www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 17(1), Jan-April 2026, 182-195. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2025.v16i4.803

Personalization of the tourism experience through machine learning techniques

Cristina Flores Amador[1], Ernesto Bolaños Rodríguez[2], Asdrúbal López Chau[3], Juan Carlos González Islas[4] and Martín Saúl Olivera Cruz[5]
[1,2,5] Universidad Autónoma del Estado de Hidalgo. Escuela Superior de Tizayuca. Carretera Tizayuca-Pachuca Km 2.5. 43800, Tizayuca, Estado de Hidalgo, México.
[3] Universidad Autónoma del Estado de México. Centro Universitario UAEM Zumpango. 55600, Zumpango, Estado de México, México.
[4] Universidad Autónoma del Estado de Hidalgo. Instituto de Ciencias Básicas e Ingeniería. Pachuca-Tulancingo Km. 4.5, Carboneras, 42184, El Álamo, Mineral de la Reforma, Estado de Hidalgo, México.
floresa@uaeh.edu.mx, ebolanos@uaeh.edu.mx, alchau@uaemex.mx, juan_gonzalez7024@uaeh.edu.mx, ol357429@uaeh.edu.mx

Abstract. This research focuses on analysing the experiences of visitors to Alameda Central during the Easter period in order to examine the degree of personalisation of the tourist experience. To this end, the study adopts a quantitative research paradigm and a non-experimental, cross-sectional design, and applies machine learning (ML) clustering algorithms. The main conclusions suggest that ML can function as a clustering tool through which visitors are grouped according to behavioural patterns, thereby potentially informing decision-making processes.

Keywords: Machine Learning, Clustering Algorithms, Tourism Personalization, Alameda Central.

Article Info
Received Jun 26, 2025
Accepted Oct 11, 2025

1 Introduction

Globalisation and technological advancements have transformed the tourism industry, enabling millions of people to explore new destinations each year. However, as competition intensifies and traveller expectations evolve, the industry faces the challenge of delivering unique, personalised experiences tailored to individual needs (Zurita et al., 2024).

Digital management has increasingly influenced tourism-related applications, enhancing their capacity to address consumer preferences and expectations (Jayawardena et al., 2023; Kim et al., 2024).

Mexico, renowned for its cultural, geographic, and gastronomic diversity, is widely recognised as a leading global tourist destination. Within this context, artificial intelligence (AI) is progressively reshaping the design and delivery of tourism experiences by adapting them to individual visitor preferences, which may improve satisfaction and overall enjoyment (Días & Silva, 2020).

In major tourist areas of Mexico City, AI is increasingly being used to support the personalisation of tourist experiences, with the aim of enhancing the quality of foreign visitors' stays (Fernández, 2019). In Alameda Central in particular, AI appears to play a relevant role in tailoring visitor experiences. AI-powered systems integrated into mobile devices or information kiosks offer real-time recommendations on events, activities, nearby restaurants, and cultural exhibitions aligned with tourists' interests. At the same time, AI-driven data collection allows for the analysis of visitor behaviour and preferences, which can support the ongoing adjustment of tourism offerings (Fernández, 2019).

Previous research indicates that, by applying machine learning (ML) techniques and collecting data from tourists through different technologies, tourist behaviour can be identified and, to some extent, anticipated. Such insights may facilitate demand forecasting, contribute to the enhancement of tourism services, and support more efficient resource management (Nazar et al., 2021).

Although ML offers considerable opportunities for personalised tourism experiences, it also raises ethical and operational challenges. Concerns related to data privacy, algorithmic fairness, and transparency require careful consideration to ensure responsible AI implementation. Accordingly, tourism organisations are expected to adopt ethical and accountable practices when developing and deploying AI-based solutions (Forero-Corba & Negre, 2024).

The main objective of this study is to explore the application of AI in tourism through a survey conducted with tourists visiting Alameda Central. The collected data were analysed using unsupervised ML techniques to identify groups of tourists with similar behavioural patterns. Two tourism experts subsequently examined these clusters, generating insights that highlight the potential of AI for enhancing tourism personalisation.

The research questions addressed in this study are:
1. Why is tourism personalization important?
2. How can tourism personalization be achieved using ML techniques?

The integration of AI and ML in tourism presents a transformative opportunity to enhance visitor experiences through personalization and data-driven decision-making. As these technologies continue to evolve, they offer significant potential to improve customer satisfaction, optimize resource allocation, and foster sustainable tourism development. However, their implementation must be approached with caution, ensuring ethical considerations and data privacy protections are upheld.

2 Machine learning for personalization of the tourism

Tourists often share interests or exhibit common characteristics, such as preferences for visiting specific places or preferred ways of receiving travel recommendations. AI can assist in identifying groups of customers or users with similar characteristics. In the context of machine learning, user data are commonly referred to as objects or instances. To address this objective, unsupervised ML methods, known as clustering algorithms, are applied.

Several clustering methods are commonly used. Among the most widely adopted is K-means, which typically assumes that the distribution of each group of similar objects has a spherical shape. Hierarchical clustering allows groups to be visualised through a dendrogram, facilitating the interpretation of cluster relationships. Density-based approaches, such as DBSCAN, are able to identify clusters with complex geometries. Model-based methods, including Gaussian Mixture Models (GMMs), generally assume that the data represent a mixture of Gaussian distributions and are commonly applied when large volumes of data follow this probabilistic structure. Spectral clustering relies on the eigenvalues of similarity matrices to perform dimensionality reduction prior to applying a conventional clustering technique, such as K-means (Ezugwu et al., 2022).

