The impact of the COVID-19 pandemic on airlines' passenger satisfaction

Francisco Pereira [a], Joana Martinho Costa [b,*], Ricardo Ramos [b,c], António Raimundo [d]

[a] *Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal*
[b] *Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal*
[c] *Instituto Politécnico de Coimbra, ESTGOH, Rua General Santos Costa, 3400-124, Oliveira do Hospital, Portugal*
[d] *Instituto de Telecomunicações (IT-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), 1649-026, Lisboa, Portugal*

A R T I C L E   I N F O

A B S T R A C T

This study aims to understand airline passengers' satisfaction trends by analyzing the most influential factors on satisfaction before and during the COVID-19 pandemic. The sample consists of a dataset with 9745 passenger reviews published on airlinequality.com. The reviews were analyzed with a sentiment analysis tool calibrated for the aviation industry for accuracy. Machine learning algorithms were then implemented to predict review sentiment based on airline company, travelers' type and class, and country of origin. Findings show passengers were unhappy before the pandemic, aggravated after the COVID-19 outbreak. The staff's behavior is the main factor influencing passengers' satisfaction. Predictive modeling showed that it is possible to predict negative review sentiments with satisfactory performance rather than positive reviews. The main takeaway is that passengers, after the pandemic, are most worried about refunds and aircraft cabin cleanliness. From a managerial standpoint, airline companies can benefit from the created knowledge to adjust their strategies in agreement and meet their customers' expectations.

1. Introduction

The sudden outbreak caused by the novel coronavirus has brought unprecedented challenges to many industries, severely impacting the commercial aviation industry (Monmousseau et al., 2020). The impact can be explained due to the strict travel restrictions imposed by several countries to diminish the spread of the virus (Monmousseau et al., 2020). As the disease spread worldwide, everything from businesses to schools moved towards online alternatives, travel restrictions were put in place, unemployment rates skyrocketed, and people became uneasy about traveling due to the highly contagious nature of the virus. Air travel began to drop throughout the globe by mid-March 2020 (Iacus et al., 2020), with seat availability dropping as much as 90% in April compared to the same period the year before (Suau-Sanchez et al., 2020) dramatically. The pandemic brought the aviation industry to a standstill for months, and airlines worldwide faced huge revenue losses (Hotle and Mumbower, 2021).

Thanks to the indefinite timeline for the end of social distancing and travel restrictions, much uncertainty remains on how long the pandemic will endure and how long until the air transportation sector recovers (Sobieralski, 2020). However, it is expected that the impacts are for the long term, and recovery is to take at least three to six years. The air transportation sector needs to understand what is happening and what might occur to airlines to prepare themselves to adjust to the uncertain future (Tuchen et al., 2020).

The airline industry has become a place where fierce competition is the norm. To survive and distinguish themselves from each other, airlines must manage their passengers' relations effectively to guarantee and retain customer satisfaction, with the ultimate goal of driving future income (Park et al., 2020; Sezgen et al., 2019). Siering et al. (2018) say that customer feedback is a critical factor for business growth and performance mainly for product and service innovation and improving customer experience. With that in mind, it is vital to understand how passengers evaluate airlines and identify which dimensions of satisfaction are the most important for passengers (Park et al., 2004).

During pandemic situations, it is essential to understand the air transportation system (Piccinelli et al., 2021). For example, studies have addressed how diseases propagate inside airplanes (Namilae et al., 2017) and how the pandemic and travel restrictions affected airline employment (Sobieralski, 2020). However, no study in the literature assesses how the changes brought by the COVID-19 pandemic affected the passengers' travel experience, mainly their satisfaction with airlines.

With that in mind, we aim to understand the differences in customer satisfaction between the pre-COVID-19 and during the COVID-19 pandemic. Moreover, we intend to discover what factors influence customer satisfaction then and now.

To achieve the proposed objective, customer feedback must be gathered from passengers who have flown before and during the pandemic. We (and airlines altogether) can gather customer feedback through questionnaires and forms (Guo et al., 2017). However, Rane and Kumar (2018) mention that these tend to be very time-consuming and often involve substantial human resources that come at a cost in analyzing them. Additionally, information collected using questionnaires is often inaccurate and inconsistent because people do not enjoy filling out forms or do not have the patience to take the surveys seriously.

Online reviews from the Air Line Quality website were collected to overcome such issues. The reviews were collected before and during the COVID-19 period. The data were analyzed using the text mining tool Semantria. The tool, calibrated with an aviation-specific dictionary to improve accuracy, will identify the most mentioned satisfaction dimensions by the passengers and measure the sentiment polarity of each review (Lexalytics, 2021a, 2021c).

With such created knowledge, it is expected that airline stakeholders will be able to understand the market and their customers, adjusting their strategies accordingly.

2. Literature review

2.1. Airline passenger satisfaction

One output that results from the purchase of a product or the use of a service is satisfaction. It develops from the contrast between benefits, cost, and expectations (Sezgen et al., 2019). Customer satisfaction can be measured by cumulating from products/services (Churchill and Surprenant, 1982).

The literature has widely accepted the approach to customer satisfaction by Oliver (1980), who defines it as a function of expectation and expectancy disconfirmation. The theory says that consumers develop an expectation about a specific product or service before its purchase, which will be the standard for said product/service. Once the customer uses the product/service, it will compare the experience with the pre-purchase expectations. Three scenarios may emerge: the customer is satisfied if the perceived performance matches the expectation. If the expectations are exceeded, the customer is also satisfied. However, dissatisfaction occurs if the expectations are not fulfilled (Punel et al., 2019). For an airline passenger, when the service quality attributes that the passenger values are met or excelled, the passenger tends to be satisfied (Chow, 2015). Those attributes depict the various dimensions of satisfaction (Guo et al., 2017; Moro et al., 2020).

The airline industry, by nature, is very dynamic, so it is challenging to distinguish airlines from one another and describe each one in a uniformly (Mason and Morrison, 2008). However, Zeithaml (1988) explains that from the passengers' perspective, expectations and perceptions of an airline service may differ according to the different business models between a low-cost and a full-service airline.

Zeithaml (1988) adds that it would be reasonable to expect that passengers from low-cost airlines have different expectations from passengers traveling with a full-service airline. In fact, Forgas et al. (2010) suggested that low-cost passengers' satisfaction is mainly influenced by monetary cost and service quality. For full-service airline passengers, the cabin crew was essential for the passengers' satisfaction.

Zeithaml (1988) also mentions that even passengers from the same airlines might form different expectations. For example, passengers flying in a premium cabin as opposed to those flying in economy. This could be explained since the consumer utility expectations increase proportionally to the amount paid. A study conducted by Lucini et al. (2020) found that passengers traveling in different classes had different

expectations. In the study, it was found that customer service was paramount to passengers traveling in First Class.

Conversely, Economy Class passengers gave more importance to airport prices, waiting times, luggage checking, and delays. It was also possible to find similarities among different nationalities. For example, it was concluded that Americans and Canadians exhibit the same behavior when writing about satisfaction dimensions which, in turn, contrasts with the writing of the British and Australians. Finally, the passenger type (e.g., solo traveler) had a minimal impact on the customer satisfaction dimensions (Lucini et al., 2020).

Many studies concluded that customer satisfaction could ultimately affect customers' loyalty. If satisfied, that will translate into positive reviews, product recommendations, and returning customers (Forgas et al., 2010; Guo et al., 2017; Mattila, 2004). If not, however, in the case of airlines, it might result in reconsidering using the same airline in the future (Namukasa, 2013) and even promoting negative word of mouth, causing damage to the airline's reputation (Blodgett and Li, 2007).

Zhang et al. (2016) suggested that airlines have predominantly positive or predominantly negative reviews, meaning that passengers tend to praise or complain about an experience rather than write a neutral comment.

The question that remains is what dimensions influence satisfaction on airline passengers. Table 1 summarizes articles that extracted satisfaction dimensions expressed by the passengers.

Lacic et al. (2016) tried to understand which satisfaction dimensions influence airline passengers the most and the extent to which satisfaction can be predicted. The authors used a pre-made dataset from Skytrax and explored four different review categories: airport, lounge, airline,

Table 1

Satisfaction dimensions and their influence on passengers.

Satisfaction Dimensions	Influence satisfaction/ sentiment	Source
Friendliness and helpfulness of staff/ Customer Service	✔	Sezgen et al. (2019); Lucini et al. (2020); Song et al. (2020); Lacic et al. (2016)
Hassle-free customer experience	✔	Sezgen et al. (2019); Lucini et al. (2020)
Comfort of the seat	✔	Sezgen et al. (2019); Lucini et al. (2020); Song et al. (2020); Zhang et al. (2016); Lacic et al. (2016)
Value	✔	Sezgen et al. (2019); Lucini et al. (2020); Song et al. (2020); Zhang et al. (2016); Lacic et al. (2016)
Food and Beverage	✔	Sezgen et al. (2019); Lucini et al. (2020); Song et al. (2020); Zhang et al. (2016); Lacic et al. (2016)
In-flight service	✔	Sezgen et al. (2019); Lucini et al. (2020); Song et al. (2020); Zhang et al. (2016); Lacic et al. (2016)
Airplane characteristics	✔	Lucini et al. (2020); Lacic et al. (2016)
Airport Lounge	✔	Lucini et al. (2020); Lacic et al. (2016)
Delays	✔	Lucini et al. (2020)
Terminal infra-structure	✔	Lucini et al. (2020); Song et al. (2020); Lacic et al. (2016)
Check-in procedures	✔	Lucini et al. (2020)
Checking luggage	✘	Lucini et al. (2020)
Cabin flown	N/A	Zhang et al. (2016)
Wi-Fi & Connectivity	✔	Zhang et al. (2016); Lacic et al. (2016)
Ground Service	N/A	Zhang et al. (2016)
Route	N/A	Zhang et al. (2016)
In-flight Entertainment	N/A	Zhang et al. (2016); Lacic et al. (2016)

and seat reviews. A feature analysis was performed in which the review rating was correlated to the overall sentiment. It was concluded that the most impacted passenger satisfaction was the queuing time, lounge comfort, cabin crew quality, and seat legroom.

