

Percepção do turista durante a Pandemia da COVID-19 sobre o setor Hoteleiro de Aracaju/SE utilizando o método Eletronic Data Scraping (EDS)

Tourist perception during the COVID-19 Pandemic on the Hotel sector of Aracaju / SE using the Electronic Data Scraping (EDS) method

Percepción del turista durante la Pandemia COVID-19 en el sector Hotelero de Aracaju / SE utilizando el método Electronic Data Scraping (EDS)

Recebido: 14/10/2020 | Revisado: 22/10/2020 | Aceito: 25/10/2020 | Publicado: 27/10/2020

Luiz Diego Vidal Santos

ORCID: <https://orcid.org/0000-0001-8659-8557>

Universidade Federal de Sergipe, Brasil

E-mail: vidal.center@academico.ufs.br

Catuxe Varjão de Santana Oliveira

ORCID: <https://orcid.org/0000-0001-7276-4681>

Instituto Federal de Educação, Ciência e Tecnologia de Sergipe, Brasil

E-mail: catuxe.oliveira@academico.ifs.br

Francisco Sandro Rodrigues Holanda

ORCID: <https://orcid.org/0000-0003-3575-8105>

Universidade Federal de Sergipe, Brasil

E-mail: fholanda@infonet.com.br

Carlos Eduardo Celestino de Andrade

ORCID: <https://orcid.org/0000-0002-8362-6177>

Universidade Federal de Sergipe, Brasil

E-mail: eng2carlos.eduardo@gmail.com

Maraiza Santana dos Santos

ORCID: <https://orcid.org/0000-0001-7778-6978>

Universidade Federal de Sergipe, Brasil

E-mail: mara.iza.santana@hotmail.com

Wilsonita de Melo Ubirajara

ORCID: <https://orcid.org/0000-0001-8643-6347>

Universidade Federal de Sergipe, Brasil

E-mail: will_ubirajara@yahoo.com.br

Resumo

Esta pesquisa se pauta na compreensão de dados de opinião coletados nas Online Travel Agencies referente à percepção dos usuários hospedados na cidade de Aracaju, capital do Estado de Sergipe, por meio de uma abordagem mista de análise exploratória, com abordagem quantitativa. Assim, foi executado o método de Eletronic Data Scraping (EDS), com estudo posterior do discurso dos usuários por meio do método de análise lexical dos conteúdos com o software IRaMuTeQ. As maiores notas, para os ofertantes de vagas de hospedagem, estavam entre 8 e 10, em um escore que vai de 1 a 10, demonstrando uma boa aceitação geral dos serviços oferecidos. Entre os pontos positivos mais declarados pelos usuários tem-se o baixo custo, a polidez no atendimento, segurança e a facilidade no deslocamento. Já os pontos negativos e dificuldades mais observados foram estrutura física e conforto térmico. Os resultados obtidos por este método podem ser facilmente utilizados por empresas da cidade de Aracaju para prospecção comercial dentro do turismo doméstico, bem como por órgãos públicos que desejam atrair mais visitantes para capital Sergipana.

Palavras-chave: Dados eletrônicos; Turismo; Ciência de dados; Hospedagem urbana; Cidade.

Abstract

This research is based on the understanding of opinion data collected at Online Travel Agencies regarding the perception of users staying in the city of Aracaju, capital of the State of Sergipe, through a mixed approach of exploratory analysis, with a quantitative approach. Thus, the Electronic Data Scraping (EDS) method was performed, with further study of the users' discourse using the lexical analysis method of the contents using the IRaMuTeQ software. The highest scores for accommodation providers were between 8 and 10, in a score ranging from 1 to 10, demonstrating a good general acceptance of the services offered. Among the positive points most declared by users is the low cost, politeness of service, security and travel. The most observed negative points and difficulties were physical structure and thermal comfort. The results obtained by this method can be easily used by companies in the city of Aracaju for commercial prospecting within domestic tourism, as well as by public agencies that wish to attract more visitors to the capital of Sergipe.

Keywords: Electronic data; Tourism; Data science; Urban accommodation; City.

Resumen

Esta investigación se basa en la comprensión de los datos de opinión recogidos en las Agencias de Viajes Online sobre la percepción de los usuarios que se alojan en la ciudad de Aracaju, capital del Estado de Sergipe, a través de un enfoque mixto de análisis exploratorio, con enfoque cuantitativo. Así, se realizó el método de raspado electrónico de datos (EDS), con un estudio posterior del discurso de los usuarios a través del método de análisis léxico de los contenidos con el software IRaMuTeQ. Las puntuaciones más altas, para los oferentes de las vacantes de alojamiento, fueron entre 8 y 10, en una puntuación que va de 1 a 10, lo que demuestra una buena aceptación general de los servicios ofrecidos. Entre los puntos positivos más declarados por los usuarios se encuentra el bajo costo, la cortesía en el cuidado, la seguridad y la facilidad de viaje. Los puntos negativos y dificultades más observados fueron la estructura física y el confort térmico. Los resultados obtenidos por este método pueden ser fácilmente utilizados por las empresas de la ciudad de Aracaju para la prospección comercial dentro del turismo nacional, así como por las agencias públicas que desean atraer a más visitantes a la capital Sergipana.

Palabras clave: Datos electrónicos; Turismo; Ciencia de datos; Vivienda urbana; Ciudad.

1. Introdução

No Brasil o turismo receptivo não é tão explorado se comparado com países desenvolvidos. Por exemplo, enquanto no Brasil tivemos 5,2 milhões de turistas estrangeiros em 2010, a França foi visitada por 76,8 milhões, e países como Estados Unidos, China e Espanha receberam entre 50 e 60 milhões (Beni, 2019). Os meios de hospedagem no setor turístico são os responsáveis por acomodar os turistas e, dessa forma, tem forte influência na avaliação final da experiência vivida na viagem. Por isso, é importante que os hotéis e pousadas se adequem aos padrões dos potenciais turistas que irão visitar o Brasil em seus respectivos destinos e segmentos turístico em suas diversas localidades para que estejam sempre preparados durante os próximos anos (Telles & Cardoso, 2019).

O estado de Sergipe já vem se destacando no cenário nacional com relação ao aumento da demanda turística, apesar de ainda não ser um dos principais destinos turísticos. Em janeiro de 2012, a média registrada pelos meios de hospedagem atingiu 92% de ocupação (Aragão, 2017). O estado de Sergipe ainda possui um potencial de crescimento turístico grande porque seu fluxo turístico é, majoritariamente, intrarregional (60,4%). Os estados da Bahia e de São

Paulo são responsáveis por 49,2% dos turistas que visitaram o estado, destes 31,5% são baianos e 17,7% paulistas (Morais et al., 2015).

De acordo com a importância do setor turístico e consequente aumento do uso de TI, ou seja, o uso da internet e suas ferramentas que fazem ser possível um feedback, como também identificar informações dos usuários sejam eles turistas, hóspedes ou visitantes que explora ou já exploram os serviços ofertados pela hotelaria (Buhalis, 1998; Richard et al., 2018)^[A1].

