BEYOND VISUALISATION OF BIG DATA: TOWARDS DYNAMIC DATA-DRIVEN CITY PLANNING IN SINGAPORE (1068)

Zdravko Trivic[1*], Aditya Sinha[1], Kai Ma[2], Kim Huat Goh[2]

[1] Department of Architecture, College of Design and Engineering, National University of Singapore, Singapore, [*]akizt@nus.edu.sg

[2] Division of Information Technology & Operations Management, Nanyang Business School, Nanyang Technological University, Singapore

Abstract. Cities are complex systems shaped by simultaneous dynamic processes. While an increased number of studies focus on analysing urban big data, the attempts predominantly propose means for data visualisation, and fall short in data interpretation and guidance. In response, this paper aims to systematically understand dynamic urban big data, trends and anomalies in city functioning. We outline a comprehensive framework and "DataCube-CityScan" platform that harness on dynamic economic, societal, environmental, health and attitudinal data available in Singapore (e.g., people movement, public transport use, shopping behaviour). We use GIS alongside multiple outlier detection algorithms to analyse and identify specific trends and anomalies in real-time and alert city officials to respond and make informed and timely decisions, monitor changes, plan actions and strategies, and maximise their resources.

Keywords: dynamic urban data; big data visualisation; city pulse; city planning; Singapore.

1. Introduction

Cities are complex systems shaped by numerous simultaneous dynamic processes, which either work in harmony and keep the system healthy and thriving or converge into conflicts and cause anomalies and disruptions. The ways cities have developed historically and how these processes have manifested were considerably dependent on technological advancements, which arguably became more rapid in recent decades. According to Smith and Lobo (2019), contemporary cities can be described as both the devices of innovation and growth, and the containers of intensive social relations, the so-called 'e*nergized crowding'*. Such concentrations of structure, infrastructures, information, people and processes that are constantly changing require more robust, dynamic and resilient urban management mechanism to support city planning and growth.

With the advancement of information and communication technology (ICT) in cities, enormous amount of urban data is generated by the minute, concerning various aspects of everyday city's functioning. By using this dynamic real-time data alongside the conventional static and semi-static data, planners and city officials can bridge the gap between different temporal scales of urban activities (Batty and Kandt, 2021), which can allow for deep understanding of complex city processes and urban structures. This in turn helps to create timely and more refined policies, guidelines and actions to manage these processes.

Unsurprisingly, the questions of *'When, Where and How'* to use such vast, high frequency dynamic urban data reflect critical challenges. While there have been numerous studies that focused on a single domain or type of dataset to analyse, visualize variables, and solve a particular set of problems or hypothesis, very few studies have focused on using multiple datasets of different characteristics of the city simultaneously (Puiu et al., 2016). Moreover, recent studies tended to propose smart new means for big urban data visualisation, while also falling short in data interpretation and planning guidance.

In response, this paper outlines part of a larger study done in Singapore to systematically approach the vast expanse of urban big data and develop an alert system for the city officials and agencies to understand the underlying anomalies in the everyday city functioning. Singapore being one of the smartest cities in the world provides a range of open data at various scales and velocities (Lee et al., 2016). This paper presents a multi-perspective "Data Cube" framework to collect, aggregate and visualize dynamic economic, societal, environmental, health and attitudinal data at various granularities (city, regions, planning areas and planning subzones) in Singapore. It also presents a "DataCube-CityScan" platform, which harnesses on GIS, AI, data mining and outlier detection algorithms to identify specific trends and anomalies/outliers in real-time and alert city officials to respond, monitor changes, plan their actions and maximise their resources timely.

2. Literature Review

Development of digital technology has allowed cities to become much more intelligent, integrated and smarter in managing their resources. This has been possible by employing urban analytics to understand the unique set of challenges and opportunities in a city. Miranda (2017) finds analogies between everyday functions in a city with pulse in a human body. He argues that activities in a city are unpredictable in magnitude and location, and they vary with time, similar to a human pulse. These activities can generate from any part of the city at any time of the day which in turn defines the nature of that location at a particular time. In line with this, many researchers have tried to capture and measure the pulse of a city and to understand the dynamics of the city through big data analysis (Puiu et al., 2016; Vaccari et al., 2010; Froehlich, Neumann and Oliver, 2009; Kuemper et al., 2015, Liu and Biljecki, 2022). Most of the cities today provide access to city data, which allows research institutes to analyse them and provide innovative solutions to urban problems. Governments across the world have also shown interest in using big data to increase the innovation capacities of their urban centres in recent decades. UK government has set up pilot cities with Office of Data Analytics (ODA) for tacking everyday challenges (Eaton and Camilla, 2018). Main functions of ODA include taking data driven initiatives and creating reusable code of ethics and data standards (Templatizing) to be used across UK. Qualitative research done for cities, such as London and Barcelona, has shown that while these cities have big ambitions regarding the development in applied data-informed technologies, they differ in readiness of infrastructure and competence to process and analyse data to provide meaningful insights (Bibri and Krogstie, 2020).