On service failure and disruptions, Song et al. (2020) suggested that flight delays negatively affect passengers' sentiments. The study adds that passengers were unsatisfied with airline compensation mechanisms after flight delays, and attention to service aspects tends to increase after a disruption. In contrast, Xu et al. (2019) found that passenger compensation after service disruption positively affects customer emotions. However, if the compensation is for a future trip, it does not influence emotion positively, even if it is monetary compensation. Airlines are advised to provide either monetary or non-monetary compensations (e.g., upgrades, priority boarding, or complimentary meals) to ease the passengers' frustrations.

Additionally, employee attitude toward dissatisfied or complaining passengers affects the passengers' emotions (Xu et al., 2019). Service failure has more impact on full-service airline passengers than those traveling on low-cost airlines. This is explained by the higher fare that full-service passengers pay, which comes with higher expectations. For the same reason, the type of cabin flown also impacts the emotions regarding service failure. Business-class passengers that pay higher airfares are more affected than economy passengers.

The COVID-19 pandemic significantly impacted the airline sector, leading to new business models and measures to ensure passenger health and security. These measures may have had different impacts on passengers' satisfaction. For instance, as stated by Bauer et al. (2020), ultra-long-haul flights appear to be a notable development accelerated by the pandemic. This model involves non-stop flights over extended distances, possibly cutting down on layovers and reducing exposure to crowded airports, which could indirectly enhance passenger satisfaction in the context of pandemic-related anxieties (Bauer et al., 2020). This shift came against the backdrop of an industry grappling with significant challenges even before the pandemic, including high capital intensity, fluid supply, low entry barriers, and a high degree of competition. The pandemic significantly amplified these challenges.

Moreover, airline companies and airports implemented multiple security and health measures to protect passengers, such as mandatory masks, frequent airplane disinfection, and social distancing (Rita et al., 2022). These measures provided excellent safety and well-being, increasing satisfaction. One other measure was the flexibility in the cancelation and reimbursement politics. These changes permitted passengers to change or cancel their reservations without additional costs or receive credits for future flights (Amankwah-Amoah, 2020). These flexible politics may have also provided a feeling of tranquility while booking at an uncertain time, leading to satisfaction (Piccinelli et al., 2021).

Despite the turmoil in the industry, some unexpected positives emerged. Interestingly passenger satisfaction reached record highs during the pandemic, surpassing pre-pandemic levels (Williams, 2020). A key reason for this increased satisfaction could be attributed to less crowded flights and fewer passengers, which enhanced the comfort and convenience of air travel. However, as we move towards the post-pandemic era, it remains to be seen how these new business models, and the changes brought about by the pandemic, will continue influencing passenger satisfaction. Future research should focus on this area to comprehensively understand these dynamics.

3. Methodology

For this study, the sample consisted of passengers that flew on a European airline with an attributed COVID rating (Table 2) and published a review on Airlinequality.com. Airlinequality. com is the top review site for airlines, airports, and associated air travel reviews (Skytrax, 2021a). It is owned by Skytrax, a recognizable brand for its Airline and Airport Star Rating, the World Airline Awards, and Airport

Table 2
Selected airlines and respective COVID-19 rating on Skytrax.

Airline	COVID-19 Rating
Aegean Airlines	4
Air France	4
airBaltic	5
AnadoluJet	3
Blue Air	3
British Airways	4
easyJet	4
Iberia	4
KLM Royal Dutch Airlines	4
LOT Polish	3
Lufthansa	4
Pegasus Airlines	3
Ryanair	4
Turkish Airlines	4
Vueling	4
Wizz Air	3

Awards (Skytrax, 2021a). Skytrax is an international air transport rating organization that permits passengers to share their personal experiences and service evaluation (Song et al., 2020). It summarizes the passengers' comments and ratings and provides the overall service quality performance of the airline or airport from one to five stars. Five stars reflect that the airline implements strict safety protocols that enhance passenger and staff safety (Skytrax, 2021b). Airlinequality.com prides itself on being an independent customer forum with no financial association with any of the airlines or airports featured (Skytrax, 2021a). In 2020, Skytrax performed the world's only assessment and certification of the health and safety measures taken by the airlines during the pandemic. Skytrax online reviews have been used in previous studies to understand clients' satisfaction. For instance, Song et al. (2020) used Skytrax online reviews to understand passengers' satisfaction and emotion regarding flight delays. In turn, Rita et al. (2022a,b) aimed to understand how airline companies managed the impact of COVID-19 and handled issues such as cancellations and customer satisfaction using Skytrax reviews.

3.1. Data collection

We resorted to a *web scraper* to collect all the existing reviews efficiently and effectively from the selected airlines available on Airlinequality.com. A web scraper is a tool that can extract specific data from web pages. For this task, Octoparse.com was used to collect the reviews, following the approach of previous studies (Hamada and Naizabayeva, 2020).

Several fields were collected from each review, enriching this study's findings (Table 3).

A total of 16,583 reviews were collected. However, reviews before 2016 did not have consistent data. Some fields had missing values and were discarded. The final number of reviews in the dataset is 9,743, dating from January 2016 until February 2021.

Table 3
Data on each observation of the dataset.

Review	
Airline_Name	Airline flown by the reviewer
Rating	Rating given by the reviewer, on a scale from 1 to 10, being 10 the best and 1 the worst
Text	Text of the review
Country	Country from which the reviewer originates from
Class	Class of travel, being the possible values "First Class", "Business Class", "Premium Economy" and "Economy Class"
Travel_Date	Date of the flight, on a DD/MM/YYYY format
Recommend	Boolean (YES/NO) indicating if the reviewer recommends the airline flown.
ID	Unique identifier of the review

3.2. Data analysis

Text Mining is a technique where structured and unstructured data is processed and analyzed (Ramos et al., 2019). More recently, with the increasing amount of text data generated on websites, social media, and news, more studies about text mining have been conducted (Furtado et al., 2022; Ramos et al., 2022). This study will focus on sentiment analysis to analyze the gathered data. Sentiment analysis identifies the sentiment within a subjective statement or opinion and can be classified as positive, negative, or neutral (Rita et al., 2022). Sentiment analysis is an analytic method of big data that identifies the polarity of sentiment in expressions or judgments made by consumers (Rita et al., 2020). It results from artificial intelligence, natural language processing, information extraction, and information retrieval (2017).

There are four approaches: dictionary-based, machine learning, statistical, and semantic (Tsytsarau and Palpanas, 2012). As the tool used in this study relies upon the dictionary-based method, we will focus on that approach. A dictionary-based technique generally relies on a dictionary containing words and phrases that have attributed scores ranging from +1 (strongly positive) to −1 (strongly negative) (Lexalytics, 2021c). When calculating the sentiment for a specific document, the content is evaluated to match the words in the dictionary. The polarity of a document will result from the sum of the polarities of the individual words or phrases (Devika et al., 2016). Sometimes the weight of a particular word must be adjusted because of the modifier that accompanies it (Lexalytics, 2021c). The most common modifiers are negators (for example, never or not) and intensifiers (much and very). A negator usually reverses the word's score in the dictionary, while an intensifier might raise or lower the score.

Semantria was used to calculate the sentiment. Semantria is a text and sentiment analysis tool developed by Lexalytics (2021b) with an "industry pack" for aviation. In other words, Semantria contains an industry-specific dictionary. An "industry pack" calibrates the sentiment engine to be more accurate to a specific subject, meaning that the sentiment score will be precise, contributing to more accurate results (Lexalytics, 2021a).

Each score represents the polarity of the sentiment that is present in a text. The polarity in Semantria ranges from −2 to 2, and Table 4 describes the default classification scheme used in this study.

At this point, defining when the COVID-19 period begins is essential. It was reported that European airlines began reacting to the COVID-19 pandemic as early as January 2020 (Albers and Rundshagen, 2020). In late January, European carriers also saw the first COVID-19-related flight cancelations (IATA, 2020). Considering the facts, January appears to be the initial period in which passengers felt for the first time the COVID-19 restriction. For that reason, January 2020 was considered the beginning of the COVID-19 period.

Finally, the dataset was divided into two parts to prepare for the sentiment analysis. One corresponds to the pre-COVID-19 period (before January 2020), and the other corresponds to the post-COVID-19 period (after January 2020).

In addition to text sentiment analysis, it was possible to evaluate the relationships between the variables of Table 3 performed to the customers' provided rating and whether they would recommend the airline flight. A Python library named 'Dython' was used, and Cramer's V technique was employed to perform correlations between categorical/numerical correlations (McHugh, 2018). The strength of a relationship between two nominal variables was measured using this method.

Table 4
Sentiment classification according to polarity score.

