KSÜ Mühendislik Bilimleri Dergisi, 29(3), 2026
Araştırma Makalesi

KSU J Eng Sci, 29(3), 2026
Research Article

Kahramanmaras Sutcu Imam University
Journal of Engineering Sciences

Geliş Tarihi : 15.03.2026
Kabul Tarihi : 06.05.2026

Received Date : 15.03.2026
Accepted Date : 06.05.2026

A GRAPH-BASED CLUSTERING AND LARGE LANGUAGE MODEL-ENHANCED ASPECT-BASED SENTIMENT ANALYSIS FRAMEWORK FOR MODELING PASSENGER SATISFACTION IN EUROPEAN CRUISE ITINERARIES

AVRUPA KRUVAZİYER ROTALARINDA YOLCU MEMNUNİYETİNİN MODELLENMESİNE YÖNELİK ÇİZGE TABANLI KÜMELEME VE BÜYÜK DİL MODELİ DESTEKLİ ÖZELLİK TABANLI DUYGU ANALİZİ YÖNTEMİ

Ahmet Cumhur ÖZTÜRK[1] (ORCID: 0000-0002-2677-3269)
Gözde TIĞA ÖZTÜRK[2*] (ORCID: 0000-0002-8204-1206)

[1] Aydın Adnan Menderes Üniversitesi, Veri Tabanı, Ağ Tasarımı ve Yönetimi Bölümü, Aydın, Türkiye
[2] Milli Eğitim Bakanlığı,Aydın, Türkiye

*Sorumlu Yazar / Corresponding Author: Gözde TIĞA ÖZTÜRK, gzd_tiga@hotmail.com

ABSTRACT

Cruise tourism offers travelers the opportunity to visit multiple destinations while enjoying various onboard services. It has emerged as one of the fastest growing sectors in the tourism industry. Online reviews have become an essential resource for both travelers and cruise operators, with the rapid increase in user-generated content on online platforms. This study presents a novel framework for analyzing cruise passenger satisfaction across European itineraries by integrating transitive clustering, topic modeling, and Aspect Based Sentiment Analysis (ABSA). Cruises were grouped into four major geographic regions based on the ports they visited. Eight key aspects from online reviews were identified where three of them are itinerary related. The GPT-4o mini large language model, accessed through the OpenAI API, was used for fine-grained sentence-level aspect-based sentiment analysis. Results show that cruises in the Eastern Mediterranean received the highest satisfaction scores across itinerary related aspects while those in the Western European and Atlantic Coastal scored lowest. This study highlights the effectiveness of the GPT-4o mini large language model in extracting nuanced sentiment from informal online reviews. It also identifies the most and least satisfying cruise regions in Europe based on customer feedback, while providing an aspect×region knowledge map for decision support.

Keywords: Aspect based sentiment analysis (ABSA), cruise tourism, GPT-4o mini, itinerary clustering, online cruise reviews.

ÖZET

Kruvaziyer turizmi, seyahat edenlere birden fazla destinasyonu ziyaret ederken çeşitli gemi içi hizmetlerden yararlanma olanağı sunmaktadır. Turizm endüstrisinin en hızlı büyüyen sektörlerinden biri hâline gelmiştir. Çevrimiçi platformlardaki kullanıcı kaynaklı içeriklerin artmasıyla birlikte çevrimiçi yorumlar, hem yolcular hem de kruvaziyer işletmecileri için önemli bir bilgi kaynağı hâline gelmiştir. Bu çalışma, Avrupa kruvaziyer rotalarındaki yolcu memnuniyetini analiz etmek için geçişli kümeleme, konu modelleme ve Hedef (Aspect) Bazlı Duygu Analizi (ABSA) yöntemlerini entegre eden özgün bir çerçeve sunmaktadır. Kruvaziyerler, ziyaret ettikleri limanlara göre dört ana coğrafi bölgeye ayrılmıştır. Çevrimiçi yorumlardan sekiz temel boyut belirlenmiş olup bunların üçü doğrudan seyahat rotası ile ilişkilidir. OpenAI API üzerinden erişilen GPT-4o mini büyük dil modeli, cümle düzeyinde hedef bazlı duygu analizinin gerçekleştirilmesinde kullanılmıştır. Bulgular, Doğu Akdeniz rotasındaki kruvaziyerlerin rota ile ilgili boyutlarda en yüksek memnuniyet puanlarını aldığını, Batı Avrupa ve Atlantik kıyı bölgelerindekilerin ise en düşük puanları elde ettiğini göstermektedir. Bu çalışma, GPT-4o mini büyük dil modelinin çevrimiçi kullanıcı yorumlarındaki incelikli duyguları yakalamadaki etkinliğini vurgulamaktadır. Ayrıca, müşteri

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

geri bildirimlerine dayalı olarak Avrupa'daki en memnun ve en memnuniyetsiz kruvaziyer bölgelerini belirlemekte ve karar destek amacıyla bir hedef × bölge bilgi haritası sunmaktadır.

Anahtar Kelimeler: Hedef bazlı duygu analizi (ABSA), kruvaziyer turizmi, GPT-4o mini, rota kümeleme, çevrimiçi kruvaziyer yorumları.

INTRODUCTION

Cruise tourism is one of the fastest-growing segments of the global tourism industry, with the number of cruise passengers projected to reach 37.1 million in 2025, 38.8 million in 2026, and 39.7 million by 2027 (CLIA, 2023). This rapid growth intensifies competition within the cruise industry, compelling cruise ship operators to expand their market share to increase revenues. To ensure long-term market-share growth, cruise ship operators need to sustain dynamic communication with consumers, develop marketing strategies to foster customer loyalty, and implement an effective revenue management system. These managerial decision making needs can be fulfilled by building pipelines that are able to extract, represent, and reason knowledge from large-scale, noisy text data such as online customer reviews. Online customer reviews have become a valuable source of information that provides opportunities for enhancing operational efficiency (Jiao et al., 2024), with these opportunities including accurately identifying consumer needs and preferences, offering personalized products or services, and creating customer value through a consumer-oriented approach. These reviews attract the attention not only of cruise ship operators but also of cruise travelers themselves. In particular, destination choices and booking decisions are increasingly influenced by the firsthand experiences documented in these reviews (Bahja et al., 2019; Castillo-Manzano et al., 2024). Unlike structured survey data, which are constrained by predefined questions and response options (Huang & Liu, 2023), online reviews also offer rich, unsolicited insights into passenger experiences and emerging service themes (Bryman,2016).

Sentiment Analysis is the process of identifying the emotional polarity of text by classifying content as positive, negative, or neutral, and has been widely applied in tourism research. A more fine-grained approach, Aspect-Based Sentiment Analysis (ABSA), examines sentiments tied to specific aspects or attributes of an experience. Traditional ABSA techniques based on rule-based or classical machine-learning approaches often demonstrate limited effectiveness with ungrammatical, informal, or domain-specific language features prevalent in user-generated content. Recent advances in Large Language Models (LLMs), such as GPT (Generative Pretrained Transformer), provide more robust tools for conducting ABSA, particularly in informal texts. These models, trained on vast text corpora, are capable of handling complex linguistic structures including sarcasm, mixed opinions, and informal phrasing. Unlike traditional models, LLMs require minimal feature engineering and demonstrate higher accuracy and adaptability. Studies have shown that GPT-based ABSA outperforms earlier methods in both precision and scalability (Mughal et al., 2024).

Despite the increasing reliance on online reviews, there remains a significant gap in cruise tourism research; specifically, no prior study has applied LLMs like GPT to construct ABSA on cruise reviews at the itinerary level. This study addresses this gap by integrating multiple data-mining techniques: graph-based clustering, unsupervised topic modeling, and GPT-4o-mini-based ABSA to identify regional patterns of passenger satisfaction across European cruise itineraries. Accordingly, the following research questions are adressed:

RQ1. How can European cruise itineraries be clustered based on shared port visits to define distinct regional cruise groups?

RQ2. What are the most frequently mentioned aspects in online cruise reviews, and how can they be identified?

