

Use of Natural Language Processing to understand users' perspective on the Art Places/Places of Interest in Global Cities Singapore and Hong Kong

Xinyu Zeng^{a,*}, Frederick Peter Ortner^a and Bige Tuncer^b

^a *Singapore University of Technology and Design (SUTD), Singapore City, Singapore*

E-mails: xinyu_zeng@mymail.sutd.edu.sg, peter_ortner@sutd.edu.sg

^b *Eindhoven University of Technology, Eindhoven, The Netherlands*

E-mail: b.tuncer@tue.nl

Received 20 May 2024

Accepted 20 June 2024

Abstract. The Art Places/Places of Interest (POI) are increasingly important for Singapore and Hong Kong in their bid to be Global Cities. The design and operation of such Art Places are often led by the museum owners, city government and panel of experts from a top down approach, as well as funded by national governments for public benefit as part of long-term planning. The inputs from the actual users are often neglected. Public participation in Art Places/POI is often limited by individual visits and interactions. The diverse views and feedback on the design and operation of Art Places are difficult to capture accurately. In order to understand the perceptions of the users, extensive and expensive surveys and interviews need to be undertaken. Despite this, there is still a challenge of selection bias and interpretation bias. This paper explores the use of technology and big data to understand the similarities and differences between well-liked and disappointing areas of Art Places/POI in Singapore and Hong Kong. Public reviews on Art Places/POI in Singapore and Hong Kong will be examined using Natural Language Processing tools including the prevalent topic modelling method, namely Latent Dirichlet Allocation. The study revealed common strengths and weaknesses among artistic venues in Singapore and Hong Kong. “Place and experience” emerged as a common strength, while “price and content” were identified as a shared weakness. Singapore’s Art Places were distinguished by a unique strength in their “kid-friendly element,” whereas Hong Kong excelled in “food and shopping.” However, Singapore faced a unique weakness in “racial enclaves,” whereas Hong Kong’s distinctive weakness lay in “service.” These insights can aid urban planners and operators in comprehending and addressing areas of improvement highlighted by negative reviews, thereby enhancing overall performance.

Keywords: Art Places, Places of Interest, Singapore, Hong Kong, Global City, Natural Language Processing, Latent Dirichlet Allocation

* Corresponding author. E-mail: xinyu_zeng@mymail.sutd.edu.sg.

1. Introduction

This paper seeks to understand the positive and negative aspects of Art Places/Places of Interest (POI) through users' perspectives in Singapore and Hong Kong. Art Places is a means of putting narrative or arts in place to bring local communities and visitors together, making it also POI [10,20,25,26]. User-generated content, via the online review content of 220 Art Places/POI in Singapore and Hong Kong, are analyzed by Natural Language Processing (NLP) to identify the topics for positive and negative reviews. The results of 220 Art Places/POI would be consolidated under different cities for discussion. Individual Art Places that are of interest (comprising of places with a higher number of reviews, places with high positive reviews as well as places that have the worst reviews) are also discussed. By delving into the authentic voices of users in Singapore and Hong Kong, we aspire to contribute to the field of cultural management and participatory design, offering valuable feedback to urban planners to formulate better-designed Art Places/POI as part of their informed decision-making processes.

The development and promotion of arts and culture are increasingly crucial to drive economic growth and social development [6,19]. Cities are bidding to be Global Cities that foster innovation and creativity across different fields, such as arts, culture, technology, and economics [4,16]. Having the understanding that arts and culture are important pillars to drive economic growth and social development, the means and ways to execute the plans are to be explored. The planning for arts and cultural spaces is traditionally a top-down approach but there is a need to have further support for an arts-led urban rejuvenation in cultural planning in balance with the traditional formal governance structure [11]. This is aligned with the principles drawn from public participation in urban planning as the public goods are intended for public consumption [7,8]. However, there is limited research on the elements of desirable Art Places/POI from public perception. In summary, the research gap primarily concerns the lack of a comprehensive understanding of Art Places/POI from the user-centred approach. Bridging this gap can help to enhance cultural planning from the traditional top-down approach to one that aligns with the preferences and needs of the user. This also supports the bid for cities aspiring to be Global Cities and optimizing the resource allocation to one that effectively meets users' needs.

Singapore and Hong Kong, both economic powerhouses and former British colonies, are our chosen cities for this exploration. The reasons for choosing both Singapore and Hong Kong as case studies are due to the similarities in their economic and demographics and yet seemingly differences in their approaches to art development. Despite their economic similarities, these cities exhibit contrasting strategies for nurturing art spaces [1,5,14]. For instance, both cities have invested significantly in cultural institutions such as the National Gallery Singapore and M+ Hong Kong, yet they diverge widely in societal trajectories, including levels of art education, acceptance of art spaces, art trading practices, attitudes towards cultural and social spaces, policies governing art space development, and governmental approaches to art spaces. Therefore, by exploring the two cities, this paper can extrapolate the common themes for desired and undesired aspects of Art Places/POI from users' perspectives as well as understand the unique aspects of Art Places of different cities.

This paper seeks to address two questions: (1) the key elements of well-liked art spaces based on public opinions, and (2) the converging and diverging areas that mould Art Places within the dynamic contexts of Singapore and Hong Kong.

To answer these questions, public input on their experience on the Art Places/POI is needed. Traditional methods, such as extensive surveys and interviews are not only resource-intensive but also prone to selection biases and interpretation biases [29]. The emergence of technology, however, heralds a transformative era as the level of user-generated content increases, providing more insights into the services from the users' perspectives [24]. For example, the mobile digital footprint was used to understand the user visitation patterns for Art Places in Singapore [31]. Another perspective could be tapping into the social media platforms as valuable sources of people's opinions for knowledge discovery and user behaviour analysis [13,27]. Platforms like Google Maps Reviews now empower users to spontaneously contribute their perspectives [17]. This methodology finds support in its cost-effectiveness, speed, and equitable access based on other similar studies on city parks in Chicago and thematic parks in Bandung, Indonesia, which harnessed Google Maps reviews to understand public perceptions and inform urban branding strategies [12,22]. Google Maps reviews was used together

with the user star rating to understand the areas of service to be improved for restaurants in the Philippines [9].

Having the massive data of Google Reviews for each Art Place/POI, the challenge is to analyze the underlying positive and negative elements discussed for each Art Place/POI. NLP will be used to analyze the data. This approach is preferred over traditional or manual analysis which was impractical for processing such large datasets; NLP can harness the technological breakthrough in computational power to process Big Data and it also circumvents human biases [17,21]. The use of NLP to analyze online reviews was well-documented and it presented a comprehensive review of the different NLP ways to do data scraping, text cleaning, tagging, sentiment analysis and text analytics [2]. There was further summarized and discussed different text representation strategies, techniques and models for tourism sector analysis [18]. This also provides an understanding that NLP can be useful for analyzing the tourism sector by doing tourist profiling and understanding the demand and preferences of the consumer/market.

Among NLP models, Latent Dirichlet Allocation (LDA) is used to unearth hidden topics within the database, providing granular insights into user perspectives. Similar studies have successfully utilized LDA to gain a nuanced understanding of user-generated content, from hotel reviews' aspects [23] to Airbnb experiences in diverse cities [15]. Another application used LDA to understand the online reviews of popular restaurants [30]. There was a summary of different topic modelling applications used for social media analysis and it was found that the NLP helps and LDA was one of the most popular applications used for topic modelling [17].

In summary, this paper seeks to use technological advancement in NLP/LDA to gain insights into key aspects of these Art Places/POI to provide feedback to urban planners and related stakeholders to improve the overall user experience in their interactions with Art Places/POI. By improving the users' experience, the spillover effect could hopefully strengthen the cities' bid as Global Cities and have better economic and social development.