Sentiment polarity range	Classification
[-2, −0.05 [	Negative
[-0.05, 0.22 [	Neutral
[0.22, 2]	Positive

The heatmap below shows the direct correlation between all the variables presented in Table 3, except for the columns 'Text,' 'Travel Date,' and 'ID'. A Pearson correlation is not achievable because all the supplied variables (except for 'Rating') are categorical (non-numerical).

According to Fig. 1, the relationship between Rating' and 'Recommended' is strongly positive, showing that when a client recommends an airline, it usually results in a high-rated review. The remaining features present a low correlation.

We used supervised machine learning algorithms to determine whether the country, type of traveler and class are predictors of customer satisfaction based on their potential review sentiment,. In other words, this means that we can anticipate possible review sentiment (Positive or Negative) without any text analysis, rating evaluation, or yes/no recommendation. Machine learning techniques can be beneficial for airline companies. Through them, managers can identify to what extent the customer's features determine their overall satisfaction with the airline company. They can also improve the target customer retention procedures, marketing campaigns, and customer user experience.

To perform the algorithm, first, we created a new binary variable named 'Review Sentiment', dependent on an existing variable called 'Rating' (see Table 3), following the below criteria.

- If a review's rating value were less than or equal to 5 (Rating = 5), the 'Review Sentiment' variable would be set to 'Negative'.
- If a review's rating value were greater than 5 (Rating >5), the 'Review Sentiment' variable would be set to 'Positive'.

After defining the Review Sentiment as the target variable (the one to predict), we used three supervised machine learning methods to generate the predictions: Support Vector Machines (SVM), Logistic Regression (LR), and Decision Trees (DT). These methods are often used to build predictive models for classification and regression problems (Zhang et al., 2022). Since we are predicting whether a review will be positive or negative, this consists of a binary classification problem. All the algorithms discussed were implemented using the Python library 'Scikit-Learn'.

4. Results and discussion

Sample characterization is described in Tables 5 and 6 by class and country for pre and during the COVID-19 periods.

It is apparent that most passengers wrote review travels in economy class in both periods (78.53% and 85.58%, respectively) and are from the United Kingdom (36.01% and 28.15%, respectively).

Semantria calculated the sentiment polarity for each review. In addition, other relevant information regarding each review was analyzed. Fig. 2 demonstrates, during the pre-COVID-19 period, the distribution of the reviews regarding the polarity attributed to each one, the rating given, and if it recommends the airline.

It is possible to acknowledge that most reviews classified with negative polarity have lower ratings and negative recommendations (NO). Similarly, most positive reviews exhibit higher ratings and recommendations (YES). As expected, passengers who positively recommend the airline are more likely to leave a positive review with a high rating. The opposite also happens, confirming the findings of Xu et al. (2019), which mentioned that positive emotions increase satisfaction and negative emotions decrease satisfaction.

From the pre-COVID-19 dataset, 24.18% of the reviews were classified as positive, 19.92% as neutral, and 55.90% as unfavorable. This means that more than half of the reviews on Skytrax from January 2016 to December 2019 are most likely complaints or dissatisfaction with the airline's service. It is also worth mentioning that according to the expectancy-disconfirmation theory formulated by Oliver (1980), passengers recommending the airline must have had their expectations met. Otherwise, they do not recommend an airline. The results show that

Fig. 1. Correlation between variables.

Table 5
Passengers by class.

Travel Class	Pre-COVID-19		COVID-19	
	Absolute frequency	Relative frequency	Absolute frequency	Relative frequency
Business Class	1372	15.48%	100	11.35%
Economy Class	6959	78.53%	754	85.58%
First Class	154	1.74%	7	0.79%
Premium Economy	377	4.25%	20	2.27%
Total	8862	100.00%	881	100.00%

Table 6
Top 20 countries of the sample.

Country of origin	Pre-COVID-19		Country of origin	COVID-19	
	Absolute frequency	Relative frequency		Absolute frequency	Relative frequency
United Kingdom	3191	36.01%	United Kingdom	248	28.15%
United States	1104	12.46%	United States	124	14.07%
Germany	500	5.64%	Germany	53	6.02%
Canada	359	4.05%	Canada	37	4.20%
Netherlands	310	3.50%	Netherlands	30	3.41%
Australia	245	2.76%	Spain	28	3.18%
France	217	2.45%	France	26	2.95%
Greece	183	2.06%	Ireland	24	2.72%
Spain	182	2.05%	Portugal	23	2.61%
Switzerland	176	1.99%	Australia	17	1.93%
Ireland	157	1.77%	Greece	17	1.93%
Italy	151	1.70%	Switzerland	16	1.82%
Poland	116	1.31%	Italy	16	1.82%
Portugal	96	1.08%	Poland	15	1.70%
Belgium	94	1.06%	Czech Republic	13	1.48%
Turkey	94	1.06%	Sweden	11	1.25%
Singapore	90	1.02%	Norway	9	1.02%
Sweden	82	0.93%	Singapore	9	1.02%
Romania	74	0.84%	Russian Federation	9	1.02%
Czech Republic	74	0.84%	Turkey	9	1.02%
Others	1367	15.43%	Others	147	16,69%
Total	8862	100.00%	Total	881	100.00%

most passengers are not having their expectations met.

We can also recognize that the ratings 7, 8, 9, and 10 explain 22.46% of the 24.18% population of positive reviews classified by Semantria. In the same way, ratings 1, 2, 3, and 4 explain 49.25% of the 55.90% population with reviews classified as negative. This indicates that the Semantria algorithm accurately identifies sentiment since the results align with the rating classification system created by the passengers. Finally, the reviews are predominantly negative (55.90%) because customers tend to complain or praise an experience rather than leave a neutral review (Zhang et al., 2016). This also explains why neutral ratings occur in the sample.

Fig. 3 illustrates the same data mentioned above but during the COVID-19 period.

The overall distribution appears to be similar to the pre-COVID-19 period. However, it is apparent that the number of negative reviews has increased, suggesting that the COVID-19 restrictions worsened the travel experience. In this dataset, 15.45% are positive reviews, 17.81% are neutral, and 66.75% are negative. Unlike the previous dataset, most neutral reviews appear to have a rating of 1, further emphasizing the overall negative attitude toward the airline industry.

Lastly, the passengers' ratings appear condensed on the scale's extremities. Rating 1 explains 47.88% of the population of 66.75% of negative reviews, and ratings 9 and 10 explain 10.14% of the 15.45% of positive reviews. These results further emphasize the findings of Zhang
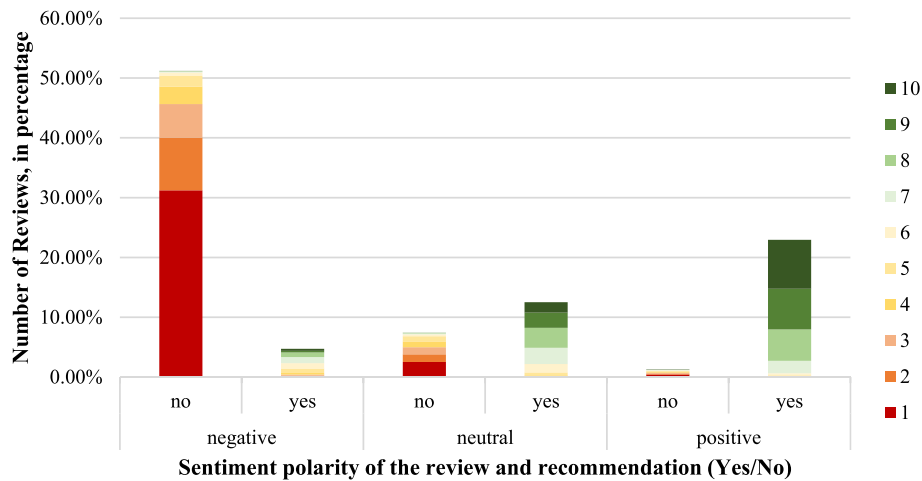
Fig. 2. Review rating and sentiment polarity distribution, pre-COVID-19 period.
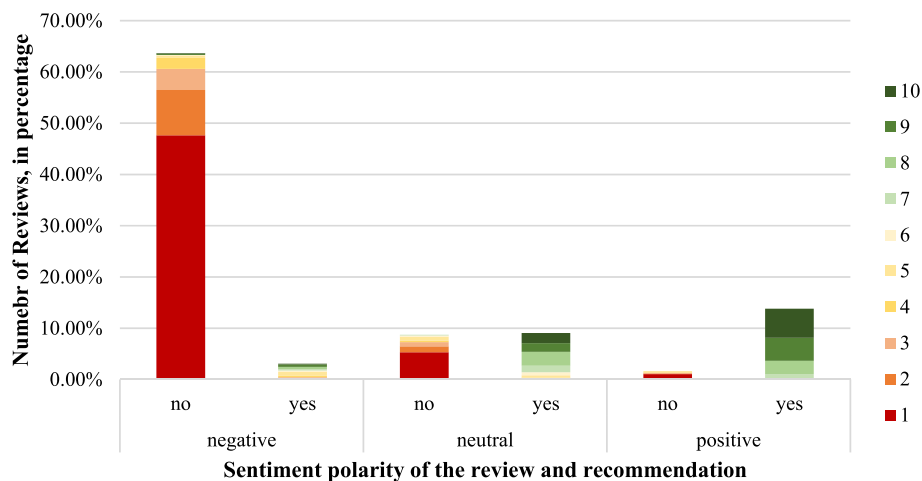

Fig. 3. Review rating and sentiment polarity distribution, COVID-19 period.

et al. (2016), which stated that people instead praise or complain about an experience rather than leave a neutral review.