RQ3. How does passenger sentiment toward itinerary-related aspects vary across the identified cruise regions?

To answer these research questions, the study follows five main methodological steps: (1) data collection, (2) identification of European cruise regions, (3) aspect identification, (4) aspect-based sentiment analysis, and (5) comparison of geography-related satisfaction. First, the itineraries of cruises operating in Europe and their corresponding online customer reviews were collected from c*ruiselines.com*. Second, cruises were grouped into clusters based on the ports they visited in common using a graph-based transitive clustering technique. Third, Latent

Dirichlet Allocation (LDA), an unsupervised topic modeling technique, was used to identify key aspects discussed in the reviews. As a result, eight distinct aspects were determined, three of which were identified as directly related to geographical travel experiences. Fourth, the GPT-4o mini model was used to perform aspect-based sentiment analysis within each itinerary group. Finally, geography-related satisfaction dimensions within each itinerary group were calculated and compared across itinerary groups.

The study contributes to the literature in several ways. First, it identifies four dominant cruise itinerary regions in Europe through graph-based clustering. Second, it reveals eight key customer experience aspects using topic modeling. Third, it demonstrates the effectiveness of the GPT-4o mini large language model in conducting fine-grained sentiment analysis in the cruise tourism context. Lastly, it identifies the most and least satisfying European regions based on the sentiments of itinerary-specific aspects.

LITERATURE REVIEW

Customer Satisfaction and Online Cruise Reviews

The Expectancy Disconfirmation Theory suggests that customer satisfaction depends on the alignment between perceived performance and pre-purchase expectations (Oliver, 2014). Satisfaction arises when performance meets or exceeds expectations (positive disconfirmation), whereas dissatisfaction emerges when performance falls short (negative disconfirmation). This evaluation is shaped by factors such as prior experiences, brand image, and marketing messages. Thus, managing expectations and enhancing perceived value are crucial for sustaining satisfaction and loyalty.

Online reviews have become a vital source of information in tourism, as they offer real-time insights into tourist preferences and sentiments, and they play a key role in destination decision-making and demand forecasting (Li et al., 2023). The information obtained from online platforms is widely regarded as objective, authentic, and insightful, and serves as a valuable source of implicit knowledge for both consumers and businesses (Sánchez-Franco & Rey-Tienda, 2024). Bahja et al. (2019) highlight that online reviews are among the most influential factors affecting customers' cruise vacation choices and emphasize their significance in the decision-making process. On the cruise service-provider side, online reviews are crucial as they serve as a resource for improving service quality, identifying key satisfaction drivers, and enhancing customer experience management (Jiao et al., 2024).

With the rapid development of online cruise-review platforms such as Cruise Critic and Cruise Lines, cruise passengers increasingly share their travel experiences, leading to an enormous volume of publicly available cruise reviews. Traditional analytical methods often prove inadequate for deriving meaningful, actionable insights from large-scale, unstructured datasets such as online reviews, where manual interpretation of individuals' emotions and opinions presents substantial challenges. As a solution, Natural Language Processing (NLP) techniques have emerged as powerful tools for analyzing online reviews and uncovering hidden sentiments in customer feedback, enabling businesses to gain more accurate and flexible insights compared to traditional methods and ultimately enhancing overall service quality and strategic decision-making (Al Sari et al., 2022).

According to the multi-attribute theory (Keeney, 1975), a product or service is composed of various attributes that are interpreted differently by customers based on their unique expectations and perceptions. Consequently, each attribute needs to be examined individually. In the context of cruise tourism, customers often consider factors such as dining, onboard experiences, shore excursions, and entertainment when making decisions related to their cruise experience. By analyzing the impact of these attributes on customer satisfaction, cruise companies can focus on areas that require improvement, allocate resources more efficiently to the aspects that most influence satisfaction, and differentiate themselves by excelling in the services that passengers value most.

In this study, aspect-based sentiment analysis (ABSA) is used to evaluate the influence of aspects identified through the LDA method on customer satisfaction. This approach enables the analysis of customer feedback based on specific attributes, allowing for the assessment of sentiment associated with each individual feature. Thus, the impact of the identified aspects on customer satisfaction can be measured in a more detailed and accurate manner.

Sentiment Analysis

Sentiment analysis is the process of extracting the emotion in a given text as positive, negative, or neutral. Regarding sentiment analysis on cruise reviews, Castillo-Manzano et al. (2024) analyzed online cruise reviews to uncover relationships between cruise-related features and sentiment words. Shi et al. (2025) applied sentiment analysis to

Chinese-language reviews of Royal Caribbean Cruises collected from Ctrip.com to identify and rank user requirements influencing cruise satisfaction. Ma et al. (2024) performed sentiment analysis on online cruise reviews from Cruise Critic to capture both explicit and implicit user demands related to cruise services and to assess satisfaction with Royal Caribbean International.

There are three levels of sentiment analysis; document level, sentence-level and aspect level (Lu & Zheng, 2021). Document-level sentiment analysis evaluates the overall sentiment expressed in an entire document; sentence-level analysis focuses on the sentiment of individual sentences; and aspect-level sentiment analysis targets sentiments associated with specific features or aspects within the text. While document level provides a general overview and sentence-level captures sentence-specific emotions, both can fail to detect fine-grained opinions tied to particular attributes or elements of a service or product (Wankhade et al., 2022).

Identification of aspects or attributes mentioned within documents is the fundamental step of ABSA. In the tourism literature, various methods have been used for identifying aspects, such as syntax based approaches (Yang et al., 2025), supervised machine-learning approaches (Afzaal et al., 2019) and unsupervised machine learning approaches (Chen et al., 2025). Unsupervised methods rely on algorithms such as clustering or statistical co-occurrence to discover hidden structures in data without labeled examples (in contrast to supervised learning). Syntax-based methods utilize linguistic patterns such as dependency parsing and part-of-speech tagging to extract relevant terms or phrases; because they depend on grammatical structure, they can be sensitive to informal, user-generated content like online reviews. Supervised approaches require large labeled datasets to train models for aspect identification, which is often costly and time-consuming to create. Unsupervised methods offer domain flexibility by not relying on labeled data, but they may require careful tuning to avoid noisy or overly generic aspects.

Latent Dirichlet Allocation (LDA) (Blei et al., 2003) is a widely used unsupervised probabilistic topic modeling technique for identifying latent themes within textual data. It represents documents as mixtures of topics and topics as distributions over words. In this study, LDA is employed to extract aspects from online cruise reviews, as it is a commonly used technique in ABSA tasks to uncover latent aspects from text (Hua et al., 2024).

Large Language Models (LLMs) in Sentiment Analysis

Large Language Models (LLMs) are artificial-intelligence–based modeling frameworks designed to understand and generate human language. These models can perform NLP tasks while requiring minimal or no task-specific training data. Pre-trained language models like the Generative Pre-trained Transformer (GPT) have the ability to capture contextual relationships and sentiment expressions across domains. GPT-based models can identify aspects and their associated sentiments without using any labeled data (Brown et al., 2020). Recent advances in LLMs such as GPT-3.5 and GPT-4 (OpenAI, 2023) have reduced reliance on human labeling, thereby increasing the scalability and reliability of the training process. In the tourism context, LLMs have recently begun to be applied for sentiment analysis. Roumeliotis et al. (2024) conducted a comparative analysis evaluating NLP models such as BERT against advanced LLMs for sentiment analysis of hotel reviews; they found that GPT-4o outperformed traditional NLP models in accuracy and contextual understanding, highlighting the potential of LLMs for tourism-related sentiment analysis. Jeong et al. (2024) performed a ChatGPT-based, aspect-level summarization of TripAdvisor hotel reviews to identify service failures and generate summaries across ten predefined hotel attributes.

In this research, aspect-based sentiment analysis (ABSA) is applied to sentences extracted from online cruise reviews using the GPT-4o mini large language model. Rather than analyzing full review documents, the reviews were split into sentences before performing ABSA for two main reasons. First, sentence-level ABSA has been shown to produce more accurate results than document-level analysis (Pontiki et al., 2016), as it focuses on smaller units of text with reduced noise. Second, prompt-based large language models are more reliable on shorter texts in ABSA tasks due to their limited context windows (Zhao et al., 2023).

