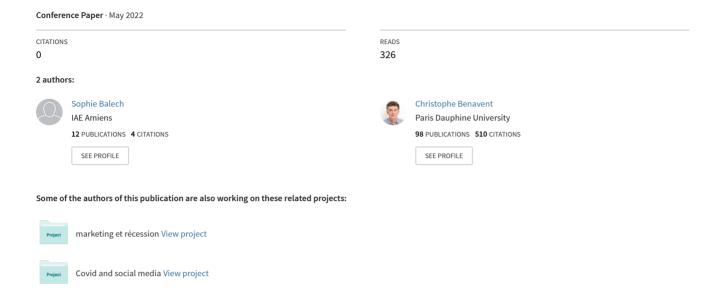
### LE ROLE DES DIMENSIONS DE L'EXPERIENCE DANS LA SATISFACTION CLIENT : UNE APPLICATION AU CAS DE L'INDUSTRIE HOTELIERE EN POLYNESIE FRANÇAISE



## LE ROLE DES DIMENSIONS DE L'EXPERIENCE DANS LA SATISFACTION CLIENT : UNE APPLICATION AU CAS DE L'INDUSTRIE HOTELIERE EN POLYNESIE FRANÇAISE

#### SOPHIE BALECH

MAITRE DE CONFERENCES

UNIVERSITE PICARDIE JULES VERNE, CRIISEA, 80000 AMIENS, FRANCE

SOPHIE.BALECH@U-PICARDIE.FR

#### CHRISTOPHE BENAVENT

PROFESSEUR DES UNIVERSITES

UNIVERSITE PARIS DAUPHINE-PSL, DRM, 75016 PARIS, FRANCE

CHRISTOPHE.BENAVENT@DAUPHINE.PSL.EU

## LE ROLE DES DIMENSIONS DE L'EXPERIENCE DANS LA SATISFACTION CLIENT : UNE APPLICATION AU CAS DE L'INDUSTRIE HOTELIERE EN POLYNESIE FRANÇAISE

#### Résumé:

Dans cette communication nous poursuivons deux buts : le premier est d'examiner la manière dont les consommateurs jugent leur expérience en comprenant mieux comment les dimensions de l'évaluation captée dans les commentaires laissés en ligne affectent la note donnée en utilisant le cadre de la théorie bi-factorielle de la satisfaction ; le second est méthodologique, il s'agit de proposer un processus d'analyse qui combine un modèle de topic pour identifier les attributs d'évaluation à un modèle de régression GAM pour en tester les effets sur la satisfaction. Les données utilisées représentent 24 031 avis et notes de satisfaction associées laissés sur TripAdvisor concernant 52 hôtels des îles de Polynésie Française. Les résultats montrent une dichotomie du rôle des attributs expérientiels et des attributs fonctionnels sur la satisfaction et mettent en lumière l'intérêt d'analyser les commentaires clients à travers des méthodes NLP pour les gestionnaires d'hôtels.

#### Mots-clés:

NLP; satisfaction; recommandation; théorie bi-factorielle; expérience hôtelière

## THE ROLE OF EXPERIENCE DIMENSIONS IN CUSTOMER SATISFACTION: AN APPLICATION TO THE HOTEL INDUSTRY IN FRENCH POLYNESIA

#### Abstract:

In this paper we pursue two goals: the first is to examine how consumers judge their experience by better understanding how the dimensions of evaluation captured in online reviews affect the rating given using the framework of the bi-factorial theory of satisfaction; the second is methodological, it is to propose an analysis process that combines a topic model to identify evaluation attributes with a GAM regression model to test the effects on satisfaction. The data used represent 24,031 reviews and associated satisfaction ratings left on TripAdvisor concerning 52 hotels in the French Polynesian islands. The results show a dichotomy of the role of experiential and functional attributes on satisfaction and highlight the interest of analyzing customer reviews through NLP methods for hotel managers.