2. Methodology

This section discusses the methodology used to understand users' perception of Art Places/POI in Singapore and Hong Kong, including how the data is collected and processed, and the specific tools and methods used to analyze the data. The methodology is based on NLP techniques, in particular the LDA for topic modelling. It comes in the following steps:

- **Data collection:** using online reviews from Google Reviews for 220 Art Places/POI which are based on local authorities' perspective and the tool used for Data Scraping.
- **Data processing:** sorting the reviews into different users' rating categories and using NLP libraries, such as NLTK for stopwords, and wordnetlemmatizer.
- **Data analysis:** using LDA for topic modelling.

The results are the topics/keywords generated for the positive as well as negative reviews for each Art Place.

The methodology consists of three distinct steps (Fig. 1) including (i) data collection, (ii) data treatment and analysis, and (iii) data output. First, there is a flowchart to explain how the Art Places are defined and shortlisted, followed by how the online reviews for each Art Place are collected before sending for data treatment and analysis. The second part of the flowchart represents the data treatment after data collection, in the different steps, such as sorting review based on the users' score, followed by data treatment before putting it into the analysis. The third part shows the summary of the data collected as well as the results from the analysis.

2.1. Data collection and processing

The process of collecting data from the online reviews of 220 Art Places/POI has multiple complex implications for the type of processing and analysis subsequently conducted.

As the definition of Art Places can be broad and varied, a structural approach based on the perspective of local authorities is adopted. Another reason for using the definition from local authorities or government is because

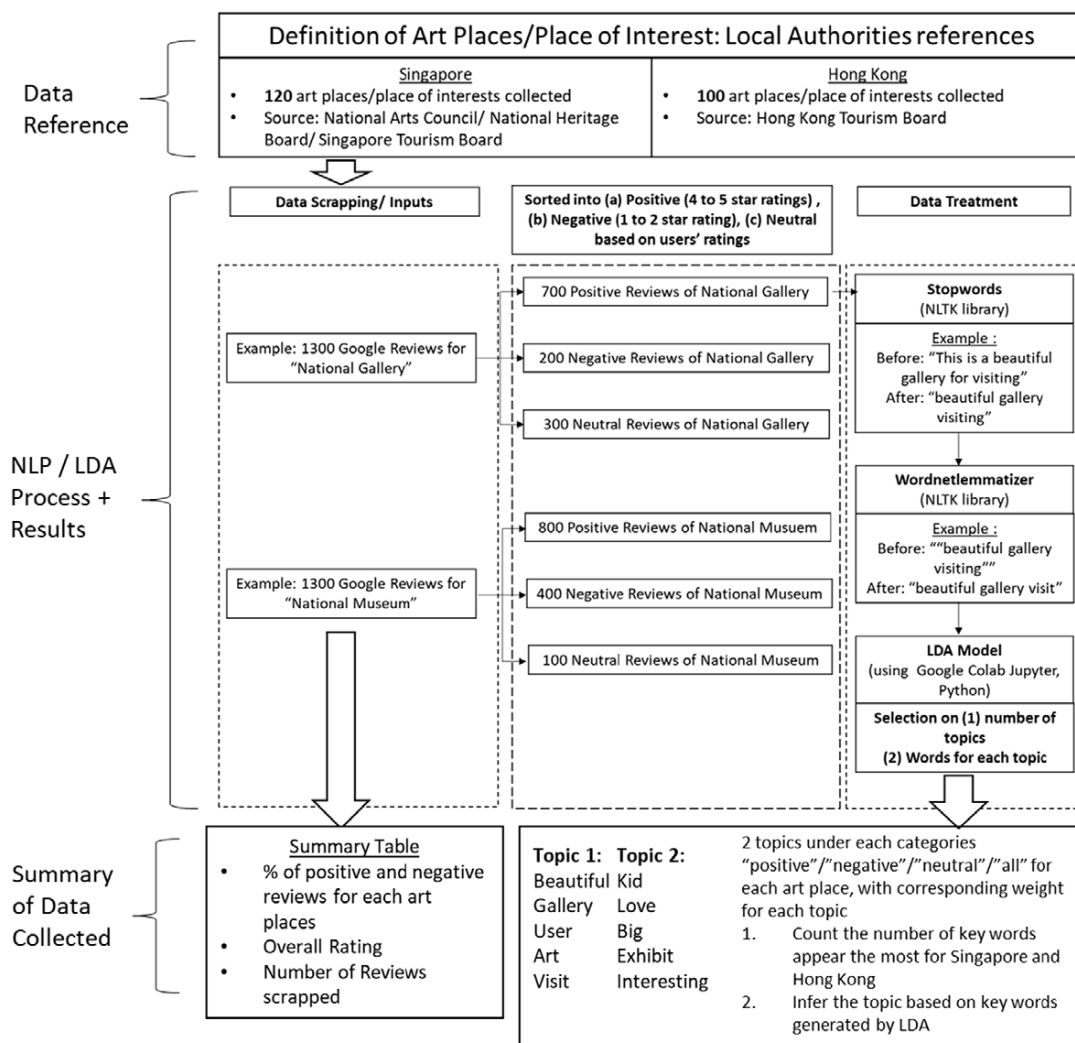


Fig. 1. An overview of the proposed framework.

the study covered two different cities and it is easier to rely on official understanding to make an apple-to-apple comparison. There are a total of 220 Art Places in Singapore and Hong Kong. Google reviews of these 220 Art Places are then collected using the Web-based Application Program Interface (API) of Data Scraper. There are approximately 200,000 reviews for these 220 Art Places. Each review has approximately 100 to 200 words. These big data are then analyzed using NLP in particular the topic analysis tool of LDA.

2.1.1. List of Art Places/POI in Singapore

The list of Art Places/POI in Singapore was created by the National Arts Council (NAC), National Heritage Board (NHB) and visitsingapore.com by the Singapore Tourism Board (STB). Based on the POI recommended by visitsingapore.com, the categories under "art museums and institutions", "performance and exhibitions", "culture – place of worship" and "culture-heritage discovery" are singled out. As for the sites under NAC, the categories under "education", "art space/venue", and "art housing scheme" are singled out. All the sites listed Under the "museum roundtable" by the National Heritage Board are included in the list of Art Places.

