Geojournal of Tourism and Geosites
ISSN 2065-0817, E-ISSN 2065-1198

Year IX, vol. 64, no. 1, 2026, p.457-464
DOI 10.30892/gtg.64140-1691

TOURISM ROUTE DEVELOPMENT USING ASSOCIATION RULE MINING AND GOOGLE MAPS IN SONGKHLA, THAILAND

Thanapoom MAAKSORN [1], Numtip TRAKULMAYKEE [2*], Nattapong TONGTEP [3]

[1] Prince of Songkla University, Faculty of Graduate School, Songkhla, Thailand; thanapoom.maak@gmail.com (T.M.)
[2] Prince of Songkla University, Faculty of Science, Songkhla, Thailand; n.trakulmaykee@gmail.com (Nu.T.)
[3] Prince of Songkla University, College of Computing, Phuket, Thailand; nattapong.t@phuket.psu.ac.th (Na.T.)

Citation: Maaksorn, T., Trakulmaykee, N., & Tongtep, N. (2026). Tourism route development using association rule mining and Google Maps in Songkhla, Thailand. *Geojournal of Tourism and Geosites*, 64(1), 457–464. https://doi.org/10.30892/gtg.64140-1691

Abstract: This research aims to explore tourist behaviour patterns and develop data-driven tourism route recommendations based on Association Rule Mining (ARM), particularly using the Frequent Pattern Growth (FP-Growth) algorithm. The study focuses on identifying co-occurrence patterns among tourist attractions based on real behavioral data, with the goal of supporting tourism route design and enhancing recommendations for self-guided tourists. The study employed a structured data mining methodology. Raw data was collected from real-world tourist check-ins and GPS traces in Songkhla City, Thailand; encompassing 45 attraction points. The study used the Cross-Industry Standard Process for Data Mining (CRISP-DM) for research methodology. The FP-Growth algorithm was used to discover association rules, using parameters set at support > 0.10, confidence > 0.5, lift > 1.0, and conviction > 1.0 to ensure the relevance and reliability of the extracted rules. The evaluation phase included interpretation of rule quality metrics such as support, confidence, lift, and conviction indicators. Visualizations presented the route recommendations on geographic maps. The results revealed 13 valid association rules involving only 5 out of the 45 locations, suggesting strong behavioral patterns and repeat co-visitation. Interestingly, some of the top-visited locations based on frequency alone did not appear in any of the 13 rules, indicating that frequency-based route planning may not reflect actual tourist behavior. Further analysis showed that all route combinations extracted from the rules had a total travel distance of under 88 kilometers. With Google Map data, this study demonstrates that ARM with FP-Growth can provide insightful patterns for route planning and destination clustering. The proposed method supports smart tourism development by bridging behavioral analytics and operational tourism planning. Stakeholders can apply this approach to periodically update route recommendations to ensure alignment with current tourist behaviors and preferences.

Keywords: association rule, Google Maps, FP-Growth, machine learning, recommendation route, Thailand, tourist

* * * * * *

INTRODUCTION

Tourism has emerged as a significant contributor to economic development, particularly in regions where cultural, historical, and natural attractions abound. As the volume of tourist data generated through digital platforms increases through online review systems such as Google Maps, there is a growing opportunity to extract actionable insights to enhance tourist experiences and optimize travel planning (Chandiramani et al., 2021).

Traditional route planning approaches rely heavily on static travel guides or expert knowledge, and often fail to capture the dynamic and contextual preferences of modern travelers. To better understand and respond these preferences, data-driven techniques are increasingly being adopted to uncover hidden patterns in tourist behavior. Integrating geolocation technologies into tourist route planning can help to increase tourist satisfaction and commercial efficiency (Amangeldi et al., 2025; Czyz & Javed, 2025; Seidualin et al., 2024). Association Rule Mining (ARM), a prominent technique in machine learning, enables the discovery of meaningful co-occurrence relationships among items within large datasets (Arreeras et al., 2019, Chang, 2025; Jiang & Dai, 2024). In the context of tourism, ARM can be used to identify frequently visited combinations of attractions, which can be transformed into recommended travel routes. Among the various ARM algorithms, the Frequent Pattern Growth (FP-Growth) algorithm stands out for its efficiency and scalability, especially when dealing with large and complex datasets (Duan, 2025).

This study proposes a novel approach to tourist route development by leveraging ARM using the FP-Growth algorithm on user-generated content from Google Maps reviews. Google Maps constitutes a rich repository of real-time tourist data, including location check-ins and tourist logins, which collectively reflect actual tourist movement patterns and preferences. By analyzing this data, the study aims to develop recommendation routes using ARM to extract frequent patterns of tourist site visits and construct optimized route recommendations that align with authentic tourist behavior.

* Corresponding author

LITERATURE REVIEWS

1. Association Rule Mining and Tourism

The evolution of the tourism industry has benefited significantly from technological advancements, particularly in the development of route recommendations (Ruko et al., 2024; Xiong & Zhang, 2024). Route recommendations have emerged as a crucial tool for enhancing tourist experiences. They address the challenge of information overload that tourists face when planning their trips. Route recommendations offer curated suggestions that align with the interests and constraints of individual tourists or tourist groups. The integration of data mining techniques, particularly ARM, has shown promising results in addressing these limitations by uncovering hidden patterns in tourist behaviors and preferences. ARM has gained significant attention in tourism research due to its ability to discover meaningful relationships between different attractions, activities, and destinations (Liu et al., 2021). This technique enables the identification of frequent patterns in tourist behaviors, which can be leveraged to generate relevant recommendations for similar user profiles (Chang, 2025).