#### Key-words:

NLP; satisfaction; recommendation; bi-factorial theory; hotel experience

# LE ROLE DES DIMENSIONS DE L'EXPERIENCE DANS LA SATISFACTION CLIENT : UNE APPLICATION AU CAS DE L'INDUSTRIE HOTELIERE EN POLYNESIE FRANÇAISE

#### **Introduction:**

L'environnement digital donne l'occasion de revisiter des concepts établis, ou ancien, à l'aune de nouvelles méthodes et de nouvelles données. Celles-ci sont désormais largement disponibles dans les réseaux sociaux et plus spécifiquement dans les sites d'avis de consommateur.

Dans cette communication nous poursuivons deux buts : le premier est d'examiner la manière dont les consommateurs jugent leur expérience en comprenant mieux comment les dimensions de l'évaluation captée dans les commentaires affectent la note donnée en utilisant le cadre de la théorie bi-factorielle de la satisfaction ; le second est méthodologique, il s'agit de proposer un processus d'analyse qui combine un modèle de topic pour identifier les attributs d'évaluation à un modèle de régression GAM pour en tester les effets sur la satisfaction.

A travers l'étude des commentaires laissés sur la plateforme TripAdvisor concernant 52 hôtels de Polynésie Française, les résultats confirment le rôle différencié des attributs relevant du facteur expérientiel de ceux relevant du facteur fonctionnel associés à la consommation d'un service d'hébergement hôtelier.

#### Revue de littérature :

L'approche bi-factorielle de la satisfaction. La théorie bi-factorielle de la satisfaction propose que satisfaction et insatisfaction sont deux construits indépendants, par opposition à l'approche de l'infirmation des attentes qui conçoit satisfaction et insatisfaction comme étant les deux extrêmes d'un même continuum (Oliver, 1993; Oliver and DeSarbo, 1988). La théorie bi-factorielle repose sur les contributions des différents attributs d'un produit ou service sur la satisfaction du consommateur (Audrain, 2000). Cette théorie concurrente à la théorie de l'infirmation des attentes a reçu une attention croissante des chercheurs, pour établir sa validité (par exemple, Maddox (1981) montre que cette théorie ne s'applique pas à toutes les catégories de produits) ou discuter des méthodes de collecte et de traitement des données utilisées (Mikulié, 2007; Vargo et al., 2007).

Dans leur travaux sur les raisons de la satisfaction au travail, Herzberg et al. (1959) mettent en évidence deux groupes d'attributs qui vont jouer sur l'insatisfaction (les facteurs d'hygiène) ou sur la satisfaction (les facteurs de motivation) chez les employés. La méthodologie qu'ils emploient est celle de la technique des incidents critiques qui visent à faire éliciter naturellement aux travailleurs les raisons principales qui les conduisent à ressentir de la satisfaction ou de l'insatisfaction au travail. Cette méthode repose sur la saillance des attributs qui viennent naturellement à l'esprit des répondants (Romaniuk and Sharp, 2004). Cette théorie développée dans le domaine du travail et des ressources humaines a suscité l'intérêt du marketing, avec les travaux fondateurs de Kano dans le domaine de l'innovation (Mikulić, 2007; Mikulić and Prebežac, 2016) et de Cadotte & Turgeon (1988) sur le rôle des attributs dans la satisfaction.

Ces prolongements mettent en évidence 4 classes d'attributs (Vargo et al., 2007) : ceux qui conduisent à la satisfaction mais si leur performance sont faibles ne conduisent pas à

l'insatisfaction (attractive ou satisfiers selon les terminologie); ceux qui conduisent à l'insatisfaction mais si leur performance sont élevées ne produisent pas de satisfaction (les must-be ou satisfiers); ceux qui n'ont pas d'impact sur la satisfaction ou l'insatisfaction (indifferent ou neutrals); ceux qui ont un impact fort sur la satisfaction ou l'insatisfaction selon leur niveau de performance (one-dimensional ou criticals)<sup>1</sup>.