METHODOLOGY

Proposed Method

The present study combines three data-mining techniques; topic modeling, clustering, and aspect-based sentiment analysis (ABSA) to uncover how European cruise destinations affect customer satisfaction. The main steps are illustrated in Figure 1. First cruise names and itinerary of all cruises in Europe were collected from Cruise Lines (Step 1). Cruises are then clustered according to their route similarities (Step 2). Next itinerary specific online cruise reviews within each group were collected (Step 3) where these reviews are combined according to their respective

itinerary clusters (Step 4). ABSA is performed on these reviews using a Large Language Model (LLM), focusing on the main aspects extracted through topic modeling (Step 5). Finally, customer satisfaction and dissatisfaction are compared across different itinerary-based regional clusters (Step 6).

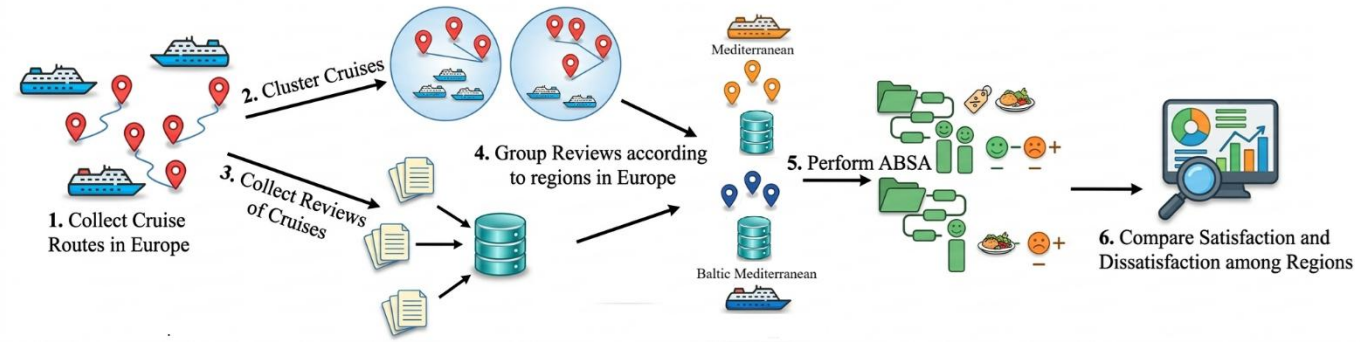


Figure 1. Workflow of the Current Study.

Data Collection

The data used in this study were collected from the online cruise-review website Cruise Lines (cruiseline.com) using the Selenium web-scraper library in Python. Cruise Lines is an online platform that provides a trusted resource for cruise travelers, detailed reviews, and expert advice to assist tourists in making informed decisions. As a popular resource, the website is widely used in various cruise-related studies (Castillo et al., 2025). The itineraries of cruise journeys in Europe between January 1, 2025, and December 31, 2025, were collected. It was observed that these journeys were offered by 53 different cruise ships belonging to 18 distinct cruise companies. A review dataset was then generated by collecting online reviews specific to these journeys. The Cruise Lines website offers itinerary-specific user reviews, enabling users to view feedback tailored to distinct cruise journeys. For the online-review analysis, a total of 3,400 reviews covering the period from January 1, 2018, to December 31, 2024, were collected. The cruises and, consequently, their online reviews were separated into four geographic regions in Europe. The regions, along with the number of cruise companies, cruises, and itinerary-specific online reviews, are as follows: (I) Western European and Atlantic Coastal regions: 13 cruise companies, 23 cruises, and 448 online reviews; (II) Eastern Mediterranean: 15 cruise companies, 26 cruises, and 1,449 online reviews; (III) Northern Europe and the Baltic Sea: 9 cruise companies, 12 cruises, and 258 online reviews; (IV) Central Mediterranean: 9 cruise companies, 13 cruises, and 1,245 online reviews.

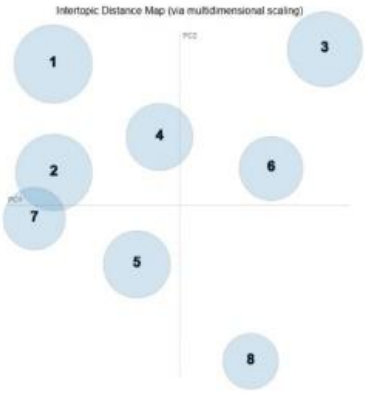
Aspect Identification with Topic Modeling and Determining Itinerary Specific Aspects

The Latent Dirichlet Allocation (LDA) (Blei et al., 2003) was employed to extract aspects from online customer reviews, enabling the identification of key discussion points relevant to cruise experiences. A total of 3,400 online reviews written for 53 different cruises were used to determine the topics mentioned in the reviews. Because LDA does not automatically determine the optimal number of topics that best represent themes within a dataset, the number of topics must be specified in advance. To address this, the Intertopic Distance Map provided by the LDAvis library in Python was used to visually assess topic separation and overlap, thereby guiding the selection of an appropriate number of topics. In this map, each topic is displayed as a circle: distances between circles indicate how similar or different the topics are (closer circles are more similar), and circle size indicates the prevalence of the topic in the dataset (larger circles are more dominant). The Intertopic Distance Map of cruise online reviews with eight topics is shown in Figure 2 (a). The corresponding cluster IDs, aspect names, and top 10 words for these topics are shown in Figure 2 (b). In that figure, the column “Aspect (LDA Topic)” provides descriptive labels assigned to each topic, and the column “Top 10 Frequent Words” lists the ten most frequent words within each topic in descending order of frequency.

By leveraging the Inter Topic Distance Map, eight distinct aspects were identified: Itineraries and Ports, Recommendations and Preferences, Dining and Food, Onboard Activities and Entertainment, Shore Excursions and Sightseeing, Cabin and Hospitality, Booking and Reservations, and Complaints and Problem Resolution. Among the identified aspects, three are specifically related to the geographical experiences of cruise customers: Itinerary and Ports, Shore Excursions and Sightseeing, and Recommendations and Preferences. Although the aspect name

Recommendations and Preferences may not explicitly suggest a geographical component, previous studies (Baker & Fulford, 2016; Hosany & Witham, 2010) have shown that positive experiences with destinations and itineraries significantly influence passengers' willingness to recommend.

a



b

Cluster ID	Aspect (LDA Topic)	Top 10 Frequent Words
1	Itineraries & Ports	cruise, port, great, ship, experience, overall, itinerary, time, good, enjoyed
2	Recommendations & Preferences	cruise, ship, time, one, norwegian, recommend, line, definitely, ever, epic
3	Dining & Food	food, good, service, room, ship, dining, restaurant, entertainment, excellent, great
4	Onboard Activities & Entertainment	day, staff, excursion, crew, ship, tour, great, everything, royal, every
5	Shore Excursions & Sightseeing	time, pool, package, excursion, country, drink, dinner, port, day, tour
6	Cabin & Hospitality	ship, staff, show, really, cabin, friendly, always, everyone, lovely, clean
7	Booking & Reservations	cruise, year, book, night, first, back, last, got, booked, suite
8	Complaints & Problem Resolution	service, special, reservation, customer, complaint, work, bed, allure, week, provided

Figure 2.a. Intertopic Distance Map **b.** Descriptive Name and Top Ten Most Frequent Words for each Topic.

Cruise type refers to the service classification of a cruise and it may influence various aspects of passenger satisfaction. However, the segment of the cruises such as luxury, premium, and standard was not considered in the current study as itinerary specific aspects are not significantly influenced by the cruise segment. Bahja et al., 2019 found that itineraries consistently rank among the top decision making factors regardless of service tier. Similarly, Jiao et al., 2024 showed that shore excursions predict perceived value equally across all cruise segments with no significant variation. These findings align with earlier studies (Petrick, 2004; Weaver, 2005) which argue that destination based experiences are valued primarily for their geographic and cultural appeal and remain independent of onboard luxury or amenities.