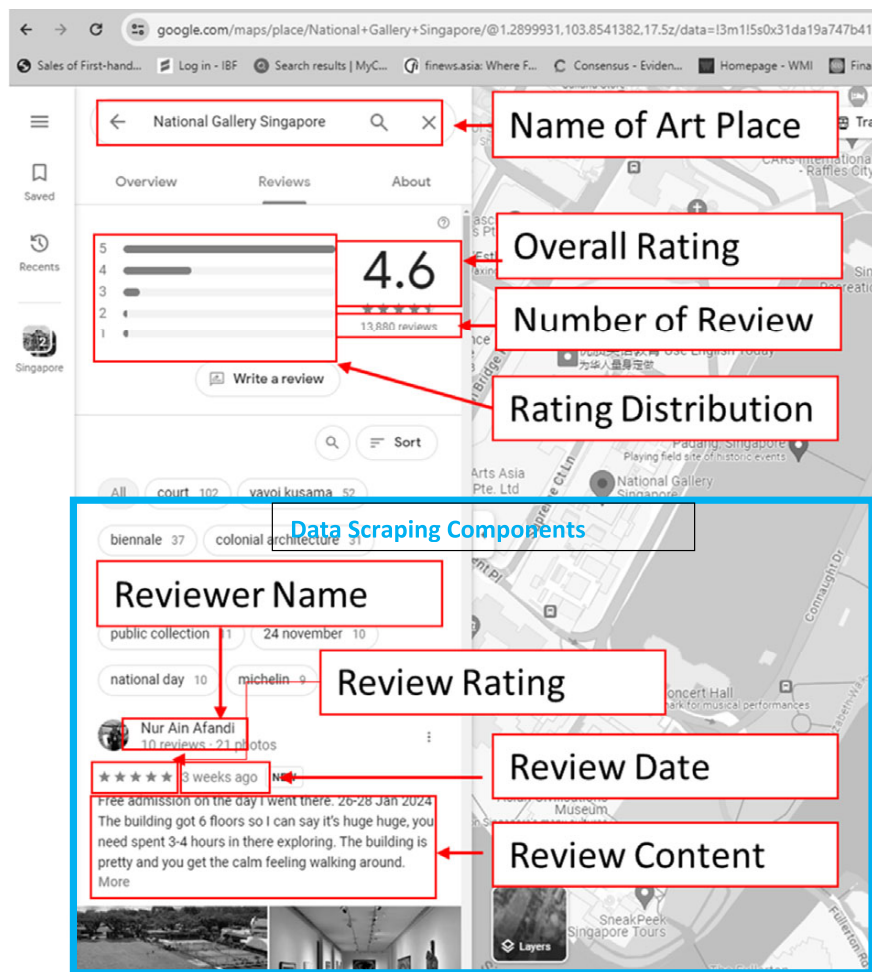


Fig. 2. Data scraping component (example for the National Gallery Singapore).

2.1.2. List of Art Places/POI in Hong Kong

The list of Art Places/POI in Hong Kong was created by discoverhongkong.com by the Hong Kong Tourism Board. All the Art Places recommended by the section “art” are singled out. Examples of the lists of Art Places are “Places to Discover Cantonese Opera”, “Hong Kong’s Iconic Neon Signs Are Becoming an Art Form”, “where to Enjoy Performing Arts in Hong Kong”, “arts by the Harbour”, “arts down south”, “a taste of the arts”, “iconic film locations in Hong Kong” and many others.

2.1.3. Data scraping of Google Reviews of Art Places/POI

A web-based Application Program Interface (API) Scraper [28] for data scraping is adopted to collect the content of the reviews for different Art Places/POI on the Google Maps reviews. This particular API used HTTP JSON as its programming language to automatically download large amounts of website data. The data scraping is based on the website addresses of the Google Maps reviews for each of the 220 Art Places. The information collected includes (1) the reviewer’s name, (2) their review rating (from 1 star to 5 stars), (3) the review date and (4) their review content. (see Fig. 2) This data will then be stored in spreadsheets as raw data for further analysis.

Based on the limitations of the web scraper, the total number of reviews can only be a maximum of 1300. For Art Places with more than 1300 reviews, only the top 1300 relevant comments are collected. The criteria for a Google Maps review to be considered relevant are (1) length, such as the number of words in the review, (2) keywords, such

as the mention of keywords like the name of business (3) specific, such as a story or example rather than generic content (4) local, such as it is from a user based locally and from users with more reviews and (5) time, such as it is posted more recently. The reference data was collected in July 2022.

2.2. Data treatment

After data collection, the data is treated prior to analysis, as summarized in Fig. 1. Each review was then sorted into positive, negative or neutral reviews based on the users' ratings. The review content can be in different languages, and it is translated using Google translation into English. The non-English content is then manually deleted before putting into the NLP for analysis.

Instead of using Sentiment Analysis to identify the positive and negative reviews, this study used the individual users' ratings as they believed this is a more accurate representation. For instance, a negative review of said 1 star with a comment like "this Art Place is sooooo wonderful and cheap at \$50/pax for just 5 mins visit" might be classified wrongly as a "positive review" as the Sentiment Analysis is unable to identify sarcasm.

Each group of reviews with a different classification, for example the positive reviews, then underwent data treatment for "stopwords" and "wordnetlemmatizer" to remove irrelevant words as well as to prepare the basic form (for example, from "eating" to "eat") of words for analysis.

In the LDA model, the authors have then identified the number of topics and the number of words to be associated with each topic. The results of this group, for example, positive reviews for the National Gallery, are then completed with the LDA model providing the words associated with the 2 topics identified. For each topic under the positive classification, the LDA model will provide its weightage. This will illustrate how important it is in the particular Art Place. For example, for the positive review of the National Gallery, there are two topics: (1) experience, and (2) place. LDA model will also provide the weight for (1) experience, said 80%, and (2) place, said 20%. This implies that 80% of the topics for the positive reviews for the National Gallery are on its experience, and the remaining 20% is on the place.

2.3. NLP model: LDA for topic modelling

The large volume of data available on the Art Places/POI, of approximately 200,000 reviews with approximately 100 to 200 words for each review, makes manual analysis unfeasible. NLP with LDA modeling makes it possible to computationally analyze the underlying topics for each Art Place/POI.

LDA, introduced by Blei, Ng and Jordan [3], is a generative probabilistic model that assumes each topic is a mixture of an underlying set of words, and each document is a mixture of over a set of topic probabilities. LDA represents topics by word probabilities. The words with the highest probabilities in each topic usually give a good idea of what the topic represents. The topics for each Art Place can be further split based on the positive reviews or negative reviews to understand what the topics are discussed in good or bad reviews for individual Art Places.

Based on [3], LDA is the generative process for each document w in a corpus D :

- (a) Choose a multinomial distribution φ_t for topic t ($t \in \{1, \dots, T\}$) from a Dirichlet distribution with parameter β
- (b) Choose a multinomial distribution θ_d for document d ($d \in \{1, \dots, M\}$) from a Dirichlet distribution with parameter α .
- (c) For a word w_n ($n \in \{1, \dots, N_d\}$) in document d ,
 - i. Select a topic z_n from θ_d .
 - ii. Select a word w_n from φ_{z_n} .

Given the parameters α and β , the joint distribution of a topic mixture θ , a set of N topics z , and a set of N words w is given by:

$$\rho(\theta, z, w | \alpha, \beta) = \rho(\theta | \alpha) \prod_{n=1}^N \rho(z_n | \theta) \rho(w_n | z_n, \beta)$$

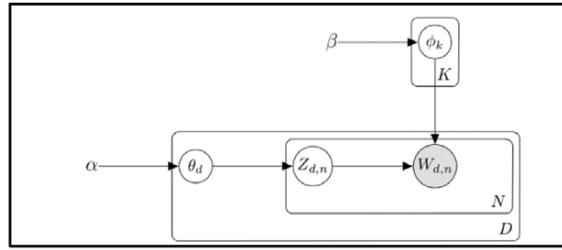


Fig. 3. Graphical representation of Latent Dirichlet Allocation.

It is also represented as a probabilistic graphical model in Fig. 3. There are three levels to LDA representation. The parameters α and β are corpus-level parameters, assumed to be sampled once in the process of generating a corpus. The variables θ_d are document-level variables, sampled once per document. Finally, the variables $Z_{d,n}$ and $W_{d,n}$ are word-level variables and are sampled once for each word in each document.

LDA does not highlight the actual topics, but it extracts the keywords under a potential topic group. The researcher can determine the number of topics and the number of keywords for such analysis. In this research, each Art Place will have 2 topics based on all reviews, 2 topics based on positive reviews, 2 topics based on negative reviews and 2 topics based on neutral review. The key reason why 2 topics are selected is (1) the words generated for the topics under different combinations between 1 to 4 topics seem to be more meaningful under 2 topics and (2) most of the comments are relatively short with 1 to 2 sentences implying generally 2 topics. Based on the keywords generated in the LDA model, the topic is determined by the researcher based on their own understanding.