The K-Means algorithm clusters data into K clusters based on minimizing intra-cluster variance. Its goal is to assign each data point to the nearest centroid through an iterative optimization approach. This algorithm aims to partition these points into K clusters $C = \{C_1, C_2, \dots, C_K,\}$, minimizing the sum of squared errors within each cluster, which is represented mathematically as shown in the following equation.

$$J = \sum_{i=1}^{K} \sum_{x_j \in C_i} \left\| x_j - \mu_i \right\|^2$$

DBSCAN is a density-based clustering algorithm that identifies clusters as regions of high point density separated by sparse areas. Unlike K-Means, DBSCAN does not require the number of clusters to be pre-specified and is robust to noise and outliers. This method clusters the data based on two hyperparameters:

- ε: The radius within which points are considered neighbors.
- The ε -neighborhood of a point x_i is defined as

$$N_\varepsilon(x_i) = \{x_j \in X | d(x_i, x_j) \leq \varepsilon\}$$

where $d(x_i, x_j)$ is the Euclidean distance.
minPts : The minimum number of points required to form a dense region

Gaussian Mixture Models (GMMs) are a probabilistic clustering technique that models the data as a mixture of multiple Gaussian distributions. This approach provides a soft clustering framework, in which each data point is assigned a probability of belonging to each cluster. As a result, GMMs are often considered suitable for representing complex, overlapping, and non-spherical cluster structures.

A GMM typically assumes that the observed data are generated from a mixture of K Gaussian distributions, each characterised by its own mean vector and covariance matrix. Model parameters are estimated using the Expectation–Maximisation (EM) algorithm, which iteratively refines the parameters of the Gaussian components. Accordingly, the probability density function is expressed as the weighted sum of K Gaussian distributions, as shown in the following equation.

$$p(x) = \sum_{i=1}^{K} \pi_i \, N(x|\mu_i, \Sigma_i)$$

where is π_i the mixing coefficient for cluster i, satisfying $\sum_{i=1}^{K} \pi_i = 1$ and $0 \leq \pi_i \ \leq 1$; $N(x|\mu_i, \Sigma_i)$ is the Gaussian probability density function. The EM algorithm maximizes the likelihood function $L(\pi_i, \mu_i, \Sigma_i)$, defined as shown below.

$$L(\pi_i, \mu_i, \Sigma_i) = \prod_{j=1}^{n} \sum_{i=1}^{K} \pi_j N(x_i|\mu_i, \Sigma_i)$$

The EM algorithm iterates through two steps (the expectation step, and the maximization step) until a convergence criterion is satisfied.

Spectral Clustering can detect arbitrarily shaped clusters by leveraging the connectivity structure of the data, that is represented as an undirected graph G=(V,E), where V is a set containing the vertices, and E is a set with the edges. The similarity matrix W is computed as follows.

$$W_{ij} = exp(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}$$

Using W, the diagonal matrix D called degree matrix is built using the following equation.

$$D = \sum\nolimits_{j=1}^{n} W_{ij}$$

The normalized, unnormalized or random walk graph Laplacian is computed from the matrices D and L = D-W, and then eigenvalues and eigenvectors are used to reduce dimensionality of data. Finally, a clustering algorithm, such as K-means, is used to cluster the data in this reduced space.

Each of the above methods finds different clusters, due to the criteria used to measure the similarity between objects and the strategy used to identify the clusters. In addition, clustering algorithms can identify different groups of objects in each run, depending on the parameter values used at the beginning of their execution.

Therefore, it is necessary to objectively evaluate the quality of the clusters found, to determine the best grouping of instances (Ulu & Türkan, 2024). The most popular evaluation metrics for clustering are the following.

- Silhouette Score. It measures how similar the instances within the same group are compared to those in other groups. The value of this metric ranges from -1 to 1, where values close to 1 indicate well-separated and well-defined groups, values close to -1 indicate that the points are poorly clustered.
- Davies-Bouldin index. It is the average of the maximum ratio of the sum of the dispersion within each group and the distance between the centroids of the groups. Lower values of this index indicate more compact and better separated groups.
- Dunn Index. This is the ratio of the minimum distance between clusters to the maximum distance within clusters. Higher values indicate better separated and more compact clusters.
- Explained Variance Index. Measures the proportion of the total variance that is explained by the identified clusters. Higher values indicate a greater ability of the clusters to explain the variance in the data.

To fit parameters in machine learning algorithms, including clustering algorithms, the grid-search technique is commonly

$$n = \frac{Z_\alpha^2 \cdot N \cdot p \cdot q}{i^2(N - 1) + Z_\alpha^2 \cdot p \cdot q} \quad (1)$$

Where:
n = Sample
N = Size of the population or universe. Having 363,287 national tourists in Mexico City (Secretary of Tourism, 2022).
i = Unforeseen or uncontrolled error, takes values from 1 to 10% p= Probability of success
q = Probability of failure
Z = Gaussian Normal Curve statistic, at 95% confidence and 5% error. Z = 1.96 (Lind et al., 2019).

The sample consisted of 80 tourists, to whom the data collection instrument (designed questionnaire) was applied.

Within the research, simple random probability sampling without replacement is taken into account, which involves the collection of data from a sample at a single point in time (Hernández-Sampieri & Mendoza, 2018).

Within the research, simple random probability sampling without replacement is taken into account, which involves the collection of data from a sample at a single point in time (Hernández-Sampieri & Mendoza, 2018).