The FP-Growth algorithm constructs a compact tree structure that enables faster pattern discovery, making it particularly suitable for data analytics on sparse datasets (Jie et al., 2024). When used for ARM in tourism applications, FP-Growth has also shown efficiency in handling large datasets (Duan, 2025) and several studies have demonstrated the superior performance of FP-Growth in tourism contexts (Chang, 2025; Duan, 2025; Ritthipakdee et al., 2024). Furthermore, our literature review showed that most previous research works integrated FP-Growth into other approaches (Duan, 2025; Li, 2024; Song & He, 2023). The most recent studies of ARM used in tourism contexts are presented in Table 1.

Table 1. Previous studies using ARM in Tourism

Authors	Summary	Route	Apriori	FP-Growth	Other Algorithm
Chang (2025)	Compared the performance of Apriori and a proposed algorithm based on FP-Growth, and provided several route themes	/	/	/	
Duan (2025)	Proposed a new prediction model based on Random Forest and FP-Growth, to recommend preferred attractions	/	X	/	Random Forest
Jie et al. (2024)	Proposed an algorithm based on FP-Growth for air traffic control, and compared the proposed algorithm with Apriori and FP-Growth algorithms	X	/	/	
Li (2024)	Proposed a hybrid tourist attraction recommendation model using Random; Forest, ANN, and FP-Growth algorithms, and recommended attractions	X	/	/	ANN, Random Forest
Ritthipakdee et al. (2024)	Developed a recommendation system for cultural attractions based on FP-Growth, and generated cultural attractions along the BTS Skytrain	/	X	/	Decision Tree, Random Forest
Song & He (2023)	Compared the performances of the Apriori algorithm and three machine learning algorithms for tourist attraction recommendations	/	/	X	Decision Tree, Linear Regression, Logistic Regression

2. Integration of Review Data with Association Rule Mining

Google Maps has emerged as a valuable source of user-generated tourism data, containing rich information such as ratings, reviews, timestamps, and geolocations (Xiong et al., 2025; Zhang & Zhang, 2025).

These reviews provide deep insights into tourist experiences, preferences, and satisfaction levels, offering real-world perspectives that complement traditional survey-based approaches (Zhao et al., 2025). The combination of Google Maps data with ARM represents a novel approach to tourism route recommendation. This integration enables the extraction of preference-based patterns that reflect actual tourist experiences rather than theoretical assumptions.

The rich metadata available in review data, including ratings, timestamps, and textual content, provides multiple dimensions for pattern analysis. Despite the growing body of research on tourism recommendation systems and route management presented in Table 2, few studies have combined ARM with user-generated data from platforms like Google Maps to develop routes (Hassan et al., 2025), especially using the FP-Growth algorithm (Liu et al., 2024).

Therefore, this study aims to bridge this gap by leveraging FP-Growth to mine frequent co-visit patterns from unstructured review data, thereby generating data-driven, realistic, and route recommendations.

Table 2. Previous studies using Google Maps

Authors	Summary	Context	Route	ARM	Others
Hassan et al. (2025)	Proposed a hybrid recommendation system using the Apriori and K-mean algorithms	Tourism	/	/	K-mean
Xiong et al. (2025)	Utilized Google Maps reviews to examine users' firsthand experiences with BRI infrastructure and how this infrastructure functions in users' daily lives	City	/	X	Sentiment
Zeng et al. (2025)	Proposed a knowledge graph-based Hidden Markov Model (HMM) for personalized travel routing using spatial and semantic data in urban environments	Smart City	/	X	Graph-based HMM
Zhang & Zhang (2025)	Explored how Google Maps effectively shapes conceived and perceived space from user feedback of place ratings and reviews.	Tourism	/	X	Qualitative & Quantitative Analysis

Authors	Summary	Context	Route	ARM	Others
Liu et al. (2024)	Explored the risky factors of Autonomous Vehicle (AV) crashes using the Apriori algorithm	AV-crashes	/	/	
Lou et al. (2023)	Recommended tourist attractions and routes based on tourist posts, text annotations, and photo sharing	Tourism	/	X	TF-IDF Algorithm
Lou (2022)	Proposed a travel destination recommendation system based on association rule mining	Tourism	/	/	
Karaş et al. (2021)	Investigated a mobile route planning and navigation application for tourism activities based on GIS	Tourism	/	X	
Nagy et al. (2022)	Examined the correlation and causal relationship between real-time mobility data and statistical data on tourism	Tourism	/	X	Statistical & Neural Networks
Nitu et al. (2021)	Improved travel recommendation system based on tweets data with sentiment analysis, and a classification approach	Tourism	/	X	Sentiment, Classification

RESEARCH METHODOLOGY

This research aims to develop optimal tourism route recommendations by applying ARM techniques to user-generated data extracted from Google Maps. The methodological framework follows the Cross-Industry Standard Process for Data Mining (CRISP-DM), which provides a structured approach to the knowledge discovery process in real-world tourism data. The process includes six main phases: business understanding, data understanding, data preparation, analysis, evaluation, and deployment.