En plus de la classification des attributs selon leur rôle dans la satisfaction, le modèle de Kano met en évidence leurs effets non-linéaires. Notamment, l'étude de Mittal et al. (1998), fondée sur la théorie des prospects Kahneman & Tversky (1979), met en évidence la nature de la réponse asymétrique et non-linéaire de la performance perçue d'un attribut sur la satisfaction. Dans la lignée de ces travaux, notre recherche vise à mettre en évidence la contribution des attributs clés de l'expérience hôtelière à la satisfaction.

Les dimensions de l'expérience et de la satisfaction d'un séjour à l'hôtel. De nombreux travaux dans le domaine du tourisme et de l'hospitalité se sont intéressés aux attributs clés de l'expérience hôtelière depuis les travaux de Cadotte & Turgeon (1988). Dans une revue de littérature, Dolnicar & Otter, (2003) mettent en évidence plus d'une vingtaine d'études sur ce sujet entre 1984 et 2000, définissant plus de 173 attributs regroupés en 9 catégories : l'image, le prix, l'hôtel, la chambre, les services, le marketing, la restauration, la sécurité et le lieu. Depuis, de nombreux travaux se sont intéressés à ce sujet, élargissant le champ d'étude aux locations de pair-à-pair type Airbnb et les sources de données aux commentaires laissés par les clients sur les sites d'avis en ligne (UGC, User Generated Content) (par exemple Ding et al., 2021; Guo et al., 2019; Park et al., 2021; Tussyadiah & Zach, 2015; Xiang et al., 2015). Certaines de ces recherches s'appuient sur la théorie bi-factorielle de la satisfaction (comme Alegre & Garau, 2010; Park et al., 2021; Xiang et al., 2015).

Les différentes recherches visant à déterminer les dimensions de l'expérience hôtelière s'accordent sur la multiplicité des attributs, notamment liée au contexte de la visite, et les distinguent selon leurs caractères expérientiels ou fonctionnels.

Les attributs fonctionnels correspondent à ceux relatifs à la délivrance du service de base : les chambres sont-elles propres, le paiement s'est-il bien passé, les interactions avec le personnel ont-elles été courtoises. Chaque activité, chaque événement qui participe à délivrer ce qui constitue le service de base fait littéralement fonctionner le service. Et si tout a bien fonctionné on s'attend à ce que personne n'en parle. A contrario, si un incident s'est produit, il risque selon sa gravité de retenir l'attention, et d'être saillant à l'esprit du consommateur.

Les attributs expérientiels correspondent à ce qu'ont éprouvé les consommateurs au cours du séjours : les émotions vécues, les images suscitées et les activités qui ont été éprouvées (think, feel, act (Schmitt, 1999)). Dans un tel contexte, l'expérience négative est peu probable : rares sont les tempêtes ou les tremblements de terre, parfois on peut se blesser en faisant du sport, ou en faisant une mauvaise rencontre. L'ordinaire c'est plutôt une expérience positive : réaliser un rêve, éprouver un grand bien-être, faire des rencontres, faire des choses qui sortent de l'ordinaire. Le risque principal, c'est en la matière le risque de l'ennui, de la non-expérience.

1

<sup>&</sup>lt;sup>1</sup> Ces éléments peuvent être rapprochés de ceux identifiés dans les travaux de Llosa (1997) concernant l'expérience de service. Même si l'approche se concentre sur la satisfaction et n'est pas bifactorielle au sens strict de Maddox (1981), Llosa (1997) propose un modèle tétra-classe des éléments de l'expérience de service sur la satisfaction des clients, en classifiant les attributs selon leur rôle sur le niveau de satisfaction et leur niveau de performance perçue.

En conséquence on peut formuler l'hypothèse suivante relative à la composition des thématiques et attributs de l'expérience vécue et racontée. Elle suppose que, lorsque les attributs fonctionnels posent problèmes, ceux-ci sont saillants et masquent les attributs expérientiels, leur satisfaction libérant alors l'expression de ces derniers.