3.1 Data collection techniques and instruments

In developing the questionnaire as the data collection instrument, the studies by Fernández (2019) and Matovelle and Salas (2018), which include relevant items related to experience and personalisation in the tourism context, are used as reference points. The questionnaire consists of a total of 24 questions, together with a section for final comments, and was administered to 80 respondents. Each question addresses specific dimensions of inquiry.

Several key dimensions are incorporated into the design of the questionnaire. First, the immersive dimension considers the tools used by individuals when booking a trip. Second, the social dimension encompasses aspects such as marital status, travel motivations, travel companions, gender identity, occupation, and nationality, and also examines tourists' perceptions during their

Number of clusters	Mode Q1	Mode Q2	Mode Q3	Mode Q4	Mode Q5	Mode Q6	Mode Q7	Mode Q8	Mode Q9	Mode Q10	Mode Q11	Mode Q12	Mode Q13	Mode Q14	Mode Q15	Mode Q16
0	3	1	2	3	1	1	1	2	3	1	1	2	3	2	1	1
1	3	1	2	1	2	3	1	1	3	1	1	2	2	2	1	1
2	3	1	1	1	2	1	4	2	2	1	1	2	2	1	1	1
3	3	1	2	1	2	3	1	2	2	1	1	2	2	2	1	1

4.2 Cluster analysis

A total of 24 variables from the CSV file were encoded using One-Hot Encoding, resulting in 88 binary attributes and 80 samples.

Parameter Tuning (Grid-Search)
The grid-search technique was applied to optimize the parameters of five clustering algorithms. The best values found are:

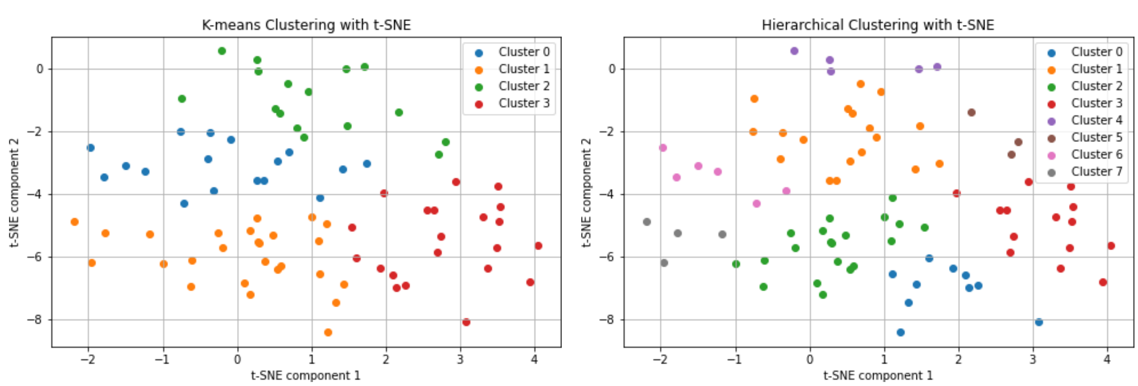
1. K-Means
    - init: KMeans++
    - n_clusters: 4
    - n_init: 35
    - Best Silhouette Score: 0.32
2. Hierarchical Clustering
    - linkage: ward
    - n_clusters: 8
    - Best Silhouette Score: 0.21
3. Spectral Clustering
    - affinity: nearest_neighbors
    - n_clusters: 4
    - n_neighbors: 5
    - Best Silhouette Score: 0.17
4. Gaussian Mixtures (GMM)
    - covariance_type: full
    - n_components: 4
    - Best Silhouette Score: 0.20
5. DBSCAN
    - eps: 0.1
    - min_samples: 3
    - Best Silhouette Score: -1.0

The best-performing algorithm was K-Means with a Silhouette Score of 0.32.
The worst performance was observed with DBSCAN, which obtained a Silhouette Score of -1.0.

4.2.1 Identified clusters

The Silhouette Score values obtained are relatively low, which may indicate that the identified clusters are not clearly defined. To support cluster visualisation, t-SNE is applied and the clusters are represented in a two-dimensional space. Figure 1 illustrates the groupings produced by each algorithm, where each point represents an object in the dataset and each colour denotes the cluster to which the object is assigned.

Among the methods considered, K-means appears to distinguish four groups more clearly when compared with the other algorithms, whereas DBSCAN assigns all observations to a single cluster. It should be noted that Figure 1 provides only a simplified representation of the data, as it is generated after applying dimensionality reduction to the original set of 88 variables. The resulting clusters are labelled using numerical identifiers ranging from zero to three.