While big data opens endless avenues for experimentation, one of the common issues it faces is the lack of direction (Batty, 2013). It is argued that as the amount of data increases, the number of correlations increases as well. Thus, without a theory such correlation can be considered a diversion. In his book, Bibri (2020) tried to combine the analytical capacities involving big data and context-aware technologies to provide directed solutions for the urban form.

2.1. Using Big Data in Smart Cities

Through the years, smart cities have been associated with different nomenclatures, such as intelligent cities, information cities, and virtual cities, amongst others (Batty, 2013). Urban analytics has become synonymous with smart city to provide solutions to urban problems by drawing on theoretical perspectives and identifying causal relationships. It offers a set of methods that can be applied to digital infrastructure to ensure good data processing and inform the decision makers of accurate insights into and about the urban systems (Kandt and Batty, 2021). Urban analytics can help us answer numerous questions, such as 'What has happened?', 'Why the thing has happened?', 'What will happen?', and 'What should we do?' (Barbero et al., 2016). Recent research has also tried to apply the scientific base of urban analytics to define and measure urban pulse by creating novel framework that can allow the city officials and citizens alike to take part in the decision-making process (Puiu et al., 2016).

Some of the studies have focused on providing solutions to common issues associated with using big data, such as data acquisition, semantic interoperability and real-time analysis, to name a few. For instance, one such research developed data annotation and aggregation, event detection, data federation, context filtering and decision support modules to support decision making process by working on some of the technical issues associated with handling of big data (Puiu et al., 2016). Another research employed data collected from sensors to map and measure movement and aggregation points for special city events that attracted people from all around the world, such as the 2009 presidential Inauguration (Vaccari et al., 2010). Similar study had tried to define the rhythm of the city by capturing spatio-temporal variations of city 'activities' (Miranda et al., 2017), while another analysed usage of bicycle stations in Barcelona to gain insights into the dynamics of the city (Froehlich et al., 2009). Such research approaches represent only an example of the endless capacity that big data holds. Being able to provide the necessary direction to the use of big data is a feat of human ingenuity and has the capacity to optimize a wide spectrum of city's functions.

2.2. Visualisation of Smart City Datasets

Visualisation has a great impact on how we perceive and interpret data, often allowing us to see things that we might ignore if not properly presented. Interpretation and understanding of data are often dependent on effective visualisation. With big data in picture, it becomes even more critical to convey relevant messages across the board. It is important to have clear understanding of what needs to be visualised at the start of every project (Midway, 2020). In the past few decades several business intelligence platforms such as Tableau, Power BI, Qlik sense, Looker and ArcGIS dashboard among others have emerged. These platforms provide a range of tools to visualize spatial or non-spatial big data with capacity to do real time analysis as well. Apart from cities directly using these platforms, many cities have designed their own dynamic dashboards such as Dublin, London, Hawaii and New York to name a few. Topics such as economy, safety, health, diversity, education and planning modules amongst other points of interests are presented in these dashboards. These city dashboards provide interfaces that are able to deliver complex information to multiple user types in a more intuitive manner (Young and Kitchin, 2020). Using intuitive dashboards to convey technical city data using simple visuals proves to be a powerful instrument to gaining insights on city dynamics.

3. SCoRe Methodology

3.1. Conceptual Framework and Data Cube

Based on a comprehensive review of the literature, we developed SCoRe (Societal Comprehensive Reflective Estimate) Methodology to systematically and holistically capture and measure the dynamics of the city, its pulse. In reference to Singaporean context, we put emphasis on five key perspectives, namely: Economic, Health, Environmental, Societal and Attitudinal.

The data is collected along the lines of these five perspectives at various granularities, temporal variations and formats. Some of the data that have been collected are highly dynamic in nature, such as movement and behavior related indicators while other are static or semi-dynamic with quarterly or annual updates such as economic, health, household and spending habits related data. The datasets vary in granularity as well with some captured at national level which provides macro-view of the urban makeup to subzone level which can be used to analyse micro-situations. A number of datasets have been captured in spatial formats, such as point of interest, movement and administrative boundaries while others are aggregated in form of tables which are sorted at national, planning area or subzone levels. A representation of how the three dimensions interact with each other is given as a data cube in the Figure 1 below.