Com a ampla difusão do sistema interligado e global de redes de computadores, tornou-se cada vez mais dinâmico o trânsito de informações alojadas em diversificados sítios eletrônicos, com este efeito é possível compartilhar experiências de estilo de vida que denotam um modo de comportamento (Marx, 2018). Inseridos neste escopo, estão os endereços eletrônicos especializados em viagens, ou seja, Online Travel Agencies (OTA) (Chang et al., 2019).

Ocorre ainda que a implementação de estratégias de comercialização vem se consolidando como uma das diretrizes mais sensatas no disputado campo virtual. Dada esta perspectiva, as apresentações dos produtos turísticos sofreram mudanças significativas na forma como são percebidos e vendidos. Esta quebra de paradigmas alterou o comportamento do consumidor viajante, provocando uma transformação no modo como se planeja uma viagem (Pinto & Castro, 2019).

Nota-se contemporaneamente o surgimento de variados endereços eletrônicos especializados em viagens, isto é, negócio mediado por uma organização virtual em que vários comerciantes se inscrevem e oferecem seus produtos, o chamado Marketplace (Watanabe, 2020). Além disso, é forte a percepção de que o ambiente virtual atual é delineado por uma gama variada de possibilidades de hospedagens, que variam desde residências domiciliares até grandes conglomerados de hotéis (Talwar et al., 2020).

Inserido neste contexto, ainda é possível notar a racionalização de informações relacionadas à opinião dos usuários sobre o produto utilizado (M. B. de O. Silva, Arruda, Souza, & Mariani, 2019), denotando este aspecto como sendo uma forte fonte de dados de experiência do usuário (Ashkezari-Toussi, Kamel, & Sadoghi-Yazdi, 2019).

Com tal característica, estas informações têm potencial para serem analisadas para aferir a experiência dos usuários das redes de hotéis de uma determinada região, indicando o respectivo perfil hoteleiro (Palazzo, 2021). O acesso a este tipo de informação é bastante prático e passou a ser critério relevante no momento de escolher o destino e o local de hospedagem, pois auxiliam a tomada de decisão pautada em critérios mais seguros (Suntikul

& Jachna, 2016).

Esse tipo de Informação Geográfica Voluntária, sigla em inglês (VGI), quando usado para coleta de dados, facilita o alcance do público objeto da pesquisa, reduz o custo e o tempo da fase de coleta de dados maximizando a correção de problemas, assim como facilita o processo de amostragem (J. Yan, 2019); para mais, traz como uma grande vantagem a maior veracidade da opinião disponibilizada, pois a realização do comentário é, em sua maioria, disposta nas páginas por vontade própria do usuário (Soares et al., 2018).

Esta tendência gerada a partir da opinião de um usuário tem, cada vez mais, se tornado a força motriz para o setor de turismo investir em planejamento específico, para que assim possa disponibilizar produtos mais atrativos e serviços focados na experiência do usuário, atendendo desta forma as pretensões de seus visitantes elevando sua capacidade de competir no mercado (Junior et al., 2020).

Esse tipo de Informação Geográfica Voluntária (IGV), quando usado para coleta de dados, facilita o alcance do público objeto da pesquisa, reduz o custo e o tempo da fase de coleta de dados, possibilita a melhora da recuperação de problemas, assim como facilita o processo de amostragem (Y. Yan et al., 2017); além disso, tem como grande vantagem a veracidade da opinião disponibilizada, pois o comentário é, em sua maioria, realizado nas páginas por vontade própria do usuário (J. C. Soares et al., 2018).

Estas plataformas incentivam os comércios ligados ao setor de turismo a investirem na melhor disponibilização das suas informações na web e, conseqüentemente, na melhoria dos produtos e serviços oferecidos, tornando-os mais atrativos e focados na experiência do usuário, atendendo, desta forma, os anseios de seus visitantes e ganhando competitividade no mercado (Costa & Galina, 2016).

Nesta lógica, a mineração de dados vem evoluindo e disponibilizando técnicas, procedimentos e softwares computacionais automatizados adequados para auxiliar a extração de informações úteis contidas em um elevado montante de dados. De acordo com Baskaran & Ramanujam (2018), existem diversos usos para coleta de dados abertos, que incluem: mineração de opinião, análise de sentimento, modelagem de tópicos e análise de tendências.

Para todos estes métodos a literatura reporta a necessidade de se estabelecer uma coleta inicial através de algum método, aqui foi escolhido o Eletronic Data Scraping (EDS), sigla em inglês para Raspagem Eletrônica de Dados. Essa técnica é desenvolvida segundo três etapas distintas: na primeira etapa, são realizados a coleta e a cópia dos dados conforme parâmetros escolhidos pelo pesquisador, para em seguida realizar o salvamento (segunda etapa) e finalmente suceder com a análise dos dados na terceira etapa (Ahamad, Mahmoud, &

Akhtar, 2017).

Neste cenário, este trabalho se pauta na compreensão de dados coletados a partir das opiniões registradas nas Online Travel Agencies, de maneira a estudar a percepção dos usuários turistas que já se hospedaram na rede hoteleira da cidade de Aracaju, capital do Estado de Sergipe, por meio de uma abordagem heterogênea de análise exploratória, com abordagem quantitativa, empregando a técnica de software Eletronic Data Scraping (EDS) para extrair informações de websites, bem como a utilização da ferramenta IRaMuTeQ para análise lexical dos conteúdos para a construção de indicadores. Para isso a pesquisa foi delineada em dois momentos: no primeiro momento foi realizada uma raspagem de dados de opiniões contidas nos principais sites de oferta de vagas para hospedagem; enquanto no segundo momento foi realizada uma análise do discurso utilizando a análise quantitativa de dados textuais através da utilização de Unidades de Contexto.

2. Metodologia

O fluxo de trabalho do estudo de caso começa com a coleta de dados nos sites de oferta de vagas Airbnb, Booking.com e Tripadvisor, assim como o site de geolocalização “Google Local Guides Platform”, limpeza de dados e pré-processamento. Posteriormente foi realizada a análise dos dados macros, objetivando a organização dos dados e a análise de opinião baseada no discurso buscando alcançar a compreensão quantitativa e qualitativa dos dados. As seções a seguir detalham os métodos supracitados.