H1 : La saillance des attributs expérientiels est associée positivement à l'évaluation, la saillance des attributs fonctionnels est associée négativement à l'évaluation.

A cette hypothèse de distribution nous pouvons en ajouter une seconde directement issue des modèles bi-factoriels, qui supposent des relations non-linéaire entre les attributs et l'évaluation générale : les attributs clés entretiennent une relation concave avec la satisfaction, l'effet est décroissant avec la valeur de l'attribut : le fait de le juger positivement à un effet marginal sur la satisfaction globale. Les attributs « plus » ou de « ravissement » (delightfull) n'affectent pas le jugement s'ils sont absents, ils le rendent enthousiaste quand ils sont largement présents. La relation dans ce cas est convexe.

H2 : la saillance des attributs relevant des facteurs expérientiel/fonctionnel ont une relation convexe/concave avec l'évaluation.

#### Méthodologie:

En nous appuyant sur les études antérieures, nous proposons d'étudier le rôle des dimensions de l'expérience hôtelière sur la satisfaction client, à partir des UGC concernant une destination *sun and sand*, les îles de Polynésie Française. Les dimensions retenues dans l'étude concernent le cadre du séjour, les activités réalisées et les services offerts par l'hôtel pour les facteurs expérientiels, et les caractéristiques de la chambre, l'accueil et le prix pour les facteurs fonctionnels. Nous cherchons à évaluer l'impact de ces dimensions sur la satisfaction client, mesurée par les notes laissées sur la plateforme d'UGC. Utiliser les UGC comme source de données permet de s'approcher de la méthode des incidents critiques, dans la mesure où ces jugements sont produit à l'occasion d'une expérience et de lever les difficultés de son application relevées par Vargo et al. (2007).

Le jeu de données. Le jeu de données utilisé a été constitué par scrapping de la plateforme TripAdvisor, leader dans les avis clients concernant les activités touristiques (TripAdvisor.com, 2021). Il concerne les hôtels de Polynésie Française, une destination *sun and sand*, et contient 24 031 avis clients et notes associées. Ces avis ont été laissés entre 2001 et 2020 sur 52 hôtels. Les notes obtenues s'échelonnent de 3 à 5 étoiles, sur une échelle allant de 1 à 5, avec une moyenne à 4,55, une médiane à 4,5 et un écart-type de 0.33. L'annexe 1 présente le détail du jeu de données.

Méthodes. La méthode employée pour identifier les attributs s'appuie sur modèle de topic de type LDA (Latent Dirichlet Allocation) semi-supervisé (Lu et al., 2011). Par rapport au modèle initial de Blei (Blei et al., 2003), le modèle prend en entrée un dictionnaire préétabli par les chercheurs pour identifier les topics. Ici, le dictionnaire retenu contient 6 topics relatifs aux dimensions de l'expérience et de la satisfaction hôtelière (le paysage qui concerne le cadre du séjour, les activités qui représentent une partie de l'expérience vécue durant le séjour, le service qui concerne tous les services proposés par les hôtels, l'accueil qui s'intéresse aux conditions de l'accueil du voyageur, la chambre pour les attributs physiques du lieu du séjour et le prix qui comprend également le rapport qualité-prix). Il est appliqué sur le corpus qui a été

prétraité pour ne retenir que les lemmes des noms communs, noms propre, adjectifs et verbes. On retient les paramètres *thêta* du modèle pour estimer la saillance des attributs.

Puis, on teste l'effet des topics sur la note de satisfaction avec un modèle GAM ( $generalized\ additive\ model$ ), en utilisant les probabilités qu'un topic k soit associé à un document i (résultats du modèle de topic semi-supervisé). Cette méthode permet de saisir les effets non-linéaires, anticipés dans la formalisation des modèles asymétriques de la satisfaction. Le modèle GAM est un modèle linéaire d'une fonction qui n'est pas forcément linéaire ; ici, on utilise des splines qui se prêtent aisément à des fonctions fortement non-linéaires.

