

Classification of Business and Leisure Travelers in Hotel Reviews: A Comparative Analysis Using TF-IDF with Logistic Regression and Support Vector Machine

Xinyu Hou

Data Science Institute, Vanderbilt University

DS 5780: Natural Language Processing

Author Note

ChatGPT 4.0 was utilized to correct grammar mistakes and refine word choices for academic writing. This research was supported by the Data Science Institute at Vanderbilt University. Correspondence concerning this article should be addressed to Xinyu Hou, Data Science Institute, Vanderbilt University. Email: xinyu.hou.1@vanderbilt.edu.

Abstract

This study explores the distinct preferences and expectations of leisure and business travelers within the hotel industry by analyzing hotel reviews from booking.com. Although previous research has utilized Term Frequency-Inverse Document Frequency (TF-IDF) to examine various aspects of review data, it has not adequately categorized reviews into business or leisure categories. To address this gap, this research employs logistic regression and Support Vector Machine (SVM) models, leveraging TF-IDF as an analytical tool to quantitatively assess differences in language and thematic emphasis between these two traveler types. Both models are used to classify the reviews, enhancing understanding of how textual nuances correlate with traveler categories. By identifying the better performed model, the study will provide actionable insights for hotels, aimed at enhancing customer satisfaction by catering to the diverse needs of different traveler groups and refining service offerings to assist targeted marketing strategies within the hospitality industry.

Keywords: *hotel review, TF-IDF, Natural Language Processing, marketing targeting, business travel, leisure travel*

1. Introduction

In the hospitality industry, understanding the diverse preferences of various traveler demographics is crucial for delivering personalized services. Leisure and business travelers present different sets of expectations and demands, which are often subtly reflected in the language of their online hotel reviews (Chu, R., & Choi, T., 2000). Such customer feedback offers invaluable insights into guest satisfaction and offer a valuable database for improving service offerings. This research explores the language used in hotel reviews on booking.com to determine whether it can effectively distinguish the top priorities of leisure versus business travelers.

To achieve this, this study applies the Term Frequency-Inverse Document Frequency (TF-IDF) method—a powerful statistical tool that

assesses the significance of words within a collection of hotel reviews. This technique helps identify specific linguistic patterns that correspond to different traveler preferences. Additionally, logistic regression model and SVM model are applied to categorize reviews based on the inferred travel purpose, thereby connecting each review to its respective traveler type. By integrating these two analytical approaches, this research aims to uncover whether leisure and business travelers focus on different aspects of their hotel experiences. This study not only contributes to academic discussions but also offers practical solutions for enhancing customer engagement and service differentiation in the rapidly evolving tourism sector, as expected to provide actionable insights that can help the hospitality industry refine their service offerings and marketing strategies.

2. Research Question

The purpose of this study is to analyze and quantify the different linguistic expressions in online hotel reviews from leisure and business travelers. Using the TF-IDF statistical measure, this research will assess word prominence in these reviews.

Central to this endeavor are the following research questions:

- Is there a discernible difference in the distribution of high TF-IDF score words between hotel reviews posted by leisure travelers and those by business travelers, suggesting distinct areas of emphasis within their respective experiences?
- Can either a logistic regression model or a Support Vector Machine (SVM) model, informed by the TF-IDF analysis, accurately classify hotel reviews into leisure and business categories? Furthermore, how effective are these models as tools for understanding and addressing the specific needs of each traveler type?

3. Literature Review

Segmentation remains a cornerstone strategy in the hospitality industry, crucial for understanding and catering to the diverse needs

of guests. This strategy employs various approaches ranging from benefits sought to behavioral classifications. Seminal works by Goldsmith and Litvin (1999) identify key guest types such as business/convention visitors and vacation/family/leisure travelers, further categorizing them into heavy and light travelers. Moreover, the industry, as characterized by STR, segments guests from a revenue perspective into three principal categories: transient, group, and contract. These categories encompass a variety of rate types and purposes, from individual leisure or business bookings to group and negotiated company rates (Leeuwen & Koole, 2022).

Jurowski and Reich (2000) define the essence of segmentation as the identification of similarities across geographical, behavioral, and demographic features. However, contemporary studies, such as those by Mandhachitara and Gulid (2019), note a shift towards focusing on customer preferences regarding service characteristics and attributes, rather than segmenting customers based on these preferences.