The LDA analysis is supplemented with information drawn from the user ratings. Regarding the earlier discussion on the Google Maps Reviews, the data for each Art Place includes the users' rating (1 to 5 stars). For those with 4 and 5 stars, they are regarded as positive reviews. As for those with 1 and 2 stars, they are regarded as negative reviews. Lastly, reviews with 3 stars are regarded as neutral. For each topic, there will be 10 keywords. The general, positive, negative and neutral reviews for each Art Place will then be compared to each other to understand the key features that users like or dislike. In other papers, the positive and negative ratings are provided by the NLP based on the review content. However, in our research, the reviewers' ratings are used instead to determine if they are positive or negative about the Art Place. This is to ensure that we keep classified content's level of satisfaction from the lens' of users instead of the content itself. This also ensures sarcasm is not misinterpreted; for example, the user gave a negative rating and yet in its comments "The staff's attitude so good! Like real!". In some NLP analyses, this might be deemed as positive feedback.

3. Results

This section presents the results of analysis of the dataset of 200,000 user comments, providing (1) identification of the most and least well-reviewed Art Places; (2) extraction of the most prominent words associated with positive and negative reviews; (3) the extraction of topics using LDA for the most and least well reviewed Art Places.

As there are a total of 220 Art Places/POI and over 200,000 user comments, this paper will analyze in detail only the Art Places receiving the highest number of reviews, focusing on two groups: the most positively reviewed Art Places, and (3) the most negatively reviewed places for each city.

3.1. Summary of Google Reviews of Art Places for Singapore and Hong Kong

In an initial analysis, the list of all Art Places is ranked according to the number of reviews received and the percentage of positive and negative scores. This analysis is applied to each city separately. There is a summary of the Art Places for each city, providing (1) a percentage of positive, negative or neutral reviews based on scores

Table 1

List of Art Places/Places of Interest (POI) in Singapore sorted by the greatest number of reviews and with conditional formatting on the review score, with green as the top 5 highest rated Art Places and with the red colour as the bottom 5 lowest rated Art Places/POI

Art Place/POI	% Positive	% Negative	% Neutral	Rating	Count
Buddha Tooth	95.7%	1.1%	3.2%	4.69	1310
Science Centre Singapore	88.6%	3.2%	8.2%	4.46	1310
Art Science Museum	87.2%	6.5%	6.3%	4.42	1310
Red Dot Design Museum	80.8%	7.6%	11.6%	4.27	1310
Haw Par Villa	80.2%	5.6%	14.3%	4.22	1310
Sultan Mosque	96.3%	0.8%	2.9%	4.77	1308
Sri Veeramakaliamman Temple	94.0%	2.5%	3.5%	4.66	1307
National Museum of Singapore	94.2%	1.9%	3.9%	4.62	1300
National Gallery	95.7%	1.2%	3.2%	4.70	1290
Esplanade Theatre	96.7%	1.1%	2.3%	4.68	1290
Asian Civilization Museum	92.6%	1.6%	5.7%	4.56	1289
Sri Mariamman Temple	88.8%	2.7%	8.5%	4.48	1278
Chinatown Heritage Centre	90.2%	2.3%	7.5%	4.47	1087
Museum of Ice Cream	63.8%	22.6%	13.6%	3.71	243
Mint Museum of Toys	65.0%	13.3%	21.8%	3.80	234
Singapore Musical Box Museum	91.4%	1.4%	7.1%	4.71	70
The Gem Museum	93.2%	0.0%	6.8%	4.77	44
NTU Centre for Contemporary Art Singapore	65.1%	11.6%	23.3%	3.86	43
Sustainable Singapore Gallery	97.6%	0.0%	2.4%	4.79	42
Fuk Tak Chi Museum	25.0%	65.0%	10.0%	2.75	40
Army Museum of Singapore	74.2%	9.7%	16.1%	3.87	31

provided by the reviewers, (2) the overall score for each Art Place, as well as (3) number of reviews scraped from Google Maps. For subsequent analysis (Sections 3.2), only the Art Places with 1000 reviews or more are studied.

From the Singapore dataset total of 69 Art Places/POI are extracted from the original list of 120 (57.5%). From the Hong Kong data set a total of 51 Art Places/POI are extracted from the original list of 100 Art Places/POI (51%). For both datasets the Art Places dropped from the list are those with less than 20 reviews each. For presentation purposes, only the Art Places/POI with 1000 reviews or more, and the 5 highest (highlighted in green) and 5 lowest rated Art Places/POI (highlighted in red) are presented in Tables 1 and 2 (for Singapore and Hong Kong respectively), sorted by the number of reviews.

Thirteen Art Places (13/69 or 18% of the initial set) in Singapore have received 1000 or more usable reviews. (For the purposes of this study, only reviews with written comments are considered usable. Reviews without comments are excluded from analysis.) In Hong Kong 10 Art Places (10/51 19.5%) have received 1000 or more usable reviews.

In the Singapore dataset the 69 Art Places have an average star rating of 4.40, implying most users felt positive about all the Art Places/POI. 64 Art Places/POI are rated 4 stars and above. From the Hong Kong Dataset of 51 Art Places/POI the average star rating is 4.16, of which only 14 Art Places/POI that have an average score of below 4 stars.

The top 5 most positive reviewed Art Places in Singapore are found to be (1) the Sustainable Singapore Gallery, (2) The GEM Museum, (3) Sultan Mosque, (4) Singapore Musical Box Museum, and (5) National Gallery. In Hong Kong the top 5 most positively reviewed Art Places/POI were found to be (1) the Art and Culture Outreach (ACO), (2) Lion Rock, (3) Bound By Hollywood, (4) Africa Coffee & Tea, (5) Muze Pens.

A total of 122 datasets (69 from Singapore, 51 from Hong Kong, plus 2 combined dataset for all reviews in the city), Art Places were in a next step subjected to topic analysis through LDA modelling and presented in Section 3.2.

Among the poorly reviewed Art Places, only 1 Art Place/POI (Fuk Tak Chi Museum in Singapore) has a rating of 2.75 stars, below 3 stars. In Singapore only 4 Art Places/POI have a rating between 3 to 4 stars. In contrast, in Hong Kong 14 Art Places/POI have been identified with an average score of below 4 stars (Table 2). The Art Places/POI

Table 2

List of Art Places/Places of Interest (POI) in HK sorted by highest number of reviews and with conditional formatting on the review score

Art Place Name	% Positive	% Negative	% Neutral	Rating	Count
Tai Kwun	92.1%	1.5%	6.3%	4.5	1310
West Kowloon Cultural District	82.1%	6.6%	11.3%	4.21	1308
IFC	87.5%	2.7%	9.8%	4.36	1305
Goldfish Market	75.2%	6.4%	18.4%	4.05	1305
Hong Kong Heritage Museum	88.3%	2.3%	9.4%	4.4	1304
Hong Kong Cultural Centre	86.6%	3.4%	10.0%	4.35	1304
The Mills	81.2%	4.6%	14.2%	4.17	1304
PMQ	80.6%	4.8%	14.6%	4.15	1294
Man Mo Temple	83.4%	2.4%	14.1%	4.28	1287
Xiqu Centre	68.0%	13.8%	18.1%	3.8	1230
Chun Yeung Market	58.6%	15.2%	26.2%	3.65	466
Leighton Centre	56.0%	14.1%	29.9%	3.6	291
Kwun Tong Ferry Pier, Kowloon	62.6%	17.4%	20.0%	3.71	230
Lion Rock	96.6%	0.0%	3.4%	4.72	206
Africa Coffee & Tea	91.4%	3.9%	4.6%	4.55	151
Bound by Hillywood	93.7%	2.1%	4.2%	4.63	95
Muze Pens	90.4%	3.6%	6.0%	4.55	83
CITY IN TIME – Tai Po Road Rest Garden	41.7%	11.1%	47.2%	3.42	72
COFFEEHOLIC	54.4%	22.8%	22.8%	3.47	57
ACO (Art and Culture Outreach)	97.8%	2.2%	0.0%	4.76	46

with the lowest review scores in both Singapore and Hong Kong also have a low number of reviews. For example, in Hong Kong, the places with ratings below 3.8 all had fewer than 500 reviews (Table 2). These Art Places/POI with low review scores also have a low number of reviews between 31 and 243.