#### Résultats:

Les résultats du modèle LDA semi-supervisé sont présentés dans l'annexe 2, à travers les mots les plus associés à chaque topic, tandis que les résultats détaillés du modèle GAM se trouvent dans l'annexe 3. La figure 1 ci-dessous représente les contributions de chaque topic à l'évaluation de la satisfaction.

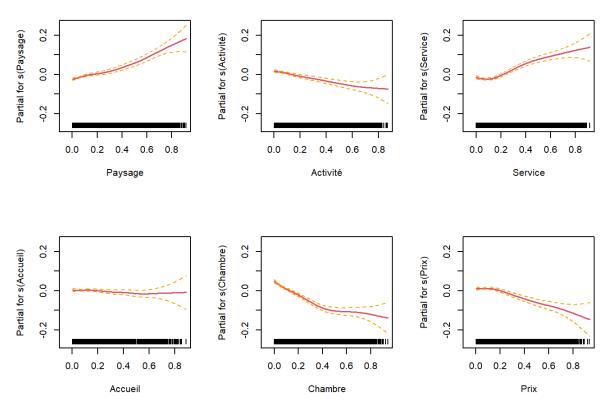


Figure 1 : les contributions partielles des topics à l'évaluation de la satisfaction Lecture : L'abscisse représente la proportion du topic dans le commentaire et l'ordonnée l'évolution relative de la note associée. Pour le premier graphique, plus le topic du « paysage » est présent dans un commentaire, plus la note sera élevée, relativement à la moyenne.

L'analyse des topics croisée avec les résultats de la partie linéaire du modèle GAM nous montre que tous les attributs ont un impact significatif sur l'évaluation de l'expérience touristiques. Les facteurs expérientiels du service et du paysage sont associés positivement à l'évaluation, tandis que les attributs fonctionnels du prix et de la chambre sont associés négativement à l'évaluation, l'accueil l'étant de manière moins évidente. De manière plus surprenante, l'attribut expérientiel de l'activité est associé négativement à l'évaluation ce qui va à l'encontre de l'hypothèse H1. Ainsi, nous pouvons donc partiellement accepter H1.

L'analyse des résultats de la partie non-linéaire du modèle GAM nous montre que les attributs de l'activité et de l'accueil sont correctement représentés par l'approche linéaire, les effets non-linéaires sont non-significatifs. Les quatre autres attributs ont un impact non-linéaire sur l'évaluation. La courbure des courbes montre une relation convexe entre le service et l'évaluation et une relation concave entre la chambre et l'évaluation. Par contre, concernant les attributs du paysage et du prix, la courbure des courbes semble aller à l'encontre de notre hypothèse. Nous ne pouvons donc pas valider l'hypothèse H2 sur l'ensemble des attributs.

#### Discussion et implications managériales :

Les résultats présentés permettent de valider partiellement l'hypothèse 1, selon laquelle la saillance des attributs expérientiels/fonctionnels dans les UGC est associée positivement/négativement à l'évaluation du séjour hôtelier. Concernant l'hypothèse 2, selon laquelle la saillance des attributs relevant des facteurs expérientiels/fonctionnels ont une relation convexe/concave avec l'évaluation, les résultats sont peu concluants et ne permettent pas de valider l'hypothèse.

On peut s'interroger sur le rôle de l'attribut activité, relevant de la dimension expérientielle. Plus cet attribut est saillant dans les discours, plus l'évaluation de l'expérience est négative, ce qui semble à première vue contradictoire avec la théorie. Cependant, les éléments composant ce topic ne sont pas directement reliés aux hôtels et à leur service, mais évoquent les activités que l'on réalise en dehors de ce contexte. Finalement, une hypothèse pour expliquer ce résultat pourrait être que la présence de cet attribut dans le discours concernant l'hôtel est un marqueur qu'aucun élément grave n'a dysfonctionné dans l'expérience, mais que rien de marquant ne s'est non plus produit pour laisser au client un souvenir mémorable de son séjour à l'hôtel. La présence de ce topic serait alors un indice pour identifier une expérience ni bonne, ni mauvaise qui conduirait à une note relativement moyenne.