Notably, the behaviors and preferences of these segments diverge significantly. Business and leisure travelers not only exhibit different characteristics but also evaluate service attributes distinctly (Kashyap & Bojanic, 2000; Yavas & Babakus, 2003). For instance, leisure travelers show a marked preference for ancillary services such as cross-reservation with restaurants and entertainment venues, and value personalized touches like fruit baskets or newsletters. Conversely, business travelers prioritize functional hotel features and personalized service, underscoring the importance of exclusive room amenities and expertise of the staff (Aufreiter, Elzinga, & Gordon, 2003). Moreover, business travelers demonstrate less brand loyalty compared to leisure travelers, a phenomenon studied by Skogland and Siguaw (2004).

Such studies provided meaning insights for hotel operators to adjust and tailor their services and make precise promotional strategies. The critical

importance of effectively applying customer knowledge in the development of marketing strategies is well recognized in hotel management (Mandhachitara & Gulid, 2019). This view is supported by the Industry Profile Report, which underscores the significant impact of business and leisure travelers on the global hotel market's demand. The report identifies a typical distribution of 40% business travelers to 60% leisure travelers in the USA as of 2015 (First Research, 2015).

4. Methodology

4.1 Data Preparation and Analysis

This study utilizes a comprehensive dataset from Kaggle, comprising over half a million reviews from 1,493 upscale hotels across Europe, spanning August 2015 to August 2017. The dataset includes a mix of textual and quantitative data that offers insights into hotel guests' satisfaction and experience. Each review comes with detailed metadata such as the hotel's location, review date, reviewer's nationality, travel purpose, stay duration, and geographic coordinates, enabling a rich analysis of consumer preferences.

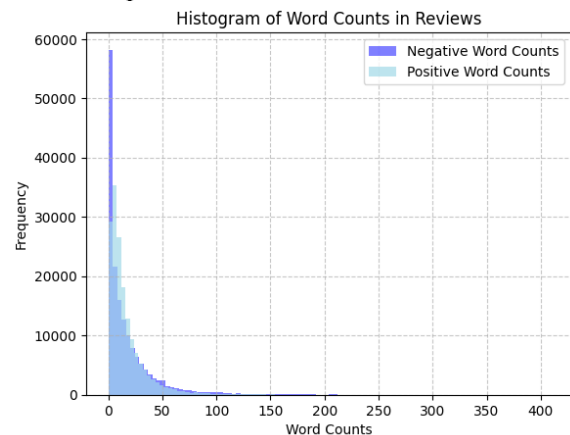


Figure 1, word count of positive / negative reviews

Most reviews contain a smaller number of words, with negative word counts slightly more prevalent in the lower range. The frequency of reviews decreases dramatically as the word count increases, with very few reviews exceeding 50 words in either category. This indicates that reviewers tend to express their opinions concisely, and extensive usage of either negative or positive words is uncommon.

RUNNING HEAD TITLE (SHORTENED)

For effective comparison between leisure and business travelers, reviews were categorized based on the 'tags' column, with labels assigned as 'leisure trip' or 'business trip'. Reviews lacking these specific tags were categorized as 'other' and excluded from the analysis.

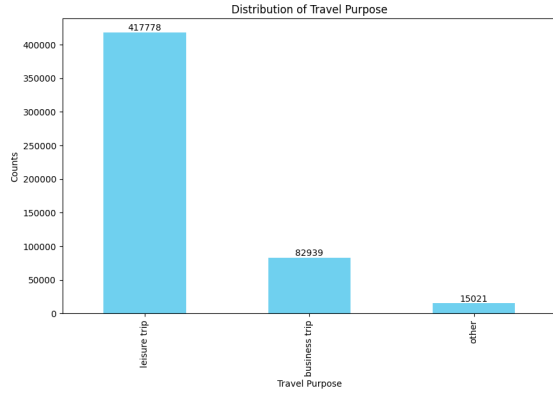


Figure 2, Distribution of travel purposes

Due to a significant imbalance between the number of leisure and business trip reviews, down sampling was implemented. This involved randomly selecting entries from the larger category to match the size of the smaller category, ensuring balanced representation, and enhancing the validity of our logistic regression model and TF-IDF analysis.