2.1 Coleta de dados, limpeza e pré-processamento

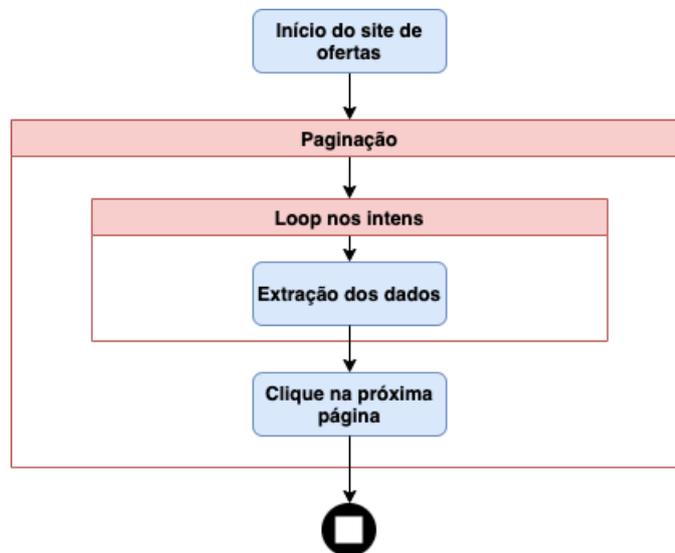
Os dados de opinião sobre hospedagem no setor hoteleiro da cidade de Aracaju-SE foram selecionados a partir dos comentários de opinião realizados pelos usuários turistas em OTAs. Para delimitar o ambiente de estudo, foram adotados como indicadores os termos “Hospedagem em Aracaju-SE” e “Vagas de hotéis em Aracaju-SE”. Como linha temporal foram selecionados os comentários realizados de 2018 a 2020.

Dos empreendimentos pesquisados na cidade de Aracaju-SE constam: hotéis e pousadas na sede do município; resorts e cabanas de aluguel no município; apart-hotel; apartamentos de moradia por temporada e casas de veraneio. Objetivando preservar a identidade das empresas envolvidas neste estudo, estas serão descritas como: (HT) Hotéis, (PO) Pousadas, (AP) Apartamentos, (CA) Cabanas e Hostels, (RE) Resorts e (CV) Casas de

Veraneio, seguido pela sua posição numérica de classificação no escore geral.

Na fase inicial deste estudo, foi configurada, para a coleta automatizada, a raspagem web Octoparse (versão 7.1.2, Octopus Data Inc., Walnut, CA, USA). Este software é um sistema de Eletronic Data Scraping que permite a raspagem dos dados presentes em sites abertos. O software foi programado para extrair os comentários sobre a percepção da estadia, desta forma, o processo de extração foi realizado em cinco passos, a saber: acesso ao site de ofertas; paginação dos locais de hospedagem; loop de raspagem dos itens; extração dos dados e abertura de novas páginas. Este processo está descrito na Figura 1. Durante a coleta dos dados, os comentários que se mostraram em duplicidade foram suprimidos.

Figura 1 - Espaço de trabalho para raspagem de comentários em sites de ofertas de vagas para hospedagens em Aracaju-SE.



Fonte: Os autores (2020).

2.2 Análise quantitativa de dados textuais

Após a seleção dos comentários diretamente relacionados ao tema proposto nesta pesquisa, foram realizadas lexicométricas e análises de discursos nos textos presentes com o auxílio do software IRaMuTeQ (Conde, 2015; da Silva Cezar, Bandeira, Dornelles, de Mello Barcellos, & Corso, 2018). Através da utilização de Unidades de Contexto Iniciais (UCIs) na construção do modelo para análise, foi realizada uma prévia correção no conjunto de dados brutos mantendo as palavras analisáveis, os termos caracterizados como substantivos, adjetivos e verbos. As palavras que geralmente não são relevantes para uma análise, como

preposições e advérbios, foram excluídas. Assim, ao trabalhar com corpus textuais cada conjunto de texto deve compor uma UCI. Um conjunto de UCI é conhecido como corpus de análise, os quais o software segmenta em textos de aproximadamente três linhas, chamados de Segmento de Texto (ST) (Fernandes, 2018). Após a segmentação dos textos, o software realiza uma Categorização Hierárquica Descendente (CHD) dando origem a classes lexicais caracterizadas pelo vocabulário e a segmentos de textos que partilham o mesmo vocábulo.

A Classificação Hierárquica Descendente (CHD) consiste em um tipo de análise de conglomerado que categoriza as palavras obtidas em classes lexicais (eixos). A análise avalia a constância e os arranjos das palavras ativas que estão no corpus textual usando os dados das tabelas de contingência das palavras existentes (Carvalho, Mota, & Saab, 2020; Mendes, Tonin, Buzzi, Pontarolo, & Fernandez-Llimos, 2019). Neste sentido, as diferentes classes resultantes representam o espaço de sentido das palavras narradas, surgindo, assim, elementos pertencentes aos temas observáveis da pesquisa (Gavasso, Fernandes, & Andrade, 2016). Por fim, é importante ressaltar que a avaliação da metodologia proposta neste estudo foi mediada pela técnica desenvolvida por Costa et al. (Costa, Zoltowski, Koller, & Teixeira, 2015), tendo como base a Assessment of Multiple Systematic Reviews (AMSTAR).

2.3 Análise quantitativa de dados textuais

Cada site minerado continha uma escala de escore que variava de 1 a 5 e 1 a 10. Para manter uma homogeneidade dos conteúdos analisados, foi realizada uma conversão dos dados médios presentes em cada estabelecimento encontrado para o escore que mantém a margem de pontos entre 1 a 10. A fórmula está apresentada na Equação 1:

$$Mp = \frac{\sum_{i=1}^n (p_i * X_i)}{\sum_{i=1}^n p_i} \quad (1)$$

Onde:

Mp = Média aritmética ponderada,

Pi = Pesos para o escore adotado e

Xi os valores dos dados.

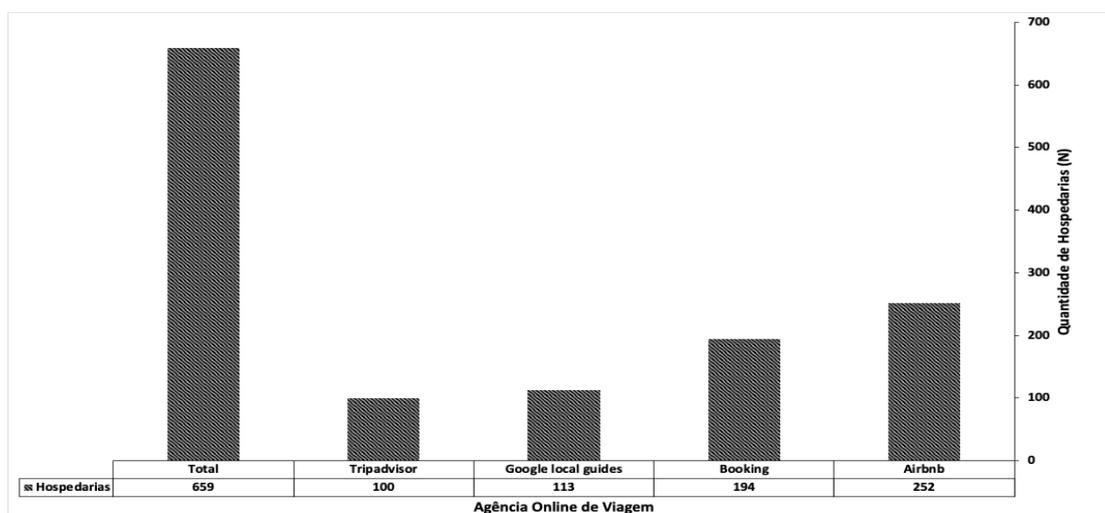
3. Resultados e Discussão

3.1 Avaliação da raspagem dos dados e compreensão dos resultados

A Figura 2 resume as estatísticas descritivas dos quatro principais bancos de dados on-line sobre oferta de vagas de hospedagens utilizados nesta pesquisa. Das quatro OTAs estudadas, a Airbnb apresentou um maior volume de informações inéditas disponíveis sobre o objetivo da pesquisa (64%), Google Local Guides (17%), seguido pelo Booking.com (11%) e o TripAdvisor (8%), de um total de 659 locais disponíveis.

