Article

Analyzing Restaurant Customers' Evolution of Dining Patterns and Satisfaction during COVID-19 for Sustainable Business Insights

Susan (Sixue) Jia

School of Finance and Business, Shanghai Normal University, Shanghai 200234, China; s_jia@shnu.edu.cn

Abstract: Observing and interpreting restaurant customers' evolution of dining patterns and satisfaction during COVID-19 is of critical importance in terms of developing sustainable business insights. This study describes and analyzes customers' dining behavior before and after the pandemic outbreak by means of statistically aggregating and empirically correlating 651,703 restaurant-user-generated contents posted by diners during 2019–2020. Twenty review topics, mostly food, were identified by latent Dirichlet allocation, whereas analysis of variation and rating-review regression were performed to explore whether and why customers became less satisfied. Results suggest that customers have been paying fewer visits to restaurants since the outbreak, assigning lower ratings, and showing limited evidence of spending more. Interestingly, queuing, the most annoying factor for restaurant customers during normal periods, turns out to receive much less complaint during COVID-19. This study contributes by discovering business knowledge in the context of COVID-19 based on big data that features accessibility, relevance, volume, and information richness, which is transferable to future studies and can benefit additional population and business. Meanwhile, this study also provides practical suggestions to managers regarding the framework of self-evaluation, business mode, and operational optimization.

Citation: Jia, S. Analyzing Restaurant Customers' Evolution of Dining Patterns and Satisfaction during COVID-19 for Sustainable Business Insights. *Sustainability* **2021**, *13*, 4981. <https://doi.org/10.3390/su13094981>

Academic Editor: Hong-Youl Ha

Received: 28 March 2021

Accepted: 26 April 2021

Published: 29 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: COVID-19; restaurant customer; satisfaction; user-generated content; latent Dirichlet allocation

1. Introduction

The COVID-19 pandemic has fundamentally changed the world in all walks of life [1–4]. The restaurant industry has been among the hardest hit industries by coronavirus such that daily restaurant demand reduced by 0.06% for every 1% increment in daily new COVID-19 cases [5]. This has posed both threat and opportunity to the sustainability of the restaurant industry. On the threat side, the reduced number of dining-out customers undoubtedly harms restaurants' revenue, whereas the precautionary measures by restaurants may also compromise customer satisfaction [6]. However, on the opportunity side, because customers' need for tasty food, cozy environment, and premium service never actually disappear, restaurants have to be fully prepared for the resumption of customer demand in the long run [7] while sparing no effort to better understand and serve their customers in the short run. It is therefore meaningful to observe and interpret restaurant customers' evolution of dining patterns and satisfaction during COVID-19 to come up with sustainable business insights.

Researchers have been studying the impact COVID-19 on consumers' dining behavior shortly after the outbreak and have proposed preliminary findings. For example, in the cyber space, Mayasari et al. [8] reported that people were submitting fewer Google queries about "restaurant" but searching more about "delivery" and "take-away". In the real world, Yang et al. [5] discovered that both daily new COVID-19 cases and stay-at-home orders had negative impact on restaurant consumption, especially for full-service establishments.

Meanwhile, Kim et al. [9] observed sharp increase in daily food-delivery orders along with growing COVID-19 case and death daily report.

Unfortunately, existing studies have been limited by data accessibility, data relevance, data volume, and information richness such that it could be problematic to transfer the methodology or knowledge from extant studies to future ones and to benefit additional population or business. For example, although foot traffic and card transaction data [5] are extremely reliable, they are not unconditionally open to the research community. Google query data [8], though fully accessible, can only measure customer purchase intention instead of actual purchase. Likewise, Kim and Lee's [10] discovery that customers prefer private dining spaces during COVID-19 was based on questionnaires in a virtual setting. To achieve stronger relevance, Chen et al. [11] obtained user posts on the internet written by customers who had actually visited the restaurants, but the relatively small dataset might weaken the generalizability of related findings. Kim et al.'s [9] large dataset of delivery sales during COVID-19 revealed the hierarchical effect of the pandemic on food-delivery orders in terms of restaurant type, but no mediating effect was addressed because the dataset contains no detailed description regarding how a meal was consumed or whether the meal was satisfactory.

Indeed, obtaining restaurant dining data during COVID-19 with authenticity, great volume, and rich information is a tough task, but it is not without solution. Consumers are more and more willing to share their experience and point of view about restaurants on the internet, collectively referred to as user-generated contents (UGCs). UGC contains important information regarding whether a customer is satisfied with a service as well as the underlying reason [12,13]. For example, a five-star rating surely indicates satisfactory, whereas the accompanying review "the steak is incomparable" provides the explanation. Under this rationale, Jia [14,15] exhibited the feasibility and achievement of quantifying the causation between restaurant rating and review to discover the knowledge hidden within.

Based on the necessity of observing and interpreting restaurant customers' evolution of dining patterns and satisfaction during COVID-19 as well as the availability of online UGCs and analyzing tools, this study aims at describing and understanding customers' dining behavior before and after the pandemic outbreak by means of statistically aggregating and empirically correlating restaurant UGCs posted by diners during 2019–2020. More specifically, two key research questions are examined:

RQ1: Did customers visit restaurants less frequently because of COVID-19?

RQ2: Were customers less satisfied with restaurants because of COVID-19? If so, what were the reasons?

The remainder of this paper is thus organized as follows. The literature review section provides a brief overview of UGC analysis techniques and concerns. The data and method section describes in detail the source of UGC, the processing of review data, and the modeling approach. The results section reports the statistical and empirical findings with the discussion section responding with explanations. The final section concludes the paper with future study suggestions.

The contribution of this study is two-fold. On one hand, this study reveals the change of dining patterns and satisfaction of restaurant customers due to COVID-19 and its underlying mechanism. It complements existing meal service literature with rich facts and findings uncovered during COVID-19, a crisis situation. On the other hand, this study provides instant and practical suggestions to restaurant managers, helping them reflect on their pain and gain during the pandemic and getting them better prepared for the upcoming resumption of customer dining-out demand.

2. Literature Review

2.1. UGC Analysis Techniques

Online rating and review are the two most typical forms of UGC [16]. Rating, usually based on an interval scale, quantifies whether and to what extent a consumer is satisfied with a service [17]. Review is rather a textual comment, describing the service experience of

the consumer in a qualitative style [18] that can explain why the consumer is happy or not. Rating and review together construct a complete description of customer satisfaction [19].