Semantria, thanks to the built-in topic detection function, can also classify each sentence of the reviews into airline industry-related categories. Figs. 4 and 5 illustrate the ten most mentioned airline-related categories during the pre-COVID-19 period for Low-Cost Carriers (LCC) and Full-Service Carriers (FSC), respectively. It is possible to know
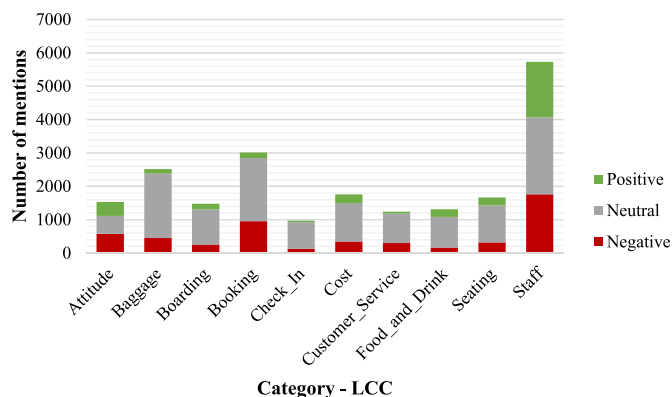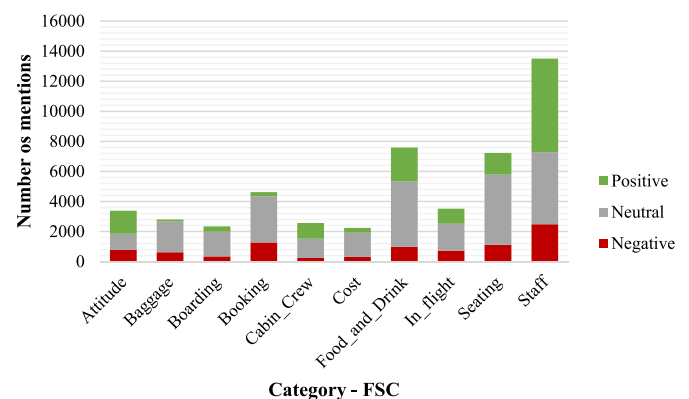

Fig. 5. Number of mentions, by category for FSCs during the pre-COVID-19 period.

how many positive, neutral, and negative mentions for each category.

It becomes apparent that Staff is the most mentioned aspect by the LCC and FSC passengers. This indicates that passengers pay attention to how the airport staff and cabin crew treat them whether they are helpful. This result confirms previous studies that suggest that the airport staff and cabin crew are among the factors that most influence passenger satisfaction (e.g., Lacic et al., 2016; Lucini et al., 2020; Sezgen et al.,

Fig. 4. Number of mentions, by category for LCCs during the pre-COVID-19 period.

2019; Song et al., 2020). Sezgen et al. (2019) acknowledge that staff attitude is among all passenger groups' most critical satisfaction and dissatisfaction attributes. In other words, passenger satisfaction varies proportionally with the performance of this attribute.

Seating, Food_and_Drink, Baggage, and Booking are important factors influencing passenger satisfaction since these are the most mentioned. It is also noteworthy that most of the dimensions of satisfaction identified in Figs. 3 and 4 were also identified in other studies, as shown in Table 1.

Figs. 6 and 7 show the mean sentiment polarity for each category for LCCs and FSCs for the same period. The colors of the bars represent the sentiment polarity, red for negative (score under −0.05), grey for neutral (between −0.05 and 0.22), and green for positive (above 0.22).

LCC passengers have primarily negative experiences across most satisfaction dimensions. In contrast, FSC passengers have mixed experiences but mostly positive ones. The exact dimensions are mentioned between LCCs and FSCs. The only differences are that LCC passengers mention negative experiences with Customer_Service and Check-In. In contrast, FSC passengers mention a positive experience with Cabin_Crew and a neutral experience with In_Flight aspects.

Forgas et al. (2010) suggested that FSC passengers value the personnel's professionalism and LCC passengers' quality of service, justifying why the dimensions Cabin_Crew appear only in the FSCs sample while Customer_Service in the LCCs sample. Lastly, Booking and Baggage are the aspects that contribute to a negative experience in both LCCs and FSCs, suggesting that these aspects are to be improved by the airlines.

Researchers have explored the possibility that the type of cabin flown may influence passenger satisfaction and sentiment differently (Lucini et al., 2020; Sezgen et al., 2019). Economy class and Premium Economy passengers demonstrate mostly positive polarity towards Cabin_Crew-Attitude and Staff-Helpfulness, which aligns with Sezgen et al. (2019). Economy cabin passengers value a Friendly-helpful staff and a Hassle-free customer experience. By the number of mentions, these types of passengers also give importance to Cost, specifically Baggage_Cost (luggage fees, for example) and Food_and_Drink-Cost. These results mean that cost-conscious passengers are only interested in getting from point A to point B (Lucini et al., 2020). Business and First Class passengers also seem to praise Cabin_Crew-Attitude and Staff-Helpfulness. However, they do not exhibit a significant number of mentions of Cost, which is also in line with Lucini et al. (2020)'s findings that customer service is paramount to passengers traveling in First Class. They focus on Seating_Quality, In-flight_Quality, Lounge, and

Food_and_Drink-Quality, appearing to be a type of passenger that appreciates the airline's product.

Lastly, Table 7 shows that First Class passengers are the most satisfied and with higher sentiment polarity, followed by Business, Premium Economy, and Economy Class. The average rating corroborates the sentiment polarity, which is expected as Lacic et al. (2016) suggest that the review rating correlates to the overall sentiment.

Regarding the COVID-19 period, Figs. 8 and 9 illustrate the ten most mentioned satisfaction dimensions.

By comparison with the pre-COVID-19 period, there are few changes. The staff remains the central aspect that passengers discuss. Mentions about Booking appear to have increased for FSC passengers and Customer_Service, which previously did not make the top ten categories of FSC passengers. Regarding LCC passengers, the only difference is that the Food_and_Drink disappeared from the top ten, giving its place to In_Flight. This is to be expected since it is known that due to COVID-19 restrictions, airlines reduced, and some even completely suspended, the onboard food service (Food & Wine, 2020).

Similarly to the pre-COVID-19 period, LCC passengers have a primarily negative sentiment towards most categories (Figs. 10 and 11). There are no categories with a positive sentiment. Compared to the pre-COVID-19 period, the categories of Seating and Cost became negative. Forgas et al. (2010) mentioned that LCC passengers are sensitive to monetary cost, and airfares have increased due to the pandemic (Barrons, 2020), explaining the decrease in the sentiment polarity of Cost. The negative sentiment in Seating might be explained due to some airlines occupying the middle seats with passengers, disregarding the guidelines of social distancing (Nytimes, 2020).

Regarding the FSC passengers, the categories Staff, Attitude, and Food_and_Drink became neutral. As FSC passengers value the professionalism of airline employees, the protocols in place to contain the pandemic might have impacted how the employees perform their job, resulting in a worse sentiment towards this specific aspect. Regarding the Food_and_Drink sentiment decrease, it can be explained by some airlines reducing or even suspending the food offerings onboard (Food & Wine, 2020). Finally, Cost became negative, understandably for the same reason mentioned above: the airfares have risen (Barrons, 2020).

The only positive sentiment is towards Cabin_Crew, which remained the same. This is to be expected since cabin crew functions remained the same in flight during the pandemic. They still greet and serve the passengers while ensuring the passengers' safety.

Overall, the factors influencing the passengers' satisfaction have not changed during the COVID-19 period. However, there is a noticeable
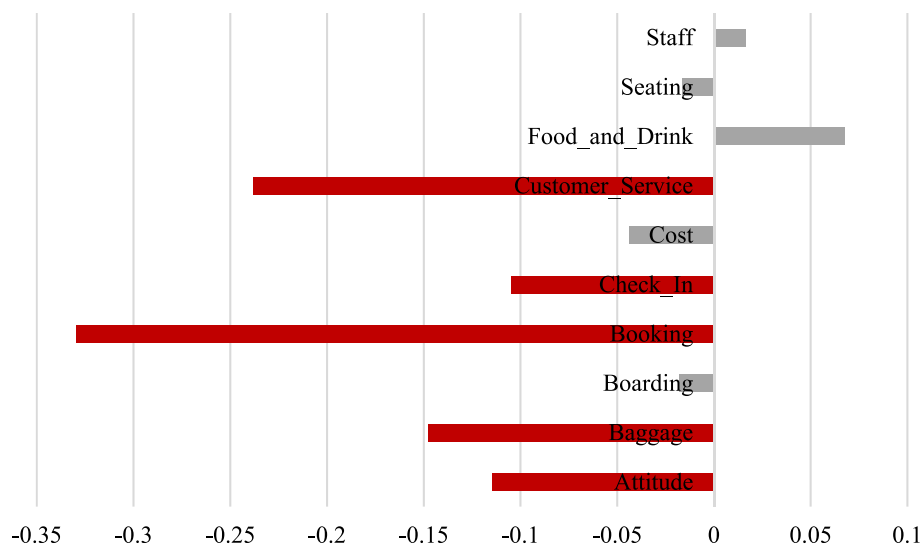