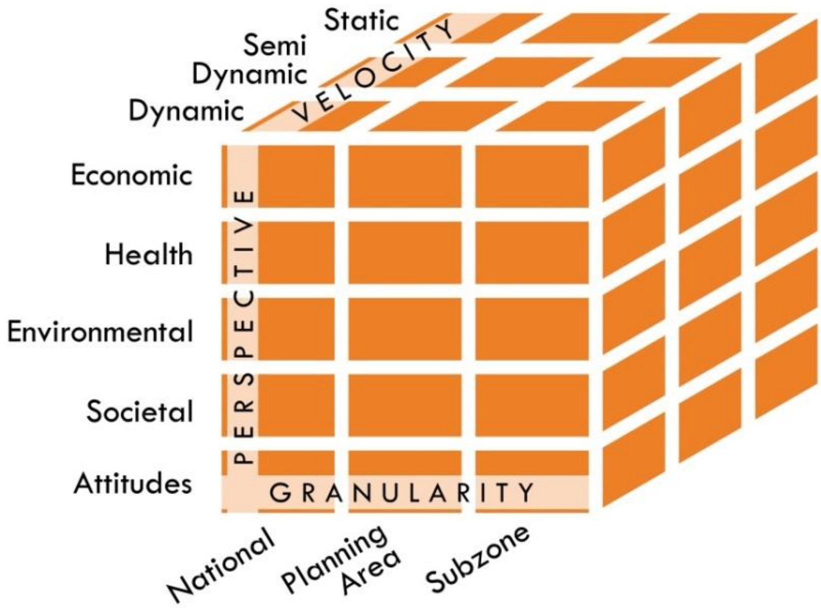

Figure 1. Data Cube

The data cube provides the basis for analysing dynamic data and distilling the anomalies in the datasets. The process consists of three steps, as shown in Figure 2. Firstly, the data that we gathered from various government sources and collaborators is cleaned and standardised as per spatial and analytical standards. Indicators such as building footprint, entropy, green cover, etc. have also been calculated at this stage for analysis. Secondly, the data is fed into the system for outlier analysis. At this stage the stakeholder has the option to select multiple fields and datasets for outlier analysis other than the pre-defined sets of combinations that are based on extensive literature review. As there can be many outliers, the top 1-5 percent of the results are then further analysed. Finally, the outliers that have been detected will be visualized along with the supporting data to understand the issues in depth.

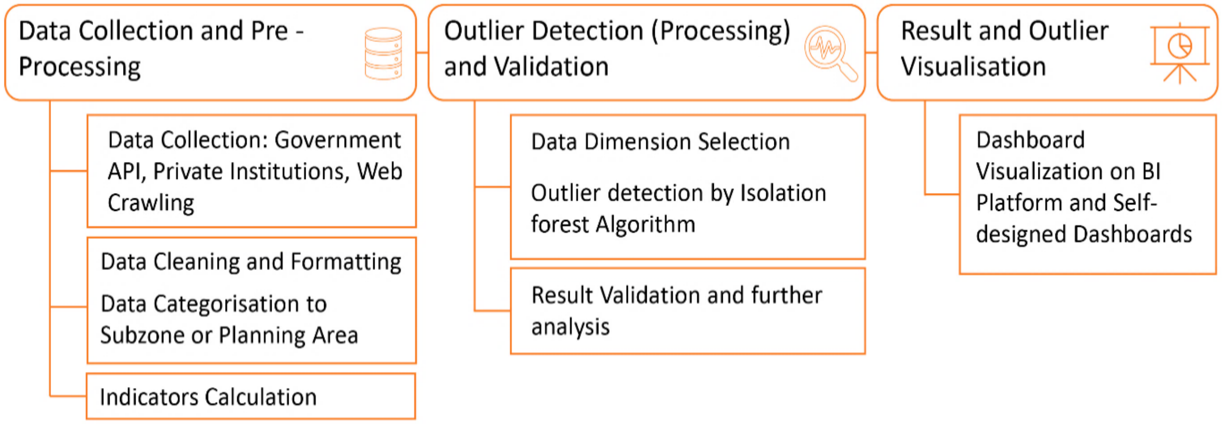

Figure 2. Summary of Outlier detection process and Visualisation

For extracting the outliers, we are using Isolation Forest algorithm due to its merits over other outlier detection systems. First, it has the ability to handle unsupervised datasets. Second, there is no need to establish criteria for defining the datasets. Third, it can handle multiple variables at the same time and with high dimensionality and finally, it can work with few samples as well as extremely large datasets and requires less computational power than some of its other counterparts. However, in order to provide additional flexibility, there will be a provision to use other algorithms as well if needed for validation purposes.

3.2. Data Collection

Data collection is one of the most important parts of this project as it is employed to provide a comprehensive view of the pulse of the city. Datasets are collected from various sources, including government institutions and private organizations, such as healthcare providers and a major credit card network provider. Some of the datasets are gathered as static files while most of the dynamic data has been aggregated by using APIs and by web crawling. All datasets collected have been fully anonymized before use to ensure residents' privacy, which is deemed critical in this study. To ensure data privacy, which is critical in this study, all private data collected is fully anonymized, i.e., do not contain any variables that will lead to the individual's identity. More precisely, the healthcare data gathered has been fully anonymised in accordance with the 'Health Insurance Portability and Accountability Act of 1996' (HIPAA). The payment transaction data is geo-aggregated and indexed to ensure that no identifiable data can be traced. Other private data from third parties are also handled with similar protocols, and ethics approval is obtained when necessary. Additionally, for city planning purposes, individual-level data is not required; hence, the data captured is aggregated at a subzone level, making anonymization plausible. Figure 3 shows different sources we have gathered the data from.