Tais dados demonstram que houve alto índice de adeptos à plataforma digital AirBnB no município pesquisado, ocasionando um aumento importante no número de oferta de vagas de hospedagem alternativas, entendendo-se aqui a hospedagem alternativa como todo o tipo de alojamento consolidado no contexto da economia compartilhada, à exceção dos hotéis (Chiarelli, 2019). Este comportamento pode estar ligado às mudanças no processo cultural, através da melhoria da apropriação econômica das classes que anteriormente não possuíam poder aquisitivo para realização de viagens (Aquino et al., 2020), à busca por ambientes mais sustentáveis (Rahman, Reynolds, & Svaren, 2012), assim como ao incentivo ao empreendedorismo local do setor, gerando maiores possibilidades de implementação de negócios onde antes era apenas uma moradia. Atualmente, essa categoria de estabelecimentos é conhecida também como low-cost por se tratar, em sua maioria, de hospedagens de baixo custo.

Figura 2. Porcentagem de comentários válidos por OTA.



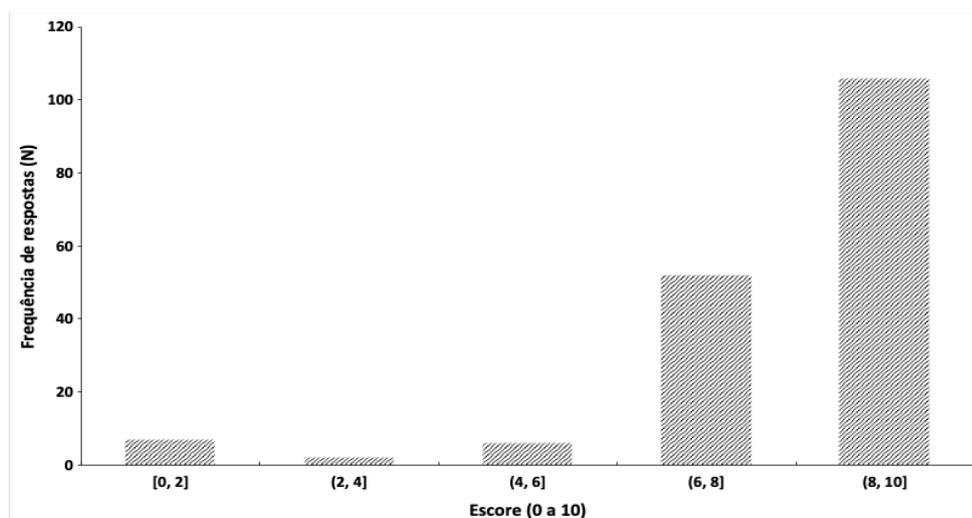
Fonte: Os Autores (2020).

Outro fator atípico para este comportamento pode ser explicado também quando observamos as mudanças econômicas ocorridas durante o período da oferta das vagas, período este acometido pela pandemia do SARS-CoV-2. Muitos países limitaram as viagens para conter a pandemia, essas medidas restringiram as ofertas de vagas formais de hospedagem abrindo espaço para as vagas alternativas (Devi, 2020). Segundo a Nota Informativa produzida pelo Ministério da Economia do Brasil (ME, 2020), a epidemia pode ter provocado comportamentos precaucionais na população como diminuição de viagens, reduções de jornadas de trabalho, ou ainda adoção de hospedagens como ocupação *hosting-office*, fatores que podem ter levado à redução de ofertas de vagas em locais formais, uma vez que estes foram forçados ao controle mais robusto quanto aos contatos pessoais e à proximidade com os hóspedes. Farmaki et al. (2020) em sua pesquisa sobre a percepção de anfitriões de hospedagens alternativas observou que alguns anfitriões estavam otimistas, pois a pandemia trouxe oportunidades que transformaram positivamente o setor de acomodação.

Ademais, a pandemia reforçou a demanda por acomodações alternativas, pois os hotéis correm maior risco de infecção e, assim, as pessoas preferem ficar em tipos mais isolados de acomodação com menos contato pessoal.

Observando o escore para avaliação geral do serviço prestado pelos locais, conforme apresentado na Figura 3, é possível notar uma maior frequência de pontos acima de 6 (variando de 6 a 10). Assim, quando observados os pontos $p < 6$ é possível notar uma baixa adesão (percepção ruim do local).

Figura 3. Escore geral da percepção dos usuários da rede de hospedagem.



Fonte: Os Autores (2020).

3.2 Análise lexical dos comentários minerados

O programa IRaMuTeQ possui a vantagem de organizar, codificar e categorizar as informações, o que proporcionou a localização ágil de todo o corpus utilizado na pesquisa qualitativa dos comentários coletados nesta pesquisa. Para o alcance dos resultados, buscando realizar a CHD, o software percorre quatro etapas: a codificação, a preparação textual inicial e CHD, assim como a interpretação das classes finais (Amaral-Rosa et al., 2019).

Buscando construir todos os conjuntos de textos necessários ao desenvolvimento do corpus de análise, a preparação inicial dos textos se deu com a inclusão em um único corpus construído com os comentários depositados nas OTAs selecionadas. Assim, dos 833 comentários resultantes, formou-se 1 Unidade de Contexto Inicial (UCI). Cada Segmento de Texto (ST) foi separado na UCI por uma linha contendo somente uma variável (n), descrita conforme um número dado a cada artigo estudado (**** *COMENTARIO_n).