Fig. 6. Mean sentiment polarity, by category for LCCs, during the pre-COVID-19 period.
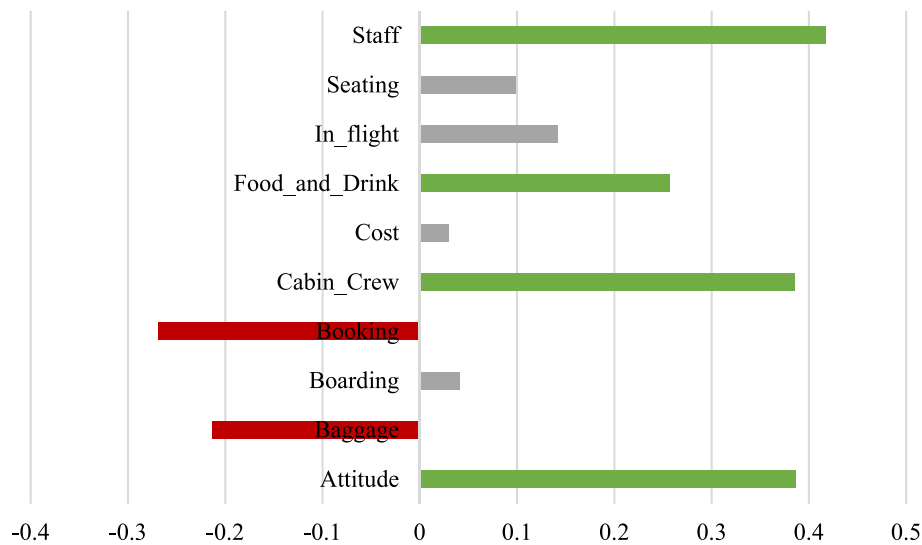
Fig. 7. Mean sentiment polarity, by category for FSCs, during the pre-COVID-19 period.

Table 7
Average rating and sentiment polarity, by travel class, pre-COVID-19.

Travel Class	Average rating	Average sentiment polarity
Business Class	5,707,016,191	0,049,027,327
Economy Class	4,17,890,182	−0,101,735,103
First Class	6,263,157,895	0,136,772,101
Premium Economy	4,842,406,877	−0,009,125,104
Total	4,47,258,578	−0,07,112,772

surge of mentions in Customer_service-refunds and In_flight-Cabin-Cleanliness in Economy Class and Business Class passengers.

The surge of Customer_service-refunds, as mentioned by Dada (2021), can be explained because airlines have been known to intentionally hinder the refund process, making passengers wait long periods and, to an extreme, not answering the passenger's contact attempts. Some airlines are processing refunds through vouchers that the passenger can redeem later. However, despite the financial stress that airlines worldwide are going through, they are obligated to refund the passenger. These situations are causing passengers to go on social media to complain (Dada et al., 2021).

Regarding In_flight-Cabin-Cleanliness, it can be explained simply because due to the coronavirus. Passengers pay more attention to infection prevention and disease control procedures to feel safe (Sotomayor-Castillo et al., 2021).

Overall, the rating and sentiment have worsened during the COVID-19 period for all travel classes except First Class (Table 8). The average rating of First-Class passengers has increased. However, the sentiment polarity did not.

Table 9 allows us to compare the Skytrax COVID-19 ranking with the passengers' cleanliness scores.

It is worth mentioning that this average rating only contributed to reviews that focused on in-flight cleanliness (cabin and bathroom), airport lounge cleanliness, and airport boarding area cleanliness. Although this method might not be the most accurate, it reveals that, at least for Air Baltic, the passengers notice their cleanliness protocols (proven by the highest cleanliness rating) and deserve the five-star score awarded by Skytrax.

4.1. Predictive modeling

Since we are dealing with a binary classification problem, the class
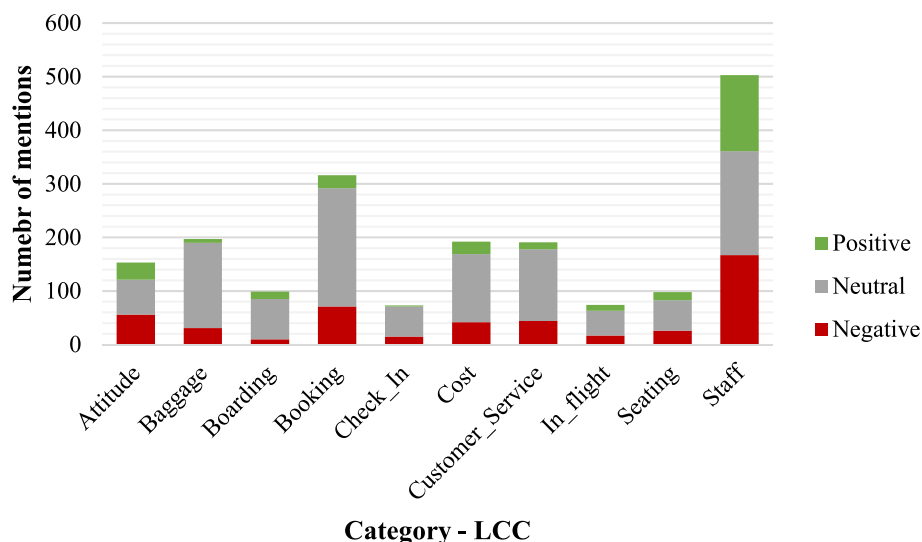

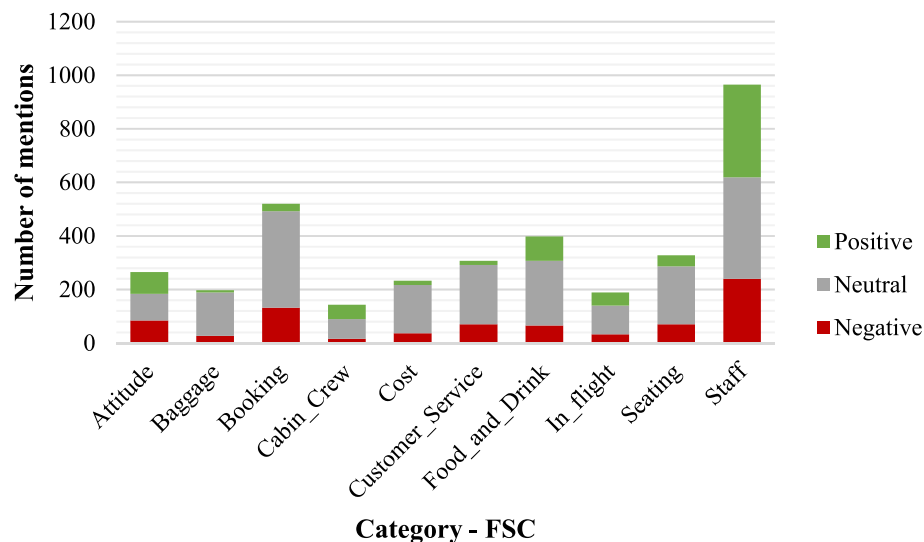**Fig. 8.** Number of mentions, by category for LCCs during the COVID-19 period.

Fig. 9. Number of mentions, by category for FSCs during the COVID-19 period.
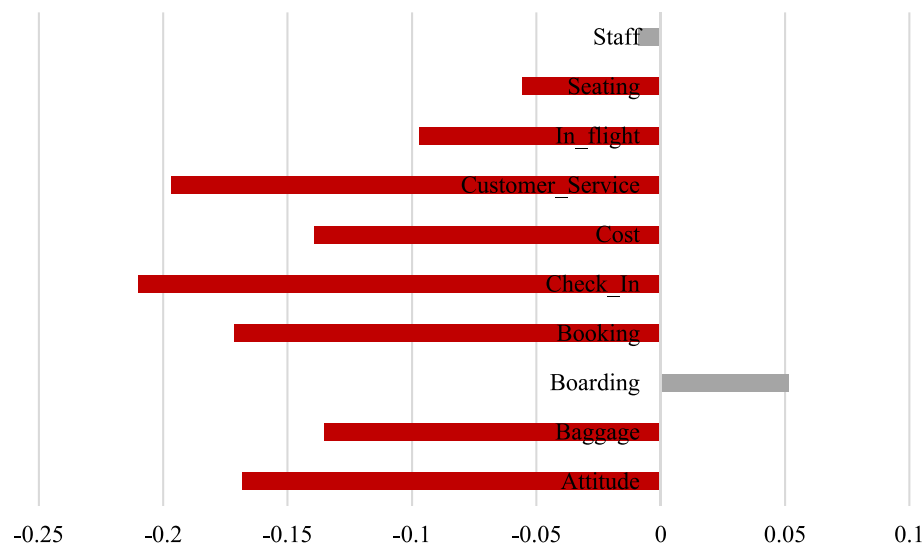

Fig. 10. Mean sentiment polarity, by category for LCCs, during the COVID-19 period.

distribution for a total of 881 occurrences of the COVID-19 period sample (see Table 5) is as follows.

- For the 'Negative' class, we have 673 instances, representing a relative frequency of 76,39%.
- For the 'Positive' class, we have 208 instances, representing a relative frequency of 23,61%.

For a machine learning algorithm to predict with higher accuracy, the dataset must be balanced, which implies that the number of classes must be distributed as evenly as possible among all instances (Zhang et al., 2022). The optimum situation for a binary classification problem would be a 50/50 ratio for each class. The issue with imbalanced datasets arises because the generated model performs better on the most frequent class than the others (applicable for binary and multiple classification problems). Depending on the airline company's strategy, the model will better predict 'Negative' rather than 'Positive' reviews.