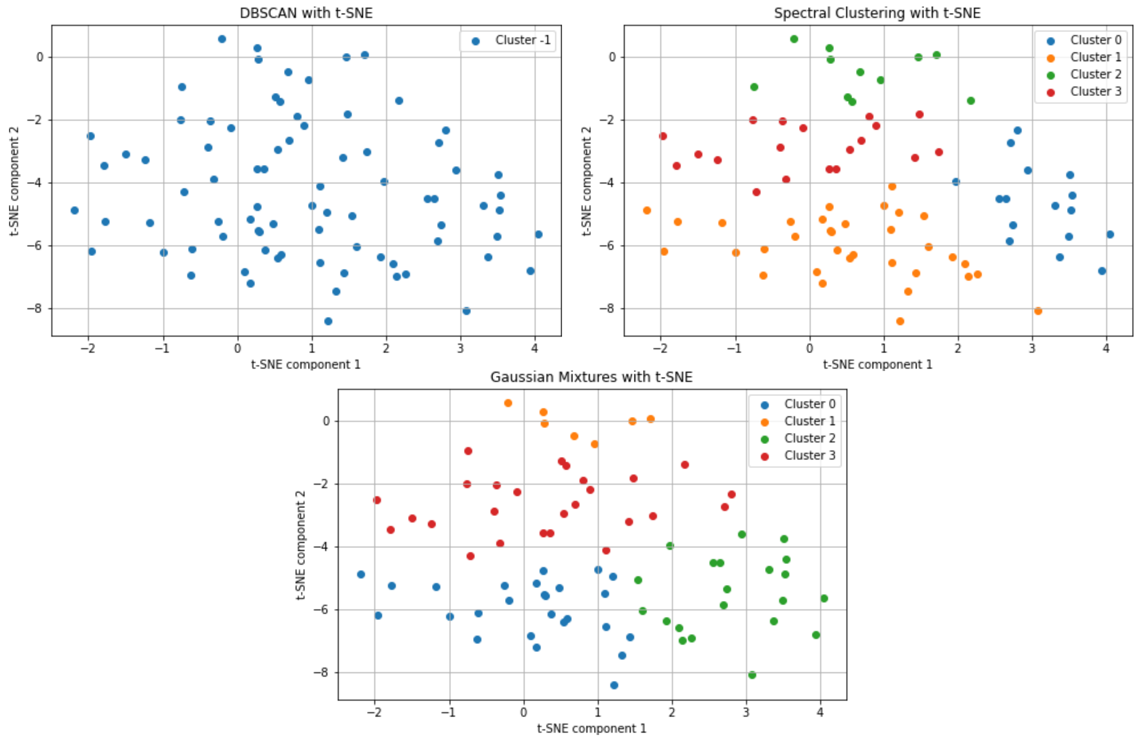


Fig. 1. Visualization of the groups identified using TSNE

The data from the respondents corresponding to each cluster were analyzed by two experts in tourism research. The findings are presented below.

4.3 Descriptive analysis

The frequencies of responses to questions 17 to 23 of users belonging to cluster 0 are shown in figure 2.

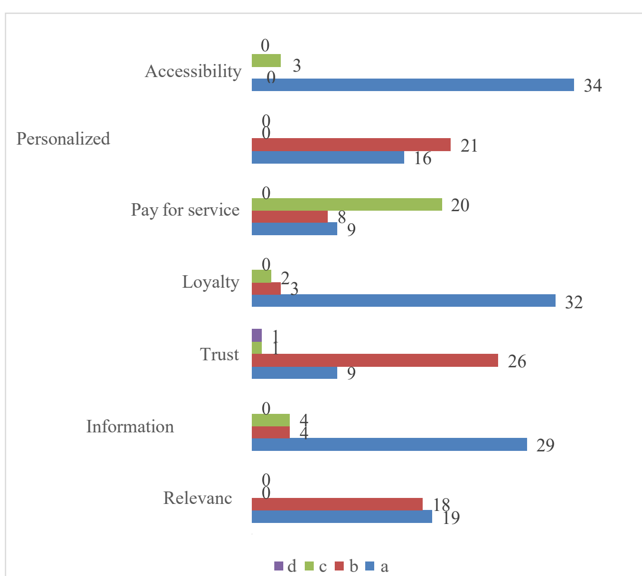


Fig. 2. Frequencies of responses for questions 17 to 23 in cluster 0

The results obtained with the application of the five ML algorithms used are presented, whose best performance is that of K-Means, with a score of 0.32, which were obtained for cluster 1 determined for questions 17 through 23, as shown in figure 3.

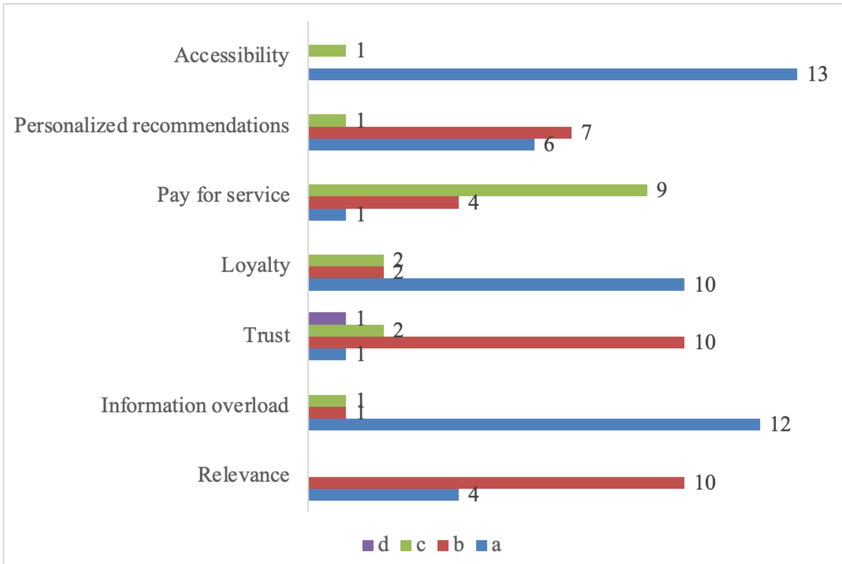


Fig. 3. Frequency of responses for questions 17 to 23 in cluster 1

The results obtained with the application of the ML algorithms used, whose best performance corresponds to K-Means with a Silhouette Score of 0.32, are shown, which are obtained for cluster 2 determined for questions 17 to 23, see figure 4.