1. Business Understanding

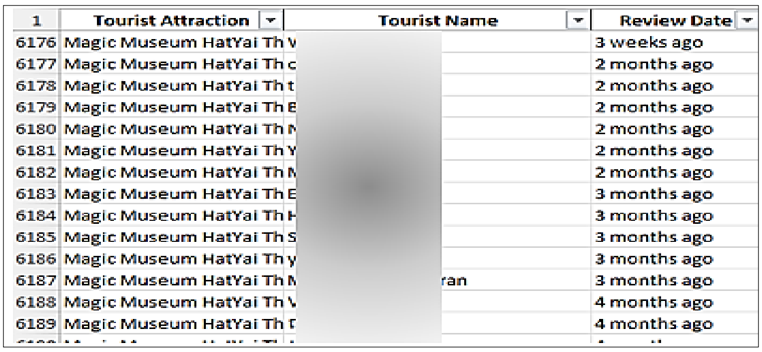
Traditional route development frequently lacks real-time adaptability, relying on static information that fails to capture emerging attractions, or evolving tourist interests. The absence of data-driven validation mechanisms results in routes that may not optimize tourist satisfaction. In addition, traditional methods often suffer from limited scalability, as manual route design cannot efficiently process large volumes of tourist preference data or adapt to rapidly changing visitor behaviors. Expert-based approaches introduce inherent biases so that the recommendations reflect individual perspectives rather than comprehensive tourist preferences derived from actual visitor patterns.

Therefore, this study identified patterns of tourist behavior and preferences by analyzing textual review data and location sequences from Google Maps. The ultimate goal was to generate intelligent recommendations for tourist routes that were context-aware and data-driven. ARM was adopted to discover frequent co-visited locations, enabling the design of tourism routes that reflect actual traveler behaviors and satisfaction levels.

2. Data Understanding

The dataset was obtained through web scraping from Google Maps, covering 19,243 reviews from 45 top tourist attractions in Songkhla City, Thailand. The attributes include tourist attractions, tourist names, and review dates as presented in Figure 1. An initial exploration was conducted to understand the number of unique tourists (8,925) and attractions (45), and frequency of tourist visits per attraction.

The top 10 attractions were Kim Yong Market (1,559), Samila Beach (1,343), Hatyai City Municipal Park (1,296), Chang Puak Elephant Camp Hat Yai (1,289), Khlong Hae Floating Market (1,139), Wat Laem Pho (1,121), Song Thale Park (1,096), Tangkuan Hill (1,048), Thale Noi Bird Watching Park (908), and Chalathat Beach (871).



1	Tourist Attraction	Tourist Name	Review Date
6176	Magic Museum HatYai Th	[illegible]	3 weeks ago
6177	Magic Museum HatYai Th	[illegible]	2 months ago
6178	Magic Museum HatYai Th	[illegible]	2 months ago
6179	Magic Museum HatYai Th	[illegible]	2 months ago
6180	Magic Museum HatYai Th	[illegible]	2 months ago
6181	Magic Museum HatYai Th	[illegible]	2 months ago
6182	Magic Museum HatYai Th	[illegible]	2 months ago
6183	Magic Museum HatYai Th	[illegible]	3 months ago
6184	Magic Museum HatYai Th	[illegible]	3 months ago
6185	Magic Museum HatYai Th	[illegible]	3 months ago
6186	Magic Museum HatYai Th	[illegible]	3 months ago
6187	Magic Museum HatYai Th	[illegible]	3 months ago
6188	Magic Museum HatYai Th	[illegible]	4 months ago
6189	Magic Museum HatYai Th	[illegible]	4 months ago

Figure 1. Example of raw data from Google Map

3. Data Preparation

Several data preprocessing operations were conducted to ensure quality and consistency. The data was first clean by removing duplicates, and incomplete entries. Then, a user-based session grouping was created by clustering reviews of the same user representing a single trip or itinerary. Each session was treated as a transaction, and the visited attractions were treated as items within the transaction. The third process was data filtering, when transactions with at least two visiting attractions were retained to ensure meaningful rule extraction. Finally, the data was encoded by applying a transaction format to convert data to true or false for frequent pattern mining as presented in Figure 2.

TransactionID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
1	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
2	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE
3	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE
4	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
5	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE
6	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
7	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
8	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
9	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
10	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
11	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
12	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
13	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
14	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
15	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Figure 2. Example of data after preparation

4. ARM Analysis

To discover meaningful patterns of co-visited tourist attractions, this study employed the FP-Growth algorithm in conjunction with an ARM process. The ARM analysis used Rapidminer software, which is a visual data mining workflow, as illustrated in Figure 3. The first operator loaded the preprocessed dataset where each tourist session was already formatted as a list of visiting attractions. The second operator was a nominal to binominal converter, since FP-Growth required binary (0/1) data representation for item presence. This operator transformed nominal values into a binominal format. The core frequent pattern mining component was the FP-Growth operator, which constructed a compact FP-Tree and extracted all itemsets that met the minimum support threshold at 0.10.

The final operator generated association rules from the frequent itemsets discovered by the FP-Growth operator. Each rule was evaluated using standard metrics such as support, confidence, lift, and conviction.