Aspect Based Sentiment Analysis with GPT

In this study, sentence-level Aspect Based Sentiment Analysis (ABSA) was applied to online cruise reviews, and therefore, before performing the ABSA task, the reviews were segmented into individual sentences using syntactic sentence boundary detection via the spaCy library in Python. Once the reviews were segmented, the aspect category and sentiment score of each sentence were analyzed using the GPT-4o mini model through the OpenAI API, implemented in Python. The OpenAI API provides programmatic access to GPT models, enabling automated processing of text data for tasks such as aspect and sentiment analysis. The effectiveness of instruction based Large Language Models (LLMs) such as GPT largely depends on the quality and clarity of the instruction prompts. A prompt refers to a structured input that guides the language model in performing a specific task. It defines the context, task requirements, expected output format, and any constraints, ensuring consistent and meaningful responses.

A structured prompt, as shown in Table 1, was developed to instruct the GPT-4o mini model to perform ABSA on cruise review data. To ensure clarity and consistency, the prompt was organized into four components: System Role, Guidelines, Output Structure, and Source Data. The System Role defined the model's function as a cruise industry analyst tasked with classifying aspects and evaluating sentiment objectively. The Guidelines provided detailed instructions for identifying aspects and assigning sentiment scores. Assigning sentiment scores enhances the model's ability to capture nuanced opinions within user generated content and it generates more precise differentiation between varying degrees of sentiment (Cahyani, 2023). The prompt was executed using OpenAI's API with gpt-4o-mini (model ID: gpt-4o-mini-2024-07-18), a cost effective version of GPT-4. This model was selected for its optimal balance between performance and efficiency in conducting aspect-based sentiment analysis on cruise reviews.

The Output Structure specified the format: Aspect: 'Sentence' (Sentiment Score) where the Aspect refers to the feature being evaluated (e.g., "Dining and Food"), the Sentence is the actual review sentence related to that aspect, and the Sentiment Score is a numerical value (e.g., -1 to +1) indicating the reviewer's emotional tone toward that

aspect. The Source Data component described the input as an online cruise review sentence. The performance of the GPT-4o mini model in performing ABSA was evaluated using a manually labeled subset of the review dataset. A random sample of 1,000 sentences was selected and manually labeled with aspect categories and corresponding sentiment scores by two independent annotators with tourism research backgrounds. The outputs generated by the GPT-4o mini model were then compared with the manual labels to compute standard evaluation metrics. The model achieved an accuracy of 91.2%, precision of 90.1%, recall of 89.6%, and an F1 score of 90.7% on the labeled test database. These results demonstrate that the GPT-4o mini model, when guided with well-structured instructions, performs ABSA with high reliability, even in the presence of informal expressions and mixed sentiments often seen in user generated cruise reviews.

Table 1. Structured Prompt for Instructing GPT-4o mini Model in the ABSA Task.

System Role	You are a cruise industry analyst responsible for analyzing online cruise reviews from cruise lines. The objective is to classify cruise related aspects and assess their sentiment score (-1 to +1.0, -0.5, 0.0, 0.5, 1.0) based on user feedback. Itineraries and Ports: Evaluating destinations, port stays, embarkation/disembarkation efficiency, and overall travel experience. Recommendations and Preferences: Assessing the suitability of the cruise for various travelers and whether it is recommended. Dining and Food: Reviewing variety, taste, and quality of food and beverages on board. Onboard Activities and Entertainment: Evaluating entertainment, shows, music, and onboard activities. Shore Excursions and Sightseeing: Reviewing the organization, quality, and value of shore excursions and local attractions. Cabin and Hospitality: Assessing the comfort and cleanliness of cabins, professionalism of the crew, and service quality. Booking and Reservations: Evaluating the efficiency of booking, customer service, and reservation processes. Complaints and Problem Resolution: Examining how effectively customer complaints and issues are resolved.
Guidelines	Classify the review based on the mentioned aspects and assign a sentiment score for each aspect using the following scale: -1.0 = Strongly Negative, -0.5 = Moderately Negative, 0.0 = Neutral, +0.5 = Moderately Positive, +1.0 = Strongly Positive. Classify the review based on the mentioned aspects and assign a sentiment score (-1 to +1.0, -0.5, 0.0, 0.5, 1.0) for each aspect. If multiple aspects are discussed, list them separately with individual sentiment scores. Do not assume all reviews are positive; consider criticisms objectively. Ensure neutrality if sentiment is unclear and assign 0.0 for neutral statements.
Output Structure	Return a structured output in the following format: Aspect: "Review text snippet" (Sentiment Score) Example: Dining and Food Experience: “The food was delicious and diverse.” (+1.0) Cabin, Crew and Hospitality: “The cabin was small and not very clean.” (-0.5) If no aspect is found, return an empty response. Ensure consistency in formatting.
Source Data	<Online Review>

Clustering Cruises According to Their Itineraries

Cruises were grouped by their itinerary to understand and examine how customer expectations and satisfaction levels vary across different European regions. By comparing these satisfaction levels across itinerary groups, the influence of regional differences on customer perceptions was assessed. The grouping process is performed by using a transitive clustering approach based on Jaccard Similarity. The Jaccard similarity measures the overlap between elements of two different sets. In this study, it is used to assess the similarity between itineraries of different cruises based on the sets of ports they visit during a journey. The visited ports of each cruise during a journey were considered as a single set for each cruise, and the Jaccard similarity between two cruises, A and B, is calculated as follows:

$$(A,B) = \frac{|Ports_A \cap Ports_B|}{|Ports_A \cup Ports_B|} \quad (1)$$

In equation (1), PortsA and PortsB are the sets of ports visited by cruise A and cruise B, respectively. The numerator, |PortsA∩PortsB|, denotes the number of ports visited by both cruises, and the denominator, |PortsA∪PortsB|, represents the total number of unique ports visited by either cruise A or cruise B. The resulting Jaccard similarity lies in [0, 1], where 1 indicates that cruises A and B have identical itineraries and 0 indicates that they have no visited ports in common.