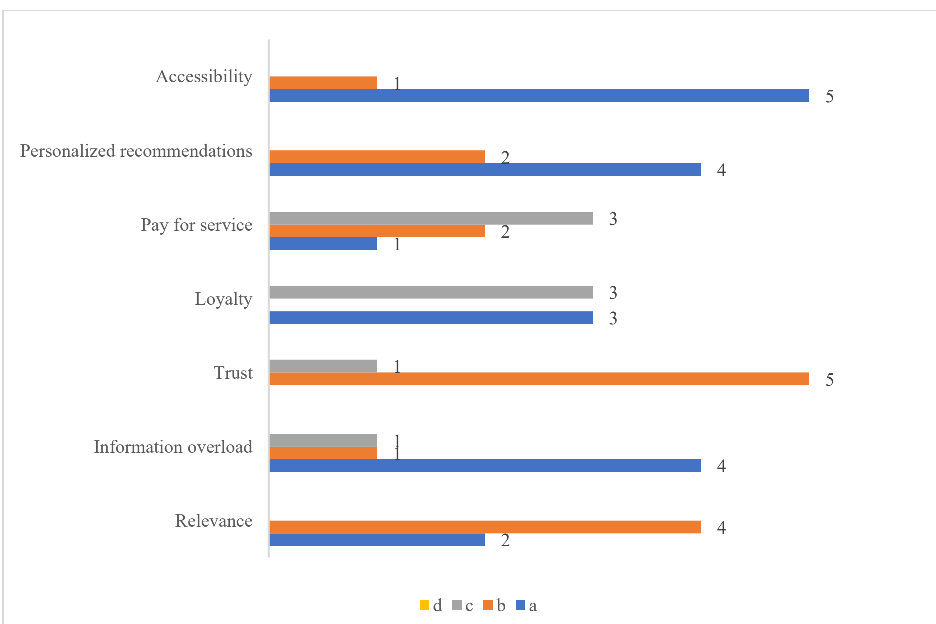


Fig. 4. Frequencies of responses for questions 17 to 23 in cluster 2

The results obtained with the application of the five ML algorithms used are reflected, whose best performance corresponds to that of K-Means, which presents a Silhouette Score of 0.32, obtained for cluster 3, determined for questions 17 to 23, as shown in figure 5.

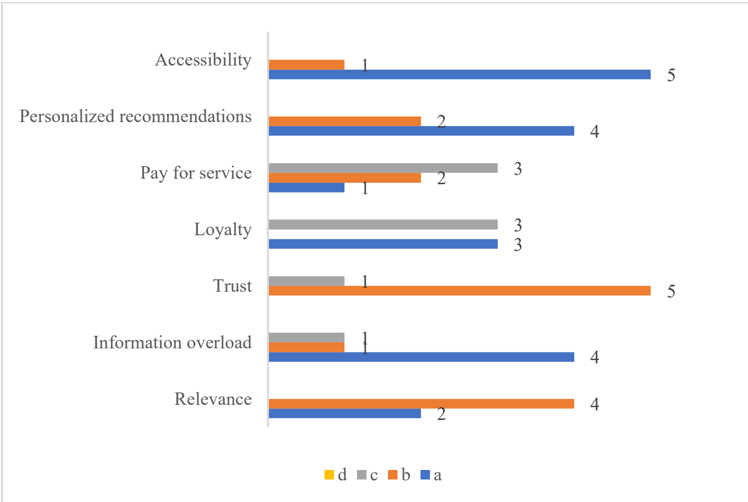


Fig. 5. Frequencies of responses for questions 17 to 23 in cluster 3

4.4 Analysis of clusters by experts

Cluster 0

In this group, the most frequently used tool for making travel reservations is mobile applications, as members are characterized by being digital natives who rely on technology to address their everyday needs, such as booking accommodations, flights, dining, and recreational or cultural activities. Predominantly Single, Male Mexican Professionals

1. Demographics & Travel Profile
    - Mostly single, male Mexican professionals.
    - Prefer traveling with friends, aiming to socialize and share experiences on social media.
2. Perception of Alameda Central
    - Good perception of Alameda Central, valuing its cultural and historical appeal, as well as its gastronomic and recreational spaces.
3. Booking Experience
    - The majority have not faced a bad booking experience.
4. Importance of Personalization & AI
    - 91.8% believe personalization is key to improving travel satisfaction.
    - Recognize AI's crucial role in personalization, offering rapid information and eliminating the need for a traditional tour guide.
    - Most have not experienced AI-driven personalization firsthand, mainly due to distrust and security concerns.
    - Generally unwilling to share personal information for AI-based recommendations, citing worries about data misuse.
5. Influence of AI on Decisions
    - Slight influence of AI on travel decisions; they still rely on smartphones to quickly solve problems and share experiences but remain cautious.
6. Planning & Key Personalization Factors
    - Prefer pre-trip personalized recommendations (price, quality, conditions) to ensure a satisfying travel plan.
    - The most valued personalization aspect is receiving recommendations of places to visit, including local attractions and conditions.
7. ML Algorithm Insights
    - Personalization based on previous experiences and preferences is highly relevant for these travelers.
    - AI helps avoid information overload by filtering key data.