There are two processes in supervised machine learning algorithms: training and testing. The training process involves creating a model in which the expected output (in our case, 'Review Sentiment') is known (hence 'supervised'). The training model is then used in the testing

process to make predictions and evaluate the algorithm's performance (Straub, 2021). To implement these supervised algorithms, generally, a dataset is split into two subsets: 'training' and 'testing'. The 'training' subset is used in the training process, while the 'testing' dataset is used in the testing process. Since the main goal is evaluating which comparing algorithms perform the best, all the results below were generated using the 'testing' dataset.

The primary dataset was split in a ¼ ratio, with 75% of all instances used for the training process and the remaining 25% on the testing process, resulting in 221 instances (169 were 'Negative' and 52 were 'Positive'). Table 10 depicts the performance results for all the used algorithms, namely the results for the most common evaluation metrics used on classification problems: Accuracy, Precision, Recall, and F1-Score.

From Table 10, we can verify that the SVM (with 77% accuracy) and LR (with 76% accuracy) were the best-performing algorithms, with SVM slightly outperforming the others. Since the Precision, Recall, and F1-Score metrics have class-wise scores, each algorithm predicts the 'Negative' better than the 'Positive' class because the dataset is imbalanced, confirming the previously discussed scenario. The SVM is the best algorithm for predicting the 'Negative' class, based on the Recall value
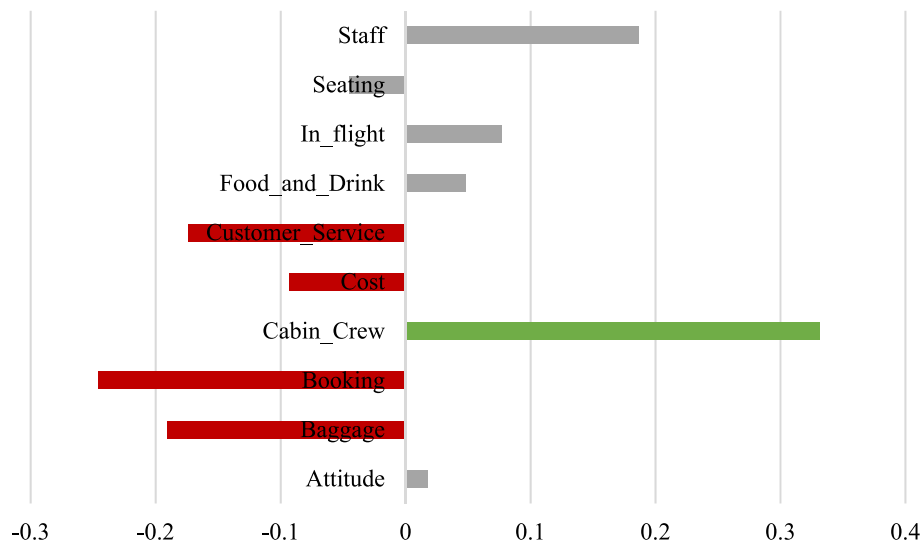
Fig. 11. Mean sentiment polarity, by category for FSCs, during the COVID-19 period.

Table 8
Average rating and sentiment polarity, by travel class, COVID-19 period.

Class	Average rating	Average sentiment polarity
Business Class	4,617,021,277	−0,036,573,094
Economy Class	3,059,065,934	−0,183,293,641
First Class	7	0,098,356,624
Premium Economy	4	−0,04,130,185
Total	3,285,377,358	−0,161,523,444

Table 9
Average cleanliness ranking vs. Skytrax COVID-19 ranking, by airline.

Airline Name	Average Rating	Skytrax COVID-19 Rating
airBaltic	10	5
Vueling Airlines	9	4
Wizz Air	8,666,666,667	3
easyJet	8,4	4
Aegean Airlines	8	4
Lufthansa	7,6	4
British Airways	7,1	4
Ryanair	7	3
KLM Royal Dutch Airlines	5,75	4
LOT Polish Airlines	5,5	3
Air France	5	4
Turkish Airlines	4,666,666,667	4

(99%), which is the ratio of all correct predictions to the sum of all predictions generated by the model on a single class (including false positives and true negatives). A correct prediction is considered when the model predicts a class (for example, 'Negative') and the actual/true value is also 'Negative.' (AlZoman, 2021). Tables 11–13 represent the confusion matrices for all three algorithms, displaying the absolute instance predictions of both class values ('Positive' and 'Negative') when compared to the actual/true class values of those same instances. All the above-stated metrics can be calculated for each algorithm using the values from its confusion matrix. (AlZoman, 2021).

Table 11 shows that SVM performs reliable results at predicting the 'Negative' class (Predicted Negative and Actual Negative - True Negative), with only two incorrect predictions (Predicted Positive and Actual Negative – False Positive). Once there were more instances of the 'Negative' class than the 'Positive' class, SVM did poorly in predicting the 'Positive' class. Table 11 confirms too many (49) False Negatives (Predicted Negative and Actual Positive).

The values from the confusion matrix depicted in Table 12 belonging to the LR algorithm are very similar those from the SVM's confusion matrix from Table 11, with a slightly low performance compared to the SVM algorithm.

Table 13 shows the confusion matrix for the DT algorithm. Analyzing these values, makes it possible to conclude that the DT algorithm was the worst-performant algorithm regarding incorrect predictions vs. correct predictions.

The importance of each feature in the generated models can also be measured, which can be very useful for providing insight into our dataset and machine learning algorithm. The feature importance is "computed as the (normalized) total reduction of the criterion brought by that feature", also known as the Gini importance (Scikit-Learn, 2022). The higher the value, the more important that feature is to our model.

When comparing the DT algorithm with SVM and LR, the feature importance has only positive values (Scikit-Learn, 2022). In this

Table 11
Confusion matrix for SVM algorithm.

Predicted vs. Actual		Predicted	
		Positive	Negative
Actual	Positive	3	49
	Negative	2	167

Table 10
Performance results of the applied machine learning algorithms.

Algorithm	Accuracy	Precision		Recall		F1-Score	
SVM	0.77	Positive	0.60	Positive	0.06	Positive	0.11
		Negative	0.77	Negative	0.99	Negative	0.87
LR	0.76	Positive	0.44	Positive	0.08	Positive	0.13
		Negative	0.77	Negative	0.97	Negative	0.86
DT	0.73	Positive	0.39	Positive	0.29	Positive	0.33
		Negative	0.80	Negative	0.86	Negative	0.83

Table 12
Confusion matrix for **LR** algorithm.

Predicted vs. Actual		Predicted	
		Positive	Negative
Actual	Positive	4	48
	Negative	5	164

Table 13
Confusion matrix for **DT** algorithm.

Predicted vs. Actual		Predicted	
		Positive	Negative
Actual	Positive	15	37
	Negative	23	146

scenario, because every machine learning algorithm only accepts numerical values as input, all the data had to be encoded to represent discrete categorical values as input to our algorithms. In this scenario, every feature was codified using the 'One-Hot Encoding' technique, comprising 96 one-hot encoded features (Seger, 2018). Figs. 12 and 13 show depict the 10 most relevant features for SVR and LR algorithms, respectively (negative and positive coefficients), while Fig. 4 only depicts the five most relevant positive features (see Fig. 14).

By analyzing Figs. 12 and 13, four common positive values appear to be the most relevant features for both SVM and LR algorithms: the traveler from Iceland, Belgium, or India and traveling in the first class. We cannot immediately conclude that travelers with one of these characteristics are more likely to send a positive review (positive coefficient leads to an increased output value, and on binary classification, towards the value of 1 – which is the 'Positive' class). On the same line, from the one common negative feature importance value in SVM and LR, we cannot conclude that travelers in the economic class are more likely to give negative reviews. Additional data and information are required to confirm such assumptions to determine whether there is a pattern directly tied to the travelers' country of origin and travel class.

5. Conclusions and recommendations

This research analyzed online reviews by airline passengers using sentiment analysis, a well-established text mining technique that extracts information hidden in unstructured text (Sezgen et al., 2019). We successfully found what factors affect passengers' satisfaction, i.e., satisfaction dimensions, before and after the COVID-19 pandemic, and the slight differences in passengers flying with LCC, FSC, and different travel classes.

Results suggest passengers were unhappy with the airline industry before the pandemic. This general feeling was aggravated even more during the pandemic.

Satisfaction dimensions were extracted, and it was determined that the most mentioned dimension, before and after the pandemic, concerned staff attitude. The results suggest that staff behavior is the satisfaction dimension that impacts all passenger groups, regardless of the airline's business model.

FSC passengers, before the pandemic, gave importance to the airline's cabin crew. In contrast, LCC passengers emphasized the airline's customer service more. However, FSC and LCC passengers were displeased with topics linked to bookings and baggage. Regarding the type of cabin, we found that Economy Class passengers value and are pleased with the attitude and helpfulness of staff and cabin crew and signs of being cost-conscious. Passengers flying in premium cabins also praise the attitude and helpfulness of cabin crew and staff. They show signs of valuing seat quality, food offerings, and flight experience.