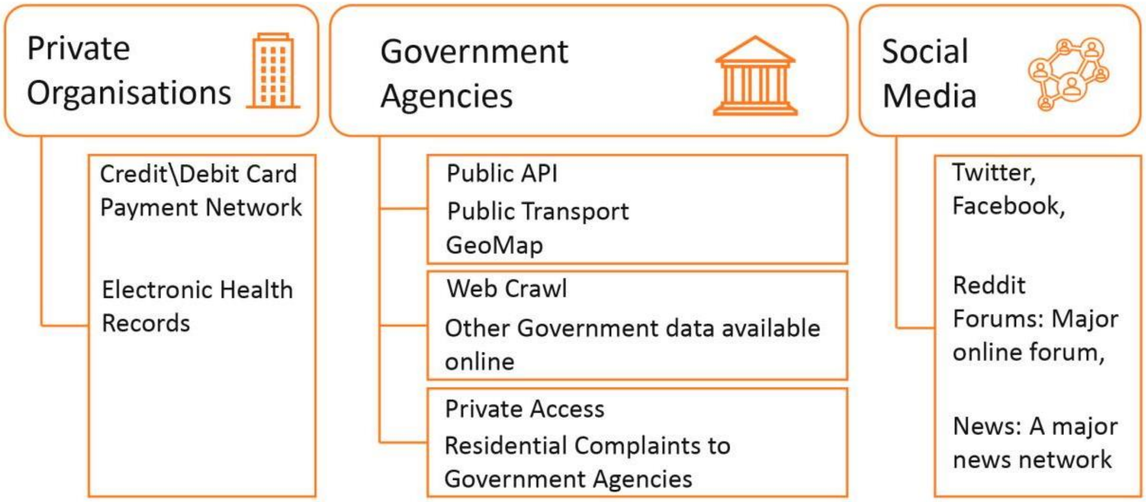

Figure 3. Sources of Data Collection

A list of data that has been collected at different granularity and velocity so far is presented in Table 1. Apart from the raw datasets collected, indices, such as Network Density, Road Coverage, Green Cover, Entropy Index, Dissimilarity Index (Bordoloi et al., 2013), Herfindahl–Hirschman Index (HHI) (Zagorskas, 2016), Land Use Distribution, Built Density, etc., were calculated to gain more holistic insights of the urban data and the urban landscapes.

Table 1. List of Datasets collected

Data Collected	**Granularity**	**Source**	**Velocity**
Environmental Perspective			
Carpark availability	Subzone	Government	Real time
PSI Index	Subzone	Government	Daily
Administrative Boundary - Singapore Boundary	National	Government	Not applicable
Administrative Boundary - Planning Area	Planning Area	Government	Not applicable
Administrative Boundary - Subzone	Subzone	Government	Not applicable
Land Use Data	Subzone	Government	Not applicable
Building Footprint Data	Subzone	Open Street Map	Not applicable
Point Of Interest Data	Subzone	Government	Not applicable

Passenger volume by bus stop	Subzone	Government	Monthly
Passenger volume by train stations	Subzone	Government	Monthly
Taxi availability	Subzone	Government	Real time
Mode of Transports to School Data	Planning Area	Government	Annually
Mode of Transport to Work Data	Planning Area	Government	Annually
Health Perspective			
Infectious diseases	Planning Area	E-Medical Record	Quarterly
Chronic	Planning Area	E-Medical Record	Quarterly
Mental Well-being	Planning Area	E-Medical Record	Quarterly
Injuries	Planning Area	E-Medical Record	Quarterly
Substance abuse	Planning Area	E-Medical Record	Quarterly
Other cause of concerns	Planning Area	E-Medical Record	Quarterly
Economic Perspective			
Residential property transacted values (Affluence)	Subzone	Payment Provider Network (PPN)	Daily
Retailer: Txn Amount	Subzone	PPN	Monthly
Retailer: Txn Count	Subzone	PPN	Monthly
Retailer: Avg Ticket Size	Subzone	PPN	Monthly
Number of distinct cards used	Subzone	PPN	Monthly
Avg ticket size per card use	Subzone	PPN	Monthly
Avg frequency of use per card	Subzone	PPN	Monthly
Avg amount spent per card	Subzone	PPN	Monthly
Economic Status Data	Subzone	PPN	Annually
Work Income for Household (Monthly)	Subzone	PPN	Annually
Income from Work Data	Subzone	PPN	Annually
Industry of Population Data	Subzone	PPN	Annually
Societal Perspective			
Dwelling Type Household Data	Subzone	Government	Annually
Dwelling Type Population Data	Subzone	Government	Annually
Household Size Data	Subzone	Government	Annually
Household Structure Data	Subzone	Government	Annually