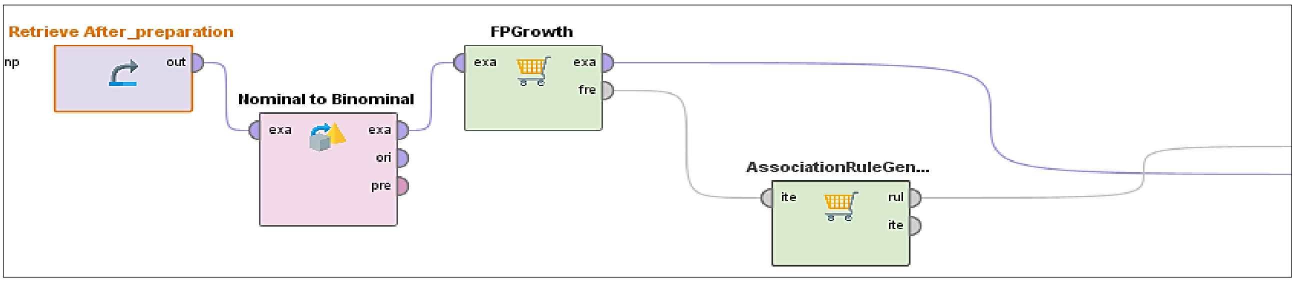


Figure 3. ARM analysis using Rapidminer

5. Evaluation

The performance of the association rules was evaluated using standard metrics such as support, confidence, lift, and conviction (Brin et al., 1997; Han et al., 2011; Liu et al., 2024; Kundakci & Nas, 2025).

Support measures how frequently the combination of antecedent and consequent occurs in the dataset. It indicates the popularity or generality of the rule. The support value is calculated as Equation (a), where A and B are itemsets, and N is the total number of transactions (Han et al., 2011; Kundakci & Nas, 2025).

$$\text{Support } (A \Rightarrow B) = \text{Count } (A \cap B) / N \qquad \text{(a)}$$

Confidence indicates the reliability of the inference made by the rule. It measures the conditional probability that the consequent appears in a transaction, given that the antecedent is already present. The confidence value is calculated as Equation (b), and the threshold should be greater than 0.5 (Han et al., 2011; Kundakci & Nas, 2025).

$$\text{Confidence } (A \Rightarrow B) = \text{Count } (A \cap B) / \text{Count } (A) \qquad \text{(b)}$$

Lift assesses the strength of the rule in comparison to random co-occurrence. It measures how much more likely it is that consequent B will appear when antecedent A is present, compared to if A and B were statistically independent. The lift value calculates as Equation (c), the threshold is greater than 1.0, indicating positive correlation (Han et al., 2011; Liu et al., 2024; Kundakci & Nas, 2025).

$$\text{Lift } (A \Rightarrow B) = \text{Confidence } (A \Rightarrow B) / \text{Support } (B) \qquad \text{(c)}$$

Conviction indicates the degree to which the presence of the antecedent implies the absence of the consequence. It compares the expected frequency of A occurring without B (assuming independence) to the observed frequency.

The lift value is calculated as Equation (d). The threshold is greater than 1.0, indicating the rule has predictive value (Brin et al., 1997).

$$\text{Conviction } (A \Rightarrow B) = (1 - \text{Support } (B)) / (1 - \text{Confidence } (A \Rightarrow B)) \qquad \text{(d)}$$

6. Deployment

The final output includes a set of frequently co-visited tourist locations that form the basis for route development. These insights can be embedded into smart tourism applications, route planning tools, or local tourism authority dashboards to assist in decision-making and enhance tourist experiences.

RESULTS

The ARM process using the FP-Growth algorithm was executed in RapidMiner and then produced a set of frequent co-visit patterns among tourist attractions in Songkhla City, Thailand. These rules reflected the behavioral tendencies of tourists who often visit multiple attractions on the same trip. A total of meaningful rules was selected based on thresholds defined during the evaluation phase (Support > 0.10, Confidence > 0.5, Lift > 1.0, conviction>1.0), a total of 13 valid association rules were generated as summarized in Table 3.

Table 3. ARM results

Rule ID	Association Rule	Support	Confidence	Lift	Conviction
10	{3} => {4}	0.143	0.994	1.076	11.820
11	{3} => {42}	0.143	0.993	1.086	13.165
2	{3} => {4, 42}	0.142	0.987	1.169	12.012
1	{11} => {4, 42}	0.116	0.966	1.144	4.622
3	{11} => {42}	0.119	0.987	1.079	6.586
8	{4, 3} => {42}	0.142	0.993	1.086	13.080
4	{4, 11} => {42}	0.116	0.989	1.082	8.040
9	{42, 3} => {4}	0.142	0.993	1.076	11.743
7	{43, 3} => {4}	0.119	0.992	1.075	9.829
13	{43, 3} => {42}	0.120	0.997	1.090	32.842
5	{43, 3} => {4, 42}	0.119	0.990	1.172	14.983
12	{4, 43, 3} => {42}	0.119	0.997	1.090	32.587
6	{42, 43, 3} => {4}	0.119	0.992	1.074	9.803

The rules revealed patterns of co-occurrence among itemsets that are indicative of common tourist preferences or behavioral tendencies. The range of support values for the rules between 0.116 and 0.143 indicated that the rules were based on frequently occurring combinations of items. The confidence levels were consistently high, ranging from 0.966 to 0.997, suggesting the strong predictive power of the antecedent for the consequent in each rule.

All rules exhibited lift values greater than 1.0, implying positive correlations between the items, and conviction values ranging from 4.622 to 32.842, further confirming the reliability and interestingness of the discovered rules.

Figure 4 shows a mapping graph of 13 rules based on the ARM results.