A prerequisite for analyzing online review is transforming the natural language within into structured data. Some studies use manual coding [20–22] or manually create a dictionary for subsequent automatic coding [23–25], both involving substantial human labor. Alternatively, the advancement in text-mining technology has increased the chances of automatically discovering the major topics in the reviews [26]. Compared with manual topic modeling, automatic topic modeling not only has the advantage of being faster when processing larger data sets, but it also incorporates less human intervention, resulting in more objective results. Latent Dirichlet allocation (LDA), a state-of-art topic modeling tool, is both capable and suitable to automatically perform the topic identification task [27–29] such that a qualitative review is thereby quantifiable by calculating the weight of topics within [30,31].

2.2. Concerns Over Fake UGC

UGC analysis is not without limitations. Perhaps the biggest concern that disturbs researchers is UGC fraud, or fake UGC, which is generated by individual fraudsters or professionally organized review campaigners for “face lift” or other purposes [32]. Fake UGCs may undermine the credibility of a UGC study [33]. Fortunately, UGC websites are now capable of automatically detecting fake UGCs thanks to advanced filtering technology [34]. Besides, extant UGC studies have agreed that fake UGCs have limited impact on the detection of major topics and satisfaction factors from UGC [19,35,36].

3. Data and Method

3.1. Data Source

This study obtained UGC data from Dianping.com (accessed on 19 April 2021), a Chinese online review community website where a number of important researches gathered data [12–15]. Furthermore, this study focused on the top 100 restaurants in Shanghai, China, with the most UGC posted up to 1 January 2021, as this could help eliminate the effect of regional difference as well as achieve best research cost efficiency [12–14]. One restaurant was removed from the top 100 list because it opened in 2020, which meant missing data in 2019. Regarding the remaining 99 restaurants, the 651,703 UGCs posted during 2019–2020 were obtained from Dianping.com (accessed on 19 April 2021). The sampling period was chosen as such because 1 January 2021 was a key time point for Chinese citizens regarding COVID-19, around which people began to know about the pandemic, then considered an epidemic, and the government started to enact stay-at-home decrees. Therefore, the first sample year 2019 was regarded as not influenced by COVID-19, whereas the second sample year 2020 was treated as influenced by COVID-19.

3.2. Ethical Considerations

This study was carefully performed to preserve the rights of the UGC website, the restaurants, and the UGC posters. Using online UGC data for academic research objectives has been an appropriate and usual practice. Yet, the data acquiring procedure was cautiously conducted to minimize the network burden on the website server. In the meantime, the brands of the restaurants were masked in order not to appreciate one restaurant over another. This study rather intends to uncover knowledge from the collective ensemble. Finally, the identities of the studied UGC posters were anonymized.

3.3. Data Processing

A typical UGC contains four pieces of information, namely: date, price per customer, rating, and review. The date of UGC helps identify whether the dining happened in 2019 or 2020. It is assumed the date difference between restaurant visiting and UGC posting is negligible. Price per customer is calculated by dividing a restaurant bill with the number of diners. Rating is an interval scale assigned by the UGC poster, with 5.0 indicating extremely

satisfied and 0.5 indicating totally unsatisfied. The above three pieces of information are all structured data that can be directly incorporated into statistical and empirical analysis.

Reviews, or textual comments, require additional data processing to convert into structured data. The reviews, mostly written in Chinese, were first broken into sentences and further into words using Jieba [37], a Chinese language processing tool. Before counting the frequency of each word in all reviews, words that carry little meaning were removed, such as “I” and “today”. Words that had a frequency of less than 3258 were also removed because they appeared on average in only 0.5% of the reviews which means that they were of limited importance. The remaining high frequency meaningful words were grouped into clusters, or topics, in order to reduce the dimension of the analysis. Latent Dirichlet allocation (LDA) [38,39] was applied to handle the clustering task such that words of similar meanings or high correlations were aggregated to form topics. Finally, weight of topic (WOT) was calculated per Equation (1) such that a review is quantified by the number of topics within and the relative frequency of each topic.

$$WOT_{ij} = \frac{TF_{ij}}{\sum_i TF_{ij} + 1} \quad (1)$$

where WOT_i is the weight of topic i in review j , and TF_{ij} is the frequency of topic i in review j . A “1” was added to the denominator in case there was a review that did not mention any topic. The data and variables of this study are thus summarized in Table 1.

Table 1. Data and variables, UGC level.

Data	Variable	Note
Date of UGC	Year	0, if the UGC was posted in 2019 1, if the UGC was posted in 2020
Price per customer	Price	In RMB, 1 RMB $\approx$ 0.15 USD
Rating	Rating	Interval scale, 0.5–5.0
Review	WOT	Per Equation (1)

3.4. Model

Analysis of variance (ANOVA) was applied to statistically explore whether customers paid fewer visits to the restaurants, whether price per customer declined, and whether the customers were less satisfied after the COVID-19 outbreak. To quantify the causation between rating and review, Equation (2) was proposed.

$$Rating_j = \alpha_0 + \sum_i \alpha_i WOT_{ij} + \beta_0 Year_j + \sum_i \beta_i WOT_{ij} Year_j + \varepsilon \quad (2)$$

where $Rating_j$ is the rating assigned by the UGC poster along with review j ; $Year_j$ is the year when review j was posted; and α_0 , α_i , β_0 , and β_i are the coefficients to be estimated. A positive α_i suggests the users are relatively satisfied with topic i [40], such as food or service. For similar reasons, a negative β_i indicates the users become less satisfied with topic i after the COVID-19 outbreak.

4. Results

4.1. Descriptive Statistics

The descriptive statistics of this study is shown in Table 2. Panel 1 aggregates the data on the restaurant level. Averagely, 42% of the visits happened in 2020, indicating a decline in annual visits from 2019. The price ranges from 30.23 RMB (a beverage store) to 422.54 RMB (a Japanese restaurant). The standard deviation for price per customer is also as large as 64.67 RMB comparing with the average value of 127.33 RMB. This suggests the necessity of addressing the hierarchical effect of restaurant price level. The mean rating is 4.34 with a small standard deviation of 0.17. This means the customers are generally satisfied with the studied restaurants, echoing the fact that the restaurants have more UGC

posted than other restaurants. Panel 2 reports the statistics on the UGC level, which is largely in agreement with Panel 1. The standard deviation of rating is as high as 0.83, making it possible and meaningful to explain the variation of customer satisfaction with review information.