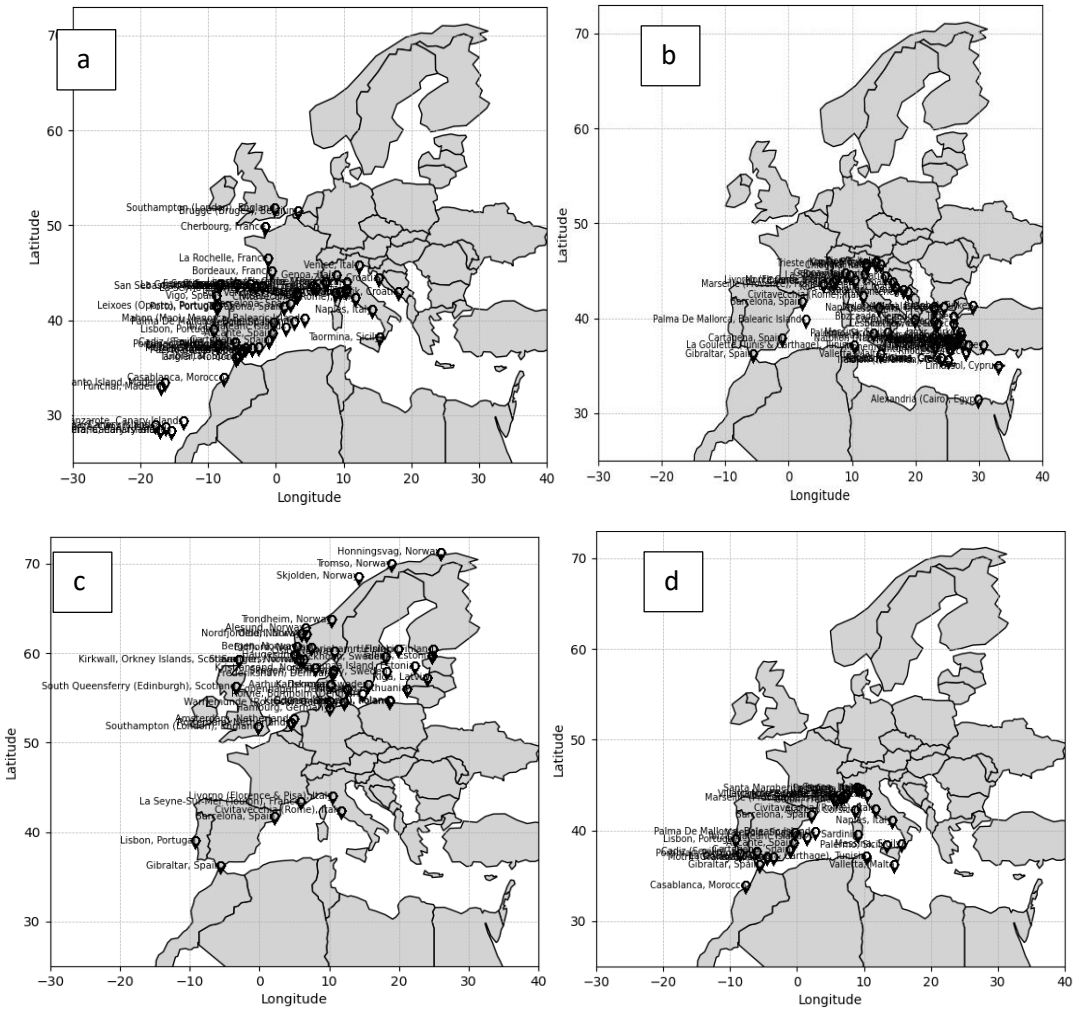


Figure 3.a. Western European and Atlantic Coastal Regions **b.** Eastern Mediterranean **c.** Northern Europe and the Baltic Sea **d.** Central Mediterranean

To model the relationships between cruise itineraries, a graph representation is used. In this graph, each node represents a cruise, and an edge (connection) between two nodes exists only if their Jaccard similarity, calculated based on the overlap of visited ports, exceeds a predefined threshold. Since similarity between two cruise itineraries is symmetric (i.e., if Cruise A is similar to Cruise B, then Cruise B is also similar to Cruise A), the edges of this graph are undirected. Additionally, as the edges only indicate the existence of similarity between nodes, the edges are unweighted, meaning all connections are treated equally, regardless of the exact similarity score.

Once the graph is constructed, transitive clustering is applied to group similar itineraries, allowing indirect connections. The key idea behind transitive clustering is that if the itineraries of Cruise A are similar to Cruise B and the itineraries of Cruise B are similar to Cruise C, then the itineraries of Cruise A, B, and C should all belong to the same cluster, even if Cruise A and Cruise C were never directly compared. Such transitive grouping allows for more flexible and realistic itinerary clusters, capturing broader regional similarities [35]. For example, consider three cruises with the following itineraries:
Itineraries(Cruise_A) = {Barcelona, Rome, Athens},
Itineraries(Cruise_B) = {Rome, Athens, Santorini},
Itineraries(Cruise_C)={Santorini, Mykonos, Rhodes}.

If cruises were grouped only based on direct similarity, two separate clusters would be formed: Cluster 1 = {Cruise_A, Cruise_B} (since Cruise_A and Cruise_B both visit Rome and Athens), Cluster 2 = {Cruise_B, Cruise_C} (since Cruise_B and Cruise_C both visit Santorini). On the other hand, transitive clustering puts all three cruises into a single cluster because the itineraries of Cruise_A are similar to Cruise_B, and the itineraries of Cruise_B are similar to Cruise_C. This clustering technique ensures that cruises with overlapping but non-identical itineraries are meaningfully grouped, capturing broader similarities between routes. Also, the use of transitive clustering is particularly important for ensuring that overlapping itineraries in different clusters do not result in the same online reviews appearing in different clusters. By applying transitive clustering, reviews from similar routes remain in a single cluster, allowing for a more accurate and comprehensive analysis of passenger experiences across itineraries.

To ensure meaningful clusters, two threshold constraints are applied. First, the minimum Jaccard similarity threshold is set to t_j=0.6, meaning that in the graph representation, an edge is formed between cruises only if at least 60% of their visited ports overlap (i.e., similarity ≥ 0.6). Second, a minimum cluster size m is enforced to prevent small, fragmented clusters from forming. These constraints help ensure that the resulting clusters represent significant groups of cruises with a high degree of itinerary similarity and sufficient size for meaningful analysis. A single cruise ship can operate multiple itinerary sets throughout a season and visit different sets of ports. Despite being the same ship, each of these distinct itineraries provides a different travel experience due to variations in destinations, shore excursions, and travel durations. Cruise Lines (cruiselines.com) provide user reviews specific to an itinerary rather than the cruise ship itself. This means that passenger reviews are tied to the particular set of destinations visited on a given sailing, rather than reflecting the ship as a whole. For instance, a traveler who took the Barcelona–Rome–Athens itinerary can comment only on their experience visiting these ports, while another traveler who took the Venice–Dubrovnik–Santorini itinerary can share insights relevant only to that route. Consequently, even though both passengers sailed on the same ship, their experiences, opinions, and feedback may differ due to the different destinations included in their respective itineraries.

The results of the clustering analysis are presented in Figure 3 (a)–(d), where each subfigure corresponds to a different cruise-itinerary region in Europe. Figure 3 (a) shows Cluster 1, in which cruises primarily visit the Western European and Atlantic coastal regions, with major ports in Portugal, Spain, France, and the UK. Figure 3 (b) shows Cluster 2, containing cruises that mainly visit the Eastern Mediterranean, covering ports in Italy, Greece, and Malta, along with some ports in North Africa. Figure 3 (c) shows Cluster 3, where cruises visit Northern Europe and the Baltic Sea, including key destinations in Norway, Sweden, Denmark, Finland, and Estonia. Figure 3 (d) shows Cluster 4, which contains cruises visiting the Central Mediterranean, with notable ports in Italy, Spain, and North Africa.

Normalizing the Sentiment Scores

In this section, customer satisfaction and expectations in four different itinerary groups are analyzed. First, a dataset called *Itinerary_Reviews* was created, containing the following columns: Review ID, Sentence, Aspect, Sentiment Score, and Itinerary Cluster. The Review ID uniquely identifies the review to which a given sentence belongs. Since each online review is split into individual sentences, a Review ID can belong to more than one sentence, each corresponding to a different aspect of the review. The Sentence column stores the textual content of the review sentence. The Aspect column represents the aspect mentioned in the review sentence, where these aspects were identified as explained in the Aspect Identification with Topic Modeling and Determining Itinerary Specific Aspects subsection. The Sentiment Score of a review sentence ranges from +1.0 (strongly positive) to −1.0 (strongly negative). A five-point sentiment scale was adopted to score the sentiments of each aspect. This scale includes −1.0 for strongly negative, −0.5 for moderately negative, 0.0 for neutral, +0.5 for moderately positive, and +1.0 for strongly positive sentiment. The Itinerary Cluster column indicates the itinerary group to which the review belongs. As mentioned in the Clustering Cruises According to Their Itineraries subsection, four different itinerary groups are generated and labeled as Cluster 1, Cluster 2, Cluster 3, and Cluster 4. For illustration, Table 2 presents an example from the *Itinerary_Reviews* dataset. In this example, the review sentence “Every stop felt like a dream, with breathtaking scenery and perfect organization” was classified under the aspect Itineraries and Ports, receiving a sentiment score of +1.0 and belonging to Cluster 1.

After the *Itinerary_Reviews* dataset was created, aspect-based sentiment scores across four itinerary groups were compared. To ensure a statistically robust comparison despite unequal numbers of reviews in each group, we implemented a four-step normalization-and-scaling procedure. This approach mitigates bias from unequal review counts while preserving interpretability. The procedure comprises the following steps:

Table 2. Example of *Itinerary_Reviews* dataset.