3.2. LDA analysis

As discussed in the methodology of the LDA model, each Art Place will have a total of 4 topics: (1) 2 topics based on positive reviews with 4 and 5 stars, (2) 2 topics based on negative reviews with 1 to 2 stars. Each topic will have 10 words identified by the LDA model.

The first part of the LDA analysis was to conduct a frequency count of the keywords generated by the LDA model for all Art Places in Singapore and Hong Kong separately. Based on the word count, key topics are identified for Singapore and Hong Kong separately. Further comparisons between Singapore and Hong Kong were conducted to identify the common positive and negative keywords, as well as their unique positive and negative keywords.

The second part of the LDA analysis was an in-depth analysis of each Art Place. However, it is difficult to examine all Art Place across the full dataset of 220 Art Place. In this paper, we will cover the Art Places that have more than 1000 Google Reviews to reflect the public opinions of the more popular Art Places. For this selected list, there is a section to discuss the top 5 most highly rated Art Places, as well as the bottom 5 least rated Art Places. To provide a clearer understanding of individual Art Places, this paper will discuss the 1 topic with the highest weight generated by the LDA analysis instead of the 2 topics generated.

3.2.1. Frequency count of keywords generated by the LDA model

Based on the keywords generated for all the Art Places/POI, the word clouds below illustrate the keywords that appear most frequently for Singapore and Hong Kong. Separately, it is also possible to illustrate the keywords via word clouds on the keywords for individual Art Places, such as the National Gallery in Singapore and West Kowloon Cultural District/M+ in Hong Kong.



Common and unique themes for Singapore and Hong Kong The keywords (see Fig. 4) derived from the reviews of Singapore and Hong Kong Art Places, can be classified into two broad categories: ‘hardware’ and ‘software’. The words related to ‘hardware’ describe experience of the physical components and spaces of the places are “place”, “museum”, and “beautiful”. The words related to ‘software’ describe the experience, service or content of the place and include, for example, “nice”, “good”, “great”, “interesting” and “history”.

The four most common positive keywords are identical for both Singapore and Hong Kong’s Art Places/POI: “place”, “nice”, “great”, and “good”. Five other shared positive keywords between Singapore and Hong Kong are identified. These Art Place generally refer to aesthetics, and “sense of place” with keywords such as “beautiful”, “place”, “building”, and “visit”. The 5 shared keywords associated primarily with negative reviews for both Singapore and Hong Kong make reference to pricing, value for money and content, with keywords such as “much”, “nothing”, “small” and related words “construction/renovation”.

As highlighted in Fig. 5, the 5 positive keywords identified as unique to Singapore Art Places emphasize experience and content, with keywords such as “museum”, “learn”, “peaceful”, “experience” and “kid”. In contrast, the 3 unique negative keywords identified for Singapore reflect judgements of value for many and include “expensive”, “long”, and “waste”. Hong Kong Art Places is often related to food and shopping, with keywords such as “shop”, “coffee”, and “food”. The unique negative topic in Hong Kong that is not found in Singapore is its service attitude, with keywords such as “service”, “staff” and “rude”.

3.2.2. Topics identified for individual Art Places

For further analysis on individual Art Places, we first filtered Art Places with more than 1000 reviews. The large number of reviews provides more context on Art Places and might also reflect the Art Places’ popularity.

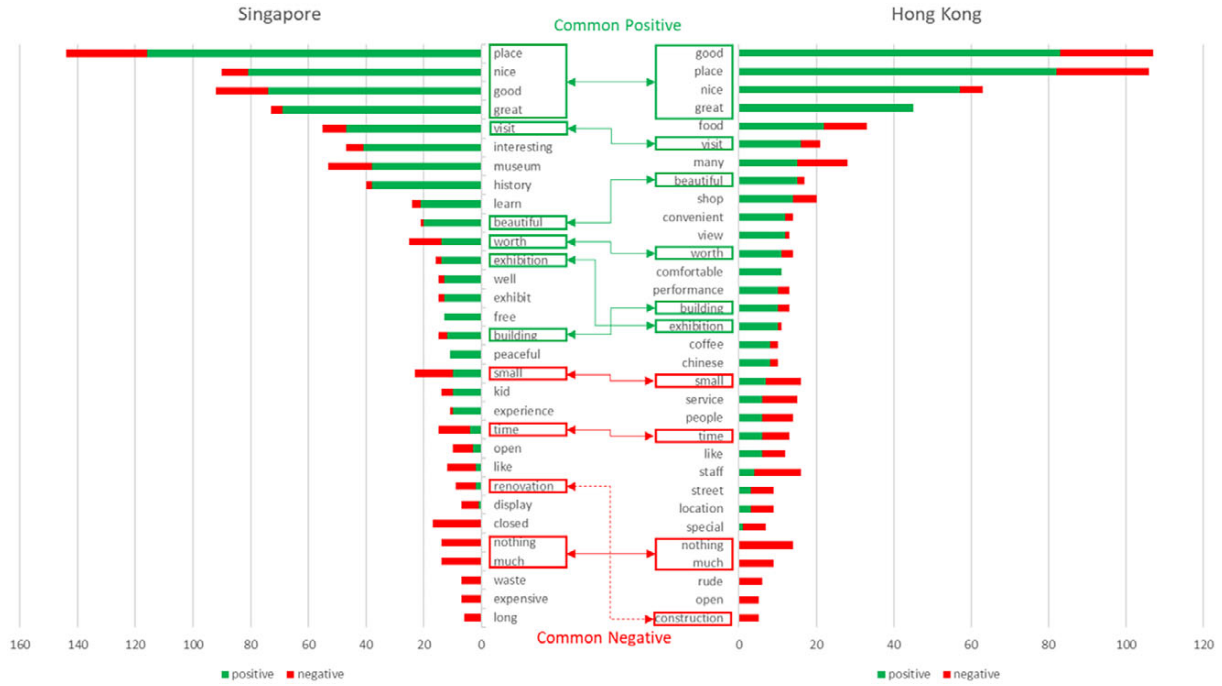


Fig. 4. Comparison of common themes between Singapore and Hong Kong.