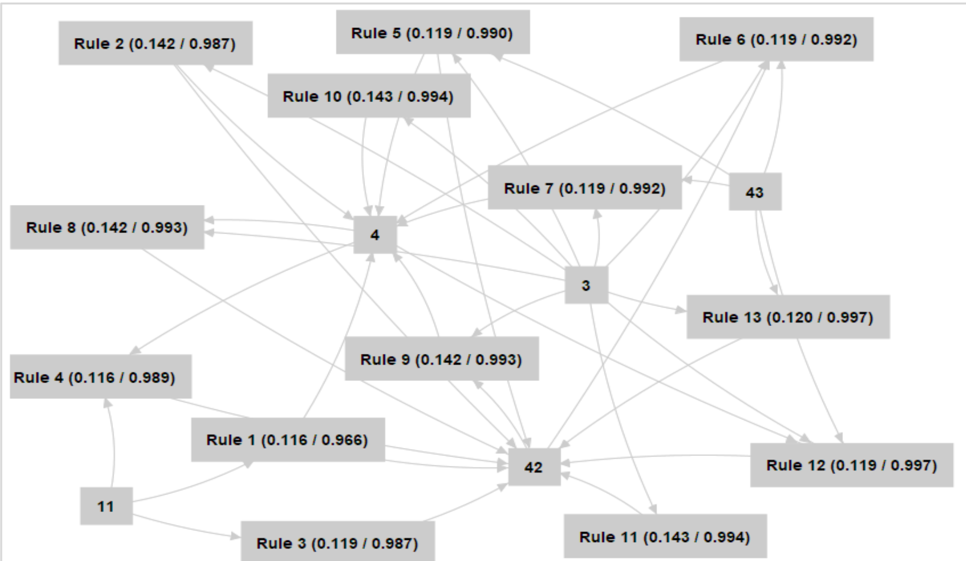


Figure 4. Mapping graph based on ARM results

Notably, these findings provided the four critical insights into visitor behavior, showing how preferences clustered around specific items.

First, Rule#10: {3} => {4} with a support score 0.143 and a confidence score of 0.994, suggested that tourists who showed interest in item 3 were highly likely (99.4%) to also be associated with item 4. The total distance of the generated recommended route was around 11 kilometers as shown in Figure 5(a).

Second, Rule#2: {3} => {4, 42} showed a slightly lower confidence (0.966) but a higher lift value (1.169), indicating that the combination {4, 42} had a stronger positive association with {3}. The total distance of the generated recommended route was around 71 kilometers, as shown in Figure 5(b).

Third, Rule#13: {43, 3} => {42} yielded the highest conviction (32.842) and confidence (0.997) scores among all the rules, highlighting a strong implication that tourists interested in both item 43 and item 3 are very likely to also be associated with item 42. The total distance of generated recommended route was around 82 kilometers, as shown in Figure 5(c).

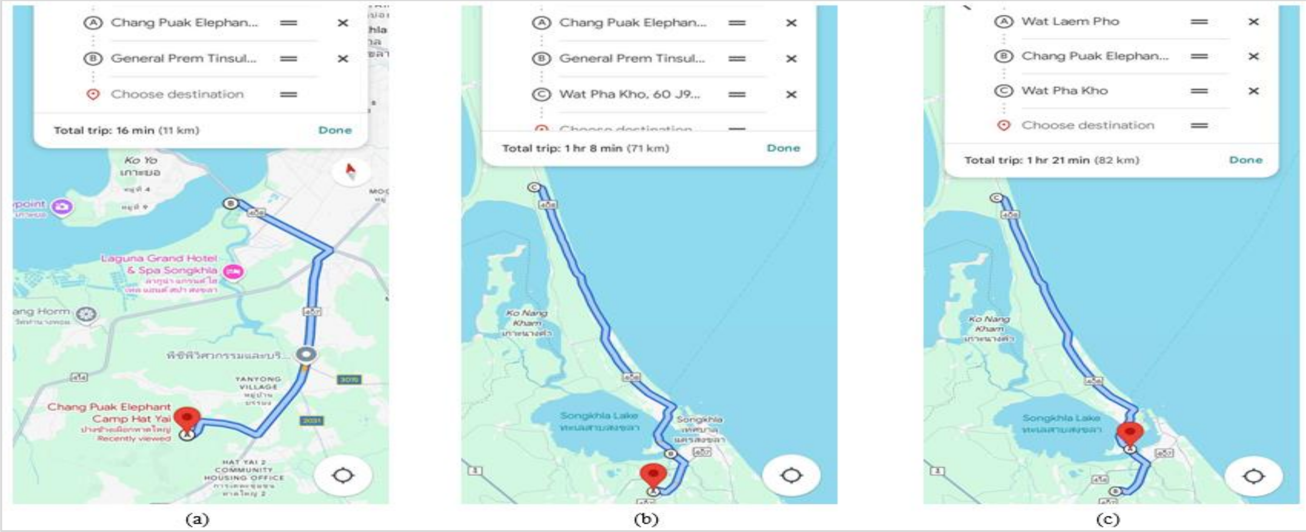


Figure 5. Recommended routes based on ARM (a) Rule#10 (b) Rule#2 (c) Rule#13

Finally, Rules with longer antecedents such as Rule#5 and Rule#12 ({43, 3} => {4, 42} and {4, 43, 3} => {42}) also demonstrated strong associations, with lift values over 1.17, confirming multi-dimensional item relationships. The total route distances associated with Rule#12 and Rule#5 are around 88 and 86 kilometers as shown in Figure 6 (a) and (b).