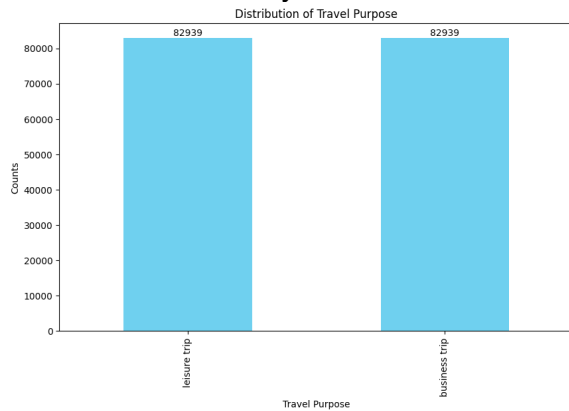


Figure 3, Distribution of travel purposes after down sampling

The dataset distinctly separates positive from negative feedback, thus obviating the need for separate sentiment analysis tools. It also includes metrics like average hotel score, total review count, review history, and additional non-textual ratings, providing a comprehensive temporal analysis of consumer behavior.

4.2 TF-IDF analysis

To process the review text, we utilized the spaCy library for detailed text analysis. The initial step involved tokenizing the text, which means breaking down the reviews into individual words or tokens. We removed all punctuation and whitespace, as these elements do not significantly contribute to the semantic content of the reviews. Lemmatization was then applied, which uses the full vocabulary of a language to perform a morphological analysis on words, reducing them to their lemma or canonical form. For example, the words "colleague" and "colleagues" are reduced to the lemma "colleague". Additionally, stop words—common but uninformative words—were excluded from the analysis.

Building on this refined tokenization process, we employed the Term Frequency-Inverse Document Frequency (TF-IDF) Vectorizer. TF-IDF is a statistical measure used to evaluate the importance of a word within a document in relation to a corpus of hotel reviews. It combines two metrics: term frequency (TF) and inverse document frequency (IDF). Term frequency measures how often a word appears in a review, reflecting its relevance.

$$TF(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d}$$

Inverse Document Frequency measures how important a term is within the corpus. The intuition behind IDF is that terms that appear in many different documents are less significant than terms that appear in fewer documents.

$$IDF(t, D) = \log \left(\frac{\text{Total number of documents } D}{\text{Number of documents with term } t \text{ in them}} \right)$$

The TF-IDF value is calculated by multiplying these two measures together for each term in each document:

$$TF-IDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

This dual consideration highlights words that are not only common in individual reviews but also distinct across the dataset, thereby providing a nuanced metric of a word's relevance and uniqueness. It could effectively quantify the significance of each term, facilitating a focused analysis of linguistic patterns that define guest

experiences. This streamlined approach enables a clearer understanding of what aspects of their stays are most impactful to guests, guiding insights into customer satisfaction and service improvement opportunities.

4.3 Classification Models

Two distinct classification models were utilized to discern the top preferences of business versus leisure travelers. Each model contributes uniquely to our understanding of textual differences in the reviews. We will compare the models' performance and choose the better one for following analysis.

- **Logistic Regression:** This model serves as a fundamental tool in our text classification strategy, identifying whether a review pertains to business or leisure travel. The model has been fine-tuned with a regularization parameter of 0.1, striking a balance between accuracy and model simplicity to prevent overfitting. Additionally, we have set a frequency cut-off, considering only those terms that appear 10 or more times across the dataset. This refinement focuses the model on more significant terms and enhances its interpretability, ensuring a more reliable analysis of the textual data.
- **Support Vector Machines (SVM):** The model applies text classification to high-dimensional datasets using Support Vector Machines (SVM) paired with TF-IDF vectorization. It refines the input data by removing English stop words and disregarding terms with less than 10 occurrences to ensure only relevant features are used. For optimal model calibration, it includes cross-validation and tunes hyperparameters, specifically the regularization parameter 'C' of the SVM as 0.1 to enhance classification performance.