Table 2. Descriptive statistics.

Variable	Mean	Standard Deviation	Max	Min
Panel 1. Restaurant level. All variables were calculated as restaurant average.				
Year	0.42	0.09	0.71	0.09
Price	127.33	64.67	422.54	30.23
Rating N = 99	4.34	0.17	4.89	3.76
Panel 2. UGC level.				
Year	0.42	0.49	1.00	0.00
Price	128.59	75.51	1000.00	1.00
Rating N = 651,703	4.34	0.83	5.00	0.50

4.2. Review Topic Identification

From the 651,703 reviews, high frequency meaningful words were first extracted by a word counting program, and then clustered by the unsupervised LDA. Altogether, 20 topics were identified, as are enumerated in Table 3. Table 3 has three panels. Panel 1 contains topics that provide information on things before a meal. Users wrote in reviews about friends or the occasion of a meal, about making appointment or otherwise having to wait for seat, and about the environment of a restaurant. It is noted here that the naming of each topic based on the word cluster, as well as the grouping of topics, were manually conducted, which is a common practice [12,13,15]. Panel 2 contains topics that provide information on the meal itself. As many as 13 food topics were identified, ranging from specific food to food styles, suggesting users' primary focus of a restaurant is food. Panel 3 contains topics that provide information on things after a meal, including services (this actually happens during the meal, but is put here for categorizing purpose), discount, and cost performance.

Table 3. Identified topics.

Topic	High Frequency Meaningful Words						
Panel 1. Topic group A: Before the meal.							
1. Friend	Friends (89,677)	Tooth (30,856)	Color (19,340)	Taste (318,279)	Hotpot (70,609)	Services (188,876)	Environment (145,046)
	Next time (47,497)	Get-together (13,947)	Smile (11,794)	Flavor (100,075)	Lo mein (23,983)	Colleagues (15,656)	Waiter (107,553)
	Eat hot pot (14,285)	Party (7884)	Place (30,174)	Candy (6267)	Cuisine (134,040)	Company (7276)	
2. Appointment	Queuing (236,235)	Advance (35,418)	Hours (64,874)	Appointment (16,991)	Wait (43,799)	Friends (89,677)	Time (41,350)
	Take a number (14,116)	Radio (10,850)	Hotpot (70,609)	Location (40,923)	Evening (38,186)	Lane (9836)	Online (9352)
	Afternoon (26,073)	Door (29,663)	Heated (12,281)	Internet celebrity (22,463)	Line (13,367)	Weekend (21,261)	

Table 3. Cont.

Topic		High Frequency Meaningful Words					
3. Queuing	Queuing (236,235)	Hours (64,874)	Noon (40,346)	Working day (30,430)	Wait (43,799)	Time (41,350)	Evening (38,186)
	Afternoon (26,073)	Minutes (20,674)	Pandemic (17,742)	Door (29,663)	Period (11,719)	Weekend (21,261)	Speed (20,687)
	Queue (9209)	Internet celebrity (22,463)	Popular restaurant (17,323)	Order (10,337)	Morning (6077)	Line (13,367)	
4. Environment	Environment (145,046)	Location (40,923)	Decoration (31,314)	Mint (25,327)	Place (30,174)	Services (188,876)	Seats (17,414)
	Easy to find (9534)	Subway (8814)	Clean (14,720)	Comfortable (11,466)	Waiter (107,553)	Traffic (6158)	Subway line No. (6138)
	Noisy (5630)	Crowded (5223)	Queuing (236,235)	Restaurant (8381)	Lights (5112)	Spacious (4251)	
Panel 2. Topic group B: The meal.							
5. Spicy chicken	Spicy (63,269)	Chicken (35,652)	Taste (318,279)	Potatoes (18,098)	Cuisine (134,040)	Tasty (44,997)	Garlic (12,572)
	Flavor (100,075)	Chicken feet (19,117)	Copper pot (8526)	Chicken meat (23,905)	Beef (84,851)	Braised in oil (6568)	Earthen bowl (6900)
	Oil (18,797)	Glutinous (24,181)	Slightly spicy (13,033)	French toast (4696)	Rice wine (6364)	String (10,745)	
6. Chicken	Taste (318,279)	Fish (25,994)	Tofu (26,085)	Bullfrog (101,224)	Cuisine (134,040)	Chicken wings (11,785)	Egg yolk (12,112)
	Tasty (44,997)	Tender (61,118)	Flavor (100,075)	Chicken feet (19,117)	Sweet and sour (7611)	Roast (26,464)	Meat (71,411)
	Mouthfeel (68,455)	Stone pot (6244)	Tenderloin (5724)	Roast chicken (8647)	Salted egg (17,782)	Meat quality (19,373)	
7. Duck	Taste (318,279)	Roast duck (15,048)	Cuisine (134,040)	Porridge (17,313)	Sweet (31,610)	Duck (17,873)	Flavor (100075)
	Osmanthus (7778)	Sweet fermented- rice (6892)	Sticky rice (8591)	Braised pork (6318)	Crab meat (21,336)	Glutinous (24,181)	Salted duck (6216)
	Mouthfeel (68,455)	Must order (49,094)	Decoct (19,496)	Salty (18,255)	Sauce (24,603)	Food stalls (7271)	
8. Beef hotpot	Beef (84,851)	Tender (61,118)	Tripe (24,014)	Paste (25,484)	Hotpot (70,609)	Fresh (69,639)	Shrimp (58,748)
	Tribute (13,265)	Meatballs (12,421)	Boiled (24,644)	Duck blood (15,395)	Meat (71,411)	Mouthfeel (68,455)	Bullfrog (101,224)
	Hotpot condiment (51,776)	Must order (49,094)	Taste (318,279)	Fresh duck blood (8632)	Fat beef (12,426)	Spicy (63,269)	
9. Bullfrog hotpot	Bullfrog (101,224)	Hotpot condiment (51,776)	Pot (35,832)	Spicy (63,269)	Taste (318,279)	Hotpot (70,609)	Frog (15,339)

Table 3. Cont.