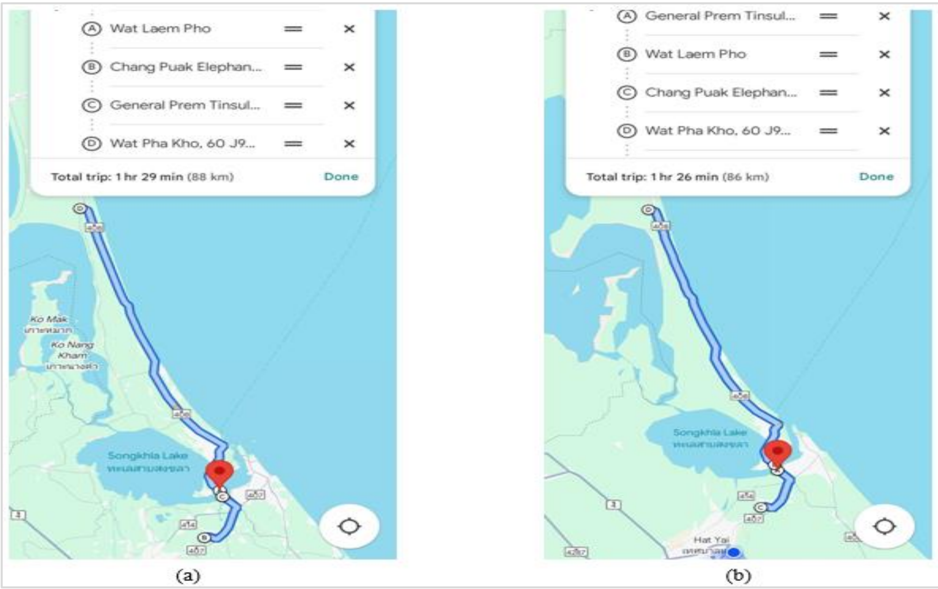


Figure 6. Recommend Routes based on ARM (a) Rule#5 (b) Rule#12

DISCUSSION

The results of this study provide valuable insights into the co-visitation patterns of tourists in Songkhla City, Thailand, as revealed through user-generated content on Google Maps. By applying the FP-Growth algorithm to mine association rules from 19,243 reviews, we successfully uncovered meaningful relationships among 45 top-rated tourist attractions. These patterns offer a foundation for data-driven route recommendations that align with actual tourist behaviors. In addition, the findings also contribute to both theoretical and practical implications.

The theoretical implications of this research contribute to the growing body of knowledge on smart tourism analytics by demonstrating how ARM can be used to infer latent behavior patterns from Google Maps data. It supports the notion that behavioral data extracted from digital footprints can be a more reliable indicator of tourist behavior patterns than traditional survey-based itinerary studies (Chang, 2025). In addition, this study demonstrates that the use of multiple evaluation metrics such as support, confidence, lift, and conviction allow for a nuanced assessment of rule quality beyond simple frequency (Liu et al., 2024; Sudirman et al., 2021). In particular high confidence and conviction values indicate strong directional relationships, while lift helps to ensure statistical significance and reduce spurious associations. Moreover, the results indicate that the use of the FP-Growth algorithm significantly improves the efficiency and scalability of rule generation, making it feasible to process large volumes of review data, as Duan (2025) suggestions.

For tourism stakeholders, this study presents several key practical implications, particularly for route design, service integration, and infrastructure planning, based on the association rules. First, with regard to data-driven tour packaging, despite the original dataset containing 45 tourist attractions, only 5 locations were involved in all 13 meaningful association rules, namely Chang Puak Elephant Camp Hat Yai (3), General Prem Tinsulanonda Historical Park Songkhla (4), Khlong Hae Floating Market (11), Wat Pha Kho (42), and Wat Laem Pho (43). This insight suggests that these five attractions exhibit strong co-occurrence in tourist visit patterns. Therefore, travel agencies can capitalize on this finding by designing and offering thematic tour packages or day-trip itineraries that integrate these highly associated attractions, increasing the likelihood of tourist satisfaction and operational efficiency. Second, with regard to infrastructure planning, the recurrence of specific attraction combinations also offers valuable insights for local government and tourism planners. Therefore, the five frequently co-visited attractions could be prioritized in signage development, shuttle routes, and tourist facility enhancements to improve both navigation and the overall tourist experience. Data-driven planning could support a more efficient allocation of resources, particularly in areas with high associative value.

Third, with regard to conscious route planning for independent tourists, the five association-based routes visualized in Figures 5 and 6 highlight that the total travel distance across all routes was never more than 88 kilometers, regardless of whether tourists visited 2, 3 or 4 attractions. These findings are consistent with the suggested maximum recommended route distance (100 kilometers) and have direct implications for the design of self-guided travel experiences on mobile apps or smart tourism platforms. Developers should be distance-aware, ensuring that independent tourists can complete the itinerary.

Finally, beyond frequency-based route planning, an important contrast emerges when comparing results from the data understanding phase. Only three of the top ten attractions by visiting frequency appeared in the derived association rules. These locations were Chang Puak Elephant Camp Hat Yai, Khlong Hae Floating Market, and Wat Laem Pho.