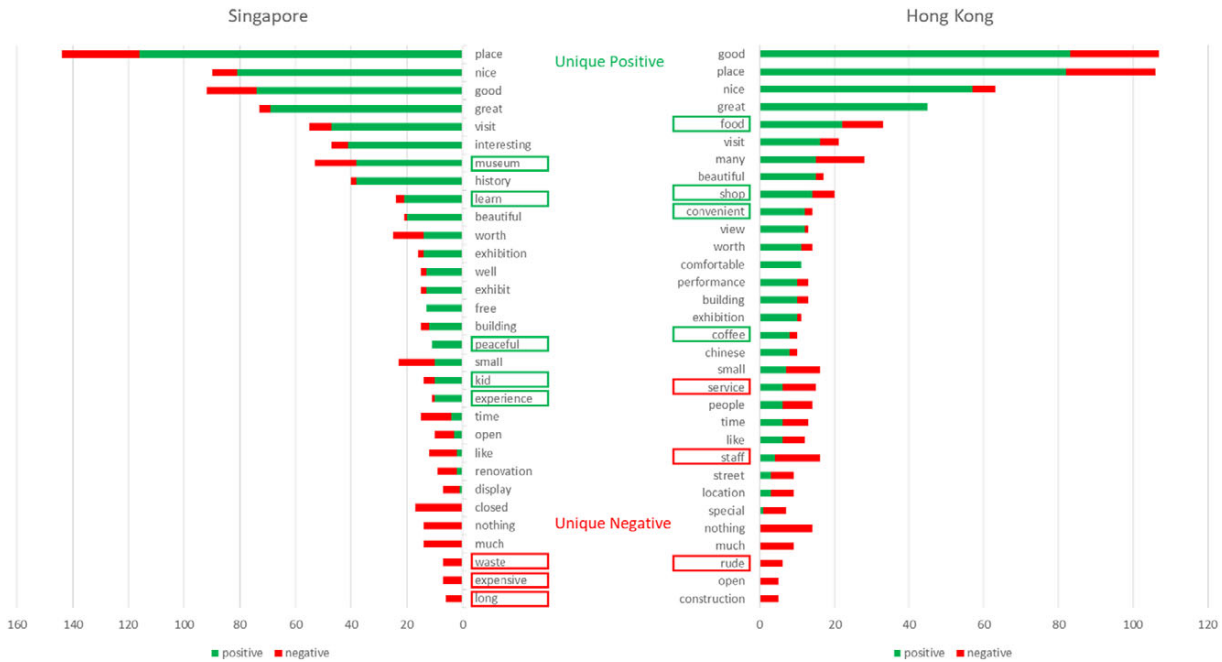


Fig. 5. Comparison of unique strengths and weaknesses between Singapore and Hong Kong.

Under this filtered list, the top 5 highest overall scores, and the bottom 5 with the lowest overall score are then discussed. The high number explores the 1 positive topic and 1 negative topic for each Art Place with higher weight from the LDA model, compared to the other topic under the same classification (positive or negative). For the top 5,

Singapore					Hong Kong				
Sultan Mosque	National Gallery	Buddha Tooth	Esplanade Theatre	Sri Veeramakaliamman Temple	Tai Kwun	Hong Kong Heritage Museum	IFC	Hong Kong Cultural Centre	Man Mo Temple
Place	Place	Aesthetic	Experience	Place	Place	Place	Shopping	Experience	Aesthetic
mosque	place	temple	place	nice	place	good	mall	place	temple
nice	gallery	place	nice	temple	nice	nice	shopping	great	place
singapore	singapore	visit	performance	good	visit	place	nice	nice	nice
beautiful	visit	beautiful	great	place	good	exhibition	place	sunset	visit
place	nice	nice	view	singapore	history	museum	good	relax	beautiful
good	building	must	free	hindu	kong	bruce	great	comfortable	hong
clean	great	singapore	singapore	love	great	hong	shop	beautiful	kong
masjid	exhibition	buddha	good	indian	hong	kong	hong	visit	good
pray	museum	peaceful	theatre	powerful	well	time	kong	view	must
food	exhibit	chinatown	esplanade	feel	building	great	high	night	chinese

Fig. 6. Top 5 Art Places for each city, topic and keywords.

only 1 positive topic for each Art Place is discussed. As for the bottom 5, only 1 negative topic for each Art Place is discussed.

In Singapore, there are 13 Art Places/POI with more than 1,000 reviews; in Hong Kong, there are 10 Art Places/POI with more than 1,000 reviews. Within this list, the top 5 rated and bottom 5 rated Art Places for Singapore and Hong Kong are selected for a deeper discussion.

LDA topic analysis for 5 best reviewed Art Places in each city Based on Fig. 6, the topics for the top 5 Art Places in Singapore are “Place”, followed by “Experience” and “Aesthetic”, whereas the topics for the top 5 Art Places in Hong Kong are “Place”, followed by “Experience”, “Shopping” and “Aesthetic”.

The rationale for assigning the topic based on keywords is to see what the potential underlying common topics for the keywords are. To illustrate the reasons for assigning the various topics, below are some selected examples:

- **National Gallery:** most of the keywords involve the place description, such as “gallery”, “place”, “building”, “museum”, “exhibition”,
- **Buddha Tooth:** A few keywords, “beautiful”, and “nice” reflect the aesthetic aspects of the place
- **Esplanade Theatre:** most of the keywords involve the experience, such as “performance”, “view”, “free”, “theatre”, “great”
- **IFC:** most of the keywords involve the shopping description, such as “shopping”, “mall”, “shop”
- **Man Mo Temple:** most of the keywords involve aesthetic description, such as “beautiful”, “good”, and “nice” and limited place description such as “place”, “temple”

In light of the positive reviews highlighting highly rated art establishments, it is evident that the success of art venues is significantly shaped by several key factors:

Physical characteristics and ambience: For places like the Sultan Mosque, National Gallery, Sir Veeramakaliamma, Tai Kwun, HK Heritage Museum, and HK Cultural Centre, the keywords primarily focus on describing the place itself. This suggests that the physical characteristics and ambience of the location play significant roles in attracting visitors and creating a memorable experience.

Experience: For venues like Esplanade Theatre and HK Cultural Centre, keywords related to the overall experience, such as performance, view, and sunset, indicate that visitors value the quality of events and activities offered at these locations.

Aesthetic: Keywords like beautiful, nice, and aesthetic aspects are particularly relevant for places such as Buddha Tooth and Man Mo Temple. This suggests that the visual appeal and architectural design of the venue contribute significantly to its attractiveness and overall experience for visitors. This focuses more on aesthetic aspects of the Art Place, as compared to a simple place description, such as place/temple.

Shopping: The IFC keyword analysis highlights the importance of shopping as a unique feature for certain locations, particularly in Hong Kong. Incorporating shopping opportunities alongside cultural and artistic offerings can enhance the overall appeal and diversity of experiences available at a venue.

Overall, these insights suggest that successful Art Places should prioritize creating a welcoming and aesthetically pleasing environment, offering engaging and memorable experiences, and potentially integrating unique features to cater to more specifically to local visitor interests and preferences, like the preference in Hong Kong for shopping integrated in Art Places.

Singapore					Hong Kong				
Chinatown Heritage Centre	Science Centre Singapore	Art Science Museum	Red Dot Design Museum	Haw Par Villa	West Kowloon Cultural District	The Mills	PMQ	Goldfish Market	Xiqu Centre
Operations	Service	Pricing	Place	Operations	Place	Shopping	Place	Shopping	Place
chinese	especially	place	nothing	place	construction	good	nothing	street	name
good	individual	ticket	small	visit	site	place	much	fish	building
interesting	feel	exhibition	much	hell	kowloon	feature	imagination	goldfish	construction
crowded	many	would	design	know	nice	time	disoriented	shop	nothing
welcome	would	clear	money	like	long	shop	nice	people	inconvenient
white	priced	adult	interesting	culture	space	shopping	small	good	people
closed	worse	price	picture	closed	good	special	good	forbidden	chinese
service	maintenance	money	museum	chinese	culture	nothing	building	longer	hong
interest	scolding	world	content	renovation	park	staff	fancy	past	money
long	family	disappointed	store	even	completed	fire	main	want	xiqu

Fig. 7. Bottom 5 Art Places for each city, topic and keywords.

LDA topic analysis for 5 worst reviewed Art Places in each city Based on Fig. 7, the topics for the worst reviewed 5 Art Places in Singapore are “Operation”, followed by “Pricing”, “Place” and “Service”, whereas the topics for the top 5 Art Places in Hong Kong, are on “Place”, followed by “Shopping”.