Topic	High Frequency Meaningful Words						
	Tomato (19,315)	Side dish (13,321)	Two-flavor hotpot (12,635)	Boiled (24,644)	Tasty (44,997)	Ice jelly (29,986)	Beautiful frog (7830)
	Slightly spicy (13,033)	Borscht (7406)	Spicy pot (10,553)	Quantity (48,424)	Charcoal fire (5170)	Soup (36,860)	
	Coffee (26,780)	Taste (318,279)	Crab cream (17,501)	Noodle (21,988)	Soy sauce noodles (11,049)	Crab meat (21,336)	Tea (18,757)
10. Crab	Baking (6704)	Crab (19,284)	Yellow croaker (8321)	Beer (7614)	Workshop (5626)	Coffee beans (5552)	Cake (26,443)
	Fish (25,994)	Bread (22,689)	Dessert (21,494)	Flavor (100,075)	Experience (22,426)	Fish noodles (4513)	
	Fresh (69,639)	Shrimp (58,748)	Ingredients (28,870)	Seafood (29,395)	Buffet (17,550)	Fruit (17,204)	Variety (15,172)
11. Seafood	Roast (26,464)	Crab (19,284)	Sashimi (8129)	Salmon (12,977)	Crayfish (7347)	Sea urchin (8233)	Ice cream (15,356)
	Taste (318,279)	Turtle (7220)	Beverage (22,673)	Liver (8558)	Quality (10,348)	Abalone (6388)	
	Cuisine (134,040)	Taste (318,279)	Xinjiang (12,976)	Mutton shashlik (11,828)	Quantity (48,424)	Large plate chicken (9691)	Mutton (11,278)
12. Xinjiang cuisine	Pack (16,670)	Next time (47,497)	Roast (26,464)	Milk tea (24,782)	Yogurt (6408)	Can't eat it all (9123)	Flavor (100,075)
	Cost performance (67,240)	Weight (17,489)	Must order (49,094)	Performance (7062)	Waiter (107,553)	Waste (6308)	
	Shrimp dumplings (24,791)	Red rice bowl (23,437)	Chicken claw (20,352)	Taste (318,279)	Fried dough sticks (16,147)	Steaming (16,058)	Sea shrimp (14,140)
13. Cantonese cuisine	Gold medal (14,330)	Jin Sha (11,463)	Shrimp meat (23,848)	Belly (10,910)	Ribs (11,118)	Mouthfeel (68,455)	Barbecued pork bun (9616)
	Snacks (28,442)	Rice rolls (7767)	Baby porridge (8248)	Must order (49,094)	Pastry (9231)	Pigeon (9958)	
	Rice (35,012)	Taste (318,279)	Oil chicken (16,353)	Pineapple bread (13,856)	Pineapple (18,056)	Milk tea (24,782)	Leaky milk (12,481)
14. Hong Kong cuisine	Salted egg (17,782)	Pork chop (11,591)	Black pepper (11,701)	Decoct (19,496)	Shrimp meat (23,848)	Soy (10,633)	Fried (24,427)
	Tender (61,118)	Hong Kong style (16,847)	Sweet (31,610)	Exclusive (6848)	Tea restaurant (16,152)	Greasy (25,533)	
	Curry (25,133)	Taste (318,279)	Soup (36,860)	Tom Yam Gong (15,527)	Thai style (14,011)	Mango (11,585)	Seafood (29,395)
15. Thai cuisine							

Table 3. Cont.

Topic	High Frequency Meaningful Words						
16. Western cuisine	Shrimp cake (11,050)	Rice (35,012)	Beef brisket (7306)	Flavor (100,075)	Water spinach (9388)	Coconut milk (6291)	Cuisine (134,040)
	Slice (9310)	Lobster (9321)	Glutinous rice (6251)	Coconut juice (7935)	Sauce (24,603)	Rich (24,795)	
	Cheese (45,741)	Beef (84,851)	Hamburger (36,148)	French fries (24,043)	Cake (26,443)	Milkshake (18,055)	Taste (318,279)
17. Dessert	Bread (22,689)	Burger (12,895)	Mushroom (12,730)	Sauce (24,603)	Cream (10,648)	Steak (20,976)	Chocolate (13,812)
	Mouthfeel (68,455)	Pizza (9601)	Greasy (25,533)	Limit (8521)	Strawberry (7599)	Salad (10,135)	
	Crisp (58,364)	Meat (71,411)	Glutinous rice cake (19,552)	Ice jelly (29,986)	Brown sugar (17,209)	Taste (318,279)	Butterfly (13,667)
Panel 3. Topic group C: After the meal.	Moon cake (10,079)	Queuing (236,235)	Fresh meat (7896)	Mouthfeel (68,455)	Sweet green rice ball (4748)	Glutinous (24,181)	Crispy (11,017)
	String (10,745)	Sweet (31,610)	Egg (10,252)	Fragrance (13,482)	Decoct (19,496)	Tender (61,118)	
18. Services	Services (188,876)	Waiter (107,553)	Environment (145,046)	Bell (21,403)	Attitude (19,811)	Little sister (18,953)	Brother (12,257)
	Active (11,016)	Service attitude (20,263)	Considerate (11,190)	Lo mein (23,983)	Mint (25,327)	Speed (20,687)	Guest (10,877)
	Thoughtful (6594)	Laddie (7331)	Personnel (5631)	Shop assistant (10,868)	Auntie (10,126)	Cuisine (134,040)	
19. Discount	RMB (32,437)	Dine-and-dasher (22,158)	Friends (89,677)	Cost-effective (13,975)	Activities (10,726)	Coupon (9470)	Cost performance (67,240)
	Snacks (28,442)	Group purchase (9015)	Birthday (10,981)	Free (11,516)	Noon (40,346)	Experience (22,426)	Voucher (7497)
	Offer (7436)	Next time (47,497)	Queuing (236,235)	Parking (5897)	Waiter (107,553)	Time (41,350)	
20. Cost performance	Price (54,334)	Taste (318,279)	Cost performance (67,240)	Flavor (100,075)	Cuisine (134,040)	Affordable (18,585)	Queuing (236,235)
	Cheap (17,768)	Expensive (16,601)	Environment (145,046)	Per capita (12,417)	Quantity (48,424)	Snacks (28,442)	Next time (47,497)
	Exquisite (8157)	Services (188,876)	RMB (32,437)	Friends (89,677)	Hong Kong style (16,847)	Weight (17,489)	

The above topics have constructed a general view of what a diner cares about concerning a restaurant. Considering the fact that among the identified topics are similar item pairs

such as appointment and queuing and discount and cost performance and the fact that there are a great number of food topics, the current result has been sufficient for subsequent analysis without further increasing the number of topics.