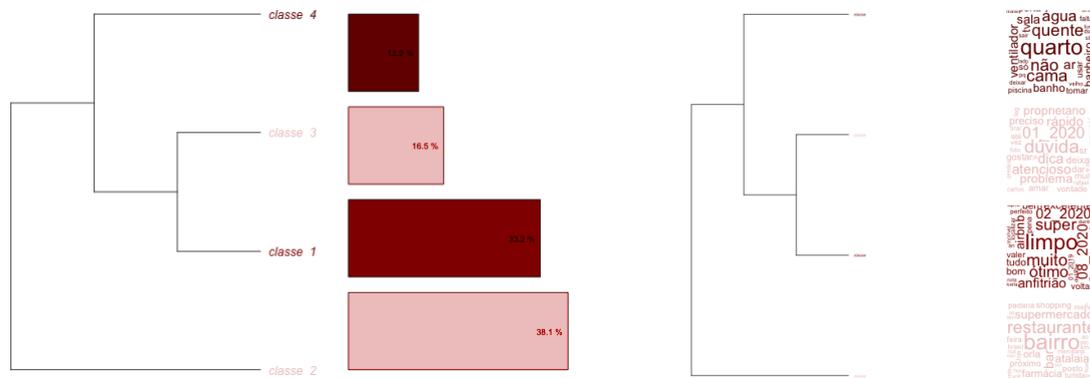
A UCI foi organizada como arquivo de texto (formato .txt), utilizando caracteres no padrão UTF-8 (Unicode Transformation Format 8bit codeunits), utilizando o software Notepad++. Foram suprimidos todos os caracteres especiais contidos no texto (*%&^\$), foi realizada também a amarração de palavras compostas, como por exemplo: “café_da_manhã”, evitando que o software as identificasse como palavras distintas.

Da UCI criada, originaram-se 503 Segmentos de textos (STs), sendo 394 válidos, os quais continham 17847 ocorrências (Palavras, Formas ou Vocábulo), das quais 3324 foram palavras distintas e 2072 hapax (palavras com uma única ocorrência). Exibindo uma frequência de 11,61% para ocorrência de palavras e uma frequência média de 62,33% de ocorrência por ST. O programa utiliza para a criação da matriz analítica de palavras o teste estatístico qui-quadrado (χ^2), o qual demonstra a força associativa entre as palavras contidas em cada cluster resultante (Camargo & Justo, 2013). Uma boa força associativa é considerada quando o χ^2 apresenta valores acima de 3,84, representando p-value <0,0001 (Souza & de Oliveira Folharini, 2019). Um valor inferior a 3,84 demonstra pouca correlação entre as variáveis.

Na Figura 4, é possível observar a Classificação Hierárquica Descendente (CHD) formada pela leitura das palavras que apresentaram os maiores valores de qui-quadrado, como também os segmentos de textos característicos com maior pontuação de score por soma absoluta do (χ^2) das palavras mais destacadas nos segmentos, possibilitando o alcance dos objetivos primários da pesquisa que foi analisar a percepção dos usuários hospedados na cidade de Aracaju-SE por meio de uma abordagem mista de análise exploratória com

abordagem quantitativa.

Figura 4. Dendrogramas resultante das classes de palavras com maior qui-quadrado (χ^2).



(A) Porcentagem de STs em cada classe de palavras

(B) Nuvens das principais palavras contidas no CHD

Fonte: Elaborado pelos autores a partir do software IRaMuTeQ (2020).

É interessante notar que dos comentários analisados surgiram dois grandes eixos e neles 4 classes lexicais caracterizadas pelo vocabulário e por segmentos de textos que compartilham o mesmo vocabulário. O primeiro eixo foi desmembrado em três subcorpus, o segundo em apenas um subcorpus textual (Figura 5).

3.2.1 Classe 1 – Polidez, bom Trato e acolhimento

A Classe 1, localizada no primeiro eixo, foi construída através de 131 STs (33,25%) do total. Pôde-se constatar, por meio da observação da representação fatorial (Figura 5), que esta classe apresenta termos que tratam de fatores positivos da hospedagem relacionados ao local e atendimento pessoal durante o período de estadia. Os trechos que melhor descrevem esta classe são: (AP 53) “No geral foi uma estadia agradável” e (CV 14) “Anfitriã muito atenciosa e receptiva super recomendo”, (HT 30) “Um Atendimento muito bom e a refeição excelente também” e (HT 31) “Achei o atendimento dos funcionários fantástico”.

Similar a esta pesquisa, em um estudo sobre as percepções dos turistas espanhóis, López & Amaya (2019) constataram que, apesar da preferência geral dos participantes pela simpatia e solidariedade, as estratégias de polidez também foram tratadas como necessárias nesse contexto de aprovação, demonstrando que tanto a simpatia quanto a polidez no tratamento aos hóspedes influenciam na boa imagem do local. Os dados aqui analisados

demonstram que as hospedarias oficiais e alternativas da cidade de Aracaju-SE, em sua maioria, agradaram os hóspedes que ali se instalaram. Patwardhan et al. (2020) verificaram também que a sensibilidade emocional para com os clientes tem efeito positivo no apego ao lugar e na lealdade do retorno ao destino.

3.2.2 Classe 2 – Localização e Segurança

A Classe 2 refere-se à localização e proximidade dos pontos turísticos da cidade, e corresponde a 38,1% do total dos STs analisados (150). Nela, os termos remetem tanto à sensação de segurança do local, quanto à facilidade de locomoção quando se busca chegar aos pontos turísticos, como a região praieira da cidade. Os comentários que melhor explicam esta classe são: (HT 04) “[O hotel está bem localizado na orla de atalaia e bem próximo à passarela do caranguejo...]”, (PO 34) “[...Pudemos comprar os passeios com conforto e segurança sem contar que os preços estavam todos dentro da média ou abaixo]”, (HT 07) “O Hotel fica entre o oceanário e a passarela do caranguejo um caminhada de mais ou menos 10 minutos para cada além de passarela para a praia e um ponto do caju bike exatamente em frente” e (HT 52) “[...Ressalto ainda que o aeroporto fica bem próximo onde você pode dispensar o traslado e pegar um carro de transporte alternativo onde você tem um gasto de no máximo R\$ 12,00]”.

O fator segurança e a localização são pontos cruciais para a escolha de um local, Tavares & Fraiz (2020) estudando a preferência dos hóspedes pelos hostels na cidade do Porto-PT, observaram em sua pesquisa que o primeiro fator a ser retratado pelos hóspedes quanto à qualidade do local foi a segurança em geral, mostrando a importância dada a esta característica. Num ambiente informal, onde as pessoas mal se conhecem, é uma preocupação e um aspecto a ser considerado, alguns parâmetros pesquisados pelos turistas são a existência de portas de segurança, a vigilância da polícia nas redondezas, a existência de funcionário na portaria, entre outros (Tavares & Brea, 2020).

Observando os comentários sobre a proximidade aos pontos turísticos, a grande maioria destes foram qualitativamente positivos, um fator importante da cidade é que a sua orla concentra a maior parte das atrações turísticas, devido ao seu comprimento em relação a toda a cidade, assim como a proximidade desta aos pontos turísticos urbanos. As atividades de lazer na praia tornaram-se os projetos mais atrativos para os turistas, o que se deve à qualidade das praias, espaços abertos, bares e restaurantes, o que dá grande apelo às cidades litorâneas focadas em turismo (Zhang et al., 2020).