Education Status Data	Subzone	Government	Annually
Ethnic Distribution Data	Subzone	Government	Annually
Language Literacy Data	Subzone	Government	Annually
Marital Status Data	Subzone	Government	Annually
Occupation Data	Subzone	Government	Annually
Age Data	Subzone	Government	Annually
Religion Data	Subzone	Government	Annually
Spoken Language Data	Subzone	Government	Annually
Tenancy Data	Subzone	Government	Annually
Formal Complaints from residents	Subzone	Government	Quarterly
Attitudinal Perspective			
Social media comments from Twitter (Geotag to the country)	Planning Area	Twitter	Real time
Social media comments from Facebook (Public Institution accounts)	Planning Area	Facebook	Real time
All news articles from the major news network	Planning Area	Major News Network	Real time
All news articles from the community news network	Planning Area	Community news network	Real time
Relevant threads from Reddit forums	Planning Area	Reddit	Real time
All threads from a major online discussion board	Planning Area	Major Online Forum	Real time

3.3. Using SCoRe Methodology and Data Cube

There are two key applications of SCoRe methodologies.

3.3.1. Exploratory Research – Detecting and Interpreting Outliers

The primary use of the methodology and the resulting data cube is to understand the city dynamics. One of the ways to achieve this is to use the data cube as an alerting system. Being able to highlight outliers or anomalies in the datasets can help us to identify points of distress in the city's pulse. By assuming outliers as *'self-emerging'* and can surface from any section of the dataset we can use outlier detection algorithm to flag out the outliers (Goh et.al., 2022). With the outlier detection mechanism in place, we have tried to solve the problem by allowing users to focus on what data is relevant in the current situation and dive deep into it.

One example of the outlier that was detected by examining the health datasets was hospitalization of unusually high percentage of residents due to asthma cases in two subzones. The number of asthma cases in these two subzones was much higher than the national incident rate of asthma cases. Our further analysis revealed that these two areas also had a relatively higher proportion of rental housing units which are typically occupied by lower-income families. Nevertheless, any potential relationships between the public housing typology and health outcomes are highly unclear and speculative, and require further study.

Despite possible limitations with approximation and speculation, such initial observations may alert city officials of specific issues and provide a path that can be followed to address the problems timely and more quickly. One thing to note is that the aim of this study is not to argue for any causal relationships that may be formed during the analysis but to be used as an alerting tool and comparing various characteristics of the city data.

3.3.2. Directed Research – Comparative Analysis and Forecasting

The data cube can be used for comparison of properties of different subzones. This can be in the form of a directed question that stakeholders might be interested in or can be used to forecast the trends in data. While the purpose of using data cube is to find anomalies in datasets, the outliers that are detected can help to form the questions that need to be asked in a focused manner.

Comparison has been one of the important tools that planners and city officials use to examine proposed guidelines (Kandt and Batty, 2021). In Singapore, much of the data is available at high granularity and up to date, which enables the user to compare multiple subzones at the same time. To achieve this, the platform is equipped to visualize all relevant datasets to aid comparison of cross-domain data (Zhang et al., 2022). To make the process more intuitive and flexible, at this stage the user will be able to select the type of data to be showcased and the time period that needs to be visualized. By providing the ability to select the data, users can be connected to the process itself and will be better equipped to look at the facts from different angles.

Figure 4 provides an example of the comparative analysis of two specific attributes, namely: 'rail (MRT) tap-in volume' and 'economic spending per active credit/debit card' at two different time periods. The months selected, namely July 2021 and July 2022, represent the periods of complete lockdown and subsequent opening of the city state once Covid-19 cases went down. In respect to movement data, one can observe that the movement in the southern part of the country has increased with easing of Covid-19 restrictions, which is an indication that people had started returning to the offices for work in the downtown region of the city (in the central south). On the other hand, credit card spendings were distributed more evenly across the island than in 2022.

