LandMarkdown: A Spatial Analysis of Business Entities in Singapore

CHEONG Wei Herng Eugene
Singapore Management University
whcheong.2018@sis.smu.edu.sg

CHOY Jing Wei
Singapore Management University
jwchoy.2018@sis.smu.edu.sg

KANG Yi Qi
Singapore Management University
yiqi.kang.2017@sis.smu.edu.sg

ABSTRACT

Various government initiatives such as Singapore's Smart Nation Initiative have resulted in large amounts of data shared by Singapore's government agencies on public online data portals. However, such data have not been explored or worked on extensively to gain a better understanding on the spatial distribution and geographic segmentation of businesses in Singapore and thus could not be effectively translated into better insights for industry-specific policymaking or business decisions.

To address this, we designed and developed LandMarkdown, an interactive dashboard which allows policymakers and business management executives to explore and analyse Singapore's corporate entity data and gain insights about the spatial distribution and clusters of businesses.

1. INTRODUCTION

As Singapore accelerated efforts to digitally transform itself, various government agencies such as the Urban Redevelopment Authority (URA) and the Land Transport Authority (LTA) has released data on public on-line data portals such as data.gov.sg and OneMap. The datasets obtained can be used to track spatial point patterns and identify potential clusters for businesses registered in Singapore. For example, corporate entity data as well as their respective SSIC classification were collected. These data contain information and patterns which can be analysed to provide greater insights on the spatial distribution of businesses and clusters of business categories in Singapore on a subzone level. The insights can possibly lead to improvements in industry policy making and provide useful information for private companies or Government-Linked Companies to make better business decisions.

In practice, the use and analysis of a spatial data tends to be limited to Geographic Information Systems (GIS). However, access to GIS applications tend to be limited. Furthermore, the use of GIS applications are specialised and might not be suited as a decision-making or policy planning tool.

To address these two issues, we designed and developed LandMarkdown - an analytics dashboard which aims to provide government policy makers and business decision makers with a tool to analyse corporate entity data from a spatial analytics point of view.

This paper showcases our research and development work into designing and implementing a dashboard to assist business decision makers and government policy makers in analysing and visualising the spatial distribution of businesses and clusters of business categories in Singapore on a subzone level. Section 1 provides a general introduction of the paper, followed by an overview of the motivation and objectives of our research in Section 2. Section 3 provides a review of the literature detailing relevant work in our research area. We provide a detailed overview of the methodology and the results behind our research efforts in Sections 4 and 5 respectively. In the next section, we discuss the insights enabled from our research efforts, before ending off by concluding how the research can be extended in the future.

2. MOTIVATION AND OBJECTIVES

Our research and development efforts were motivated by the idea of simplifying analytical tools for the discovery of spatial point patterns and clusters in corporate entity data. Our solution aims to provide government policymakers and business decision makers with an analytical dashboard to visualise the spatial distribution of corporate entities of different industry categories. Our specific objectives are:

1. To visualise the spatial distribution of corporate entities in Singapore by industry categories on an internet-based map such as OpenStreetMap.
2. To conduct statistical simulations to reveal evidence of clustering using spatial point patterns analysis.
3. To reveal characteristics of spatially constrained clusters through clustering techniques.
4. To provide a user-friendly interface which supports the use of a variety of filters and options as dynamic inputs for visualisations.
5. To display the data on-demand in a tabular format.

3. RELATED WORK

Work has been done on the spatial point patterns and industry concentration analysis in Germany. For example, clus-

ters can positively affect competition by increasing productivity and innovation as well as stimulating the formation of new businesses, through economic effects such as economies of scale [3].

However, traditional metrics for industry concentration do not account for geospatial effects. They fail to account for the formation of clusters at different geographic scales, and the degree of concentration is not expected to be independent on the reach of grouping of the firms [3]. Kosfeld et al. introduced a concentration index of the style of Besag's L function [1] that is based on the concept of the K function. While Besag's L function is intended to measure deviations from the CSR process, the new index can be applied to identify the importance of sector-specific and more general industry-specific forces inducing clustering.

In addition, Kosfeld et al. also proposed a feasible testing procedure enabling a usage of the K function approach efficiently for large study regions. For this, Diggle and Chetwynd's D function approach [2] is replaced by a spatial similarity test based on subsamples drawn from the industry under analysis and the reference population.

While concentration of the branches of industry in Germany was only available at spatial scales given by arbitrarily defined regions. While the authors illustrated their K function approach by using selected industries, the authors additionally provided concentration numbers for sixteen German industries within different distance bands.

4. METHODOLOGY

As such, we have used location quotient as a metric for industrial concentration. The formula for calculating the location quotient of industry a in subzone i is shown as follows:

$$LQ_{a,i} = \frac{\frac{i_a}{\sum_{i=0}^{n} i_a}}{\frac{i_{Singapore}}{\sum_{i=0}^{n} i_{Singapore}}}$$

where:

- $LQ_{a,i}$ refers to the Location Quotient for industry a in subzone i,
- i_a refers to the number of industry a in subzone i,
- $\sum_{i=0}^{n} i_a$ refers to the sum of all industries in subzone i,
- $i_{Singapore}$ refers to the number of industry a in Singapore,
- and $\sum_{i=0}^{n} i_{Singapore}$ refers to the sum of all industries in Singapore.

5. RESULTS

6. DISCUSSION

7. FUTURE WORK AND CONCLUSION

8. ACKNOWLEDGEMENT

The authors wish to thank Dr. Kam Tin Seong, an Associate Professor of Information Systems (Practice) in Singapore Management University for his support and guidance.

References

[1] Besag, J. 1977. Contribution to the Discussion on Dr. Ripley's Paper. Journal of the Royal Statistical Society: Series B (Methodological). 39, 2 (1977), 192–212.

[2] Diggle, P.J. and Chetwynd, A.G. 1991. Second-Order Analysis of Spatial Clustering for Inhomogeneous Populations. Biometrics. 47, 3 (Sep. 1991), 1155–1167.

[3] Kosfeld, R. et al. 2011. Spatial point pattern analysis and industry concentration. The Annals of Regional Science. 47, 2 (Oct. 2011), 311–328.