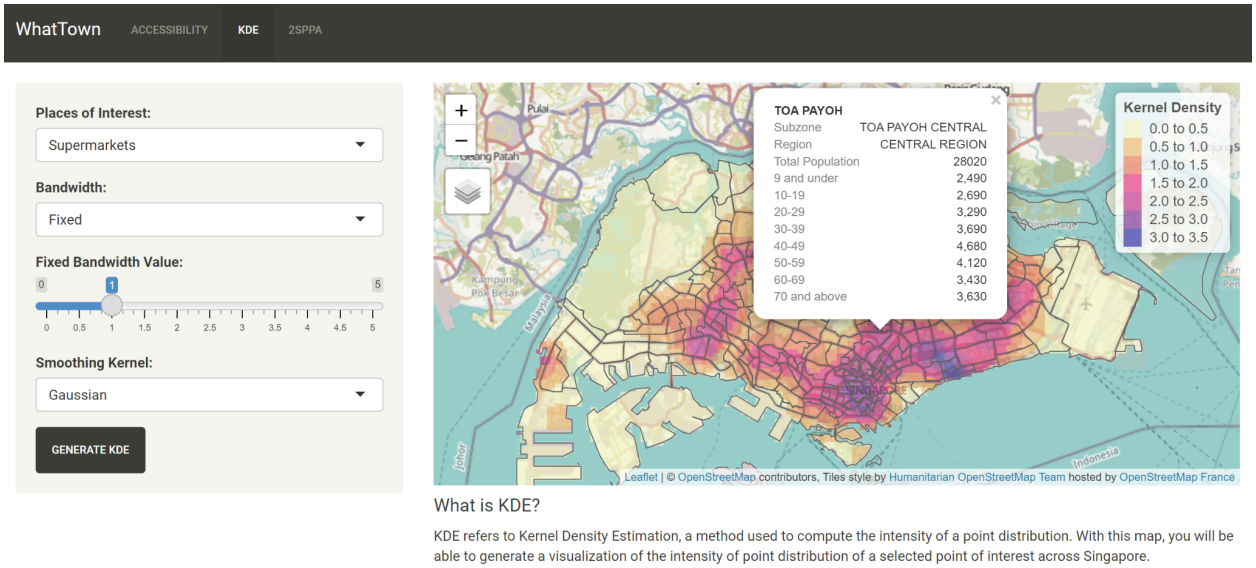
- **Modelling Family:** Select the appropriate algorithm to analyze the accessibility of amenities. Each algorithm offers a different approach, so choose one that best fits the type of analysis you wish to conduct.
- **Breaks:** Customize how the data is visualized by setting color breaks. This will help you differentiate data points based on their values more clearly.

- **Grid size:** Adjust the grid size to modify the resolution of the spatial analysis. Smaller grids provide more detail, while larger grids offer a broader overview.
- **Place of Interest:** Select the point of interest you would like to model accessibility for.
- **Distance Exponent:** Use the distance exponent to weight the analysis, simulating the decreasing likelihood of traveling long distances for amenities. This setting adjusts how distance impacts accessibility scores. For further details, you may refer to this Take-Home Exercise.
- **Subzone or Planning Area Population:** Choose the granularity of census data - either at the subzone or planning area level.
- **Grid Shape:** Choose between different shapes (e.g. Hexaganol or Square) to best fit the geographical layout of the data.
- **Colour Palette:** Select a colour palette.
- **Scaling:** Lets you choose how data categories are defined (e.g. Quantile, Fixed).
- **Capacity Multiplier:** Adjust the multiplier to scale the assumed capacity of each facility linearly.



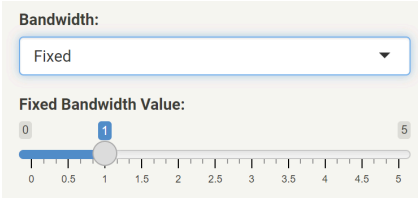
Once the map has been generated, you will be able to view the accessibility of each ‘cell’, with the color corresponding to the level of accessibility for that area, as per the color scale in the legend. For example, clusters of bright yellow hexagons indicates a concentration of highly accessible areas.