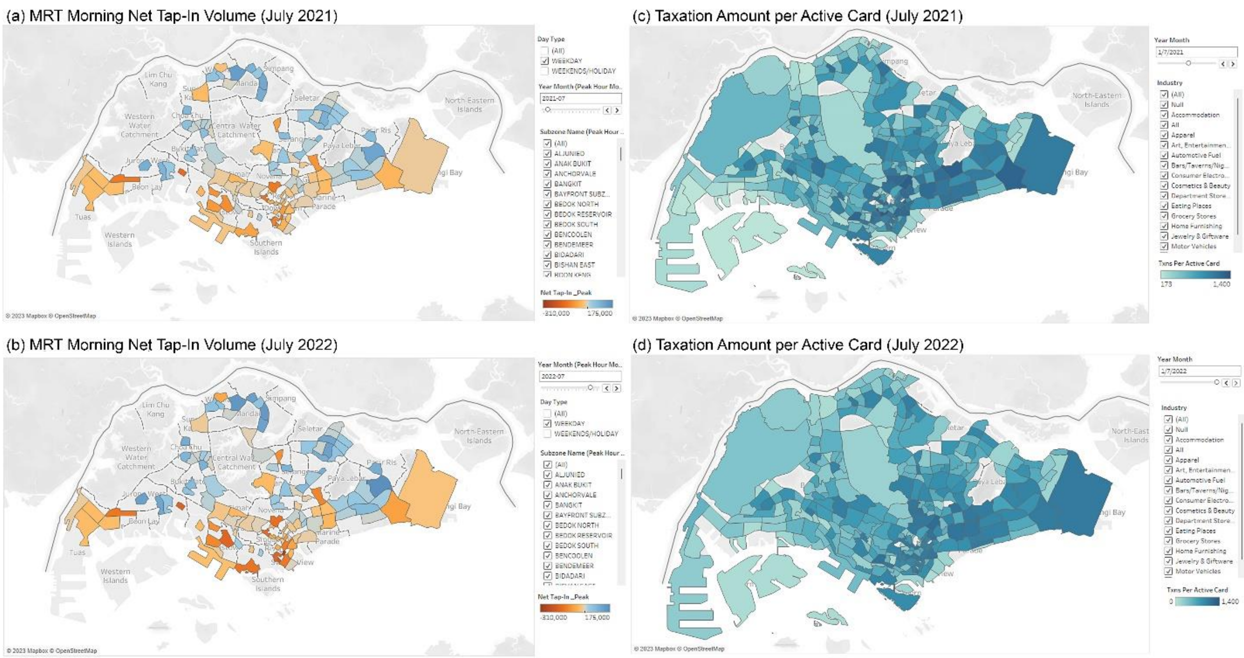

Figure 4. Movement - Economic Spending comparison: Metro rail tap-in volume compared with active card taxation at subzone level from July 2021 (when remote work was mandatory) until July 2022 (when situations were back to normal); (a) and (b) represent the MRT Net Tap-In volumes in July 2021 and July 2022 respectively; (c) and (d) represent taxation amount per active card in July 2021 and July 2022 respectively

Similarly, the user can zoom into one of the planning areas or subzones to get more details. In Figure 5, 'Chinatown' subzone is highlighted in the 'Outram' planning area to compare the Tap-In volumes (left) and active taxations (right) at different timeframes.

For performing this type of dashboarding, we used the existing BI platforms due to three main reasons. Firstly, the project needs to be handed over to the government agencies, and the used of the platform that is already in use by them would allow for a smooth knowledge transfer. Secondly, available BI platforms do not require complicated programming skills, which enables a larger group of users to use the platform easily, including editing or replacing datasets without much hassle. Lastly, the support for large number of formats and live data means that the platform can be automated to show live data analysis when required.