- A moderate trust in AI's capacity to meet their needs exists.
- Customer loyalty may increase if AI truly understands and caters to expectations.
- Willingness to pay more is contingent upon the added value provided.
- Accurate, relevant recommendations significantly improve the trip outcome.
- Affirmative that AI can enhance accessibility and inclusion for travelers with disabilities.

Cluster 1

Like Cluster 0, the preferred tool for making travel reservations is mobile applications, thanks to their convenience and wide availability. Most individuals in this group are single, which is linked to their main motivation: seeking new experiences to fulfill personal interests and enrich their lives. They also tend to travel with friends, in line with the current trend of socializing and sharing experiences on social media. These characteristics—use of mobile apps and motivation for new experiences—are similar to Cluster 0, reinforcing the ongoing dependency on technology in travel. Predominantly Female Mexican Students

1. Demographics & Travel Profile
   - Mostly female Mexican students, enthusiastic about exploring new places connected to what they have studied in history, geography, and civics.
2. Perception of Alameda Central
   - Report an excellent perception of Alameda Central, emphasizing its cultural richness.
3. Booking Experience
   - No major bad booking experiences mentioned.
4. Importance of Personalization & AI
   - 85.7% consider personalization essential for smoother trip planning within their budget.
   - Also acknowledge AI's potential in offering immediate, guide-free personalization but have not used it extensively.
5. Trust & Willingness to Share Data
   - Haven't experienced AI-driven personalization largely due to distrust, security issues, and lack of familiarity.
   - Unwilling to share personal information, citing fears of misuse for extortion or illegal activities.
6. Influence of AI on Decisions
   - AI slightly influences travel decisions because smartphones allow them to gather quick info and post reviews, but they remain skeptical.
7. Planning & Key Personalization Factors
   - Favor receiving personalized recommendations prior to the trip to align with price and quality expectations.
   - The recommendations of places to visit remain the top personalization need, paralleling Cluster 0.
8. ML Algorithm Insights
   - Personalization is relevant but less surprising for them—suggesting they have existing knowledge and are not easily impressed.
   - Uncertain that AI can curb information overload, reflecting distrust in computational tools.
   - Low confidence in AI's ability to enhance the travel experience, seeing it as a source of resistance.
   - Unsure about AI's potential for customer loyalty due to lack of personal interaction.
   - Reluctant to pay extra for AI-based services, fearing they might not be worth the investment.
   - Divided on the importance of accurate, relevant recommendations.
   - Skeptical about AI's capacity to improve accessibility for travelers with disabilities.

Cluster 2

This group also agrees that mobile applications are the most used tool for making travel reservations, highlighting the importance of technology regardless of cluster differences. Most people in this group are single and are primarily motivated by cultural exploration, indicating a strong interest in immersing themselves in the traditions and customs of their destination. Predominantly Female European Professionals (Solo Travelers)

1. Demographics & Travel Profile
   - Mostly foreign (European) female professionals who prefer traveling alone, valuing independence and privacy.
2. Perception of Alameda Central
      - Hold a good perception, albeit with higher standards given broader cultural exposure.
3. Booking Experience
   - No significant bad booking experiences reported, consistent with Clusters 0 and 1.
4. Importance of Personalization & AI

- Unanimous agreement that personalization significantly enhances satisfaction and ensures certainty.
- 50% have not experienced AI-based personalization, citing distrust, lack of knowledge, or resistance to new tools.

5. Trust & Willingness to Share Data
    - 66.6% are willing to share data for better AI-driven recommendations.
    - A strong "yes" to AI's influence on travel decisions sets them apart from Clusters 0 and 1.
6. Planning & Key Personalization Factors
    - Similar to other clusters, they value pre-trip recommendations on activities and attractions.
    - Recommendations of places to visit also rank highest in importance.
7. ML Algorithm Insights
    - Personalization is relevant, though less novel for them due to significant travel background.
    - Uncertain about AI's ability to reduce information overload (similar to Cluster 1's view).
    - Moderate trust in AI's power to enhance tourism experiences.
    - Not fully convinced AI can foster destination loyalty due to lack of personal interaction.
    - Willing to pay more only if there is notable added value, akin to Cluster 0's stance.
    - Mixed opinions on whether AI's recommendations are entirely accurate.
    - Yes, to AI's potential for accessibility and inclusion, aligning with Cluster 0 and differing from Cluster 1's skepticism.

Cluster 3

Consistent with the other three clusters, mobile applications are the most popular method for booking trips in this group, given their accessibility and ease of use. Most members are single and focus on the pursuit of new experiences. They also show a preference for traveling individually, which slightly differs from other clusters that emphasize traveling with friends or that did not specify a particular preference. Predominantly Female Mexican Students

1. Demographics & Travel Profile
    - Mostly female Mexican students, open to new experiences and direct contact with local culture.
2. Perception of Alameda Central
    - Good perception, finding it historically and culturally appealing.
3. Booking Experience
    - No significant bad booking experiences identified, consistent across all clusters.
4. Importance of Personalization & AI
    - 95.6% prioritize personalization for better budgeting and meeting travel needs.
    - 52.1% agree AI plays a pivotal role, matching other clusters' recognition of AI's immediacy in providing information.
5. Trust & Willingness to Share Data
    - Despite recognizing AI's importance, many remain unwilling to share personal information, citing fears of misuse.
6. Influence of AI on Decisions
    - AI has a slight influence, tied to the common use of smartphones for instant problem-solving and sharing opinions.
7. Planning & Key Personalization Factors
    - Favor pre-trip personalization (like recommendations on price and quality), matching other clusters.
    - Recommendations of places to visit remain the top request, mirroring previous clusters.
8. ML Algorithm Insights
    - Similar to Cluster 0 in that personalization is highly relevant, especially when clearly matching traveler expectations.
    - AI reduces information overload by summarizing the most crucial details.
    - Show a moderate level of trust in AI's ability to deliver a successful experience.
    - Uncertain about AI's role in building loyalty to specific destinations.
    - Willingness to pay more hinges on added value provided.
    - Emphasize the importance of relevant, accurate recommendations for a positive travel outcome.
    - Positive about AI's capacity to improve accessibility for travelers with disabilities, aligning with Cluster 0.