5. Results

5.1 Logistic regression model

The output for the logistic regression model can be summarized as follows:

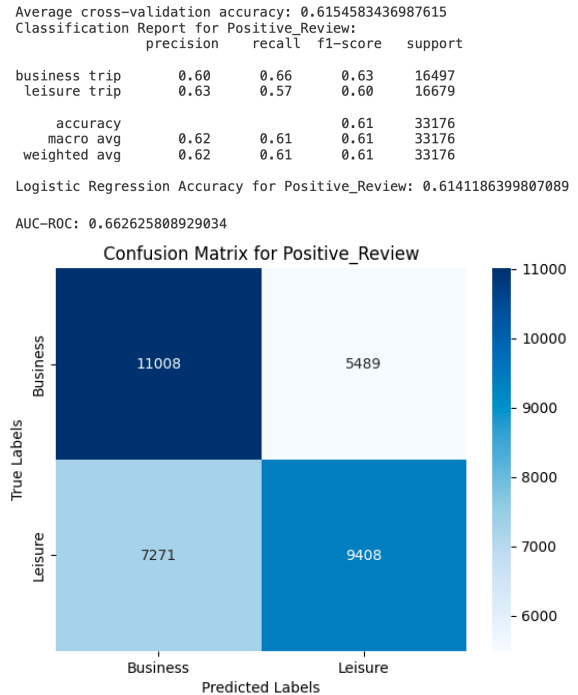


Figure 4, logistic regression matrix for positive reviews

For positive reviews, the logistic regression accuracy is 0.6141

Business trips were correctly identified 60% of the time, and actual business trip reviews were retrieved 66% of the time.

Leisure trips were correctly identified 63% of the time, with a 57% retrieval rate for actual leisure reviews.

Average cross-validation accuracy: 0.579799576745267
Classification Report for Negative_Review:

	precision	recall	f1-score	support
business trip	0.60	0.47	0.53	16496
leisure trip	0.57	0.70	0.63	16680
accuracy			0.58	33176
macro avg	0.59	0.58	0.58	33176
weighted avg	0.59	0.58	0.58	33176

Logistic Regression Accuracy for Negative_Review: 0.5819266939956595
AUC-ROC: 0.6128110102848856

RUNNING HEAD TITLE (SHORTENED)

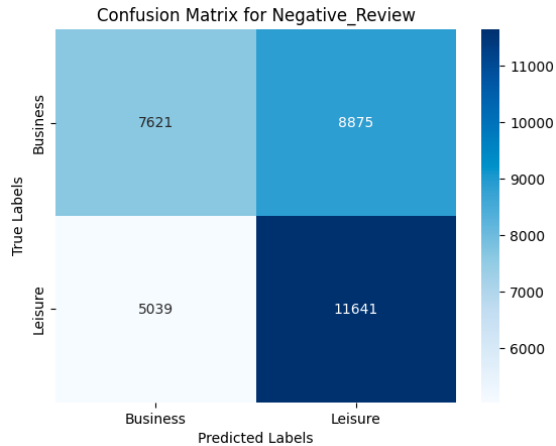


Figure 5, logistic regression matrix for negative reviews

For negative reviews, the logistic regression accuracy is 0.5819.

Business trips were correctly identified 60% of the time, but only 47% of actual business trip reviews were retrieved.

Leisure trips showed a correct identification rate of 57%, with a better retrieval rate of 70% for actual leisure reviews.

5.2 SVM model

Best parameters set found on development set:
{'classifier__C': 0.1}

Classification Report for Positive_Review:

	precision	recall	f1-score	support
business trip	0.61	0.65	0.63	16497
leisure trip	0.63	0.58	0.60	16679
accuracy			0.62	33176
macro avg	0.62	0.62	0.62	33176
weighted avg	0.62	0.62	0.62	33176

SVM Accuracy for Positive_Review: 0.617283578490475

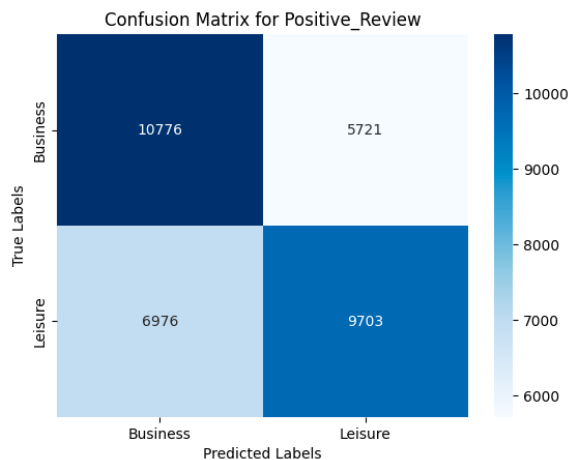


Figure 6, SVM matrix for positive reviews

For positive reviews, the SVM model achieved an accuracy of 0.6173.