3.2.3 Classe 3 – Acomodações e Custo Benefício

Ainda no contexto de espaço físico tem-se a Classe 3 (eixo 1), esta classe se constitui por 65 STs, 16,5% do total. As principais STs tratam de estrutura física, preço e acessibilidade econômica, ela pode ser melhor explicada pelos comentários: (PO 04) “Agradável custo benefício bom” e (HT 17) “Hotel com custo benefício muito bom localizado bem na orla de atalaia café da manhã diversificado sempre repondo os alimentos bom atendimento banheiros e quartos bem limpos e a piscina é legal também” e (HT 17) “Apenas o café da manhã que poderia ter mais opções, atendeu as expectativas e é recomendável em relação ao custo benefício”.

Esta classe apresentou maior volume de estabelecimentos formais, principalmente para as classificadas como Pousadas (PO), uma característica importante deste comportamento é que as hospedagens do tipo pousada são bastante procuradas por hóspedes que buscam maior custo benefício durante uma rápida estadia, tais como o “Pernoite”, formando assim um nicho para este tipo de hospedaria. Mesmo que o valor seja indiretamente comentado, nesta classe é importante observar que o termo mais utilizado foi “custo benefício”, demonstrando que o valor monetário pela hospedagem por si só não é um fator determinante para obter uma boa classificação pelos hóspedes, e sim o custo benefício deste valor (Dušek, Štumpf, & Vojtko, 2019). Ahn & Kwon (2020) estudaram o custo, a expectativa e a intenção de retorno dos clientes de um hotel, e observaram que os clientes consideram mais o valor percebido do que o custo percebido, e que o valor e a boa aceitação estão estreitamente atrelados à qualidade do produto/serviço.

3.2.4 Classe 4 – Deficiências dos locais

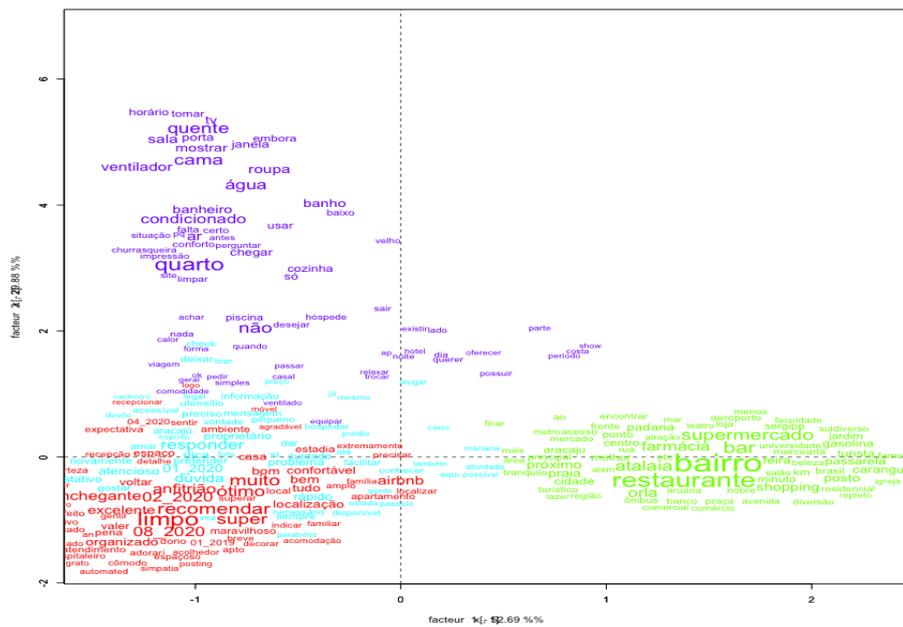
Na Classe 4 emergiram 48 STs, correspondendo a 12,18% do total. É possível observar na Figura 4 a distância desta classe para as demais, ela retrata os pontos negativos observados pelos hóspedes durante a estadia. Os comentários que melhor retratam esta classe são: (AP 16) “[...Além do banheiro sem porta uma varanda sem ralo e o prédio é novo”], (PO 10) “O café da manhã atendeu bem apesar de não ter muitas variedades”, “Só achei um pouco desorganizado porque ainda estava em reforma” e (HT 48) “Meu quarto não tinha telefone a solução da moça da recepção foi me dirigir à recepção se precisar de algo restaurante do hotel não tava funcionando”.

A influência negativa por parte de acomodações para hospedagem no Brasil também

foi observada por Schallhorn (2019) que estudou a percepção de turistas hospedados na Cidade do Rio de Janeiro-RJ durante a copa do mundo da FIFA ocorrida no Brasil em 2014 e durante as olimpíadas de 2016.

De modo geral, observando a harmonia das palavras expressas na Análise Fatorial de Correspondência (AFC) na Figura 5, é possível observar que houve uma maior frequência de comentários positivos quanto à estadia na rede hoteleira de Aracaju-SE, e que por mais que sejam apontados pontos negativos, estes são em sua maioria deficiências de treinamento e problemas pontuais das acomodações, que quando citados eram rapidamente compensados por vantagens inerentes da região, tais como: proximidade à praia, zelo, polidez, acolhimento, etc.

Figura 5. Análise Fatorial de Correspondência (AFC) das quatro classes formadas pelos comentários presentes nos OATs.



Fonte: IRaMuTeQ (2020).

Assim, o método misto de pesquisa aqui utilizado se mostrou efetivo à medida que emergiram classes de textos expressando bem a percepção dos hóspedes que utilizaram os diversos meios de hospedagens presentes na cidade de Aracaju-SE. Foi traçado o perfil geral do sistema hoteleiro e locação low cost, suas características positivas e negativas, e seus pontos a serem melhorados. Esta pesquisa se mostra importante, pois atualmente a literatura genérica de turismo e hospitalidade carece de informações sobre operações de pequenos negócios (Moufakkir & Alnajem, 2017).

Apesar de sua popularidade entre turistas nacionais e internacionais, ainda faltam informações sobre acomodações de albergues. Analisar esta frequência de palavras pode ser uma maneira interessante de destacar também as inclinações dos gostos pessoais e a aceitação dos locais pelos clientes. Por fim, a categoria que mais foi representada na análise temática de conteúdo foi a Classe 1 (Polidez, bom trato e acolhimento) que refere-se aos pontos positivos inerentes ao acolhimento e ao espaço no geral, e são fatores importantes para atração de novos clientes (Brochado, Rita, & Gameiro, 2015; C. C. B. da Silva & Calic, 2013). É importante destacar que por meio dos 394 STs válidos, 64 hóspedes (16,24%) citaram que voltariam a se hospedar nos mesmos locais e 102 (25%) retornariam à cidade de Aracaju.

4. Considerações Finais

Dado o caráter exploratório do presente estudo, o seu desenvolvimento mostrou que turistas estão relativamente satisfeitos com os meios de hospedagens oferecidos pela cidade de Aracaju-SE, no que concerne à localização, segurança e atendimento local. Cada categoria observada na Classificação Hierárquica Descendente apresentou resultados diferentes para palavras relacionadas quanto à percepção do turista que esteve hospedado na rede hoteleira na cidade de Aracaju-SE.