Based on the identified topics, WOT was calculated for each review. Table 4 reports the statistics of WOT on the UGC level. The mean values of the majority of WOT_i are around 0.050, a value suggesting generally even distribution of topics in reviews, albeit some topics are more frequently mentioned than others, e.g., $WOT_{20} > WOT_{14}$ ($p < 0.001$).

Table 4. Statistics on weight of topics, UGC level.

Topic Name	Variable	Mean	Standard Deviation	Max	Min
1. Friend	WOT_1	0.069	0.064	0.966	0.000
2. Appointment	WOT_2	0.048	0.064	0.833	0.000
3. Queuing	WOT_3	0.042	0.064	0.800	0.000
4. Environment	WOT_4	0.049	0.059	0.800	0.000
5. Spicy chicken	WOT_5	0.041	0.050	0.923	0.000
6. Chicken	WOT_6	0.047	0.049	0.857	0.000
7. Duck	WOT_7	0.040	0.049	0.857	0.000
8. Beef hotpot	WOT_8	0.054	0.063	0.933	0.000
9. Bullfrog hotpot	WOT_9	0.042	0.060	0.857	0.000
10. Crab	WOT_{10}	0.037	0.069	0.944	0.000
11. Seafood	WOT_{11}	0.036	0.066	0.875	0.000
12. Xinjiang cuisine	WOT_{12}	0.050	0.054	0.909	0.000
13. Cantonese cuisine	WOT_{13}	0.031	0.061	0.923	0.000
14. Hong Kong cuisine	WOT_{14}	0.030	0.050	0.833	0.000
15. Thai cuisine	WOT_{15}	0.036	0.050	0.950	0.000
16. Western cuisine	WOT_{16}	0.037	0.072	0.952	0.000
17. Dessert	WOT_{17}	0.048	0.061	0.917	0.000
18. Services	WOT_{18}	0.043	0.066	0.889	0.000
19. Discount	WOT_{19}	0.048	0.059	0.889	0.000
20. Cost performance	WOT_{20}	0.083	0.064	0.857	0.000
$N = 651,703$					

4.3. Analysis of Variance

ANOVA was performed to test whether customers paid fewer visits to the restaurants, whether price per customer declined, and whether the customers were less satisfied after the COVID-19 outbreak. Table 5 reports that COVID-19 did reduce yearly review posting from 3830 to 2753 per restaurant ($p < 0.001$). This suggests that customers' visits to restaurants may also suffer a decline of 28% due to COVID-19. Likewise, customer satisfaction significantly dropped from 4.39 to 4.27 ($p < 0.001$). Unlike N or $Rating$, $Price$ only slightly changed by an increase of 2.3 ($p < 0.05$) and was not significant from an economic point of view.

Table 5. Analysis of variance, restaurant level.

Variables	2019		2020		p -Value (Two-Tailed t -Test)
	Mean	Standard Deviation	Mean	Standard Deviation	
N^*	3830	3509	2753	2510	0.000
Price	126.5	64.5	128.8	66.0	0.017
Rating	4.39	0.17	4.27	0.19	0.000

* N denotes the number of reviews posted regarding a restaurant in the sample year.

4.4. Rating-Review Regression

Multi-linear regression was conducted to quantify the causation between rating and review, with ordinary least squares applied to estimate the coefficients in Equation (2).

Table 6 summarizes the regression results. It is noted that although a “0.05” was introduced into Equation (1) such that Equation (2) did not suffer from perfect collinearity by ΣWOT_i , WOT_1 (i.e., Topic 1. Friend) was excluded from regression to further reduce the risk from multi-collinearity [40,41].

Table 6. Rating-review regression, UGC level.

Topic Name	Variables	Rating			
		Model 1	Model 2	Model 3	Model 4
1. Friend	WOT_1	—	—	—	—
2. Appointment	WOT_2	0.163 *** (0.022)	0.174 *** (0.022)	0.122 *** (0.029)	0.133 *** (0.029)
3. Queuing	WOT_3	−0.699 *** (0.022)	−0.707 *** (0.022)	−0.822 *** (0.03)	−0.831 *** (0.03)
4. Environment	WOT_4	−0.152 *** (0.022)	−0.148 *** (0.022)	−0.275 *** (0.029)	−0.272 *** (0.029)
5. Spicy chicken	WOT_5	−0.214 *** (0.023)	−0.222 *** (0.023)	−0.266 *** (0.031)	−0.274 *** (0.031)
6. Chicken	WOT_6	−0.080 ** (0.024)	−0.085 *** (0.024)	−0.136 *** (0.032)	−0.141 *** (0.032)
7. Duck	WOT_7	−0.364 *** (0.023)	−0.370 *** (0.023)	−0.410 *** (0.03)	−0.417 *** (0.03)
8. Beef hotpot	WOT_8	0.654 *** (0.022)	0.661 *** (0.022)	0.759 *** (0.03)	0.766 *** (0.03)
9. Bullfrog hotpot	WOT_9	0.262 *** (0.022)	0.259 *** (0.022)	0.157 *** (0.029)	0.153 *** (0.029)
10. Crab	WOT_{10}	0.292 *** (0.016)	0.288 *** (0.016)	0.252 *** (0.021)	0.246 *** (0.021)
11. Seafood	WOT_{11}	0.107 *** (0.018)	0.144 *** (0.018)	0.022 (0.023)	0.059 * (0.023)
12. Xinjiang cuisine	WOT_{12}	−0.015 (0.021)	−0.017 (0.021)	−0.06 * (0.028)	−0.062 * (0.028)
13. Cantonese cuisine	WOT_{13}	0.246 *** (0.018)	0.242 *** (0.018)	0.133 *** (0.022)	0.129 *** (0.022)
14. Hong Kong cuisine	WOT_{14}	−0.252 *** (0.022)	−0.262 *** (0.022)	−0.397 *** (0.029)	−0.406 *** (0.029)
15. Thai cuisine	WOT_{15}	0.098 *** (0.023)	0.104 *** (0.023)	0.160 *** (0.03)	0.165 *** (0.03)
16. Western cuisine	WOT_{16}	−0.030 (0.016)	−0.031 (0.016)	−0.157 *** (0.02)	−0.159 *** (0.02)
17. Dessert	WOT_{17}	0.423 *** (0.018)	0.414 *** (0.018)	0.397 *** (0.024)	0.389 *** (0.024)
18. Services	WOT_{18}	−0.165 *** (0.02)	−0.159 *** (0.02)	−0.131 *** (0.027)	−0.125 *** (0.027)
19. Discount	WOT_{19}	−0.167 *** (0.02)	−0.169 *** (0.02)	−0.110 *** (0.027)	−0.111 *** (0.027)
20. Cost performance	WOT_{20}	0.405 *** (0.018)	0.402 *** (0.018)	0.312 *** (0.023)	0.309 *** (0.023)
	Year			−0.224 *** (0.012)	−0.224 *** (0.012)
1. Friend	$WOT_1 Year$	—	—	—	—
2. Appointment	$WOT_2 Year$			0.075 (0.046)	0.074 (0.046)
3. Queuing	$WOT_3 Year$			0.357 *** (0.044)	0.361 *** (0.044)

Table 6. Cont.