Cette recherche s'inscrit dans le cadre de la théorie bi-factorielle de la satisfaction et propose une méthode pour en renouveler les recherches, en utilisant les avis consommateurs et des outils de traitement automatique du langage naturel pour identifier les attributs saillants de l'expérience de consommation. Le modèle LDA semi-supervisé employé ici permet de pallier les difficultés de mesure de la théorie bi-factorielle, en reprenant l'idée générale de la technique des incidents critiques. S'intéresser aux attributs les plus saillants, ceux racontés et partagés par les clients à l'issue de leur expérience ne permet bien évidemment pas d'identifier le rôle de chaque attribut dans la constitution de la satisfaction, et des recherches supplémentaires sont nécessaires pour approfondir ce point.

D'un point de vue managérial, cette recherche permet de mettre en évidence l'intérêt fort pour des méthodes d'analyse approfondie des UGC afin de suivre les performances et les dysfonctionnements des hôtels. Le domaine du tourisme et de l'hospitalité est largement concerné par le poids grandissant des avis clients, les méthodes proposées ici permettent d'en tirer des informations précieuses pour s'assurer de la satisfaction des clients et d'identifier rapidement les manques.

#### Références:

- Alegre J and Garau J (2010) Tourist Sastisfaction and Dissatisfaction. *Annals of Tourism Research* 37(1): 52–73. DOI: 10.1016/j.annals.2009.07.001.
- Audrain A-F (2000) Proposition d'un cadre conceptuel des déterminants de la contribution des attributs à la satisfaction. In: *Actes du 16 ème Congrès de l'Association Française de Marketing*, Montréal, Canada, 18 May 2000, p. 15.
- Blei DM, Ng AY and Jordan MI (2003) Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 3: 993–1022.
- Cadotte ER and Turgeon N (1988) Key Factors in Guest Satisfaction. *The Cornell H.R.A. Quarterly* (February): 45–51.
- Ding K, Choo WC, Ng KY, et al. (2021) Exploring Sources of Satisfaction and Dissatisfaction in Airbnb Accommodation Using Unsupervised and Supervised Topic Modeling. *Frontiers in Psychology* 12: 659481. DOI: 10.3389/fpsyg.2021.659481.
- Dolnicar S and Otter T (2003) Which Hotel attributes Matter? A review of previous and a framework for future research. In: *Proceedings of the 9th Annual Conference of the Asia Pacific Tourism Association (APTA)*, Sidney, 2003, pp. 176–188. Griffin, T & Harris, R (eds.
- Guo Y, Wang Y and Wang C (2019) Exploring the Salient Attributes of Short-Term Rental Experience: An Analysis of Online Reviews from Chinese Guests. *Sustainability* 11(16): 4290. DOI: 10.3390/su11164290.
- Herzberg F, Mausner B and Snyderman BB (1959) *The Motivation to Work*. New York: JohnWiley & Sons.
- Kahneman D and Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2): 363–391.
- Llosa S (1997) L'analyse de la contribution des éléments du service à la satisfaction : un modèle tétraclasse. *Décisions Marketing* 10: 87–95. DOI: 10.7193/DM.010.87.95.
- Lu B, Ott M, Cardie C, et al. (2011) Multi-aspect Sentiment Analysis with Topic Models. In: 2011 IEEE 11th international conference on data mining workshops, 2011, pp. 81–88. IEEE.
- Maddox RN (1981) Two-factor Theory and Consumer Satisfaction: Replication and Extension. *Journal of Consumer Research* 8(1): 97. DOI: 10.1086/208845.
- Mikulić J (2007) The Kano Model A Review of its Application in Marketing Research from 1984 to 2006. In: *Proceedings of the 1st International Conference marketing theory challenges in transitional societies*, 2007, pp. 87–96.
- Mikulić J and Prebežac D (2016) The Kano model in tourism research: A critical note. *Annals of Tourism Research* 61: 25–27. DOI: 10.1016/j.annals.2016.07.014.