Derived from the subjects highlighted in the negative reviews of the lower-rated art establishments, it seems that the following key areas should be enhanced to elevate the quality of art venues:

Operations and maintenance: Keywords like “crowded”, “closed”, “service”, “long”, “scolding”, and “renovation” suggest that visitors’ experiences are affected by the operational efficiency and maintenance of the venues. Improving operations, managing crowds effectively, and maintaining facilities in good condition are essential for ensuring a positive visitor experience.

Place description: Keywords related to the place, such as “space”, “park”, “site”, “culture”, “building”, and accessibility challenges based on keywords like “construction” and “completed”, indicate that the physical characteristics and accessibility of the venues play a significant role in their perception. Enhancing the sense of place, ensuring accessibility, and addressing any operational challenges are crucial for attracting visitors and enhancing their experience.

Pricing and value: Keywords like “price”, “money”, “disappointed”, “ticket”, and “nothing” suggest that visitors’ perceptions of value for money play a significant role in their overall satisfaction. Ensuring fair pricing and offering compelling experiences that justify the cost of admission are important for maintaining visitor satisfaction and loyalty.

Shopping experience: Keywords related to shopping, such as “shop”, “shopping”, and “goldfish”, suggest that incorporating shopping experiences into Art Places is common, again particularly in Hong Kong, and has a strong impact on visitor perception.

Operational description: Keywords like “inconvenience”, “construction”, and operational aspects mentioned in various contexts highlight the importance of efficient operations and management in delivering a positive visitor experience. Addressing operational challenges and ensuring smooth operations are crucial for maintaining visitor satisfaction and reputation.

In summary, both Singapore and Hong Kong can improve their Art Places by focusing on areas such as operations and maintenance, enhancing the sense of place, addressing pricing and value perceptions, and improving the shopping experience where applicable. By addressing these key areas, owners, operators and local authorities/stakeholders can enhance visitor satisfaction, attract more visitors, and strengthen the overall cultural and artistic appeal of their respective destinations.

3.3. Recommendations

Based on the above findings of the study, some preliminary recommendations for urban planners and operators of Art Places in Singapore and Hong Kong are presented and discussed below.

Focus on improving pricing and content: Since the LDA model identified “pricing” and “content” as common negative aspects of Art Places in both cities, efforts should be made to enhance the value proposition offered to visitors. This could involve reviewing pricing structures to ensure they align with perceived value and investing in enriching content and experiences.

Enhance service attitude: In the case of Hong Kong particularly, where the service attitude of staff working in Art Places was found to be a frequent topic in negative reviews, it may be beneficial to invest in initiatives aimed at enhancing customer service skills and fostering a positive visitor experience.

Continue catering to families and children: Given the positive topics in reviews of Singapore Art Places associated with kid-friendliness, operators should be validated to continue catering to the needs of families and children by providing amenities and activities tailored to this demographic.

Further study can identify additional areas for improvement and capitalizing on strengths of Art Places in Singapore and Hong Kong and can also be used to track the effectiveness of institutional or city-wide efforts to improve service, content or experience in Art Places. Developing methods to quickly extract understanding from large databases of online reviews can enhance the competitiveness, attractiveness of urban Art Places and optimize their contribution to their respective cities' creative ecosystems and economies.

3.4. Limitations

Three important limitations of the study, all reflecting different types of human bias, must be acknowledged. The first: LDA topics are inferred by the author's understanding of the keywords formed for each 'hidden' topic and they might be subject to interpretation bias. For example, the same set of keywords ["place", "beautiful", "temple"] can be interpreted as topic of "place" or "aesthetic". Second, the identification of Art Places is subject to each region's government bodies' understanding. The definition of Art Places might not be aligned between Singapore and Hong Kong. For example, some shopping malls in Hong Kong are also classified as Art Places, which is not the case in Singapore. Finally, online reviews may not represent a diverse range of opinions, as they are typically provided by individuals who are motivated to share their experiences, which can lead to a skewed perspective. There is also a 'digital divide', as not everyone has access to the internet or feels comfortable expressing their opinions online. This digital divide can lead to the exclusion of certain voices from online reviews. Researchers and policymakers need to be aware of these limitations and consider complementary methods, such as surveys, interviews, or focus groups, to gain a more comprehensive and nuanced understanding of public opinion.

4. Conclusion and future research

This study has generated a comprehensive list of Art Places/POI for both Singapore and Hong Kong, with a total of 220 Art Places/POI identified. NLP technology was successfully adopted to gain insights into key aspects of these Art Places/POI. This technology helped identify areas that could be improved and aspects that were well-received by users in both cities, with a focus on supporting their bid to be recognized as Global Cities. This technology enabled the analysis of extensive data from more than 200,000 reviews, each with about 100 to 200 words.

This study demonstrates the usefulness of NLP and big data analysis in uncovering trends and providing actionable insights for urban planning and management in the context of cultural and artistic attractions. The study revealed common strengths and weaknesses among artistic venues in Singapore and Hong Kong. "Place and experience" emerged as a common strength, while "price and content" were identified as a shared weakness. Singapore's Art Places were distinguished by a unique strength in their "kid-friendly element," whereas Hong Kong excelled in "food and shopping."

Urban planners can utilize this information to prioritize improvements in Art Places based on identified weaknesses highlighted by negative reviews. They can allocate resources more effectively to enhance the overall quality and appeal of these venues, thereby enriching the cultural landscape of their cities. The socio-economic benefits of Art Places can be further explored through surveys/interviews/focus group that gauge community perceptions and behaviors towards these venues. Insights from such interviews/surveys can inform strategic urban development plans aimed at fostering a vibrant and inclusive arts scene.

Art Place operators, on other hand, can leverage this information to identify areas for improvement within their venues, addressing specific concerns raised by visitors in online reviews. By responding to negative feedback and enhancing the overall experience, operators can attract more patrons and foster a loyal customer base. Further exploration of economic benefits can guide operators in developing strategies to maximize revenue generation and

sustainability. Advanced methodologies such as deep learning can provide deeper insights into the economic impact of Art Places on the local economy, aiding operators in making informed decisions about programming, pricing, and marketing strategies.

In future research, additional NLP techniques, such as Sentiment Analysis, can be utilized to gain a more comprehensive understanding of the users' perspective on the Art Places. Similarly, other existing and emerging techniques from machine learning/deep learning in NLP can be used to compare their efficacy in identifying topics relevant to the experience of urban Art Places. As highlighted earlier, there could be surveys/interviews/focus groups to complement the quantitative analysis of online reviews. Future work may improve understanding by triangulating between web-scraped data and interviews/focus groups.

Acknowledgements

The authors thank the anonymous reviewers for their useful comments and guidance on the earlier versions of this paper.

Conflict of interest

None to report.