Topic Name	Variables	Rating			
		Model 1	Model 2	Model 3	Model 4
4. Environment	$WOT_4 Year$			0.380 *** (0.045)	0.380 *** (0.045)
5. Spicy chicken	$WOT_5 Year$			0.173 *** (0.046)	0.172 *** (0.046)
6. Chicken	$WOT_6 Year$			0.176 *** (0.049)	0.177 *** (0.049)
7. Duck	$WOT_7 Year$			0.151 ** (0.048)	0.153 ** (0.048)
8. Beef hotpot	$WOT_8 Year$			−0.110 * (0.044)	−0.111 * (0.044)
9. Bullfrog hotpot	$WOT_9 Year$			0.292 *** (0.044)	0.292 *** (0.044)
10. Crab	$WOT_{10} Year$			0.139 *** (0.034)	0.142 *** (0.034)
11. Seafood	$WOT_{11} Year$			0.183 *** (0.036)	0.179 *** (0.036)
12. Xinjiang cuisine	$WOT_{12} Year$			0.139 ** (0.043)	0.139 ** (0.043)
13. Cantonese cuisine	$WOT_{13} Year$			0.280 *** (0.038)	0.279 *** (0.038)
14. Hong Kong cuisine	$WOT_{14} Year$			0.407 *** (0.044)	0.405 *** (0.044)
15. Thai cuisine	$WOT_{15} Year$			−0.080 (0.046)	−0.077 (0.046)
16. Western cuisine	$WOT_{16} Year$			0.336 *** (0.033)	0.339 *** (0.033)
17. Dessert	$WOT_{17} Year$			0.100 ** (0.037)	0.099 ** (0.037)
18. Services	$WOT_{18} Year$			0.025 (0.042)	0.024 (0.042)
19. Discount	$WOT_{19} Year$			−0.059 (0.041)	−0.061 (0.041)
20. Cost performance	$WOT_{20} Year$			0.241 *** (0.038)	0.242 *** (0.038)
	Constant	4.297 *** (0.006)	4.336 *** (0.007)	4.374 *** (0.007)	4.411 *** (0.008)
	Price control	N	Y	N	Y
	F-value	330.10	321.19	215.75	213.88
	p-value	0.000	0.000	0.000	0.000
	Adjusted R ²	0.010	0.010	0.013	0.013
	N	651,703	651,703	651,703	651,703

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Model 1 included only WOT_i without incorporating the effect of *Year* or the control of *Price*. The result suggests that in the sample window of 2019–2020, customers were generally more satisfied with some topics and less satisfied with others. More specifically, customers were more satisfied with cost performance ($\alpha_{20} = 0.405$, $p < 0.001$) as well as some food types such as beef hotpot ($\alpha_8 = 0.654$, $p < 0.001$) and dessert ($\alpha_{17} = 0.432$, $p < 0.001$). However, customers were quite unsatisfied about queuing ($\alpha_3 = -0.699$, $p < 0.001$). In addition, lower ratings are associated with environment ($\alpha_4 = -0.152$, $p < 0.001$), services ($\alpha_{18} = -0.165$, $p < 0.001$), and discount ($\alpha_{19} = -0.167$, $p < 0.001$). Model 2 controlled *Price* in an attempt to address the hierarchical effect of restaurant price level, but no significant difference was found from Model 1 in terms of estimated coefficients or adjusted R².

Model 3 incorporated *Year* and its cross terms with WOT_i . *Year* has a negative impact on *Rating* ($\beta_0 = -0.224$, $p < 0.001$), which is consistent with the ANOVA result in Table 5

regarding both sign and magnitude. It is therefore further confirmed that customers became less satisfied with restaurants after the COVID-19 outbreak. Meanwhile, the regression results for the cross terms indicate that the reduction of satisfaction is not evenly distributed among topics. On the one side are topics that have not offset the negative impact of *Year*, such as services ($\beta_{18} = 0.025, p > 0.05$) and discount ($\beta_{19} = -0.059, p > 0.05$), meaning the customers have become less satisfied partially because of these topics. On the other side are topics that not only have significant and positive coefficients, but also have coefficients that totally offset the negative impact of *Year*, such as queuing ($\beta_3 = 0.357, p < 0.001$), environment ($\beta_4 = 0.380, p < 0.001$), and western cuisine ($\beta_{16} = 0.336, p < 0.001$). This is sending a positive signal to restaurant owners, which will be further explored in the discussion section. Last but not least, the hierarchical effect of restaurant price level was again not observed in Model 4.

5. Discussion

Identifying major topics from online review using unsupervised machine learning such as LDA has been an effective approach to probe customer interest, as has been demonstrated in this study. However, a more careful examination of the identified food topics (see Table 3, Panel 2) shows that some restaurants are more emphasized whereas others are weakened. On the one extreme are several topics addressing similar concepts, such as spicy chicken with chicken, beef hotpot with bullfrog hotpot, and Cantonese cuisine with Hong Kong cuisine. On the other extreme are topics each incorporating multiple concepts, such as coffee embedded in crab, and all western restaurants, be it German or Spanish, combined in Western cuisine. Such heterogeneity is a natural consequence of the uneven distribution of the types of the sampled restaurants as well as the unsupervised rationale of the chosen machine learning protocol. Nevertheless, Table 3, with the three panels together, has provided a complete framework to analyze a restaurant from customers' perspective. Moreover, the cross sectional WOT in Table 4 demonstrates customers' relative interest on some topics over others. Restaurant managers are therefore suggested to self-evaluate following the above framework as well as track the change of customer interest in order to optimize resource allocation in operational improvement.