A análise possibilitou a identificação das características mais proeminentes para cada categoria. Apesar dos resultados promissores obtidos, mais pesquisas são necessárias para validar esse método para uso na análise de contribuições públicas.

Os dados resultantes desta pesquisa podem ser facilmente replicados. A abordagem descrita aqui pode ser usada para avaliar diversos tipos de negócios que tenham a possibilidade de coletar comentários e dados sobre a percepção dos clientes, de acordo com o interesse de cada pesquisador, estudante ou empresa.

Embora este trabalho seja capaz de descrever minuciosamente a conjuntura dos comentários sobre a satisfação dos usuários, são recomendadas futuras pesquisas em outras áreas, a fim de que sejam apuradas todas as fases e ferramentas utilizadas nesta pesquisa.

Referências

Ahamad, D., Mahmoud, D., & Akhtar, M. M. (2017). Strategy and implementation of web mining tools. *International Journal of Innovative Research in Advanced Engineering*, 4, 01–07.

Ahn, J., & Kwon, J. (2020). Green hotel brands in Malaysia: Perceived value, cost, anticipated emotion, and revisit intention. *Current Issues in Tourism*, 23(12), 1559–1574. <https://doi.org/10.1080/13683500.2019.1646715>

Aragão, I. R. (2017). Espaço e pesquisa qualitativa sobre as festas católicas e o turismo religioso em Sergipe. *Interfaces Científicas-Humanas e Sociais*, 5(3), 61–68.

Ashkezari-Toussi, S., Kamel, M. & Sadoghi-Yazdi, H. (2019). Emotional maps based on social networks data to analyze cities emotional structure and measure their emotional similarity. *Cities*, 86, 113–124.

Baskaran, U. & Ramanujam, K. (2018). Automated scraping of structured data records from health discussion forums using semantic analysis. *Informatics in Medicine Unlocked*, 10, 149–158.

Bragaglia, A. P. & Bastos, L. L. A. (2017). O valor de aprovação social no consumo e na publicidade: Uma discussão à luz da Ética. *Revista Fragmentos de Cultura - Revista Interdisciplinar de Ciências Humanas*, 27(2), 258–270. <https://doi.org/10.18224/frag.v27i2.5061>

Beni, M. C. (2019). *Análise estrutural do turismo*. Senac.

Buhalis, D. (1998). Strategic use of information technologies in the tourism industry. *Tourism management*, 19(5), 409–421.

Camargo, B. V. & Justo, A. M. (2013). IRaMuTeQ: Um software gratuito para análise de dados textuais. *Temas em psicologia*, 21(2), 513–518.

Cantante, M. C. (2018). Social media, importância no turismo de afinidades. *Dos Algarves: A Multidisciplinary e-Journal*, (32), 107–115.

Carvalho, D. N. (2020). *Efeito da Restauração Ecológica na Recuperação e Conservação da Avifauna em Fragmentos de Mata Ciliar no Rio Uberabinha, em Uberlândia, MG*. <https://doi.org/10.14393/ufu.di.2019.2597>

Cebi, N., Dogan, C. E., Olgun, E. O. & Osman Sagdic. (2018). A survey of free glutamic acid in foods using a robust LC–MS/MS method. *Food Chemistry*, 248, 8–13. <https://doi.org/10.1016/j.foodchem.2017.12.033>

Chang, Y.-W., Hsu, P.-Y., & Lan, Y.-C. (2019). Cooperation and competition between online travel agencies and hotels. *Tourism Management*, 71, 187–196.

Chang, Y.-W., Hsu, P.-Y., & Lan, Y.-C. (2019). Cooperation and competition between online travel agencies and hotels. *Tourism Management*, 71, 187–196.

Costa, T., & Galina, S. (2016). A relação virtuosa entre empreendedorismo e capital social: Um estudo em empreendimentos de turismo em espaço rural em Portugal e no Brasil. *Tourism & Management Studies*, 12(2), 57–69.

Chiarelli, M. C. (2019). Os fatores influenciadores da intenção de procura online por hospedagem alternativa (MasterThesis, Instituto Politécnico de Lisboa, Escola Superior de Comunicação Social). Instituto Politécnico de Lisboa, Escola Superior de Comunicação Social. Recuperado de <https://repositorio.ipl.pt/handle/10400.21/11506>

Conde, D. C. (2015). Lexicometria e análise do discurso. *Revista da ABRALIN*, 14(2).

Devi, S. (2020). Travel restrictions hampering COVID-19 response. *The Lancet*, 395(10233), 1331–1332. [https://doi.org/10.1016/S0140-6736\(20\)30967-3](https://doi.org/10.1016/S0140-6736(20)30967-3)

Dušek, R., Štumpf, P. & Vojtko, V. (2019). Geomarketing—Tool for consumer spending estimation in the Czech tourism & hospitality market. *Global Business and Finance Review*, 24(1), 14–26. Scopus. <https://doi.org/10.17549/gbfr.2019.24.1.14>

Morais, R. S. de, Oliveira, R. A. de, & Azevedo Paim, A. H. C. de. (2015). As Pesquisas de Demanda Turística em Minas Gerais: evolução histórica, principais desafios e perspectivas de futuro. *Revista Turismo em Análise*, 26(1), 21-37.

Galindo, A. L. C. & Santos, V. M. L. dos. (2018). Desenvolvimento Regional Sustentável: Desafios e Alternativas para Potenciais Obtentores do Selo de Indicação Geográfica. *Revista INGI - Indicação Geográfica e Inovação*, 2(3), 155–163.

Gavasso, M. S. de B., Fernandes, J. da S. G. & Andrade, M. S. de. (2016). Revisão sistemática de estudos sobre habilidades sociais: Avaliação e treinamento. *Ciências & Cognição*, 21(1). Recuperado de <http://cienciasecognicao.org/revista/index.php/cec/article/view/1050>

Google Earth. (2020). Google Earth—Mapa de Aracaju-SE [Georeferência]. Recuperado de Google Earth—Mapa de Aracaju-SE website: <https://www.google.com/maps/@-11.0683,-37.13079,23603m/data=!3m1!1e3>

Gosling, I. T. de S., Rezende, D. C. de, Gosling, M. de S., Barros, F. M. R. & Soares, R. R. (2020). Marketing de Relacionamento em Hotéis Fazenda: Um Framework a partir da Experiência. *Revista acadêmica observatório de inovação do Turista*, 14(1), 81–105.

Junior, L. F., Junior, J. S., & Lobato, F. (2020). Um olhar sobre turismo gastronômico: Um caso no TripAdvisor. *Anais do XVII Encontro Nacional de Inteligência Artificial e Computacional*, 519–530.

