

IS415 Geospatial Analytics and Applications
Take-home Exercise 2 (15%)
Do Birds of a Feather Flock Together?:
Spatial Point Patterns Analysis of Airbnb Listing in Singapore

Setting the Scene

In recent years, the emergence of numerous peer-based business models has empowered individuals across the globe to become microentrepreneurs, earning money from their idle property and spare time. This phenomenon, entitled ‘the sharing economy’, has seen unprecedented growth in terms of user numbers, enabling new avenues of economic and social interaction (Sundararajan, 2016).

Since its inception in 2008, Airbnb has expanded into over 34,000 cities across 191 countries. Uncommonly for a sharing economy company, Airbnb transitioned into profitability in 2016, demonstrating proof-of-concept for the validity of sharing economy within the global market (Stone & Zaleski, 2017).

Interestingly, Singapore is one of the global city that has yet to legalise short-term rentals offered by platforms such as Airbnb (Read more at <https://www.todayonline.com/singapore/short-term-home-sharing-remains-illegal-singapore-airbnb-disappointed>). However, during my recent visit to Inside Airbnb (<http://insideairbnb.com/about.html>), an independent, non-commercial set of tools and data that allows anyone to explore how Airbnb is really being used in cities around the world, it came to my attention that there are data sets for Singapore.

The Task

In this take-home exercise, you are task to analyse the distribution of Airbnb Airbnb listing by using appropriate spatial point patterns analysis techniques. The specific tasks are as following tasks:

Section A: Nation-wide analysis

1) Exploratory Spatial Data Analysis:

- Using appropriate tmap function, display the locations of the Airbnb listing by room type with Openstreemap of Singapore as the background. Describe the spatial patterns observed.

2) With reference to the spatial point patterns observed in (2):

- Formulate the null hypothesis and alternative hypothesis and select the confidence level.
- Perform the test by using appropriate 1st order spatial point patterns analysis technique.
- With reference to the analysis results, draw statistical conclusions.

3) With reference to the results derived in (1) and (2):

- Derive kernel density maps of Airbnb listing by room type.
- Using appropriate tmap functions, display the kernel density maps on openstreetmap of Singapore. Describe the spatial patterns revealed by the kernel density maps. Highlight the advantage of kernel density map over point map.

Section B: By planning subzones

4) Exploratory Spatial Data Analysis:

- Extract Airbnb listing by room type within Aljunied, Balestier, Lavender and Tanjong Pagar planning subzones.
- Display their distribution as point maps. Describe the spatial patterns reveal by their respective distribution.

5) With reference to the spatial point patterns observed in (4):

- Formulate the null hypothesis and alternative hypothesis and select the confidence level.
- Perform the test by using appropriate 2nd order spatial point patterns analysis technique.
- With reference to the analysis results, draw statistical conclusions.

3) With reference to the results derived in (4) and (5):

- Derive kernel density maps of Airbnb listing by room type.
- Using appropriate tmap functions, display the kernel density maps on openstreetmap of Singapore. Deduce what are the possible factors determine the spatial patterns observed..

The Data

For the purpose of this study, the *listing* data set of Inside Airbnb will be provided. This data set is downloaded from Inside Airbnb (<http://insideairbnb.com/get-the-data.html>). You are required to obtain the remaining data from the relevant open data portal.

Grading Criteria

This exercise will be graded by using the following criteria:

1. **Geospatial Data Wrangling:** This is an important aspect of geospatial analytics. You will be assessed on your ability to employ appropriate R functions from various R packages specifically designed for modern data science such as readr, tidyr, dplyr, sf just to mention a few of them, to perform the entire geospatial data wrangling processes, including. This is not limited to data import, data extraction, data cleaning and data transformation. Besides assessing your ability to use the R functions, this criterion also includes your ability to clean and derive appropriate variables to meet the analysis need. **(Warning:** All data are like vast grassland full of land mines, your job is to clear those mines and not to step on them). (15 marks)
2. **Geospatial Analysis:** In this exercise, you are expected to use the spatial point patterns analysis techniques and R functions introduced in class to analysis the geospatial data prepared. You will be assessed on your ability:
 - a. to derive the optimal kernel density maps by combining the functions and arguments provided in spatstat package;
 - b. to conduct 2nd order spatial point patterns analysis and interpret the test result; and
 - c. to discuss the analysis results of (a) and (b). (50 marks)
3. **Geovisualisation:** In this section, you will be assessed on your ability to communicate the complex spatial statistics results in business friendly visual representations. This course is geospatial centric, hence, it is important for you to demonstrate your competency in using appropriate geovisualisation techniques to reveal and communicate the findings of your analysis. (20 marks)
4. **Bonus:** Demonstrate your ability to employ methods beyond what you had learned in class to gain insights from the data. The methods used must be geospatial in nature. (15 marks)

Deliverables

- The project folder in a single zip file format. The project folder should consist of the followings:
 - The project sandbox in data sub-folder (all raw, intermediate and final data files)
 - An R Markdown file contains all code chunks used and the written statements.
 - A html report knitted from the R Markdown document.
- A published version of the report on [Rpubs](#).

Submission Instructions

- The final deliverable (e.g. R Markdown file, project sandbox and report) must be zipped in a single zip file format.
- Name the zip file according to the course code and assignment, for example: ***IS415_Take-home_Ex02***.
- The deliverable is to be submitted in softcopy. You are required to upload the zipped into the Dropbox of LMS before the stated assignment due date. Late work will be severely penalised. Students must check and confirm on LMS the assignment due date.

Due Date

17th May 2020 (Sunday), 11.59pm (midnight).

References

Sharing economy (https://en.wikipedia.org/wiki/Sharing_economy).

Airbnb (<https://en.wikipedia.org/wiki/Airbnb>)