COVID-19 has indeed introduced negative impact on the restaurant industry, with the studied restaurants collectively being no exception. Customers have been paying fewer visits to restaurants since the outbreak, assigning lower ratings, and showing limited evidence of spending more. The stay-at-home orders, together with fear of infection, have surely reduced customers' motivation of dining out. Meanwhile, the precautionary measures by restaurants may also compromise customer satisfaction, with some buffet restaurants temporarily switching to ordering mode, resulting in a freefall in dining experience. Given the systematic problem of COVID-19, restaurant managers are advised to hold a rational perspective on the temporary situation and not to expect avoiding loss in the dine-in business. They are instead encouraged to look beyond dine-in, the primary source of income during normal periods, and seek alternative revenue stream such as takeaway and delivery.

Besides challenge, COVID-19 has also brought opportunity to restaurants. Queuing, the most annoying factor for restaurant customers during normal periods, turns out to receive much less complaint during COVID-19. Of course, that is the result of fewer customers dining out. But more importantly, both the cross-sectional data and the before-and-after comparison indicate that customers are extremely sensitive to queuing. Managers are thus strongly advised to optimize waiting time as a preparation for the upcoming resumption by analyzing customers' temporal pattern of visiting as well as designing coupons that can reposition price-sensitive customers from rush hours.

6. Conclusions

Observing and interpreting restaurant customers' evolution of dining patterns and satisfaction during COVID-19 is of critical importance in terms of developing sustainable

business insights. This study has described and analyzed customers' dining behavior before and after the pandemic outbreak by means of statistically aggregating and empirically correlating 651,703 restaurant UGCs posted by diners during 2019–2020. Twenty review topics, mostly food, have been identified by LDA, and ANOVA and rating-review regression have been performed to explore whether and why customers became less satisfied. Results suggest that customers have been paying fewer visits to restaurants since the outbreak, assigning lower ratings, and showing limited evidence of spending more. Interestingly, queuing, the most annoying factor for restaurant customers during normal periods, turns out to receive much less complaint during COVID-19.

This study contributes by discovering business knowledge in the context of COVID-19, based on big data that features accessibility, relevance, volume, and information richness, which is transferable to future studies and can benefit additional population and business. To be specific, although this study obtained restaurant UGC from Shanghai written in Chinese during COVID-19, the methodology is by no means limited to specific type of service, region, language, or time period. Therefore, this study is of interest to other national audiences at any time and about any industry. Meanwhile, this study has also provided practical suggestions to managers regarding the framework of self-evaluation, business mode, and operational optimization. They are encouraged to self-evaluate under the established framework to optimize resource allocation in operational improvement, look beyond dine-in and seek alternative revenue stream such as takeaway and delivery, and reduce waiting time as a preparation for the upcoming resumption by analyzing customers' temporal pattern of visiting as well as designing coupons that can reposition price-sensitive customers from rush hours.

One of the limitations of this study is sample bias, because this study only considers restaurant customers who have chosen to post online UGC. In other words, the sampled customers may not completely represent those who have not post UGC. Hence, future studies are encouraged to incorporate additional data sources such as foot traffic and card transaction data so as to come up with a more reliable measure of restaurant visits, and on-site questionnaires to obtain offline UGC. Another limitation of this study is the fineness of the discovered topics, with all the reviews characterized by only 20 topics. Although the discovered topics have provided an overview of customer experience, they are not elaborate enough to quantify the subtle differences among individual customers who had written seemingly similar reviews but assigned diverse ratings. The relatively small adjusted R^2 suggests that the majority of the variation of customer satisfaction remains unexplained. Therefore, future studies may address this issue by improving text mining and empirical modeling.

Funding: This research was funded by National Natural Science Foundation of China, grant number 72072118.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Giudice, F.; Caferra, R.; Morone, P. COVID-19, the food system and the circular economy: Challenges and opportunities. *Sustainability* **2020**, *12*, 7939. [\[CrossRef\]](#)
2. Jian, Y.; Yu, I.Y.; Yang, M.X.; Zeng, K.J. The impacts of fear and uncertainty of COVID-19 on environmental concerns, brand trust, and behavioral intentions toward green hotels. *Sustainability* **2020**, *12*, 8688. [\[CrossRef\]](#)
3. Umaña-Hermosilla, B.; de la Fuente-Mella, H.; Elórtégui-Gómez, C.; Fonseca-Fuentes, M. Multinomial logistic regression to estimate and predict the perceptions of individuals and companies in the face of the COVID-19 pandemic in the Ñuble region, Chile. *Sustainability* **2020**, *12*, 9553. [\[CrossRef\]](#)
4. Appolloni, A.; Colasanti, N.; Fantauzzi, C.; Fiorani, G.; Frondizi, R. Distance learning as a resilience strategy during Covid-19: An analysis of the Italian context. *Sustainability* **2021**, *13*, 1388. [\[CrossRef\]](#)
5. Yang, Y.; Liu, H.B.; Chen, X. COVID-19 and restaurant demand: Early effects of the pandemic and stay-at-home orders. *Int. J. Contemp. Hosp. Manag.* **2020**, *13*, 3809–3834. [\[CrossRef\]](#)