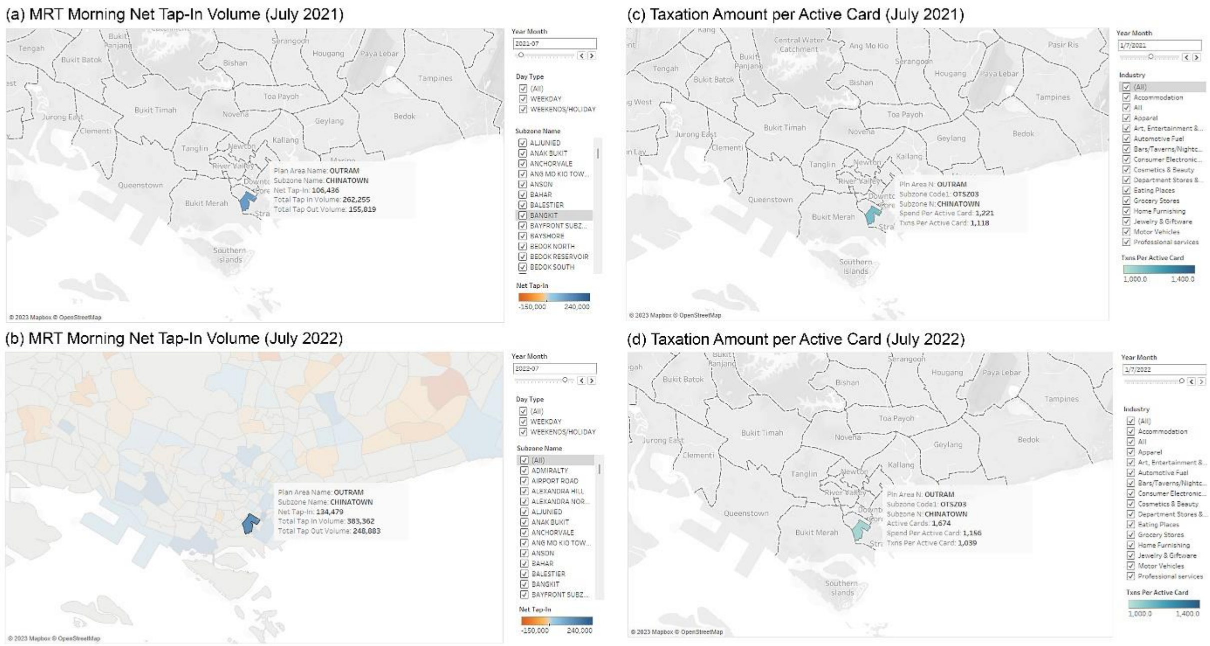

Figure 5. Movement -Economic Spending comparison - Zooming into one of the subzones in Outram Park planning area provides comparison of attributes from July, 2021 (when remote work mandatory) to July, 2022 (when situations were back to normal); (a) and (b) represent the MRT Net Tap-In volumes in July 2021 and July 2022 respectively; (c) and (d) represent taxation amount per active card in July 2021 and July 2022 respectively

4. Discussion and Future Work

In this paper, we have demonstrated the conceptual capacities of the alerting system that is being developed based on dynamic urban data understanding and analysis and the city pulse and anomalies detection. While the development of the platform for outlier detection is still in progress, the preliminary results are encouraging, indicating shorter time for issue detection, feedback and response to the community.

The response to the developments by the government agencies has been encouraging and the suggestions that were proposed by various stakeholders provided us the clarity for developing an intelligent dashboard that is intuitive and able to support real-time calculations and data visualisations. Some of the features that were appreciated by the stakeholders included: the capability to do comparative analysis (among subzones of similar characteristics), enabling the end user to select initial data and perform desired calculations in real-time (which is technologically challenging), and the ability to zoom in on particular planning areas for further analysis.

One of the requirements for being able to use the findings of urban analytics in general is adoption of data analytics techniques across the official landscape. This allows the user to be receptive towards the results of analysis. Singapore being one of the smartest cities in the world has most of its government offices well equipped with capacities of using big data in policy framework. The methodology presented here can have significant implications for city planning. The current research can have implications specifically in policy planning (understanding the current situation as an alerting system), Policy adoption and design (evidence-based design), Policy implementation and application (monitoring of policy) and future policy evaluation (Barbero et al., 2016).

Some limitations of the methodology should be acknowledged. Firstly, the quality of data available is not perfect. Data are from various sources and in different formats, which requires cleaning and pre-processing before use. As the amount of data increases so do the imperfections. More imperfections lead to increased processing time and ultimately delays in analysis. Secondly, the project depends on the availability of data at the micro-scale. Most of the data available is present at the national or regional scale, but only few datasets are present at planning area or subzone levels. This limits the granularity at which different datasets can be used for the analysis. Finally, the data are not always complete or representative. For instance, the credit card transaction and spending data that we have collected only represent part of the spectrum of the total spending by the citizens. In everyday life, much of the transactions are done using other means, such as cash and mobile payment platforms, or other credit card providers.

5. Conclusion

With this project we have tried to provide a stage for discussing the use of relevant datasets from the enormous pool of Big Data available. This will help to understand the vast expanse of city's digital footprint and by using outliers for factoring out the data, we will be able to focus on issues that needed immediate attention. Such alerting systems help to reduce the response time to issues faced by the city. By involving stakeholders in the process of development we were able to get several valuable responses that will allow us to mold the final platform towards the specific needs of the city planners and officials while retaining the original skeleton and framework intact.

Singapore is one of the smartest cities in the world and has its advantage in data readiness. On the other side of the spectrum are the cities whose digital infrastructure is not as capable. For such cities, where the availability and quality of data is in question, outlier analysis using isolation forest algorithm can still be a viable option due to its advantage of working even with small number of data points.

Acknowledgments. The content discussed in this paper is part of the research project funded under the Cities of Tomorrow R&D program (Award No: COT-CityScan-2020-1) by the National Research Foundation and Ministry of National Development of Singapore. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and Ministry of National Development, Singapore.

References

Ahlfeldt, Gabriel M. and Elisabetta Pietrostefani (2019) The economic effects of density: A synthesis. *Journal of Urban Economics, 111*, pp. 93-107. DOI: https://doi.org/10.1016/j.jue.2019.04.006.

Barbero, Martina, Jo Coutuer, Régy Jackers, Karim Moueddene, Els Renders, Wim Stevens, Yves Toninato, Sebastiaan van der Peijl and Dimitry Versteele (2016) *Big data analytics for policy making.* Brussels: European Commission. Available from: https://joinup.ec.europa.eu/sites/default/files/document/2016-07/dg_digit_study_big_data_analytics_for_policy_making.pdf Accessed [09/05/2023]

Batty, Michael (2013) Big data, smart cities and city planning. *Dialogues in Human Geography, 3* (3), pp. 274-279. DOI: https://doi.org/10.1177/2043820613513390.