Business trips were correctly identified 61% of the time, with 65% of actual business trip reviews being retrieved. F1-score is 0.63, demonstrating the balance between precision and recall for this category.

Leisure trips had a correct identification rate of 63%, with an actual retrieval rate of 58% for leisure reviews. The F1-score stands at 0.60, signaling a slightly less balanced performance than for business trip reviews.

Best parameters set found on development set:
{'classifier__C': 0.1}

Classification Report for Negative_Review:

	precision	recall	f1-score	support
business trip	0.61	0.47	0.53	16614
leisure trip	0.57	0.69	0.62	16562
accuracy			0.58	33176
macro avg	0.59	0.58	0.58	33176
weighted avg	0.59	0.58	0.58	33176

SVM Accuracy for Negative_Review: 0.5823185435254401

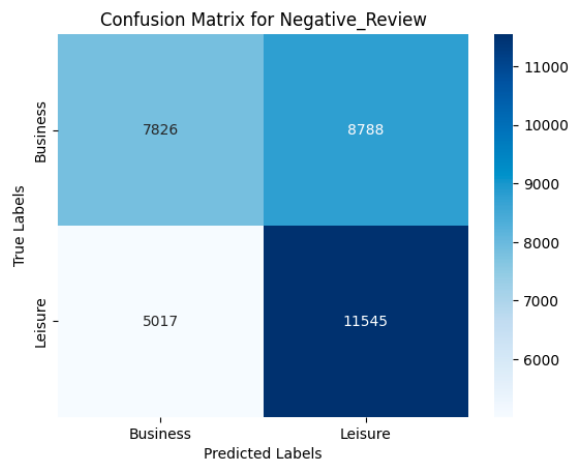


Figure 7, SVM matrix for negative reviews

For negative reviews, the SVM model demonstrated an accuracy of 0.5823. Business trip classifications were correct 61% of the time, while the actual retrieval rate for these reviews was 47%, as indicated by the recall. The F1-score for this category is 0.53, reflecting the balance between precision and recall.

Leisure trip classifications had a precision of 57%, with a higher actual retrieval rate of 69%, suggested by the recall. The F1-score for leisure trips is 0.62, indicating a more balanced

performance in this category compared to business trips.

The SVM and logistic regression models show similar performance levels in this study. While SVM excels at managing complex data structures, its advantages aren't significantly pronounced here. Logistic regression, known for its clarity and interpretability, is easier to manage, making it a preferable choice for the needs of the study. It offers a straightforward and transparent approach that aligns well with the goal of balancing performance with simplicity, facilitating ongoing model improvement, thus we chose logistic regression model for the following analysis.

5.3 TF-IDF Top Coefficients Analysis

With logistic regression model, we conduct a binary analysis to discern the impact of words in classifying hotel reviews into categories for business and leisure travelers. It pulls feature names from a TF-IDF vectorizer and their respective importance from the model's coefficients.

In this context, a negative coefficient implies a word's strong association with business traveler reviews, whereas a positive coefficient signifies a word's relevance to leisure trip reviews. The function highlights the top 20 influential words for both categories by sorting these coefficients.