KusumasariVer, B. & Prabowo, N. P. A. (2020). Utilising a Data Capture Tool to Populate a Cardiac Rehabilitation Registry: A Feasibility Study. *Heart, Lung and Circulation*, 29(2), 224–232. <https://doi.org/10.1016/j.hlc.2018.12.012>

Lefever, S., Dal, M. & Matthiasdottir, A. (2007). Online data collection in academic research: Advantages and limitations. *British Journal of Educational Technology*, 38(4), 574–582.

López, M. de la O. H. & Amaya, L. F. (2019). What makes (im)politeness for travellers? Spanish tourists' perceptions at national and international hotels. *Journal of Politeness Research*, 15(2), 195–222. <https://doi.org/10.1515/pr-2016-0060>.

Marx, D. V. (2018). *A influência dos meios de comunicação digital no crescimento das empresas*.

Me, M. da E. (2020). *O Coronavírus e seu Impacto Econômico no Brasil*. Ministério da Economia. Recuperado de <https://www.gov.br/economia/pt-br/centrais-de-conteudo/publicacoes/notas-informativas/2020/nota-informativa-coronavirus.pdf>

Mendes, A. M., Tonin, F. S., Buzzi, M. F., Pontarolo, R. & Fernandez-Llimos, F. (2019). Mapping pharmacy journals: A lexicographic analysis. *Research in Social and Administrative Pharmacy*, 15(12), 1464–1471.

Morais, R. S. de, Oliveira, R. A. de, & Paim, A. H. C. (2015). As Pesquisas de Demanda Turística em Minas Gerais: Evolução histórica, principais desafios e perspectivas de futuro. *Revista Turismo em Análise*, 26(1), 21–37.

Oliveira, N. M., Viterbo, J. & Boscaroli, C. (2020). Um estudo sobre apropriação tecnológica no fomento à hospedagem alternativa em cidades de pequeno porte. *Anais do Workshop sobre as Implicações da Computação na Sociedade (WICS)*, 122–129. <https://doi.org/10.5753/wics.2020.11043>

Paes, T. A. A., Ladeira, R. & Larocca, M. T. G. (2020). Destinos turísticos e a sua relação com o processo de cocriação de valor e o marketing de experiência: Uma revisão bibliográfica. *Caderno Virtual de Turismo*, 20(1). <https://doi.org/10.18472/cvt.20n1.2020.1757>

Palazzo, M. (2021). Urban and rural destinations on Instagram: Exploring the influencers' role in #sustainabletourism. *Land Use Policy*, 100, 104915. <https://doi.org/10.1016/j.landusepol.2020.104915>

Pinto, I. & Castro, C. (2019). Online travel agencies: Factors influencing tourists' purchase decisions. *Tourism & Management Studies*, 15(2), 7–20.

Richard, A., Casagrande, M., Jeuffroy, M.-H., & David, C. (2018). An innovative method to assess suitability of Nitrate Directive measures for farm management. *Land Use Policy*, 72, 389–401. Scopus. <https://doi.org/10.1016/j.landusepol.2017.12.059>.

Silva, C. A. da, Baltazar, M. da S., Saragoça, J. M. L., Abrantes, A., Silva, T. F. da, Santos, M. O., Viegas, M. da G. (2018). Uma análise lexicométrica dos documentos oficiais sobre a cooperação transfronteiriça Espanha-Portugal.

Silva, G. L. da, Filho, L. M. & Júnior, S. M. (2019). Análise da Percepção dos Consumidores de Meios de Hospedagem em Relação ao Uso das Online Travel Agencies (OTAs). *Revista Brasileira de Pesquisa em Turismo*, 13(1), 40–57. <https://doi.org/10.7784/rbtur.v13i1.1468>.

Silva, M. B. de O. da, Arruda, D. de O., Souza, Á. G. R. de & Mariani, M. A. P. (2019). Como os turistas percebem os atributos de atrativos turísticos em Bonito (MS)? Uma análise com base em comentários publicados no tripadvisor. *Turismo: Visão e Ação*, 21(2), 150–172.

Silva, J. A. & Santos, C. A. de J. (2015). Análise da competitividade do turismo no município de Aracaju. *Revista de Turismo Contemporâneo*, 3(2). Recuperado de <https://periodicos.ufrn.br/turismocontemporaneo/article/view/6887>.

Soares, S. V., Picolli, I. R. A., & Casagrande, J. L. (2018). Pesquisa Bibliográfica, Pesquisa Bibliométrica, Artigo de Revisão e Ensaio Teórico em Administração e Contabilidade. *Administração: Ensino e Pesquisa*, 19(2), 308–339. <https://doi.org/10.13058/raep.2018.v19n2.970>

Soares, J. C., Hoeckesfeld, L., Sarquis, A. B., Casagrande, J. L., de Lima, M. A. & Cittadin, J. (2018). Análise da popularidade de conteúdo em mídias sociais. *International Journal of Business Marketing*, 3(2), 016–030.

Souza, M. A. R. de, Wall, M. L., Thuler, A. C. de M. C., Lowen, I. M. V. & Peres, A. M. (2018). O uso do software IRaMuTeQ na análise de dados em pesquisas qualitativas. *Revista*

da Escola de Enfermagem da USP, 52(0). <https://doi.org/10.1590/s1980-220x2017015003353>.

Suntikul, W., & Jachna, T. (2016). Profiling the heritage experience in Macao's Historic Center. *International journal of tourism research*, 18(4), 308–318.

Suntikul, W. & Jachna, T. (2016). The co-creation/place attachment nexus. *Tourism Management*, 52, 276–286. <https://doi.org/10.1016/j.tourman.2015.06.026>.

Tavares, F. O. & Brea, J. A. (2020). Análise confirmatória à preferência nos Hostels na cidade do Porto, em Portugal. *Investigaciones Turísticas*, (19), 104. <https://doi.org/10.14198/INTURI2020.19.05>.

Talwar, S., Dhir, A., Kaur, P., & Mäntymäki, M. (2020). Barriers toward purchasing from online travel agencies. *International Journal of Hospitality Management*, 89, 102593.

Talwar, S., Dhir, A., Kaur, P., & Mäntymäki, M. (2020). Barriers toward purchasing from online travel agencies. *International Journal of Hospitality Management*, 89, 102593.

Telles, G. C. de Q., & Cardoso, J. de F. (2019). Turismo colaborativo em hostels Collaborative tourism in hostels Turismo colaborativo en hostels. *Revista Hospitalidade*, 16(2).

Yan, Y., Eckle, M., Kuo, C.-L., Herfort, B., Fan, H. & Zipf, A. (2017). Monitoring and assessing post-disaster tourism recovery using geotagged social media data. *ISPRS International Journal of Geo-Information*, 6(5), 144.

Porcentagem de contribuição de cada autor no manuscrito

Luiz Diego Vidal Santos – 17%

Catuxe Varjão de Santana Oliveira – 16,6%

Francisco Sandro Rodrigues Holanda – 16,6%

Carlos Eduardo Celestino de Andrade – 16,6%

Wilsonita de Melo Ubirajara – 16,6%

Maraiza Santana dos Santos – 16,6%