			Cluster 2	**Cluster 3**
Demographics	Single, male Mexican professionals	Female Mexican students	Female European professionals (mostly solo travelers)	Female Mexican students
Travel Companions	Friends	Not explicitly stated (likely groups)	Alone	Not explicitly stated (group-friendly)
Booking Experience	Mostly no negative issues	Mostly no negative issues	Mostly no negative issues	Mostly no negative issues
Perception of Alameda Central	Good cultural & historical appeal	Excellent, strong cultural connection	Good, higher expectations due to global travel	Good, open to immersive experiences
Personalization Importance	91.8% find it very important	85.7% find it very important	Universally seen as essential	95.6% find it very important
Experienced AI-driven Personalization?	Mostly no (distrust, security concerns)	Mostly no (distrust, security concerns)	50% have not used AI yet	Mostly no (distrust, security concerns)
Willing to Share Data for Personalization	Unwilling	Unwilling	66.6% are willing	47.8% are unwilling
AI, Influence on Travel Decisions	Slight	Slight	Significant (greater trust in AI, solutions)	Slight
AI, Potential to Reduce Info Overload	Yes, moderate trust	Unsure (distrust in tool)	Unsure (similar to Cluster 1)	Yes, moderate trust
Confidence in AI for Enhanced Travel	Moderate	Low	Moderate	Moderate
AI Increasing Destination Loyalty	Possible if expectations met	Uncertain	Uncertain	Uncertain
Willing to Pay More for AI-based Services	Depends on added value	Unwilling (consider it an extra expense)	Depends on added value	Depends on added value
Accuracy/Relevance of AI Recommendations	Highly important	Opinions divided	Opinions divided	Highly important
AI for Accessibility & Inclusion	Affirmative	Biased/uncertain	Affirmative	Affirmative

The research questions posed at the beginning of this article are addressed below.

Why is tourism personalisation important?

Tourism personalisation is important because it enhances visitor experiences by providing recommendations and services aligned with individual preferences, thereby potentially increasing satisfaction and engagement. The findings indicate that over 85% of respondents consider personalised recommendations to be essential for supporting trip planning and overall enjoyment. Personalised services allow tourists to save time, receive relevant suggestions, and make more informed decisions regarding their travel experiences.

In addition, personalisation contributes to improved accessibility and inclusivity, particularly for travellers with specific needs, by offering customised assistance and suitable accommodations. It also may foster customer loyalty by supporting more meaningful and enjoyable travel experiences, which can encourage repeat visits and positive word-of-mouth recommendations. Nevertheless, despite these advantages, concerns related to data security and privacy appear to constrain the broader adoption of AI-driven personalisation tools.

How can tourism personalisation be achieved using ML techniques?

Tourism personalisation can be achieved through machine learning (ML) techniques, particularly clustering algorithms, which segment tourists into groups based on shared preferences and behavioural patterns. In this study, ML techniques were applied to analyse visitor data collected at Alameda Central, resulting in the identification of four primary tourist clusters using K-means, hierarchical clustering, spectral clustering, Gaussian mixture models, and DBSCAN. Among these methods, K-means demonstrated comparatively stronger performance in categorising visitors, yielding a Silhouette Score of 0.32.

The ML-based personalization process follows these key steps:

1. Data Collection & Preprocessing:
   - Surveys and behavioral data from tourists are collected.
   - Data is encoded using techniques like One-Hot Encoding to transform categorical information into a machine-readable format.
2. Clustering & Segmentation:
   - Unsupervised ML techniques (e.g., K-Means, DBSCAN) identify groups of tourists with similar behaviors and preferences.
   - Experts analyze the clusters to derive insights into different tourist profiles.
3. Personalized Recommendation Systems:
   - AI-driven models generate real-time recommendations for tourists, such as custom itineraries, dining options, and attractions aligned with their interests.
   - These recommendations can be provided via mobile apps, smart kiosks, or interactive platforms.
4. Optimization & Continuous Learning:
   - ML models refine their recommendations over time by integrating feedback from tourists, ensuring more accurate and dynamic personalization.

ML-based personalization enables smarter decision-making in tourism management, enhancing visitor experiences while optimizing resources and services for tourism providers.

5 Conclusions

The study highlights the significance of artificial intelligence (AI) and machine learning (ML) in personalising the tourism experience at Alameda Central during the Easter period. Through the application of clustering algorithms, groups of visitors with similar behavioural patterns were identified, which can support more informed decision-making in tourism management.

The findings indicate that the most effective technique for tourist segmentation was K-means, achieving a Silhouette Score of 0.32, which may be interpreted as indicating a moderately defined clustering structure. By contrast, the DBSCAN algorithm exhibited the weakest performance, with a score of −1.0, suggesting limitations in its ability to identify meaningful clusters within the analysed dataset.