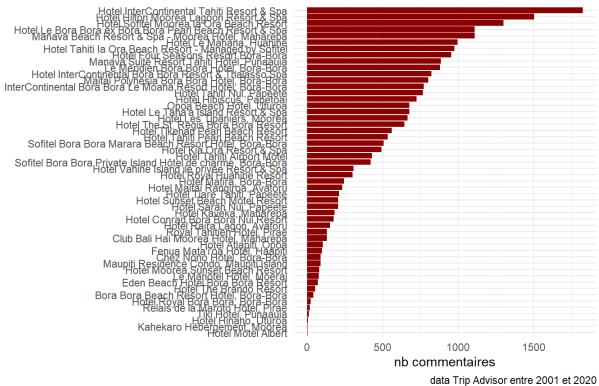
- Mittal V, Ross WT and Baldasare PM (1998) The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions. *Journal of Marketing* 62(1): 33. DOI: 10.2307/1251801.
- Oliver RL (1993) Cognitive, Affective, and Attribute Bases of the Satisfaction Response. *Journal of Consumer Research* 20(3): 418. DOI: 10.1086/209358.
- Oliver RL and DeSarbo WS (1988) Response Determinants in Satisfaction Judgments. *Journal of Consumer Research* 14(4): 495. DOI: 10.1086/209131.
- Park H, Lee M and Back K-J (2021) Exploring the roles of hotel wellness attributes in customer satisfaction and dissatisfaction: application of Kano model through mixed methods. *International Journal of Contemporary Hospitality Management* 33(1): 263–285. DOI: 10.1108/IJCHM-05-2020-0442.
- Romaniuk J and Sharp B (2004) Conceptualizing and measuring brand salience. *Marketing Theory* 4(4): 327–342. DOI: 10.1177/1470593104047643.
- Schmitt B (1999) Experiential Marketing. *Journal of Marketing Management* 15(1–3): 53–67. DOI: 10.1362/026725799784870496.
- TripAdvisor.com (2021) About Tripadvisor. Available at: https://tripadvisor.mediaroom.com/US-about-us (accessed 12 December 2021).
- Tussyadiah IP and Zach FJ (2015) Hotels vs. Peer-to-Peer Accommodation Rentals: Text Analytics of Consumer Reviews in Portland, Oregon. In: *Travel and Tourism Research Association (TTRA) 46th Annual International Conference*, Portland, Oregon, 2015. DOI: 10.2139/ssrn.2594985.
- Vargo SL, Nagao K, He Y, et al. (2007) Satisfiers, Dissatisfiers, Criticals, and Neutrals: A Review of Their Relative Effects on Customer (Dis)Satisfaction. *Academy of Marketing Science Review* 11(2): 22.
- Xiang Z, Schwartz Z, Gerdes JH, et al. (2015) What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management* 44: 120–130. DOI: 10.1016/j.ijhm.2014.10.013.

#### Annexes:

#### Annexe 1 : Présentation du jeu de données

#### Distribution des commentaires

n hôtels = 52, n commentaires = 24 031



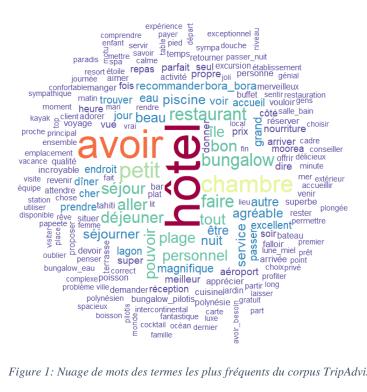


Figure 1: Nuage de mots des termes les plus fréquents du corpus TripAdvisor concernant 52 hôtels de Polynésie française