This discrepancy highlights a limitation in traditional frequency-based approaches to itinerary design. ARM provides a more behaviorally grounded basis for route recommendations, reflecting actual co-visitation patterns rather than isolated popularity metrics. As such, incorporating pattern-based insights into tour planning can better align offerings with real tourist behavior. These findings confirm the proposal of Arreeras et al. (2019).

While ARM provides valuable insights into tourist behavior, several limitations must be acknowledged. The ARM technique itself focuses solely on co-occurrence patterns and does not account for temporal, spatial, or causal relationships, which might be critical in understanding the flow and preferences of tourist movements. Additionally, external factors such as seasonality, weather conditions, transportation accessibility, and cultural events are not considered in the analysis. Another limitation lies in the temporal validity of the association rules discovered in this study. In practice, patterns of tourist behavior are not static and may evolve over time due to various factors such as the emergence of new attractions, changes in tourist preferences, or the decline in popularity of certain destinations.

Consequently, the association rules derived from the current dataset may lose relevance in the future. Nonetheless, the methodological and analytical approach presented in this study remains valuable. Travel planners, tourism authorities, and platform developers can leverage this methodology as a flexible decision-support tool, allowing for the periodic re-analysis of updated tourist data to ensure that route recommendations and tourism strategies align with evolving tourist behaviors.

CONCLUSION

In summary, this study demonstrates the potential of using web-mined data and association rule mining to support smart tourism development, offering a scalable, data-driven, and user-centric approach to route recommendation and destination planning. The application of association rule mining to tourist route development, particularly using the FP-Growth algorithm, represents a promising and rapidly evolving research area. The utilization of Google Maps review data provides unprecedented opportunities to understand tourist preferences and behavior patterns on a scale.

However, successful implementation requires careful consideration of data quality, computational efficiency, and the integration of multiple methodological approaches. Future research directions should focus on addressing current limitations while exploring innovative applications that combine traditional data mining techniques with emerging technologies. For example, augmenting rules with sentiment analysis can help prioritize not only frequently co-visited attractions, but also those associated with positive experiences.

Author Contributions: Conceptualization, T.M. and Nu.T. and Na.T.; methodology, T.M. and Nu.T. and Na.T.; software, Nu.T. and T.M.; validation, Nu.T. and T.M.; formal analysis, Nu.T. and T.M.; investigation, Nu.T.; data curation, Nu.T. and T.M.; writing - original draft preparation, Nu.T. and T.M.; writing - review and editing, Nu.T. and T.M.; visualization, Nu.T., and T.M.; supervision, Na.T. and Nu.T.; project administration, Nu.T. and T.M. All authors have read and agreed to the published version of the manuscript.

Funding: Prince of Songkla University, Hatyai Campus and Graduate School (College of Digital Science), Thailand.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study may be obtained on request from the corresponding author.

Acknowledgements: The work was supported by Graduate School (College of Digital Science), and Faculty of Science, Prince of Songkla University, Thailand. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved this paper.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Amangeldi, A., Seidualin, D., Mukanov, A., Mukatova, R., Bilisbekov, T., & Sagatbayev, Y. (2025). Innovative technologies to improve tourism experience in national parks: development of GIS maps and their value, *Geojournal of Tourism and Geosites,* 58(1), 210-224. https://doi.org/10.30892/gtg.58118-1403

Arreeras, T., Arimura, M., Asada, T., & Arreeras, S. (2019). Association rule mining tourist-attractive destinations for the sustainable development of a large tourism area in Hokkaido using Wi-Fi tracking data, *Sustainability,* 11(14), 3967, https://doi.org/10.3390/su11143967

Brin, S., Motwani, R., Ullman, J. D., & Tsur, S. (1997). Dynamic itemset counting and implication rules for market basket data. *Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data*, 255–264. https://doi.org/10.1145/253260.253325

Chandiramani, K. D., Maheswari, N., & Sivagami, M. (2021). Insights of gender based tourists preferences by constructing travel diary using social media data, *Geo Journal of Tourism and Geosites,* 38(4), 1098-1106. https://doi.org/10.30892/gtg.38415-749

Chang, Y. (2025). Weighted association rule algorithm application research in cultural tourism recommendation. *GeoJournal,* 90(2), 1-15. https://doi.org/10.1007/s10708-025-11293-w

Czyz, M., & Javed, M. (2025). Revolutionizing travel: The role of smart tourism technologies in enhancing tourist satisfaction and shaping sustainable destination images: Insights from Istanbul. *Geojournal of Tourism and Geosites,* 58(1), 446-455. https://doi.org/10.30892/gtg.58141-1426

Duan, H. (2025). Personalized self-guided tour strategy by integrating random forest preference model and attractions association model. *Informatica*, 49(5), 1-17. https://doi.org/10.31449/inf.v49i5.6002

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.