Review ID	Sentence	Aspect	Sentiment Score	Itinerary Cluster
1	Every stop felt like a dream, with breathtaking scenery and perfect organization.	Itineraries and Ports	+1.0	Cluster 1
1	The onboard entertainment was engaging, but the music selection could be better.	Onboard Activities and Entertainment	+0.5	Cluster 1
2	The food quality was inconsistent; some meals were great, while others were bland and overcooked.	Dining and Food	-0.5	Cluster 1
2	The dining service was extremely poor and unacceptable.	Dining and Food	-1.0	Cluster 1
3	The itinerary was poorly planned, with rushed port visits and very little time to explore.	Itineraries and Ports	-0.5	Cluster 2
3	The cabin was standard, nothing particularly impressive but also nothing to complain about.	Cabin and Hospitality	0.0	Cluster 2
4	The crew seemed disorganized, and service felt rushed throughout the trip.	Cabin and Hospitality	-0.5	Cluster 1
5	Every meal was a disappointment stale ingredients, long wait times.	Dining and Food	-1.0	Cluster 1

Sentiment scores range from −1.0 (strongly negative) to +1.0 (strongly positive)

Step 1. Count Unique Aspect Mentions per Review: In the *Itinerary_Reviews* dataset, directly counting review sentences to determine the number of unique reviews per aspect is insufficient for two reasons. First, multiple sentences may come from the same reviewer within a single review, which could artificially inflate the perceived importance of that aspect. Second, a reviewer can mention the same aspect more than once using different expressions within a single review. For example, in Table 2, Review with id 2 contains multiple sentences referring to the aspect Dining and Food. If sentences were counted directly, it would lead to multiple mentions from the same reviewer within a single review, and this would misrepresent the actual number of distinct individuals expressing opinions about that aspect. Therefore, for each itinerary group IG and aspect A, a unique review is counted once if a review contains at least one sentence labeled with aspect A, regardless of how many sentences in that review mention A. This count is denoted by $|UniqueReviews|_{A,IG}$. As an example, in Table 2, the number of unique reviews for Dining and Food in Cluster 1 is 2, since there are two unique reviews (Review IDs 2 and 5) mentioning this aspect in Cluster 1 despite three review sentences.

Step 2. Sentiment Normalization: For each itinerary group IG and aspect A, a raw sentiment score $RawSentiment\ Score_{A,IG}$ was computed by multiplying each discrete sentiment value (−1.0, −0.5, 0.0, 0.5, 1.0) by the number of reviews associated with that value and summing the results. While this raw score reflects the overall sentiment intensity for an aspect, it does not take into account the unequal review counts across groups. To adjust unequal review counts, a Weighted Sentiment Score (WSS) was calculated by scaling the raw sentiment score using the inverse of the square root of the total number of reviews. The inverse square root scaling is commonly used in comparative analysis to stabilize variance and avoid size related bias. The equation of WSS is as follows:

$$WSS_{A,IG} = \frac{RawSentimentScore_{A,IG}}{\sqrt{\sum UniqueReviews_{A,IG}}} \quad (2)$$

In the equation above, $WSS_{A,IG}$ is the weighted sentiment score of aspect A in itinerary group IG, and $|UniqueReviews|_{A,IG}$ denotes the total number of unique reviews corresponding to aspect A in itinerary group IG. The Weighted Sentiment Score (WSS) reflects both the sentiment intensity and the number of unique reviews while ensuring a fair comparison across itinerary groups with varying review sizes.

Step 3. Z Score Normalization: To facilitate a meaningful comparison that may have different ranges of weighted sentiment scores, Z score normalization was applied to the Weighted Sentiment Score (WSS) values of each itinerary group. The Z score of an aspect A in itinerary group IG is calculated as follows:

$$Z_{A,IG} = \frac{WSS_{A,IG} - \mu_A}{\sigma_A} \quad (3)$$

In the equation above, μ_A and σ_A are, respectively, the mean and standard deviation of WSS values for aspect A across all itinerary groups.

Step 4. Final Scaling Z to a Uniform [−1, +1] Range: For consistency and ease of interpretation, Z scores were linearly scaled to a uniform range [−1, +1]. The scaled score for each aspect in each itinerary group is calculated using:

$$Scaled\ Z_{A,IG} = \frac{Z_{A,IG}}{max(|Z_A|)} \quad (4)$$

In the equation above, Scaled $Z_{A,IG}$ is the Z score of aspect A for itinerary group IG, and max($|Z_A|$) is the largest absolute Z score for aspect A across all itinerary groups. Scaling via $Z_{A,IG}$ / max($|Z_A|$) preserves the sign of sentiment scores (negative values remain negative) while bounding all values to the interval [−1, +1]. Within this range, scores of ±1 indicate the most extreme sentiment observed for a given aspect across all itinerary groups. For example, a score of −1 for a given aspect represents the strongest negative sentiment detected for that aspect, while +1 reflects the most positive sentiment. This approach ensures interpretability while maintaining relative differences within each aspect's distribution.

The steps mentioned above offer a balance between statistical reliability and easy interpretation while allowing fair comparisons between aspects and remaining consistent with commonly accepted sentiment analysis practices (Liu, 2022; Pang & Lee, 2008). Step 1 resolves the inconsistent aspect mentioned frequency by counting unique reviews. Step 2 mitigates bias from unequal group sizes through inverse square root weighting. Step 3 standardizes sentiment comparisons via Z-score normalization. Step 4 ensures interpretability by scaling scores to a consistent [-1, +1] range.

RESULTS AND DISCUSSION

This section presents the results of the sentiment analysis conducted on the *Itinerary_Reviews* dataset, following the normalization procedures detailed in the Normalizing the Sentiment Scores subsection. For completeness, the full raw distribution of sentiment labeled review counts by itinerary group and aspect is provided in Table 3 in the Appendix. After applying the four-step sentiment score normalization procedure to the *Itinerary_Reviews* dataset, customer satisfaction and expectation fulfillment across cruise regions were evaluated based on three itinerary-specific aspects: Itineraries and Ports, Recommendations and Preferences, and Shore Excursions and Sightseeing. As these dimensions are directly tied to the cruise route and destination experience, they provide a valid and unbiased basis for comparing the performance of different itineraries.

The normalized sentiment scores across cruise itinerary groups and aspects are shown in Figure 4. Darker green shades indicate higher levels of customer satisfaction, darker red shades indicate stronger dissatisfaction, and lighter tones reflect more neutral or mixed sentiments. This visualization provides an intuitive overview of sentiment distribution across itineraries and enables a clear comparison of perceived strengths and weaknesses in both destination experiences and service components.

As reported in the Data Collection subsection, the number of online reviews is unevenly distributed across the four itinerary regions. The Eastern Mediterranean and Central Mediterranean groups are represented by relatively larger numbers of reviews, whereas the Western European and Atlantic Coastal group and the Northern Europe and the Baltic Sea group have fewer reviews. This imbalance may affect the stability of sentiment results in smaller groups. Therefore, as described in the Normalizing the Sentiment Scores subsection, a normalization procedure based on unique aspect mentions, inverse-square-root weighting, Z-score standardization, and final scaling to the [−1, +1] interval was applied to reduce review-volume bias.

Although the dataset includes sentiment scores for service-related aspects such as Dining and Food and Cabin and Hospitality, these scores may be influenced by factors unrelated to the itinerary itself. Most notably, variation in sentiment may reflect differences in service tier (e.g., luxury, premium, standard) rather than the destinations visited. To avoid confounding effects and ensure a focused comparison, we restricted our analysis to the normalized

sentiment scores of the three itinerary-specific aspects when identifying the most and least satisfying cruise itineraries.