References

- [1] A. Archer, Materialising markets: The agency of auctions in emergent art genres in the global south, *Arts (Basel)* **9**(4) (2020), 106. doi:[10.3390/arts9040106](https://doi.org/10.3390/arts9040106).
- [2] B. Batrinca and P.C. Treleaven, Social media analytics: A survey of techniques, tools and platforms, *AI & Society* **30**(1) (2014), 89–116. doi:[10.1007/s00146-014-0549-4](https://doi.org/10.1007/s00146-014-0549-4).
- [3] D. Blei, A. Ng and M.I. Jordan, Latent Dirichlet Allocation, *Journal of Machine Learning Research* **3** (2003), 993–1022.
- [4] S.W.K. Chiu, K.C. Ho and T.L. Lui (eds), *City-States in the Global Economy: Industrial Restructuring in Hong Kong and Singapore*, 1st edn, Taylor and Francis Inc., 1996.
- [5] L.W. Choy, The future was when: Art criticism and the comparative tenses of Hong Kong and Singapore, *Journal of visual culture* **6**(3) (2007), 343–348. doi:[10.1177/1470412907084510](https://doi.org/10.1177/1470412907084510).
- [6] E. Currid, How art and culture happen in New York: Implications for urban economic development, *Journal of the American Planning Association* **73**(4) (2007), 454–467. doi:[10.1080/01944360708978526](https://doi.org/10.1080/01944360708978526).
- [7] S. Damer and C. Hague, Public participation in planning: A review, *Town Planning Review* **42** (1971), 217. doi:[10.3828/tpr.42.3.b274nj807658j756](https://doi.org/10.3828/tpr.42.3.b274nj807658j756).
- [8] A.R. Davies, Hidden or hiding?: Public perceptions of participation in the planning system, *The Town Planning Review* **72**(2) (2001), 193–216, town planning review, 2001, <http://www.jstor.org/stable/40112447>. doi:[10.3828/tpr.2001.72.2.193](https://doi.org/10.3828/tpr.2001.72.2.193).
- [9] A. de Guzman, J.M. Acena, P.G. Apalisok, J.E.K. Guerrero, L.J. Yasto and B. de Castro, Examining the structural relationships of Google local guides' online review characteristics and star ratings, *Anatolia* **1**(4) (2023). doi:[10.1080/13032917.2023.2207583](https://doi.org/10.1080/13032917.2023.2207583).
- [10] J. Heilbrun, Art and culture as central place functions, *Urban Studies* **29**(2) (1992), 205–215. doi:[10.1080/00420989220080271](https://doi.org/10.1080/00420989220080271).
- [11] S.F. Hoe, Place management/making: The policy and practice of arts-centred spatial interventions in Singapore, in: *Hard State, Soft City of Singapore*, Amsterdam University Press, 2020, pp. 191–212. doi:[10.1017/9789048544004.009](https://doi.org/10.1017/9789048544004.009).
- [12] Y. Huang, Z. Li and Y. Huang, User perception of public parks: A pilot study integrating spatial social media data with park management in the city of Chicago, *Land (Basel)* **11**(2) (2022), 211. doi:[10.3390/land11020211](https://doi.org/10.3390/land11020211).
- [13] H. Jelodar, Y. Wang, M. Rabbani et al., A NLP framework based on meaningful latent-topic detection and sentiment analysis via fuzzy lattice reasoning on youtube comments, *Multimedia Tools and Applications* **80** (2020), 4155–4181. doi:[10.1007/s11042-020-09755-z](https://doi.org/10.1007/s11042-020-09755-z).
- [14] M. Jeremie, Becoming Asia's Art market hub: Comparing Singapore and Hong Kong, *Arts (Basel)* **10**(2) (2021), 28. doi:[10.3390/arts10020028](https://doi.org/10.3390/arts10020028).
- [15] K. Kiatkawsin, I. Sutherland and J.Y. Kim, A comparative automated text analysis of Airbnb reviews in Hong Kong and Singapore using Latent Dirichlet Allocation, *Sustainability (Basel, Switzerland)* **12**(16) (2020), 6673. doi:[10.3390/su12166673](https://doi.org/10.3390/su12166673).
- [16] L. Kong, C.H. Ching and T.L. Chou (eds), *Arts, Culture and the Making of Global Cities: Creating New Urban Landscapes in Asia*, 1st edn, Edward Elgar Publishing, 2015.
- [17] C.D.P. Laureate, W. Buntine and H. Linger, A systematic review of the use of topic models for short text social media analysis, *Artificial Intelligence Review* **56** (2023), 14223–14255. doi:[10.1007/s10462-023-10471-x](https://doi.org/10.1007/s10462-023-10471-x).

- [18] Q. Li, S. Li, S. Zhang, J. Hu and J. Hu, A review of text corpus-based tourism big data mining, *Applied Sciences* **9**(16) (2019), 3300. doi:[10.3390/app9163300](https://doi.org/10.3390/app9163300).
- [19] P. Marcuse, The production of regime culture and instrumentalized art in a globalizing state, *Globalizations* **4**(1) (2007), 15–28. doi:[10.1080/14747730701245574](https://doi.org/10.1080/14747730701245574).
- [20] L. Marques and G. Richards, The dimensions of Art in place narrative, *Tourism Planning & Development* **11**(1) (2013), 1–12. doi:[10.1080/21568316.2013.852990](https://doi.org/10.1080/21568316.2013.852990).
- [21] M. Mleczek and M. Desmond, Using natural language processing to construct a national zoning and land use database, *Urban Studies* **60**(13) (2023), 2564–2584. doi:[10.1177/00420980231156352](https://doi.org/10.1177/00420980231156352).
- [22] Munawir, M.D. Koerniawan and B.J. Dewancker, Visitor perceptions and effectiveness of place branding strategies in thematic parks in bandung city using text mining based on Google maps user reviews, *Sustainability (Basel, Switzerland)* **11**(7) (2019), 2123. doi:[10.3390/su11072123](https://doi.org/10.3390/su11072123).
- [23] R. Priyantina and R. Sarno, Sentiment analysis of hotel reviews using latent Dirichlet allocation, semantic similarity and LSTM, *International Journal of Intelligent Engineering and Systems* **12**(4) (2019), 142–155. doi:[10.22266/ijies2019.0831.14](https://doi.org/10.22266/ijies2019.0831.14).
- [24] A. Ray, P.K. Bala and Y.K. Dwivedi, Exploring values affecting e-learning adoption from the user-generated-content: A consumption-value-theory perspective, *Journal of strategic marketing* **29**(5) (2020), 430–452. doi:[10.1080/0965254X.2020.1749875](https://doi.org/10.1080/0965254X.2020.1749875).
- [25] M. Stuiver, P. Jagt, E. Erven and I. Hoving, The potentials of art to involve citizens in regional transitions: Exploring a site-specific performance in Haarzuilens, The Netherlands, *Community Development Journal* **48**(2) (2013), 298–312. doi:[10.1093/cdj/bss022](https://doi.org/10.1093/cdj/bss022).
- [26] J.K. Taylor, Re-envisioning community spaces in Corona, Queens, New York City, *City, Culture and Society* **14** (2017), 14–21. doi:[10.1016/j.ccs.2017.12.004](https://doi.org/10.1016/j.ccs.2017.12.004).
- [27] L. Tomarchio, P. Herthogs and B. Tuncer, Hybrid art space typologies in Singapore from social media data, *City, Culture and Society* **34** (2023). doi:[10.1016/j.ccs.2023.100519](https://doi.org/10.1016/j.ccs.2023.100519).
- [28] Webscraper.io, 2022, Accessed on 1 July 2022. Available from: <https://webscraper.io/>.
- [29] C. Winship and D.M. Robert, Models for sample selection bias, *Annual review of sociology* **18**(1) (1992), 327–350. doi:[10.1146/annurev.so.18.080192.001551](https://doi.org/10.1146/annurev.so.18.080192.001551).
- [30] W. Yang and J. Yao, Research on online reviews of popular restaurants based on latent Dirichlet allocation theme model, Conference on Artificial Intelligence and Intelligent Information processing. doi:[10.1117/12.2659912](https://doi.org/10.1117/12.2659912).
- [31] X. Zeng and B. Tuncer, Use of mobile digital footprint to understand how COVID-19 impacts the visitors to art places in Singapore, *Journal of Smart Cities and Society* **2**(4) (2023), 181–192. doi:[10.3233/SCS-230017](https://doi.org/10.3233/SCS-230017).