Bibri, Simon E. and John Krogstie (2020) The emerging data–driven Smart City and its innovative applied solutions for sustainability: The cases of London and Barcelona. *Energy Informatics, 3* (1), 5. DOI: https://doi.org/10.1186/s42162-020-00108-6.

Bordoloi, Rupjyoti, Amit Mote, Partha P. Sarkar and C. Mallikarjuna (2013) Quantification of land use diversity in the context of mixed land use. *Procedia - Social and Behavioral Sciences, 104*, pp. 563-572. DOI: https://doi.org/10.1016/j.sbspro.2013.11.150.

Cao, Xinhui, Mei Wang and Xin Liu (2020) Application of big data visualization in urban planning. *IOP Conference Series: Earth and Environmental Science, 440* (4), 042066. DOI: https://doi.org/10.1088/1755-1315/440/4/042066.

Eaton, Michelle and Camilla Bertoncin (2018) *State of Offices of Data Analytics (ODA) in the UK*, London: Nesta. Available from: https://media.nesta.org.uk/documents/State_of_Offices_of_Data_Analytics_ODA_in_the_UK_WEB_v5.pdf Accessed [09/05/2023]

Froehlich, Jon, Joachim Neumann and Nuria Oliver (2009) Sensing and predicting the pulse of the city through shared bicycling. *In IJCAI'09: Proceedings of the 21st International Joint Conference on Artificial Intelligence*, pp.1420-1426. Available from: https://www.ijcai.org/Proceedings/09/Papers/238.pdf Accessed [09/05/2023]

Goh, Kim Huat, Wai Fong Boh and Premchand Dommaraju (2022) A social science approach using big data for city planning. *In ICIS 2022 Proceedings*, 4, 1555. Available from: https://aisel.aisnet.org/icis2022/it_policy/it_policy/4 Accessed [09/05/2023]

Kandt, Jens and Michael Batty (2021) Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities, 109*, 102992. DOI:

https://doi.org/10.1016/j.cities.2020.102992.
Liu, Pengyuan and Filip Biljecki (2022*)* A review of spatially-explicit GeoAI applications in Urban Geography*, International Journal of Applied Earth Observation and Geoinformation*, 112, 102936. DOI: https://doi.org/10.1016/j.jag.2022.102936.
Miranda, Fabio, Harish Doraiswamy, Marcos Lage, Kai Zhao, Bruno Gonçalves, Luc Wilson, Mondrian Hsieh and Claudio T. Silva (2017) Urban pulse: Capturing the rhythm of cities, *IEEE Transactions on Visualization and Computer Graphics*, 23(1), pp. 791–800. DOI: https://doi.org/10.1109/TVCG.2016.2598585.
Puiu, Dan, Payam Barnaghi, Ralf Tonjes, Daniel Kumper, Muhammad Intizar Ali, Alessandra Mileo, Josiane Xavier Parreira, Marten Fischer, Sefki Kolozali, Nazli Farajidavar, Feng Gao, Thorben Iggena, Thu-Le Pham, Cosmin-Septimiu Nechifor, Daniel Puschmann and Joao Fernandes (2016) CityPulse: Large scale data analytics framework for smart cities, *IEEE Access*, 4, pp. 1086–1108. DOI: https://doi.org/10.1109/ACCESS.2016.2541999.
Smith, Micheal E. and Jose Lobo (2019) Cities through the ages: One thing or many?, *Frontiers in Digital Humanities*, 6, pp. 12. DOI: https://doi.org/10.3389/fdigh.2019.00012.
Vaccari, Andrea, Mauro Martino, Francisca Rojas and Carlo Ratti (2010) Pulse of the city: Visualizing urban dynamics of special events. *In GraphiCon'2010, Proceedings of the 20th International Conference on Computer Graphics and Vision*, September 20–24, 2010, St.Petersburg, Russia. Available from: http://hdl.handle.net/1721.1/79406 Accessed [09/05/2023]
Young, Gareth W. and Rob Kitchin (2020) Creating design guidelines for building city dashboards from a user's perspectives, *International Journal of Human-Computer Studies*, 140, 102429. DOI: https://doi.org/10.1016/j.ijhcs.2020.102429.
Zagorskas, Jurgis (2016) GIS-based modelling and estimation of land use mix in urban environment, *International Journal of Environmental Science*, 1, pp. 284-293. Available from: https://www.iaras.org/iaras/filedownloads/ijes/2016/008-0044.pdf Accessed [09/05/2023]
Zhang, Dan, L.G. Pee, Shan L. Pan and Lili Cui (2022) Big data analytics, resource orchestration, and digital sustainability: A case study of smart city development, *Government Information Quarterly*, 39(1), 101626. DOI: https://doi.org/10.1016/j.giq.2021.101626.