5.3.1 Top 20 Coefficients for Positive Reviews

	feature_low	coefficient_low	feature_high	coefficient_high
0	business	-7.569266	birthday	4.925858
1	conference	-6.272172	anniversary	4.199584
2	meetings	-4.386453	o2	3.357600
3	meeting	-4.359932	children	3.222310
4	work	-3.888628	concert	3.215276
5	exhibition	-3.431848	family	3.057815
6	excel	-3.271881	theatre	3.043252
7	university	-3.222767	honeymoon	2.933828
8	convention	-2.855681	daughter	2.863869
9	congress	-2.840864	albert	2.792925
10	rai	-2.804377	weekend	2.769899
11	colleagues	-2.596225	attractions	2.504420
12	office	-2.582453	wedding	2.484829
13	colleague	-2.151766	apollo	2.483381
14	ccib	-2.082537	hop	2.435727
15	single	-2.059399	cruise	2.412253
16	vic	-1.857692	husband	2.387783
17	olympia	-1.829579	sightseeing	2.368288
18	embassy	-1.779652	marathon	2.329957
19	iron	-1.749522	partner	2.292662

Table 1, top coefficients for leisure/business trips from positive reviews

The analysis of the top coefficient words from positive reviews distinguishes between terms associated with business and leisure travel.

For business travelers, terms with the most negative coefficients such as "conference," "meeting," and "business" underscore the professional nature of their stays. Notable mentions of "university" and "seminar" likely reflect accommodations for academic or professional development purposes. The presence of words like "colleague," "corporate," and "convention" suggests that networking and corporate events significantly influence positive reviews among this group. The term "iron" intriguingly points towards practical needs, emphasizing the importance of amenities that facilitate a professional appearance.

In leisure travel, positive hotel reviews highlight experiences tied to significant life events and personal relationships. Terms like "anniversary," "honeymoon," and "birthday" suggest a celebration of milestones, while "concert," "cruise," and "marathon" indicate the importance of entertainment and activity-based attractions. Words such as "girlfriend," "daughter," and "wedding" emphasize the intimate, relational aspects of travel.

Additionally, references to specific experiences or venues, hinted at by terms like "youtube" and "O2" as O2 Arena in London, UK show a trend towards shareable, location-based leisure activities. Interestingly, the word 'strawberry' was marked as the 7th highest frequency word. It is unusual for a specific fruit to feature prominently in hotel reviews, which suggests a few possibilities: a special dish, a seasonal festival, a gift amenity (like chocolate covered strawberry) or an iconic hotel feature—that resonates well with guests. This unexpected highlight could represent a distinctive offering or experience provided by the hotel, which may play a part in crafting memorable stays. Collectively, these terms underscore the role of hotels in enriching personal celebrations and leisure experiences.

RUNNING HEAD TITLE (SHORTENED)

This analysis demonstrates how even within positive reviews, the content can reflect different priorities: business travelers mention positive experiences related to the purpose and convenience of their trip, while leisure travelers focus on the enjoyable and recreational attributes of their stay.

5.3.2 Top 20 Coefficients for Negative Reviews

	feature_low	coefficient_low	feature_high	coefficient_high
0	colleague	-4.662106	birthday	3.417892
1	business	-4.560959	children	2.750773
2	meeting	-4.116882	family	2.700390
3	conference	-3.764834	daughter	2.568025
4	invoice	-3.488269	kids	2.299254
5	colleagues	-3.400976	adults	2.260720
6	meetings	-2.672612	attractions	2.208023
7	laptop	-2.553079	weekend	2.197480
8	company	-2.491841	sofa	2.189061
9	receipt	-2.287881	triple	2.184409
10	single	-2.122403	weren	2.167051
11	personnel	-1.875959	friends	2.147414
12	exhibition	-1.872313	child	2.055880
13	prison	-1.830313	concert	2.014563
14	airco	-1.811145	o2	1.974641
15	fitness	-1.794300	robes	1.942783
16	network	-1.784494	anniversary	1.939382
17	gym	-1.773481	cocktails	1.935543
18	shaving	-1.675104	suited	1.912319
19	impossible	-1.633288	luggages	1.848926

Table 2, top coefficients for leisure/business trips from negative reviews

Top 20 coefficients within negative reviews
The analysis of logistic regression coefficients from hotel reviews highlights words that correlate with negative and positive sentiments for business and leisure trips.