6. Zhong, Y.; Oh, S.; Moon, H.C. What can drive consumers' dining-out behavior in China and Korea during the COVID-19 pandemic? *Sustainability* **2021**, *13*, 1724. [\[CrossRef\]](#)
7. Davahli, M.R.; Karwowski, W.; Sonmez, S.; Apostolopoulos, Y. The hospitality industry in the face of the COVID-19 pandemic: Current topics and research methods. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7366. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Mayasari, N.R.; Ho, D.K.N.; Lundy, D.J.; Skalny, A.V.; Tinkov, A.A.; Teng, I.C.; Wu, M.C.; Faradina, A.; Mohammed, A.Z.M.; Park, J.M.; et al. Impacts of the COVID-19 pandemic on food security and diet-related lifestyle behaviors: An analytical study of Google trends-based query volumes. *Nutrients* **2020**, *12*, 3103. [\[CrossRef\]](#)
9. Kim, J.; Kim, J.; Wang, Y. Uncertainty risks and strategic reaction of restaurant firms amid COVID-19: Evidence from China. *Int. J. Hosp. Manag.* **2021**, *92*, 102752. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Kim, J.; Lee, J.C. Effects of COVID-19 on preferences for private dining facilities in restaurants. *J. Hosp. Tour. Manag.* **2020**, *45*, 67–70. [\[CrossRef\]](#)
11. Chen, W.K.; Riantama, D.; Chen, L.S. Using a text mining approach to hear voices of customers from social media toward the fast-food restaurant industry. *Sustainability* **2021**, *13*, 268. [\[CrossRef\]](#)
12. Jia, S. Toward a better fitness club: Evidence from exerciser online rating and review using latent Dirichlet allocation and support vector machine. *Int. J. Mark. Res.* **2019**, *61*, 64–76. [\[CrossRef\]](#)
13. Jia, S. Leisure motivation and satisfaction: A text mining of yoga centres, yoga consumers, and their interactions. *Sustainability* **2018**, *10*, 4458. [\[CrossRef\]](#)
14. Jia, S. Behind the ratings: Text mining of restaurant customers' online reviews. *Int. J. Mark. Res.* **2018**, *60*, 561–572. [\[CrossRef\]](#)
15. Jia, S. Motivation and satisfaction of Chinese and US tourists in restaurants: A cross-cultural text mining of online reviews. *Tour. Manag.* **2020**, *78*, 104071. [\[CrossRef\]](#)
16. Ye, Q.; Law, R.; Gu, B.; Chen, W. The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Comput. Hum. Behav.* **2011**, *27*, 634–639. [\[CrossRef\]](#)
17. Anderson, M.; Magruder, J. Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *Econ. J.* **2012**, *122*, 957–989. [\[CrossRef\]](#)
18. Lee, T.Y.; Bradlow, E.T. Automated marketing research using online customer reviews. *Int. J. Mark. Res.* **2011**, *48*, 881–894. [\[CrossRef\]](#)
19. Hao, H.; Zhang, K.; Wang, W.; Gao, G. A tale of two countries: International comparison of online doctor reviews between China and the United States. *Int. J. Med. Inf.* **2017**, *99*, 37–44. [\[CrossRef\]](#)
20. Pan, S.; Ting, P.; Bau, D. Perceptions of Taipei as revealed in travel blogs: A tale from two sides. *Asia Pac. J. Tour. Res.* **2014**, *19*, 700–720. [\[CrossRef\]](#)
21. London, W.R.; Moyle, B.D.; Lohmann, G. Cruise infrastructure development in Auckland, New Zealand: A media discourse analysis (2008–2016). *Asia Pac. J. Tour. Res.* **2017**, *22*, 615–633. [\[CrossRef\]](#)
22. Kim, H.; Joun, H.J.; Choe, Y.; Schroeder, A. How can a destination better manage its offering to visitors? Observing visitor experiences via online reviews. *Sustainability* **2019**, *11*, 4660. [\[CrossRef\]](#)
23. Li, H.; Ye, Q.; Law, R. Determinants of customer satisfaction in the hotel industry: An application of online review analysis. *Asia Pac. J. Tour. Res.* **2013**, *18*, 784–802. [\[CrossRef\]](#)
24. Li, N.; Tung, V.; Law, R. A fuzzy comprehensive evaluation algorithm for analysing electronic word-of-mouth. *Asia Pac. J. Tour. Res.* **2017**, *22*, 592–603. [\[CrossRef\]](#)
25. Guo, Y.; Wang, Y.; Wang, C. Exploring the salient attributes of short-term rental experience: An analysis of online reviews from Chinese guests. *Sustainability* **2019**, *11*, 4290. [\[CrossRef\]](#)
26. Miller, M.; Banerjee, T.; Muppalla, R.; Romine, W.; Sheth, A. What are people tweeting about Zika? An exploratory study concerning symptoms, treatment, transmission, and prevention. *JMIR Public Health Surveill.* **2017**, *3*, e38. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Guo, Y.; Barnes, S.J.; Jia, Q. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tour. Manag.* **2017**, *59*, 467–483. [\[CrossRef\]](#)
28. Xiang, Z.; Du, Q.; Ma, Y.; Fan, W. A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tour. Manag.* **2017**, *58*, 51–65. [\[CrossRef\]](#)
29. Nam, S.; Lee, H.C. A text analytics-based importance performance analysis and its application to airline service. *Sustainability* **2019**, *11*, 6153. [\[CrossRef\]](#)
30. Gosh, D.; Guha, R. What are we 'tweeting' about obesity? Mapping tweets with topic modeling and geographic information system. *Cartogr. Geogr. Inf. Sci.* **2013**, *40*, 90–102. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Liu, X.; Burns, A.C.; Hou, Y. An investigation of brand-related user-generated content on Twitter. *J. Advert.* **2017**, *46*, 236–247. [\[CrossRef\]](#)
32. Rahman, M.; Carburnar, B.; Ballesteros, J.; Chau, D.H. To catch a fake: Curbing deceptive Yelp ratings and venues. *Stat. Anal. Data Min.* **2015**, *8*, 147–161. [\[CrossRef\]](#)
33. Choi, S.; Mattila, A.S.; Van Hoof, H.B.; Quadri-Felitti, D. The role of power and incentives in inducing fake reviews in the tourism industry. *J. Travel Res.* **2017**, *56*, 975–987. [\[CrossRef\]](#)
34. Cardoso, E.F.; Silva, R.M.; Almeida, T.A. Towards automatic filtering of fake reviews. *Neurocomputing* **2018**, *309*, 106–116. [\[CrossRef\]](#)

-
35. Liang, Y.; Liu, Y.; Chen, C.; Jiang, Z.G. Extracting topic-sensitive content from textual documents: A hybrid topic model approach. *Eng. Appl. Artif. Intell.* **2018**, *70*, 81–91. [[CrossRef](#)]
 36. Sutherland, I.; Sim, Y.; Lee, S.K.; Byun, J.; Kiatkawsin, K. Topic modeling of online accommodation reviews via latent Dirichlet allocation. *Sustainability* **2020**, *12*, 1821. [[CrossRef](#)]
 37. Jieba 0.42.1. Available online: <https://pypi.org/project/jieba/> (accessed on 24 April 2021).
 38. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
 39. Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *84*, 55–77. [[CrossRef](#)]
 40. Jia, S. Measuring tourists' meal experience by mining online user generated content about restaurants. *Scand. J. Hosp. Tour.* **2019**, *19*, 371–389. [[CrossRef](#)]
 41. Büschken, J.; Allenby, G.M. Sentence-based text analysis for customer reviews. *Mark. Sci.* **2016**, *35*, 953–975. [[CrossRef](#)]