The results suggest that the COVID-19 pandemic brought few changes to how passengers are satisfied. Compared with the pre-COVID-19 period, the overall sentiment became more negative during the pandemic. We also verified some subtle differences in the satisfaction dimensions. Staff remained the principal dimension, but increased mentions regarding bookings and customer service within the FSC passengers were noted. However, the main takeaway is a surge of comments regarding refunds and aircraft cabin cleanliness across all traveling classes. This is expected since the pandemic raised hygiene awareness and caused many flights to be canceled.

We also determined that machine learning methods can make predictions based on non-correlated variables. We used three different algorithms in this study: Support Vector Machines (SVM), Logistic Regression (LR), and Decision Trees (DT). The SVM algorithm made the best predictions regarding accuracy (77%) and Recall for the 'Negative' class (99%). Based on these results, we can conclude that the model generated by this algorithm can be used to predict customer satisfaction, but only for the 'Negative' review sentiment. If an airline wants to provide the best customer satisfaction, it must address positive and negative reviews. Nonetheless, these models can be improved, and more satisfactory results can be obtained by creating a model with more evenly distributed target classes on the used dataset.

For academia, this research contributes to the literature by revealing the factors that influence the satisfaction of airline companies and the factors that influence such satisfaction. Moreover, this study shed some light on how the pandemic affected airline passengers, revealing that they value cleanliness more than they used to and that class and business model influence satisfaction.

From a managerial standpoint, airline companies can benefit from the created knowledge to adjust their strategies according to the created knowledge and meet their customers' expectations. Air companies must deeply understand the customer to ensure business growth and improve service innovation and customer experience (Siering et al., 2018).

Although this study encourages user-generated content, it is vital to highlight some of its limitations. Reviews were collected from only one website (airlinequality.com), which led to limited results. Additionally,
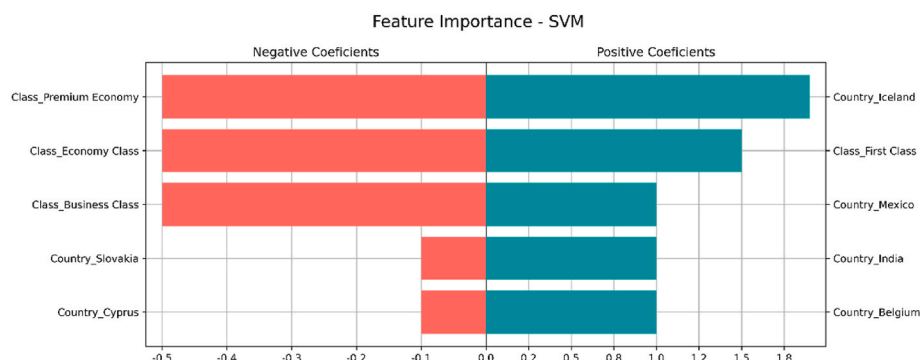

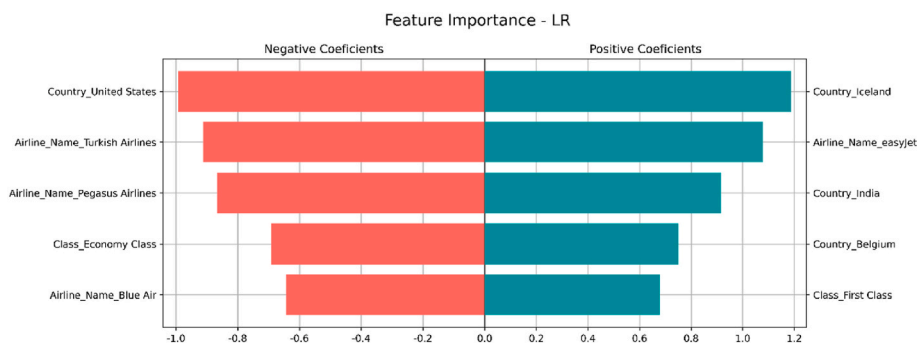
Fig. 12. Variable importance for SVM algorithm.

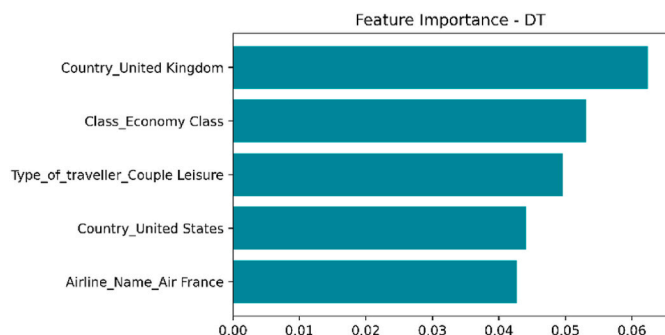**Fig. 13.** Variable importance for LR algorithm.

Fig. 14. Variable importance for DT algorithm.

the sample reviews were written in English, meaning that the opinion of passengers speaking other languages was not considered. That might explain why most of the reviews originated from English-speaking countries. Also, this study focused only on European airlines, and other airlines might pose a different reality. As mentioned throughout the study, sentiment analysis relies on identifying words. The algorithms are not prepared to deal with misspelled words, meaning that those words will not be recognized, and the final sentiment score might not be accurate. Also, the passengers introduced the metadata used to complement the research data. The information is not guaranteed to be correct as they are not subsequently validated. For future work, it is recommended that a similar study be carried out on other airline companies. In addition, other attributes should be considered besides the class of travel and airlines' business model, such as short vs. long-haul passengers or type of travel (leisure/business). To broaden and enrich the collected data, besides the website airlinequality.com, other review websites should be considered to examine if data varies significantly from website to website.

Author Statement

Francisco Pereira: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft preparation, Joana Martinho Costa: Conceptualization, Methodology, Supervision, Writing – review & editing, Ricardo Ramos: Conceptualization, Methodology, Supervision, Writing – original draft Preparation. António Raimundo: Formal analysis, Writing – review & editing

References

Albers, S., Rundshagen, V., 2020. European airlines' strategic responses to the COVID-19 pandemic (January-May, 2020). J. Air Transport. Manag. 87, 101863 https://doi.org/10.1016/j.jairtraman.2020.101863.

AlZoman, R.M., 2021. A comparative study of traffic classification techniques for smart city networks. Sensors.

Amankwah-Amoah, J., 2020. Note: mayday, Mayday, Mayday! Responding to environmental shocks: insights on global airlines' responses to COVID-19.

Transport. Res. E Logist. Transport. Rev. 143, 102098 https://doi.org/10.1016/j.tre.2020.102098.

Barrons. Coronavirus Is Causing Airfares to Rise. https://www.barrons.com/articles/airfare-bargains-coronavirus-cruises-cheap-51589549611.

Bauer, L.B., Bloch, D., Merkert, R., 2020. Ultra Long-Haul: an emerging business model accelerated by COVID-19. J. Air Transport. Manag. 89, 101901 https://doi.org/10.1016/j.jairtraman.2020.101901.

Blodgett, J., Li, H., 2007. Assessing the effects of post-purchase dissatisfaction and complaining behavior on profitability: a Monte Carlo simulation. J. Consumer Satisfaction, Dissatisfaction Complain. Behav. 20.

Chow, C.K.W., 2015. On-time performance, passenger expectations and satisfaction in the Chinese airline industry. J. Air Transport. Manag. 47, 39–47. https://doi.org/10.1016/j.jairtraman.2015.04.003.

Churchill, G.A., Surprenant, C., 1982. An investigation into the determinants of customer satisfaction. J. Market. Res. 19 (4), 491–504. https://doi.org/10.1177/002224378201900410.

Dada, O.A., Olaleye, S.A., Sanusi, I.T., Obaido, G., 2021. COVID-19 and airline refunds: an analysis of flight passengers' reviews in North America. In: 19th International Conference E-Society, 2021.

Devika, M.D., Sunitha, C., Ganesh, A., 2016. Sentiment analysis: a comparative study on different approaches. Proc. Comput. Sci. 87, 44–49. https://doi.org/10.1016/j.procs.2016.05.124.

Food & Wine, 2020. How Airline Food Service Has Changed during the COVID-19 Crisis. https://www.foodandwine.com/news/airline-food-safety-measures-coronavirus.

Forgas, S., Moliner, M.A., Sánchez, J., Palau, R., 2010. Antecedents of airline passenger loyalty: low-cost versus traditional airlines. J. Air Transport. Manag. 16 (4), 229–233. https://doi.org/10.1016/j.jairtraman.2010.01.001.

Furtado, A., Ramos, R.F., Maia, B., Costa, J.M., 2022. Predictors of hotel clients' satisfaction in the Cape Verde islands. Sustainability 14 (5), 2677. https://doi.org/10.3390/su14052677.

Guo, Y., Barnes, S.J., Jia, Q., 2017. Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent dirichlet allocation. Tourism Manag. 59, 467–483. https://doi.org/10.1016/j.tourman.2016.09.009.

Hamada, M.A., Naizabayeva, L., 2020. Decision Support system with K-means clustering algorithm for detecting the optimal store location based on social network events. In: 2020 IEEE European Technology and Engineering Management Summit, E-TEMS 2020. https://doi.org/10.1109/E-TEMS46250.2020.9111758.

Hotle, S., Mumbower, S., 2021. The impact of COVID-19 on domestic U.S. air travel operations and commercial airport service. Transp. Res. Interdisc. Perspect. 9, 100277 https://doi.org/10.1016/j.trip.2020.100277.