The segmentation process enabled the identification of four distinct types of tourists based on their preferences and attitudes towards personalisation and AI.

Cluster 0 is primarily composed of Mexican male professionals who place high value on technology for travel planning and report trust in digital recommendations. However, they express reluctance to share personal data and show a preference for receiving recommendations prior to their trips.

Cluster 1 consists predominantly of Mexican female students with a strong interest in cultural experiences. Although they acknowledge the relevance of personalisation, they remain sceptical about AI adoption and express concerns related to data privacy and potential misuse.

Cluster 2 includes European professionals travelling alone, who demonstrate greater openness to AI-based personalisation and a willingness to share data in exchange for more accurate recommendations. They perceive personalisation as enhancing satisfaction but question AI's capacity to reduce information overload.

Cluster 3 is largely composed of Mexican female students who value direct engagement with local culture and consider personalisation beneficial for travel planning. While they are hesitant to share personal data, they recognise AI's role in supporting accessibility and inclusion.

The study provides evidence that personalisation represents an important factor in enhancing the tourist experience. More than 85% of respondents consider personalised recommendations essential for supporting travel planning and satisfaction. Nevertheless, concerns related to data security and privacy persist, which appear to limit the broader adoption of AI-driven tools.

Overall, tourists demonstrate a moderate level of trust in AI, recognising its usefulness in personalising experiences and managing information volume. However, AI's influence on travel decisions remains limited, as many tourists continue to prefer independent information-seeking before relying fully on automated systems.

Finally, the findings suggest that AI has the potential to enhance accessibility and inclusion in tourism, particularly for travellers with specific needs. While AI- and ML-driven personalisation offers valuable opportunities for the tourism sector, its effective implementation is likely to depend on addressing ethical challenges and strengthening user trust in data management practices.

References

Días, I. & da Silva, I. M. (2020). The user experience in the process of acquiring information for trip planning: The case of Kayak's Chatbot. *Tourism Studies and Perspectives,* 29 (3), 792-816. https://www.redalyc.org/articulo.oa?id=180764278016

Ezugwu, A. E., Ikotun, A. M., Oyelade, O. O., Abualigah, L., Agushaka, J. O., Eke, C. I., & Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. Engineering Applications of Artificial Intelligence, 110, 104743. https://doi.org/10.1016/j.engappai.2022.104743

Fernández, A.A. (2019). Who they are, what they think and what they value: tourism in the CDMX. *Kalpana, Research Journal,* 17, 61-82. https://www.researchgate.net/publication/361222306

Forero-Corba, W., & Negre Bennasar, F. (2024). Técnicas y aplicaciones del Machine Learning e inteligencia artificial en educación: una revisión sistemática. RIED-Revista Iberoamericana de Educación a Distancia, 27(1), 209–253. https://doi.org/10.5944/ried.27.1.37491.

Hernández-Sampieri, R. & Mendoza, C. (2018). *Research methodology. Quantitative, qualitative and mixed routes* (6a ed.). Mc Graw Hill Education: México.

Jayawardena C., Ahmad A., Valeri M., Jaharadak A.A. (2023). Antecedentes de la aceptación de la tecnología en la transformación digital en la industria hotelera. *International Journal of Hospitality Management*, 108, artículo 103350. https://doi.org/10.1016/j.ijhm.2022.103350

Kim, S., Fong, L. H. N., Choi, M., & Law, R. (2024). Editorial for the Special Issue: Impacts of Future Technology on Hospitality and Tourism. Journal of Hospitality and Tourism Research, 48(6), 947-948. https://doi.org/10.1177/10963480241262478

Lim, W. (2024). What Is Quantitative Research? An Overview and Guidelines. Australasian Marketing Journal (Amj), 33(3), 325-348. https://doi.org/10.1177/14413582241264622

Lind D. A., Marchal W. G. & Wathen. (2019). *Estadística aplicada a los negocios y la economía*. (17a ed.). Ciudad de México: McGraw-Hill Interamericana.

Matovelle, R. & Salas, E. (2018). Brand building through ATL and BTL media. *Contributions to Economics*, 16 (4), 1-9. https://dialnet.unirioja.es/servlet/articulo?codigo=9025075

Nazar Neiva, D., Costa, H., & Ruas, R. (2021). Destinos turísticos inteligentes (DTI): a visão de atores chaves do turismo em Brasília (DF, Brasil). *Revista Turismo & Desenvolvimento*, 36(1), 451-464. https://doi.org/10.34624/rtd.v1i36.9167

Secretary of Tourism. (2022), Statistics of tourism activity, available in: https://www.turismo.cdmx.gob.mx/actividad-turistica/estadisticas-de-la-actividad-turistica

Ulu, M. & Türkan, Y.S. (2024). Cluster Analysis and Comparative Study of Different Clustering Performance and Validity Indices, In: Durakbasa, N.M., Gençyılmaz, M.G. (Ed.s) *Industrial Engineering in the Industry 4.0* Era. ISPR 2023: Lecture Notes in Mechanical Engineering. Springer, Cham., pp. 33-45. https://doi.org/10.1007/978-3-031-53991-6_3

Zurita, S., Jaime, V., Herrera, S. del P., Crespo, X. & Cadena, G.N. (2024). Use of Artificial Intelligence in the Organization and Planning of Travel Quito-Ecuador. *Latin Science: Multidisciplinary Scientific Journal*, 8 (1), 228-243. https://dialnet.unirioja.es/servlet/articulo?codigo=9294383