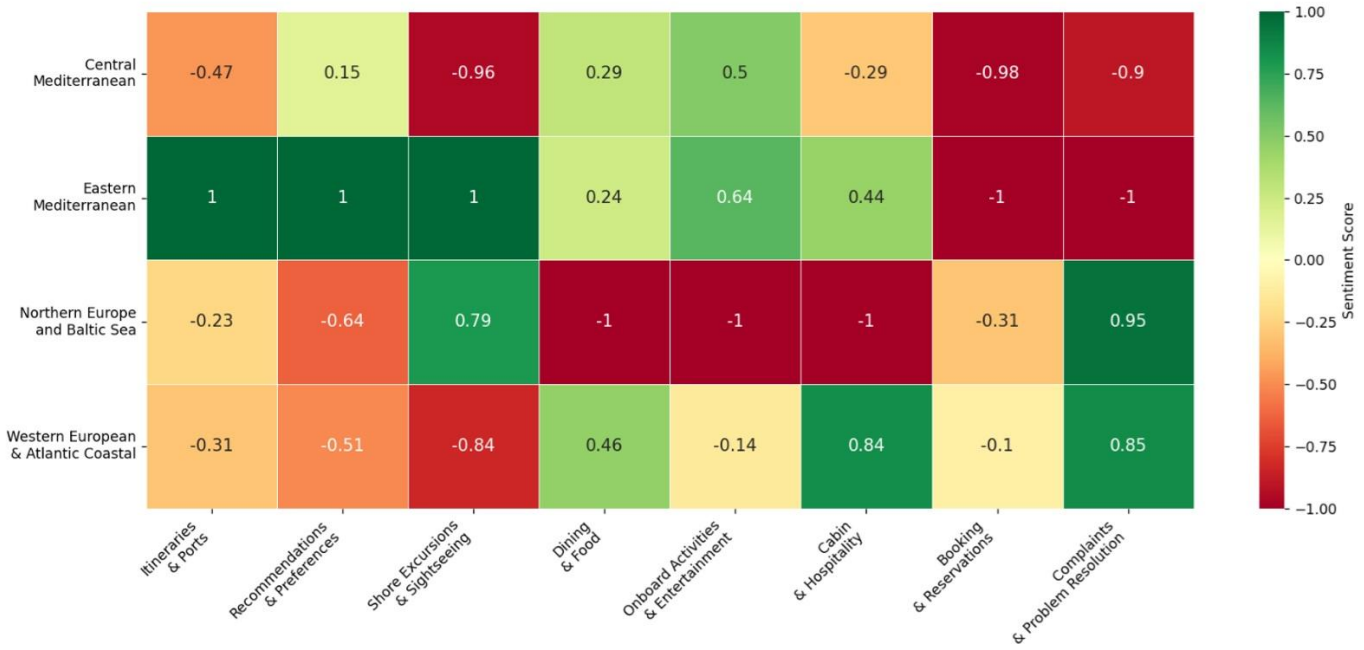


Figure 4. Aspect Sentiment Comparison Across Cruise Regions.

In Figure 4, the Eastern Mediterranean region demonstrates exceptional performance, receiving a perfect score of +1.00 across all three itinerary-specific aspects. This indicates consistently high satisfaction with route design, destination appeal, and shore excursion quality in this region. In contrast, the *Western European and Atlantic Coastal* region performs worst, with uniformly negative sentiment in *Itineraries and Ports* (−0.31), *Recommendations and Preferences* (−0.51), and *Shore Excursions and Sightseeing* (−0.84), pointing to broader dissatisfaction with the travel experience itself.

The *Central Mediterranean* presents a mixed profile, with a slightly positive Recommendations and Preferences score (+0.15) but strongly negative scores for Itineraries and Ports (−0.47) and Shore Excursions and Sightseeing (−0.96). This suggests that while some passengers find the itinerary worthwhile, many are disappointed by the actual destinations or excursion offerings. Meanwhile, Northern Europe and the Baltic Sea show a more nuanced trend: although shore excursions receive a relatively high score (+0.79), the region underperforms in Itineraries and Ports (−0.23) and Recommendations and Preferences (−0.64), implying that shore activities may be well executed even if the overall route lacks coherence or appeal from the passengers' perspective.

Overall, the normalized sentiment scores indicate that cruises visiting ports in the Eastern Mediterranean deliver high levels of passenger satisfaction with destinations, shore excursions, and the overall itinerary experience, whereas cruises operating in the Western European and Atlantic Coastal region are associated with broad dissatisfaction, particularly regarding the quality and appeal of the destination experience.

Figure 4 also displays normalized sentiment scores for service-related aspects and provides insight into the quality of the onboard experience across cruise regions. These non-itinerary aspects include *Dining and Food*, *Onboard Activities and Entertainment*, *Cabin and Hospitality*, *Booking and Reservations*, and *Complaints and Problem Resolution*. Based on these dimensions, the *Western European* and *Atlantic Coastal* region performs best, showing strong satisfaction in *Dining and Food* (+0.46), *Cabin and Hospitality* (+0.84), and *Complaints and Problem Resolution* (+0.85). The *Eastern Mediterranean* ranks second, with positive scores in *Onboard Activities and Entertainment* (+0.64) and *Cabin and Hospitality* (+0.44), though it has very low scores in *Booking and Reservations* and *Complaints and Problem Resolution* (both −1.00), indicating service-related issues. The *Northern Europe and*

Baltic Sea region ranks third, with high satisfaction in *Complaints and Problem Resolution* (+0.95) but very low ratings (−1.00) in *Dining and Food*, *Cabin and Hospitality*, and *Onboard Activities and Entertainment*, suggesting possible limitations in onboard offerings. The *Central Mediterranean* performs worst overall, with only *Onboard Activities and Entertainment* scoring positively (+0.50), while all other service-related aspects receive clearly negative evaluations.

CONCLUSION

This study presents a novel approach to evaluating cruise passengers' satisfaction across four European regions by integrating aspect-based sentiment analysis (ABSA) with the GPT-4o mini model accessed through the OpenAI API. This model was adopted for its ability to process and interpret complex linguistic structures, including mixed opinions, sarcasm, and informal expressions commonly found in online reviews. The four cruise regions were identified using a transitive clustering technique that grouped cruises by their geographically visited locations. The resulting clusters are: *Western European and Atlantic Coastal*, *Eastern Mediterranean*, *Northern Europe and the Baltic Sea*, and *Central Mediterranean*. After grouping, passenger reviews within each cluster were analyzed and compared to extract both itinerary-specific and non-itinerary satisfaction signals.

An important methodological decision in this study was not to segment cruises by service tiers (e.g., luxury, premium, standard). This choice is grounded in prior literature indicating that itinerary-related sentiments are generally unaffected by cruise type; therefore, segmentation by service tier is unnecessary for achieving this study's objectives. The results of this study provide actionable insights for both prospective cruise travelers and cruise operators. For travelers, the analysis helps identify which cruise regions consistently deliver more satisfying experiences in terms of destination appeal. For cruise companies, the findings highlight critical areas for improvement, especially for regions like the *Central Mediterranean*.

The main limitation of this study lies in the imbalance in the number of online reviews across the four itinerary regions. Although normalized aspect sentiment scores were calculated to reduce the influence of unequal review volumes, regions with fewer reviews may still be more sensitive to extreme or highly specific passenger opinions. Therefore, the findings should be interpreted as normalized comparative sentiment patterns rather than direct measures derived from equally sized regional samples.

Future work can build on this study by identifying the most and least favored port cities within each itinerary group through a focused analysis and comparison of online reviews for those ports. Such an analysis would reveal the key factors that influence passengers' satisfaction or dissatisfaction with specific port cities and provide deeper insights into what makes a destination appealing or unappealing for cruise travelers.

Artificial Intelligence Contribution Statement

The authors used OpenAI ChatGPT to assist with proofreading and language editing of the manuscript. All edits were reviewed and approved by the authors. No AI system was used to generate the study's figures or experimental results.

REFERENCES

Afzaal, M., Usman, M. & Fong, A., (2019). Predictive aspect-based sentiment classification of online tourist reviews. Journal of Information Science, 45(3), 341–363.

Al Sari, B., Alkhaldi, R., Alsaffar, D., Alkhaldi, T., Almaymuni, H., Alnaim, N., Alghamdi, N. & Olatunji, S.O., (2022). Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms. Journal of Big Data, 9, 21.

Baker, D.M.A. & Fulford, M., (2016). Cruise passengers' perceived value and willingness to recommend. Tourism & Management Studies, 12(1), 74–85.