Annexe 2 : résultats du modèle de topics semi-supervisé

0							
	corail	requin	thalasso	annulation	groisièrer	nauvais	appeler
	coraux	snorkeling	hilton	soirée	taxi	gros	' fruit
	inoubliable	tortue	lagune	bora	étage	manque	commander
5 -	chef	baleine	détendre	varié	marché	cas	plupart
	sourire	dauphin	vie	prestation	balcon	fermer	verre
	magique	nage	moana	chambre_spacieu	hambre_propre	serviette	pain
	tahaa	golf	assurer	hauteur	débordement	contre	regarder
	décorer	resto	nourriture_bon	professionnel	rue	reste	venir_chercher
10	beauté	hotel	concierge	vue_lagon	bagage	travail	perle
	grand_merci	ponton	four_seasons	décoration	marcher	éviter	là-bas
	charme	pension	couper_souffle	accès	route	direction	café
	accueil_chaleureux	beau_plage	anniversaire	personnel_agréable	hall	dommage	sortir
	chaleureux	matira	romantique	écoute	salle	coup	pearl
15	goût	snack	saint_regis	literie	port	parler	magasin
	relais	disposition	couple	redire	avion	bruit	coûter
	souvenir	location	avoir_impression	ponton	bruit	finir	chaud
	bonheur	possibilité	heureux	souriant	beau_vue	déception	marcher
	directeur	village	importer	confort	moderne	décevant	serviette
20 -	jardin_corail	sortie	voyager	animation	minute_pied	tarif	emmener
	décoration	bora	pouvoir_dire	équipé	café	impossible	chaise_long
	plaisir	eter	hâte	jacuzzi	chambre_grand	avis	bière
	attention	vue_lagon	organiser	personnel_sympathique	disposer	répondre	conduire
	charmant	contre	sûr	disposition	dormir	ménage	argent
25	atoritie	volure	plonger	tester	chambre_spacieux	froid	pizza
	Paysage	Activité	Service	Accueil	Chambre	Prix	other
				variable			

Figure 2 : les mots les plus associés aux topics

#### Annexe 3 : Résultats du modèle GAM

```
## Call: gam(formula = Note ~ s(Paysage) + s(Activité) + s(Service) +
      s(Accueil) + s(Chambre) + s(Prix), family = gaussian, data = topic2)
## Deviance Residuals:
   Min 1Q Median 3Q
## -1.55186 -0.11912 -0.01859 0.09889 0.61928
## (Dispersion Parameter for gaussian family taken to be 0.1011)
##
##
     Null Deviance: 2624.619 on 24300 degrees of freedom
## Residual Deviance: 2454.202 on 24276 degrees of freedom
## AIC: 13299.96
##
## Number of Local Scoring Iterations: NA
## Anova for Parametric Effects
##
                Df Sum Sq Mean Sq F value
                1 45.76 45.756 452.601 < 2.2e-16 ***
## s(Paysage)
                 1
                     7.56 7.561 74.787 < 2.2e-16 ***
## s(Activité)
## s(Service)
                1 80.05 80.047 791.792 < 2.2e-16 ***
                     7.41 7.411 73.308 < 2.2e-16 ***
## s(Accueil)
                 1
                 1 15.45 15.446 152.787 < 2.2e-16 ***
## s(Chambre)
                           6.764 66.910 2.98e-16 ***
## s(Prix)
                 1
                     6.76
## Residuals 24276 2454.20 0.101
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Anova for Nonparametric Effects
## Npar Df Npar F
                               Pr(F)
## (Intercept)
                  3 3.8002 0.009755 **
## s(Paysage)
## s(Activité)
                  3 0.4069 0.748061
## s(Service)
                  3 10.3591 8.262e-07 ***
## s(Accueil)
                   3 0.8779 0.451648
## s(Chambre)
                  3 10.5068 6.677e-07 ***
## s(Prix)
                   3 4.5063 0.003646 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```