Terms like "colleague," "business," "meeting," and "conference" display the most negative coefficients, thereby being strongly indicative of business trip reviews. These keywords highlight common themes in negative reviews for business-related stays, potentially reflecting issues with work-related amenities or environments, such as inadequate meeting facilities or networking spaces. "consecutive" and "untrained" emphasize the high expectation of professionalism and efficiency from business travelers. "Invoice" might point to concerns with billing or expense management — critical factors for business travelers who need to keep precise records for reimbursement. Another

interesting word, "detergent", found in the list suggesting that the cleaning products used were either inadequate or too strong, resulting in an unpleasant stay. It could be a point of inconvenience for business travelers who value efficiency and convenience and may not always have the time to replace forgotten items during a busy trip (Hoang, 2014). 'Interruption' could refer to disruptions during the stay, such as noise, construction, or service disturbances that negatively affected the guest's comfort or ability to relax.

When it comes to leisure travelers, keywords such as "birthday" and "honeymoon," emerge with high coefficients, suggesting that when issues arise during these significant personal events, they are keenly felt and impact guest satisfaction substantially. The presence of terms like "dispute" and "disgusted" underscores instances of profound dissatisfaction, perhaps where services or facilities drastically underperformed against the anticipated celebratory backdrop.

Further analysis uncovers logistical and structural grievances—terms like "downfall," "congestion," and "lounger" signal that guests faced challenges with the hotel's physical capacity or amenities, detracting from the leisure and relaxation sought. Service-related issues are also evident; "blunt" may denote perceived rudeness or lack of warmth in guest interactions, significantly affecting the overall experience.

The term "retrieve" touches upon the functional aspects of guest services, potentially pointing to the frustrations associated with retrieving lost items, a detail that can overshadow an otherwise pleasant stay. "Rambla," possibly referring to the famous street in Barcelona or a similar popular location, indicates unmet location-specific expectations, a factor critical to leisure travelers seeking enriching local experiences.

Overall, while leisure travelers look for pleasure and relaxation, these negative coefficients reflect key areas where their experience was lacking, signaling opportunities for hotels to improve service and facilities for leisure stays. Different

from business travelers that companies will cover their hotel stay, most of the leisure travelers have to pay the hotel expenses from their own pocket, which might lead to a higher expectation on the overall experiences and amenities of a hotel (JTB Business Travel, 2021).

6. Conclusion

The SVM and logistic regression models show similar performance levels in this study. With TF-IDF analysis, this study set out to understand if leisure and business travelers look for different priorities in their hotel stays by analyzing their reviews. The analysis of the top coefficients provides compelling evidence that each group places importance on different aspects of their hotel stays. For business travelers, it emphasizes the preference on functional aspects that support the purpose of their trip, such as facilities for work and meetings. Conversely, terms with high positive coefficients in reviews, aligned with leisure travel, such as "concert" "honeymoon," and "marathon" paint a picture of leisure travelers' experiences. These travelers value elements of enjoyment, accessibility, and amenities that enhance their stay.

In conclusion, the TF-IDF analysis support the hypothesis that leisure and business travelers indeed focus on different aspects of their hotel experiences. These insights affirm the necessity for a nuanced approach to customer service within the hospitality industry—where understanding and addressing the specific expectations of each traveler segment can significantly enhance guest satisfaction and loyalty.

7. Limitation and Future Study

This study utilizes TF-IDF analysis to elucidate the preferences of business and leisure travelers in luxury category hotel reviews within the European context. A limitation arises from this regional focus, which may not reflect global traveler behaviors and general preference, thus constraining the universality of our results. Furthermore, the pre-pandemic timing of the data collection does not account for the profound

changes in travel preferences brought about by COVID-19, which could influence current and future traveler behaviors.

In addition, the decision to equalize the number of reviews for business and leisure trips by downsizing the dataset could also diminish the robustness of our findings, potentially omitting vital insights due to the exclusion of data. Such a reduction may not only limit the model's generalizability but also increase susceptibility to overfitting, risking the loss of critical subtleties within the larger, original dataset.

Future research should aim to extend the analytical framework beyond the European market and pre-pandemic patterns to achieve more widely applicable insights. Incorporating a diverse set of variables—like booking behaviors, stay duration, guest nationalities, occupational information, and seasonal timing of reviews—could enrich the analysis. Utilizing metadata, such as the interval between booking and check-in or user-generated tags, may also refine the predictive accuracy of the models employed.