Bahja, F., Cobanoglu, C., Berezina, K. & Lusby, C., (2019). Factors influencing cruise vacations: The impact of online reviews and environmental friendliness. Tourism Review, 74(3), 400–415.

Blei, D.M., Ng, A.Y. & Jordan, M.I., (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022.

Brown, T. et al., (2020). Language models are few-shot learners. In: Advances in Neural Information Processing Systems (NeurIPS), 33, 1877–1901.

Bryman, A., (2016). Social Research Methods. 5th ed. Oxford: Oxford University Press.

Cahyani, A.D., (2023). Aspect-based sentiment analysis from user-generated content in Shopee marketplace platform. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI), 9(2), 444–454.

Castillo-Manzano, J.I., Castro-Nuño, M. & Pozo-Barajas, R., (2024). Cruising challenging seas: 'How was your experience?'. Current Issues in Tourism, 28(18), 2898–2905.

Castillo-Manzano, J.I., Castro-Nuño, M. & Pozo-Barajas, R., (2025). FOODIES ON BOARD! Exploring cruiser satisfaction with the culinary experience. International Journal of Gastronomy and Food Science, 41, 101238.

Chen, X., Xie, H., Tao, X., Wang, F.L., Zhang, D. & Dai, H.-N., (2025). A computational analysis of aspect-based sentiment analysis research through bibliometric mapping and topic modeling. Journal of Big Data, 12, 40.

Cruise Lines International Association (CLIA), (2023). 2023 State of the Cruise Industry Report. [Online]. Available at: https://cruising.org/-/media/clia-media/research/2023/2023-clia-state-of-the-cruise-industry-report_low-res.ashx [Accessed 18 Oct. 2025].

Hua, Y.C., Denny, P., Wicker, J. & Taskova, K., (2024). A systematic review of aspect-based sentiment analysis: domains, methods, and trends. Artificial Intelligence Review, 57(11), 296.

Hosany, S. & Witham, M., (2010). Dimensions of cruisers' experiences, satisfaction, and intention to recommend. Journal of Travel Research, 49(3), 351–364.

Huang, X. & Liu, Y., (2023). Challenges of structured surveys in tourism research: A critical review. Tourism Management, 98, 104684.

Jeong, N. & Lee, J., (2024). An aspect-based review analysis using ChatGPT for the exploration of hotel service failures. Sustainability, 16(4), 1640.

Jiao, Y., Lau, Y.-Y. & Gao, J., (2024). Exploring the factors affecting cruise passengers' perceptions of value for money expressed in online reviews. Humanities and Social Sciences Communications, 11(1), 356.

Keeney, R.L., (1975). Examining corporate policy using multiattribute utility analysis. Sloan Management Review, 17(1), 63–76.

Li, M., Zhang, C., Sun, S. & Wang, S., (2023). A novel deep learning approach for tourism volume forecasting with tourist search data. International Journal of Tourism Research, 25(2), 183–197.

Liu, B., (2022). Sentiment Analysis and Opinion Mining. Cham: Springer Nature.

Lu, Y. & Zheng, Q., (2021). Twitter public sentiment dynamics on cruise tourism during the COVID-19 pandemic. Current Issues in Tourism, 24(7), 892–898.

Ma, W., Ji, F., Liang, C., Sun, Q. & Wu, J., (2024). A deep learning and large group consensus based cruise satisfaction evaluation model with online reviews. Information Sciences, 676, 120801.

Mao, Y., Liu, Q. & Zhang, Y., (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. Journal of King Saud University – Computer and Information Sciences, 36(4), 102048.

Mughal, N., Mujtaba, G., Shaikh, S., Kumar, A. & Daudpota, S.M., (2024). Comparative analysis of deep neural networks and large language models for aspect-based sentiment analysis. IEEE Access, 12, 60943–60959.

Oliver, R.L., (2014). Satisfaction: A Behavioral Perspective on the Consumer. New York: Routledge.

OpenAI, (2023). GPT-4: OpenAI's most advanced system. [Online]. Available at: https://openai.com/index/gpt-4-research/ [Accessed 19 Oct. 2025].

Pang, B. & Lee, L., (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1–2), 1–135.

Petrick, J.F., (2004). First timers' and repeaters' perceived value. Journal of Travel Research, 43(1), 29–38.

Pontiki, M. et al., (2016). SemEval-2016 Task 5: Aspect Based Sentiment Analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), San Diego, CA, USA, June 2016, 19–30.

Roumeliotis, K.I., Tselikas, N.D. & Nasiopoulos, D.K., (2024). Leveraging large language models in tourism: A comparative study of the latest GPT Omni models and BERT NLP for customer review classification and sentiment analysis. Information, 15(12), 792.

Sánchez-Franco, M.J. & Rey-Tienda, S., (2024). The role of user-generated content in tourism decision-making: An exemplary study of Andalusia, Spain. Management Decision, 62(7), 2292–2328.

Shi, J., Chen, J., Wu, J. & Liu, Y., (2025). Evaluating cruise user satisfaction through online reviews: a method based on sentiment analysis and large-scale group decision-making. Applied Intelligence, 55(6), 418.

Wankhade, M., Rao, A.C.S. & Kulkarni, C., (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), 5731–5780.

Weaver, A., (2005). The McDonaldization thesis and cruise tourism. Annals of Tourism Research, 32(2), 346–366.

Yang, S., Liao, H. & Kóczy, L.T., (2025). Preference mining and fuzzy inference for hotel selection based on aspect-based sentiment analysis from user-generated content. Journal of the Operational Research Society, 76(7), 1414–1431.

Zhao, W.X. et al., (2023). A survey of large language models. arXiv:2303.18223. [Online]. Available at: https://arxiv.org/abs/2303.18223

APPENDIX

Table 3. Raw sentiment label counts by itinerary group and aspect.

Aspect	Sentiment Scores					Itinerary Group
	-1.0	-0.5	0.0	0.5	1.0	
Itineraries and Ports	7	29	61	60	47	Western European and Atlantic Coastal
	29	103	212	182	191	Eastern Mediterranean
	4	22	26	33	59	Northern Europe and the Baltic Sea
	23	72	185	140	123	Central Mediterranean
Recommendations and Preferences	16	22	48	52	86	Western European and Atlantic Coastal
	49	71	121	166	283	Eastern Mediterranean
	6	11	22	28	64	Northern Europe and the Baltic Sea
	55	56	121	141	217	Central Mediterranean
Dining and Food	5	25	26	41	51	Western European and Atlantic Coastal
	26	105	83	120	112	Eastern Mediterranean
	3	23	17	11	18	Northern Europe and the Baltic Sea
	19	85	91	96	103	Central Mediterranean
Onboard Activities and Entertainment	8	26	24	28	23	Western European and Atlantic Coastal
	19	72	73	100	65	Eastern Mediterranean
	2	16	12	10	12	Northern Europe and the Baltic Sea
	15	67	69	82	62	Central Mediterranean
Shore Excursions and Sightseeing	2	17	18	5	11	Western European and Atlantic Coastal
	18	40	69	37	50	Eastern Mediterranean
	1	6	12	8	9	Northern Europe and the Baltic Sea
	13	40	63	34	30	Central Mediterranean
Cabin and Hospitality	3	24	24	47	94	Western European and Atlantic Coastal
	31	105	81	113	237	Eastern Mediterranean
	5	11	10	31	48	Northern Europe and the Baltic Sea
	34	106	76	97	189	Central Mediterranean
Booking and Reservations	4	29	22	9	17	Western European and Atlantic Coastal
	33	118	74	22	42	Eastern Mediterranean
	6	14	13	3	5	Northern Europe and the Baltic Sea
	39	119	61	33	30	Central Mediterranean
Complaints and Problem Resolution	4	7	19	3	4	Western European and Atlantic coastal
	26	40	58	7	10	Eastern Mediterranean
	2	5	12	1	3	Northern Europe and the Baltic Sea
	23	34	58	4	11	Central Mediterranean