Kernel Density Estimation (“KDE”)



In this feature, you will be able to generate KDE maps for visualizing the density of selected points of interests across Singapore. Here is the parameter breakdown:

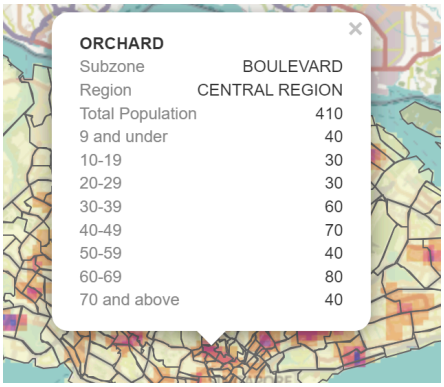
- **Places of Interest:** Select which point of interest you will like to generate a KDE map for.
- **Bandwidth:** Select the appropriate bandwidth, which determines the smoothness of the density plot. A larger bandwidth results in a smoother, more generalized density estimation, while a smaller bandwidth provides a more detailed, localized view of data points. There are 3 options:
    - **Fixed Bandwidth:** Manually set a specific bandwidth value. A slider will appear when this is selected.



    - **Adaptive Bandwidth:** Enable varying bandwidth depending on local density of points. This option is ideal for datasets with uneven distribution, allowing the KDE to adjust dynamically for areas with varying point concentrations.
    - **Automatic Bandwidth:** Select between 4 automatic bandwidth methods. A select dropdown will appear when this is selected.

- **Smoothing Kernel:** Choose the type of kernel function to smooth the data points. The kernel function influences the shape of the curve that is used to smooth the point density. Choices include:
    - **Gaussian:** The Gaussian kernel produces a smooth, bell-shaped curve around each point. It is the most commonly used kernel due to its smoothness and mathematical properties.
    - **Epanechnikov:** This kernel is parabolic and known for its efficiency. The Epanechnikov kernel is optimal in a mean integrated squared error sense, making it very effective for density estimation.
    - **Quartic:** The Quartic kernel is similar to the Epanechnikov but with a smoother tail.
    - **Disc:** The Disc kernel allocates equal weight to all points within its bandwidth and zero weight beyond. This results in a uniform density within the radius of each point.

After the map has been generated, you will be able to interact with it by clicking on specific subzones, which will provide you with demographic information (age) of that subzone.



2nd Order Spatial Points Pattern Analysis (“2SPPA”)

In this feature, you will be able to use statistical functions to analyze the spatial distribution of points like supermarkets and wet markets in Singapore subzones. These functions help us to understand how spaced out these points are. Here are the parameters:

- **Subzone:** Select the subzone to perform 2SPPA on.

- **Function:** Select between 2 statistical functions:
    - **G-Function:** Use this function to analyze the distances between each point and its nearest neighbor. This is useful for detecting local clustering or dispersion.
    - **K-Function:** Choose this function to assess clustering over a range of distances. It provides a broader view of spatial patterns across the selected subzone.
- **No. of Simulations:** This parameter sets how many simulations the software should run to establish statistical significance for the observed patterns. A higher number of simulations increases the robustness of the analysis but may require more processing time.
- **Place of Interest:** Specify the type of amenity or facility you are interested in analyzing.

Once you have configured these settings, WhatTown processes the data based on your inputs. The results will display spatial trends and patterns that can inform urban planning decisions, resource allocation, or further research into community development. Here is how we can interpret the results:

- **Graphical Output:** The analysis will produce a plot showing either the G-function or K-function curve compared to a theoretical line under Complete Spatial Randomness (CSR).
- **Statistical Significance:** The confidence envelope (grey area) derived from the simulations will help you understand whether the observed patterns are statistically significant.
    - **Above or outside the envelope:** Indicates clustering (for G-function) or more clustering than random (for K-function).
    - **Below or inside the envelope:** Suggests dispersion or less clustering than random.
    - **Within the envelope:** Supports the null hypothesis that points are randomly distributed.

In the case where the data is incomprehensible (e.g. no envelope/straight line), it suggests the data is either extremely uniform or there are too few data points. Without the envelope, we lose a benchmark for comparing the observed pattern against randomness. Consider trying out different parameters - however, we can still look at the shape of the curve to gain insights - an upward trend may indicate clustering, while a downward trend could indicate dispersion.