To enhance model performance, subsequent research could undertake more extensive fine-tuning and consider advanced models like BERT embeddings. However, this study was not able to obtain outputs from training BERT model due to the limited computing capacity. Future studies equipped with the requisite capabilities should explore these avenues to develop a more nuanced understanding of traveler classifications.

8. References

- Ashill, N. J., Rod, M., & Carruthers, J. (2008). The effect of management commitment to service quality on frontline employees' job attitudes, turnover intentions and service recovery performance in a new public management context. *Journal of Strategic Marketing*, 16(5), 437–462. <https://doi.org/10.1080/09652540802480944>
- Aufreiter, N. A., Elzinga, D., & Gordon, J. W. (2003, November 1). *Better branding*. McKinsey & Company.

RUNNING HEAD TITLE
(SHORTENED)

- <https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/better-branding>
- Babakus, E., & Yavas, U. (2003). The Effect of Management Commitment to Service Quality on Employees' Affective and Performance Outcomes. https://www.researchgate.net/publication/246926624_The_Effect_of_Management_Commitment_to_Service_Quality_on_Employees'_Affective_and_Performance_Outcomes
- Chu, R. K. S., & Choi, T. (2000). An importance-performance analysis of hotel selection factors in the Hong Kong hotel industry: A comparison of business and leisure travellers. *Tourism Management*, 21(4), 363-377. [https://doi.org/10.1016/S0261-5177\(99\)00070-9](https://doi.org/10.1016/S0261-5177(99)00070-9)
- H., G. N., Siautama, R., I. A., A. C., & Suhartono, D. (2021). Extractive Hotel Review summarization based on tf/IDF and adjective-noun pairing by considering annual sentiment trends. *Procedia Computer Science*, 179, 558–565. <https://doi.org/10.1016/j.procs.2021.01.040>
- Huang, A. (2014). Understanding the differences between business vs. leisure travelers. E-Marketing Associates. Retrieved April 4, 2024, from <https://www.e-marketingassociates.com/blog/understanding-differences-business-vs-leisure-travelers>
- Jiashen Liu, “515k Hotel Reviews Data in Europe,” Kaggle, accessed [Date], <https://www.kaggle.com/datasets/jiashenliu/515k-hotel-reviews-data-in-europe>.
- Jurowski, C., & Reich, A. Z. (2000). *An Explanation and Illustration of Cluster Analysis for Identifying Hospitality Market Segments*, 24(1). <https://doi.org/10.1177/1096348000024001>
- JTB Business Travel. (n.d.). Business vs. leisure travel: Key differences and considerations. Retrieved April 4, 2024, from <https://www.jtbbusinesstravel.com/business-vs-leisure-travel/>
- Kashyap, R., & Bojanic, D. C. (2000). A structural analysis of value, quality, and price perceptions of business and leisure travelers. *Journal of Travel Research*, 39(1), 45–51. <https://doi.org/10.1177/004728750003900106>
- Leeuwen, R. V., Koole, G., & b. (2022, September 13). Data-driven market segmentation in hospitality using unsupervised machine learning. <https://www.sciencedirect.com/science/article/pii/S2666827022000895>
- Litvin, S. W. (2000, October). Revisiting the heavy-user segment for vacation travel marketing. <https://doi.org/10.1177/13567667000060040>
- Mandhachitara, R., & Gulid, N. (2019). Stretching the performance quality-value-behavioural consequences chain: An empirical study of business and leisure segments in upscale hotels in Thailand. *International Journal of Leisure and Tourism Marketing*, 6(3/4), 279. <https://doi.org/10.1504/ijltm.2019.103565>
- Skogland, I., & Siguaw, J. A. (2004). Are your satisfied customers loyal? *Cornell Hotel and Restaurant Administration Quarterly*, 45(3), 221–234. <https://doi.org/10.1177/0010880404265231>
- STR. (n.d.). How to Use the STAR report. <https://str.com/sites/default/files/2019-07/how-to-read-star-report.pdf>. <https://str.com/sites/default/files/2019-07/how-to-read-star-report.pdf>
- Turner, M. J. (2013). Hotel Management Contracts: A particularly rich research context. *Journal of Hotel & Business Management*, 03(01). <https://doi.org/10.4172/2169-0286.1000e107>