Iacus, S.M., Natale, F., Santamaria, C., Spyratos, S., Vespe, M., 2020. Estimating and projecting air passenger traffic during the COVID-19 coronavirus outbreak and its socio-economic impact. Saf. Sci. 129, 104791 https://doi.org/10.1016/j.ssci.2020.104791.

IATA, 2020. COVID-19 Hits January Passenger Demand. IATA. https://www.iata.org/en/pressroom/pr/2020-03-04-03/.

Lacic, E., Kowald, D., Lex, E., 2016. High enough?. In: Proceedings of the 27th ACM Conference on Hypertext and Social Media, pp. 249–254. https://doi.org/10.1145/2914586.2914629.

Lexalytics, 2021a. Industry Packs. Lexalytics. https://www.lexalytics.com/technology/industry-packs.

Lexalytics, 2021b. Semantria Cloud API Text & Sentiment Analysis. Lexalytics. https://www.lexalytics.com/semantria.

Lexalytics, 2021c. Sentiment Analysis Explained. Lexalytics. https://www.lexalytics.com/technology/sentiment-analysis.

Lucini, F.R., Tonetto, L.M., Fogliatto, F.S., Anzanello, M.J., 2020. Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. J. Air Transport. Manag. 83, 101760 https://doi.org/10.1016/j.jairtraman.2019.101760. June 2019.

Mason, K.J., Morrison, W.G., 2008. Towards a means of consistently comparing airline business models with an application to the "low cost" airline sector. Res. Transport. Econ. 24 (1), 75–84. https://doi.org/10.1016/j.retrec.2009.01.006.

Mattila, A.S., 2004. The impact of service failures on customer loyalty: the moderating role of affective commitment. Int. J. Serv. Ind. Manag. 15 (2), 134–149. https://doi.org/10.1108/09564230410532475.

McHugh, M., 2018. Cramér's V coefficient. In: The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. SAGE Publications, Inc. https://doi.org/10.4135/9781506326139.n162.

Monmousseau, P., Marzuoli, A., Feron, E., Delahaye, D., 2020. Impact of Covid-19 on passengers and airlines from passenger measurements: managing customer satisfaction while putting the US Air Transportation System to sleep. Transp. Res. Interdiscip. Perspect. 7, 100179 https://doi.org/10.1016/j.trip.2020.100179.

Moro, S., Ramos, R.F., Rita, P., 2020. What drives job satisfaction in IT companies? Int. J. Prod. Perform. Manag. 70 (2), 391–407. https://doi.org/10.1108/IJPPM-03-2019-0124.

Namilae, S., Srinivasan, A., Mubayi, A., Scotch, M., Pahle, R., 2017. Self-propelled pedestrian dynamics model: application to passenger movement and infection propagation in airplanes. Phys. Stat. Mech. Appl. 465, 248–260. https://doi.org/10.1016/j.physa.2016.08.028.

Namukasa, J., 2013. The influence of airline service quality on passenger satisfaction and loyalty the case of Uganda airline industry. TQM J. 25 (5), 520–532. https://doi.org/10.1108/TQM-11-2012-0092.

Nytimes, 2020. Worried about Crowded Flights? Know where Your Airline Stands. The New York Times. https://www.nytimes.com/2020/07/21/travel/crowded-flights-coronavirus.html.

Oliver, R.L., 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. J. Market. Res. 17 (4), 460. https://doi.org/10.2307/3150499.

Park, J.W., Robertson, R., Wu, C.L., 2004. The effect of airline service quality on passengers' behavioural intentions: a Korean case study. J. Air Transport. Manag. 10 (6), 435–439. https://doi.org/10.1016/J.JAIRTRAMAN.2004.06.001.

Park, S., Lee, J.-S., Nicolau, J.L., 2020. Understanding the dynamics of the quality of airline service attributes: satisfiers and dissatisfiers. Tourism Manag. 81, 104163 https://doi.org/10.1016/j.tourman.2020.104163.

Piccinelli, S., Moro, S., Rita, P., 2021. Air-travelers' concerns emerging from online comments during the COVID-19 outbreak. Tourism Manag. 85, 104313 https://doi.org/10.1016/j.tourman.2021.104313.

Punel, A., Al Hajj Hassan, L., Ermagun, A., 2019. Variations in airline passenger expectation of service quality across the globe. Tourism Manag. 75, 491–508. https://doi.org/10.1016/j.tourman.2019.06.004.

Ramos, R.F., Biscaia, R., Moro, S., Kunkel, T., 2022. Understanding the importance of sport stadium visits to teams and cities through the eyes of online reviewers. Leisure Stud. 1–16. https://doi.org/10.1080/02614367.2022.2131888.

Ramos, Rita, P., Moro, S., 2019. From institutional websites to social media and mobile applications: a usability perspective. Eur. Res. Manag. Bus. Econ. 25 (3), 138–143. https://doi.org/10.1016/j.iedeen.2019.07.001.

Rane, A., Kumar, A., 2018. Sentiment classification system of twitter data for US airline service analysis. Compsac 1, 769–773. https://doi.org/10.1109/COMPSAC.2018.00114.

Rita, P., Moro, S., Cavalcanti, G., 2022a. The impact of COVID-19 on tourism: analysis of online reviews in the airlines sector. J. Air Transport. Manag. 104, 102277 https://doi.org/10.1016/j.jairtraman.2022.102277.

Rita, P., Ramos, R.F., Borges-Tiago, M.T., Rodrigues, D., 2022b. Impact of the rating system on sentiment and tone of voice: a Booking.com and TripAdvisor comparison study. Int. J. Hospit. Manag. 104, 103245 https://doi.org/10.1016/j.ijhm.2022.103245.

Rita, P., Ramos, R.F., Moro, S., Mealha, M., Radu, L., 2020. Online dating apps as a marketing channel: a generational approach. Eur. J. Manag. Bus. Econ. 30 (1), 1–17. https://doi.org/10.1108/EJMBE-10-2019-0192.

Scikit-Learn, 2022. Decision Tree Classifier. From Scikit-Learn.

Seger, C., 2018. An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing. Stockholm, Sweden.

Sezgen, E., Mason, K.J., Mayer, R., 2019. Voice of airline passenger: a text mining approach to understand customer satisfaction. J. Air Transport. Manag. 77, 65–74. https://doi.org/10.1016/j.jairtraman.2019.04.001.

Siering, M., Deokar, A.V., Janze, C., 2018. Disentangling consumer recommendations: explaining and predicting airline recommendations based on online reviews. Decis. Support Syst. 107, 52–63. https://doi.org/10.1016/j.dss.2018.01.002.

Skytrax, 2021a. *About Us | Airline Quality*. Skytrax. https://skytraxresearch.com/about-us/.

Skytrax, 2021b. *COVID-19 Airline Safety Rating*. Skytrax. https://skytraxratings.com/covid-19-airline-safety-ratings.

Sobieralski, J.B., 2020. COVID-19 and airline employment: insights from historical uncertainty shocks to the industry. Transp. Res. Interdiscip. Perspect. 5, 100123 https://doi.org/10.1016/j.trip.2020.100123.

Song, C., Guo, J., Zhuang, J., 2020. Analyzing passengers' emotions following flight delays- a 2011–2019 case study on SKYTRAX comments. J. Air Transport. Manag. 89 (February), 101903 https://doi.org/10.1016/j.jairtraman.2020.101903.

Sotomayor-Castillo, C., Radford, K., Li, C., Nahidi, S., Shaban, R.Z., 2021. Air travel in a COVID-19 world: commercial airline passengers' health concerns and attitudes towards infection prevention and disease control measures. Infect. Dis. Health 26 (2), 110–117. https://doi.org/10.1016/j.idh.2020.11.002.

Straub, J., 2021. Machine learning performance validation and training using a 'perfect' expert system. MethodsX.

Suau-Sanchez, P., Voltes-Dorta, A., Cugueró-Escofet, N., 2020. An early assessment of the impact of COVID-19 on air transport: just another crisis or the end of aviation as we know it? J. Transport Geogr. 86, 102749 https://doi.org/10.1016/j.jtrangeo.2020.102749.

Tsytsarau, M., Palpanas, T., 2012. Survey on mining subjective data on the web. Data Min. Knowl. Discov. 24 (3), 478–514. https://doi.org/10.1007/s10618-011-0238-6.

Tuchen, S., Arora, M., Blessing, L., 2020. Airport user experience unpacked: conceptualizing its potential in the face of COVID-19. J. Air Transport. Manag. 89, 101919 https://doi.org/10.1016/j.jairtraman.2020.101919.

Williams, J., 2020. The Hill. The Hill. December 16. https://thehill.com/policy/transportation/aviation/530441-airline-satisfaction-at-record-high-despite-pandemic/.

Xu, X., Liu, W., Gursoy, D., 2019. The impacts of service failure and recovery efforts on airline customers' emotions and satisfaction. J. Trav. Res. 58 (6), 1034–1051. https://doi.org/10.1177/0047287518789285.

Zeithaml, V.A., 1988. Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. J. Market. 52 (3), 2. https://doi.org/10.2307/1251446.

Zhang, L., Sun, Y., Luo, T., 2016. A framework for evaluating customer satisfaction. In: 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA), pp. 448–453. https://doi.org/10.1109/SKIMA.2016.7916264.

Zhang, Moro, S., Ramos, R.F., 2022. A data-driven approach to improve customer churn prediction based on telecom customer segmentation. Future Internet 14 (3), 94. https://doi.org/10.3390/fi14030094.