Hassan, B. A., Hassan, A. A., Lu, J., Ahmed, A. M., & Rashid, T. A. (2025). A proposed hybrid recommender system for tourism industry in Iraq using evolutionary Apriori and K-means algorithms, *arXiv preprint,* 1-20. https://doi.org/10.48550/arXiv.2504.08767

Jiang, R., & Dai, B. (2024). Cultural tourism attraction recommendation model based on optimized weighted association rule algorithm, *Systems and Soft Computing,* 6, 200094. https://doi.org/10.1016/j.sasc.2024.200094

Jie, M., Xusheng, G., Mingyu, Z., & Rui, S. (2025). Correlation construction and analysis of segment network based on improved FP-Growth algorithm, *Preprints.org,* 1-14. https://doi.org/10.20944/preprints202506.0911.v

Karaş, Y. K., Atay, L., Sevinç, H. K., & Anday, D. (2021). *Investigation of GIS based mobile route planning-navigation applications for tourism activities.* Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Virtual Safranbolu, Turkey. https://doi.org/10.5194/isprs-archives-XLVI-4-W5-2021-321-2021

Kundakci, B., & Nas, S. (2025). Evaluation of the anchorage area selection by using association rule mining. *Case Studies on Transport Policy*, 101544. https://doi.org/10.1016/j.cstp.2025.101544

Li, J. (2024). Tourist attraction recommendation model based on RFPAP-NNPAP algorithm. *International Journal of Advanced Computer Science & Applications,* 15(5), 608-621. https://doi.org/10.14569/IJACSA.2024.0150561

Liu, P., Guo, Y., Liu, P., Ding, H., Cao, J., Zhou, J., & Feng, Z. (2024). What can we learn from the AV crashes? –An association rule analysis for identifying the contributing risky factors. *Accident Analysis & Prevention,* 199(2024), 107492. https://doi.org/10.1016/j.aap.2024.107492

Lou, C., Yang, X., & Jin, X. (2023). *Recommendation of tourist attractions and routes based on social network data mining.* Proceedings of the IEEE International Conference on Electrical, Automation and Computer Engineering (ICEACE), Changchun, China. https://doi.org/10.1109/ICEACE60673.2023.10442229

Lou, N. (2022). Tourism destination recommendation based on association rule algorithm. *Mobile Information Systems,* 2022(1), 1-13. https://doi.org/10.1155/2022/9331178

Nagy, B., Gabor, M. R., Bacoș, I. B., Kabil, M., Zhu, K., & Dávid, L. D. (2023). Google and Apple mobility data as predictors for European tourism during the COVID-19 pandemic: A neural network approach. *Equilibrium. Quarterly Journal of Economics and Economic Policy,* 18(2), 419-459. https://doi.org/10.24136/eq.2023.013

Nitu, P., Coelho, J., & Madiraju, P. (2021). Improving personalized travel recommendation system with recency effects. *Big Data Mining and Analytics,* 4(3), 139-154. https://doi.org/10.26599/BDMA.2020.9020026

Ritthipakdee, A., Ryu, K. H., & Visutsak, P. (2024). Aligning cultural tourism with SDGS: A recommendation system based on frequent patterns, associations, and correlations. *Journal of Lifestyle and SDGs Review,* 4(2), 1-25. https://doi.org/10.47172/2965-730X.SDGsReview.v4.n02.pe02052

Ruko, S., Alcalde-Llergo, J. M., Yeguas-Bolívar, E., & Zingoni, A. (2024). *Enhancing tourist experience via automatic personalized route suggestions.* Proceedings of the 2024 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE). https://doi.org/10.1109/MetroXRAINE62247.2024.10796573

Seidualin, D., Mukanov, A., Agybetova, R., Mussina, K., Berdenov, Z. G., Babkenova, L., & Zhensikbayeva, N. (2024). Development of a geographical information system for optimizing tourist routes in the Ulytau National Natural Park, *GeoJournal of Tourism and Geosites,* 52(1), 351-359. https://doi.org/10.30892/gtg.52134-1211

Song, Y., & He, Y. (2023). Toward an intelligent tourism recommendation system based on artificial intelligence and IoT using Apriori algorithm, *Soft Computing,* 27(24), 19159-19177. https://doi.org/10.1007/s00500-023-09330-2

Sudirman, I. D., Nugraha, D. Y., Utama, I. D., & Tjoe, T. F. (2021). Comparing the relationship of tourist attractions by using association rule. *IEOM Society Internationa,* 24-31.

Xiong, S., & Zhang, T. (2024). Enhancing tourist loyalty through location-based service apps: Exploring the roles of digital literacy, perceived ease of use, perceived autonomy, virtual-content congruency, and tourist engagement. *PloS ONE,* 19(1), e0294244. https://doi.org/10.1371/journal.pone.0294244.t003

Xiong, Y., Liang, J., & Guan, C. (2025). Decoding public sentiment topics in google map reviews on urban infrastructure development of belt and road initiative. *Scientific Reports,* 15(1), 1-12. https://doi.org/10.1038/s41598-025-06451-6

Zeng, Z., Qin, J., & Wu, T. (2025). A knowledge graph-enhanced Hidden Markov Model for personalized travel routing: Integrating spatial and semantic data in urban environments. *Smart Cities,* 8(3), 1-24. https://doi.org/10.3390/smartcities8030075

Zhang, S., & Zhang, G. (2025). The impact of platform affordance on the representation of iconic architecture: a case study of Markthal across visual social media. *Archnet-IJAR: International Journal of Architectural Research, 19*(2), 390-410. https://doi.org/10.1108/ARCH-05-2024-0198

Zhao, Y., Wang, P., & Lai, Y. (2025). Route generation and built environment behavioral mechanisms of generation Z tourists: A case study of Macau. *Buildings,* 15(11), 1947. https://doi.org/10.3390/buildings15111947

Article history: Received: 25.07.2025 Revised: 21.01.2026 Accepted: 14.02.2026 Available online: 06.03.2026